

DISS. ETH NO. 15675

**Effects of PFT – mediated prenylation on the
gene network in *Arabidopsis thaliana***

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH

for the degree of

DOCTOR of NATURAL SCIENCES

presented by

ANDREAS, FÜRHOLOZ

Dipl. Natw. ETH, ETH Zürich

born 29.08.1967

citizen of Biberist, SO

accepted on the recommendation of

Prof. Dr. W. Grisse, examiner

Prof. Dr. M. Aebi, co-examiner

Prof. Dr. N. Amrhein, co-examiner

2004

Table of contents

	I
Table of contents	
List of figures	V
List of tables	VII
List of abbreviations	VIII
Summary	XII
Zusammenfassung	XIV
1. Introduction	1
1.1 Protein isoprenylation	1
1.2 The discovery of prenylation	3
1.3 Prenyltransferases	5
1.3.1 Reaction path of the protein farnesyltransferase PFT.....	6
1.3.2 Sequence dependence of protein farnesylation.....	7
1.3.3 Evolutionary history.....	9
1.4 Protein prenylation in human: Ras proteins and cancer	9
1.5 Protein prenylation in plants and the Arabidopsis mutant era1-2	11
1.6 Leaf development in Arabidopsis	15
1.7 Gene-expression profiling by using microarray technology	18
1.8 Interactive Metabolic Pathway Map	20
1.9 The aim of this study	22
2. Materials and Methods	23
2.1 Plant material	23
2.2 Gene-expression profiling	23

2.2.1 Sample preparation.....	23
2.2.1.1 Isolation of total RNA.....	23
2.2.1.2 cRNA preparation.....	24
2.2.1.3 Array hybridization, staining and scanning.....	24
2.2.2 Pre-analysis of the expression data.....	25
2.2.2.1 Detection call: absent, marginal and present.....	26
2.2.2.2 Data normalization.....	27
2.2.2.3 Fold change and log transformation.....	28
2.2.3 Data mining: Introduction to bioinformatics.....	29
2.2.3.1 Information extraction by data clustering.....	29
2.2.3.1.1 Hierarchical clustering.....	30
2.2.3.1.2 Self organizing maps.....	31
2.2.3.1.3 Combination of SOM and hierarchical clustering.....	33
2.3 Interactive Metabolic Pathway Map.....	34
2.3.1 Overall design of the metabolic pathway map.....	34
2.3.1.1 Database construction for biochemical pathways in <i>A. thaliana</i>	34
2.3.1.2 Integration of compartment specific gene annotation.....	37
2.3.2 Visualization Software.....	37
2.3.2.1 Outlook.....	40
2.4 Additional databases.....	41
2.4.1 <i>A. thaliana</i> database for regulatory elements.....	41
2.4.2 <i>A. thaliana</i> database for sub-cellular genomes.....	42
2.5 Database information extraction.....	42
2.5.1 Java applications.....	42
2.5.1.1 Expression profile analyzer 1.0.0 for plants.....	43
2.5.1.2 Functional group analyzer.....	44
2.6 Biocomputational approach to screen the Arabidopsis protein-sequence database for putative PFT targets.....	45
2.7 Electronic Northern as a control for expression status of putative PFT targets.....	48
3. Results.....	50
3.1 Kinematics.....	50
3.1.1 Kinematics analysis.....	50
3.1.1.1 Leaf development in WT.....	51
3.1.1.2 Leaf development in <i>era1-2</i>	52
3.2 Gene expression analysis using Affymetrix GeneChips®.....	54

3.2.1 <i>Global analysis of gene expression data</i>	54
3.2.1.1 <i>Cluster analysis of co-regulated genes</i>	55
3.2.1.1.1 <i>Global gene expression analysis during leaf development in WT</i>	56
3.2.1.1.2 <i>Global gene expression analysis during leaf development in era1-2</i>	56
3.2.1.1.3 <i>Differences between era1-2 and WT gene expression profiles</i>	57
3.2.1.2 <i>Analysis of genes involved in leaf development</i>	59
3.2.1.2.1 <i>Genes involved in leaf expansion</i>	59
3.2.1.2.2 <i>Genes involved in cell proliferation and cell expansion</i> ...	60
3.2.2 <i>Gene-specific analysis of gene expression data</i>	61
3.2.2.1 <i>Specific analysis of genes for biochemical pathways</i>	63
3.2.2.1.1 <i>Changes in gene expression at the first transition stage in leaf development</i>	63
3.2.2.1.2 <i>Changes in gene expression at the second transition stage in leaf development</i>	64
3.2.2.1.1 <i>Changes in gene expression at both the first and the second transition stage in leaf development</i>	67
3.2.2.2 <i>Specific analysis of genes for signal transduction pathways</i>	69
3.2.2.3 <i>Genome specific analysis</i>	70
3.3 <i>Biocomputational approach to identify putative PFT targets</i>	71
4. Discussion	73
4.1 <i>An additional proliferation phase is responsible for the increase in mature leaf size in era1-2</i>	73
4.2 <i>Transcriptional control of genes involved in cell cycle regulation in era1-2</i>	73
4.3 <i>Farnesylation is involved in the transcriptional regulation of chloroplast metabolism</i>	75
4.4 <i>Farnesylation leads to redox-mediated control of chloroplast metabolism</i>	75
4.5 <i>Regulation of isoprenoid biosynthesis controls cell cycle progression</i>	76
4.6 <i>Biocomputational approach to identify putative PFT targets</i>	77
4.6.1 <i>Cytochrome b561</i>	77
4.6.2 <i>The transcription factor Hap5a</i>	

5. References.....	82
Acknowledgements.....	103
Curriculum vitae.....	104

List of figures

Figure 1:	The mevalonate dependent isoprenoid pathway (MVA-pathway).....	3
Figure 2:	Mechanism of farnesylation.....	7
Figure 3:	Hierarchical order defined for <i>A. thaliana</i> metabolic database.....	36
Figure 4:	General overview of <i>Arabidopsis</i> metabolic pathways.....	38
Figure 5:	Interconnection view of metabolic pathways within the biochemical network...	38
Figure 6:	Individual pathway view.....	39
Figure 7:	Overview of <i>A. thaliana</i> metabolic pathways with gene expression data.....	40
Figure 8:	Keyword database for <i>A. thaliana</i>	41
Figure 9:	Expression profile analyzer for plants.....	44
Figure 10:	Functional group analyzer for plants.....	45
Figure 11:	Detection call distribution within a WT gene expression profile.....	49
Figure 12:	Kinematics I: The leaf area.....	50
Figure 13a:	Kinematics II: Overview of the leaf development in WT.....	52
Figure 14:	Kinematics II: Overview of the leaf development in <i>era1-2</i>	53
Figure 15:	Raw data analysis in four steps.....	55
Figure 16:	Cluster gram of gene expression data during leaf development.....	58
Figure 17:	Combination of a Venn-diagram cluster gram analysis.....	62
Figure 18:	Venn-diagram of genes involved in metabolic pathways.....	63

Figure 19: Model for the regeneration of reduced ascorbate in the chloroplast..... 78

Figure 20: Hypothetical model which shows the relationship between loss of
farnesylation and enlarged leaf size observed in *era1-2*..... 79

List of tables

Table 1:	Protein prenyltransferases.....	6
Table 2:	Protein prenyltransferase genes in model organisms.....	9
Table 3:	Table for putative target prediction according to Moores <i>et al.</i> , 1991.....	46
Table 4:	Pairs of combined experiments.....	55
Table 5:	Genes involved in leaf expansion.....	59
Table 6:	Genes involved in developmental processes.....	61
Table 7:	Pathway genes regulated during the first transition stage.....	64
Table 8a:	Pathway genes regulated during the second transition stage.....	65
Table 8b:	Pathway genes regulated during the second transition stage.....	66
Table 9:	Pathway genes regulated during the first and the second transition stage.....	67
Table 10:	Regulation of the Calvin cycle, chlorophyll biosynthesis and MEP-pathway....	68
Table 11:	Nuclear encoded genes coding for thioredoxins and thioredoxin-reductase....	69
Table 12:	Plastid encoded genes.....	70
Table 13:	Keyword search for protein involved in the plastidic translational machinery...	71
Table 14:	Putative PFT-targets found by a biocomputational analysis.....	72

List of abbreviations

ABA	Abscisic acid
<i>aba</i>	Abscisic acid deficient mutant
<i>abi</i>	Abscisic acid insensitive mutant
<i>ag</i>	Agamous mutant
<i>an</i>	Angustifolia mutant
ANT	<i>AINTEGUMENTA</i> gene
ASC	Ascorbate
ATP	Adenosine triphosphate
<i>axr</i>	Auxin-resistant mutant
BMP	Bitmap
BR	Brassinosteroid
<i>bri</i>	Brassinolide insensitive mutant
BY-cells	Bright yellow cells
CAB	Chlorophyll a/b-binding protein
CAK	Cyclin dependent kinase activator
<i>cam</i>	Carbohydrate accumulation mutant
CaM53	Calmodulin 53
<i>cbb</i>	Cabbage mutant
CBF	CCAAT binding factor
CDK	Cyclin dependent kinase
cDNA	Complementary deoxyribonucleic acid
ChipID	Affymetrix chip identifier
<i>clf</i>	Curly leaf mutant
cRNA	Complementary ribonucleic acid
cTP	Chloroplast transit peptide

CTP	Cytosine triphosphate
CYCD	Cyclin D
Cyt b561	Cytochrome b561
<i>det</i>	De-etiolated mutant
<i>DIM</i>	<i>DIMINUTO</i> gene
DMAPP	Dimethylallyl-pyrophosphate
DNA	Deoxyribonucleic acid
<i>DWF</i>	<i>DWARF</i> gene
DXP	1-Deoxy-D-xylulose 5-phosphate
<i>era</i>	Enhanced response to ABA mutant
FOM	Figure of merit
FPP	Farnesyl-pyrophosphate
FTA	Protein-farnesyltransferase α -subunit
FTB	Protein-farnesyltransferase β -subunit
GcpE	Hydroxymethylbutenyl diphosphate synthase
GGPP	Geranylgeranyl-pyrophosphate
GGT1A	Protein-geranylgeranyltransferase α -subunit
GGT1B	Protein-geranylgeranyltransferase β -subunit
GGT2A	Ras-protein-geranylgeranyltransferase α -subunit
GGT2B	Ras-protein-geranylgeranyltransferase β -subunit
GPP	Geranyl-pyrophosphate
GPS	Geranyl-pyrophosphate synthase
GTP	Guanine triphosphate
HMG	3-Hydroxy-3-methylglutaryl
HMGR	3-Hydroxy-3-methylglutaryl reductase
hr	Hour

<i>hy</i>	Long-hypocotyl mutant
IPP	Isopentenyl-pyrophosphate
<i>kn1</i>	Knotted1 mutant
<i>kor</i>	Korrigan mutant
LEC	LEAFY COTYLEDON gene
MEP	2-C-Methyl-erythritol 4-phosphate
MM	Miss match
mRNA	Messenger ribonucleic acid
mTP	Mitochondrial transit peptide
MVA	Mevalonic acid
NCBI	National center of biotechnology information
PCR	Polymerase chain reaction
PFT	Protein-farnesyltransferase
PGGT	Protein-geranylgeranyltransferase
<i>phy</i>	Phytochrome mutant
<i>plp</i>	Pluripetala mutant
PM	Perfect match
r	Pearson correlation coefficient
Ras-PGGT	Ras-protein-geranylgeranyltransferase
RBCS	Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit
RNA	Ribonucleic acid
RNAi	Interfering ribonucleic acid
<i>rot</i>	Rotundifolia mutant
RT-PCR	Real time polymerase chain reaction
SAPE	Streptavidin-phycoerythrin
SOM	Self organizing maps

SOTA	Self organizing tree algorithm
SP	Signal peptide
STP	Stunted plant
TAIR	The Arabidopsis Information Resource
T-DNA	Transfer deoxyribonucleic acid
TXT	Text delimited
UTP	Uridine triphosphate
WT	Wild type

Summary

Protein-prenylation is a posttranslational modification, in which isoprenoid moieties are covalently attached to the C-terminal end of a target protein. The covalent attachment of the C-15 farnesylpyrophosphate (FPP) molecule is catalysed by protein-farnesyltransferase (PFT). Protein-prenylation is well studied in the context of cancer therapy. For instance, PFT inhibitors have been shown in mice experiments to antagonize the development of certain types of tumors.

The *Arabidopsis thaliana* mutant *era1-2* (enhanced response to ABA) has a deletion in the gene coding for the β -subunit of PFT. As described by Yalovsky *et al.* (2000), *era1-2* shows different phenotypical features, caused by the lack of farnesylation. One of those is the enlarged size of mature leaves. The aim of my work was to study the relationship between farnesylation and leaf development.

era1-2 was analysed on the cellular level using a kinematics approach as well as on a molecular level by profiling global gene expression. The kinematics approach showed that the enlarged leaf size in *era1-2* is mainly caused by a second proliferation phase after 21 days of development. In agreement with the kinematics study, the highest changes in global gene expression were found during later stages of leaf development. These changes were mainly caused by stronger up- or down-regulation of gene expression in wild type plants (WT) than in *era1-2*. A detailed analysis of gene expression changes revealed that genes coding for proteins involved in cell cycle and cell wall modification processes were up-regulated in *era1-2* compared to WT at 21 days. In addition, genes known to be induced after exogenous brassinosteroid (BR) application were also found to be up-regulated in *era1-2*. This led to the conclusion that cell cycle and cell wall modification could be induced again late in leaf development as the result of continued or increased synthesis of phytosterols in *era1-2* as compared to WT. The loss of farnesylation resulted also in an unscheduled up-regulation of genes for metabolic pathways in the chloroplast in parallel with an

up-regulation of genes coding for chloroplast-targeted thioredoxins. Therefore, we can propose a model that protein-farnesylation could have an important and novel role in the cytosol-chloroplast communication. A specific farnesylated protein could serve as a cytosolic isoprenoid flux sensor and, together with a redox-regulated system, could control the metabolic status in the chloroplast. In case of *era1-2*, our results suggest that the loss of farnesylation would lead to an increased contribution of plastid isoprene intermediate for phytosterol synthesis in the cytoplasm, which could promote unscheduled cell proliferation processes and therefore result in enlarged mature leaves.

Zusammenfassung

Die Prenylierung der Proteine gehört zu den posttranslationalen Modifikationen, wobei Isoprenoidmoleküle an den C-Terminus der zu modifizierenden Proteine kovalent gebunden werden. Die Protein-Farnesyltransferase (PFT) katalysiert die Übertragung des Farnesylrests von Farnesylpyrophosphat (FPP), einer ungesättigten, verzweigten C-15 Verbindung. Prenylierungsprozesse werden seit langem intensiv studiert, denn durch ihre Hemmung konnte man in Mäusen zeigen, dass die Weiterentwicklung von karzinogenen Zellen gestoppt werden konnte.

Die *Arabidopsis thaliana* Mutante *era1-2* (enhanced response to ABA) enthält eine Deletion im Gen, das für die β -Untereinheit von PFT kodiert. Damit ist *era1-2* nicht mehr fähig, Protein zu farnesylieren. Wie bereits von Yalovsky *et al.*, (2000) beschrieben, wurde bei *era1-2* unter anderen phänotypischen Merkmalen eine starke Vergrößerung der ausgewachsenen Blätter festgestellt. Der Zusammenhang zwischen der Farnesylierung und der Blattentwicklung stellt das Hauptaugenmerk meiner Dissertation dar. Die Blattentwicklung wurde sowohl auf zellulärer Ebene, durch eine kinematische Analyse, als auch auf molekularer Ebene durch die Erstellung eines globalen Genexpressionsprofils, untersucht.

Durch die kinematische Analyse konnte gezeigt werden, dass eine zusätzliche Zellteilungsphase in *era1-2* für die Vergrößerung der Blätter verantwortlich ist. Zur selben Zeit wurden auch die meisten Veränderungen im Genexpressionsprofil zwischen *era1-2* und der entsprechenden Wildtypkontrolle (WT) festgestellt. Die Änderungen gehen hauptsächlich auf stärkere Genexpressionsveränderungen innerhalb des WT, im Vergleich zur Mutante, zurück. Somit scheint die Farnesylierung eine grössere Rolle bei späten Entwicklungsprozessen zu spielen.

Im Weiteren wurde eine Akkumulation von mRNA Transkripte, die für Zellzyklusproteine, sowie für Proteine der Zellwandmodifikationen kodieren, gemessen. Ebenfalls konnte gezeigt werden, dass auch die Transkripte von

Brassinosteroid (BR) induzierbaren Genen akkumuliert vorlagen. Dies führt zur Vermutung, dass die zusätzliche Zellproliferationsphase in *era1-2* mit einer möglichen Zunahme der Brassinosteroidsynthese im Zusammenhang steht.

Durch die globale Genanalyse wurde weiter festgestellt, dass eine Mehrheit von plastidären Synthesewegen in *era1-2* induziert wurde. Ebenfalls wurde auch eine Akkumulation der Transkripte für plastidär lokalisierte Thioredoxine beobachtet. Dies führt zum Schluss, dass der Chloroplastenmetabolismus ausgehend von der Farnesylierung über die Redoxkontrolle gesteuert wird. Das Signal zur Kommunikation wird möglicherweise über Cytochrom b561, das eine mögliche Farnesylierungsseite enthält, übermittelt. Vorläufermoleküle (Isopentenylpyrophosphat IPP) für die Phytosterolsynthese werden sowohl im Zytoplasma als auch im Chloroplasten hergestellt (Laule et al., 2003). Da nun der IPP Beitrag aus den Chloroplasten wesentlich gesteigert zu sein scheint, könnte dies zur Akkumulation von Phytosterolen, insbesondere von BRs führen, die ihrerseits zur Induktion einer zusätzlichen Zellproliferationsphase führen könnte. Dies stellt ein hypothetisches Modell zur Erklärung der Blattvergrößerung in der Mutante *era1-2* dar.

1. Introduction

1.1 Protein isoprenylation

Proteins make up most of the dry mass of a cell, and they play a predominant part in most biological processes. Proteins have precisely engineered moving parts whose mechanical actions are coupled to chemical events. It is this coupling of chemistry and movement that gives proteins the extraordinary capabilities that underlie all the dynamic processes in living cells. Proteins are further involved in a wide array of different functions: as chemical catalysts, as sophisticated transducers of motions, as components of multisubunit protein machines and as signal integrators. Covalent modifications of proteins are known to occur as part of important signaling processes. Posttranslational modifications are known to alter markedly the conformation and function of a particular protein. Proteolytic cleavage, glycosylation and phosphorylation are well established mechanisms for regulating the activity and the cellular localization of proteins. Since years, increased interest has focused on lipid modification of proteins. Lipid modifications, in conjunction with other posttranslational changes, often cause proteins to undergo extensive intracellular translocation.

Four major classes of covalently-bound lipids are known: N-myristoylation is the linkage of myristic acid by an amide bond to a N-terminal amino acid residue of a protein. Palmitoylation refers to the thioester bonding of palmitic acid or another fatty acid to a cysteine residue specifically located within a protein. The C-terminus of a glycosylphosphatidylinositol anchored protein is linked through ethanolamine phosphate and a linear chain of four sugars to inositol phospholipids firmly embedded in the external face of the plasma membrane. Protein prenylation is the posttranslational attachment of an isoprenoid moiety via a thioether linkage (-C-S-C-) to a cysteine at or near the carboxy terminus of a protein (reviewed by Thompson and Okuyama, 2000).

Isoprenoids constitute an array of compounds formed from isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), which contain

the basic five-carbon building block termed an isoprenoid unit. Isoprenoid moieties transferred to proteins are derived from the isoprenoid biosynthetic pathway which is also called the mevalonic acid (MVA) pathway, named after the product of the first committed step in the pathway (Figure 1). During early steps of the MVA pathway, sequential condensation of three acetyl-CoA units generates 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA). HMG-CoA is converted to mevalonate, a 6-carbon molecule, in an irreversible reaction catalyzed by HMG-CoA reductase (HMGR). Because of its irreversible nature, this reaction is a rate-limiting step in the isoprenoid pathway (Goldstein and Brown, 1990; Chappell, 1995). It has been shown, at least in mammalian systems, that this is the major regulatory point for biosynthesis of cholesterol, an end product of the pathway. IPP is formed by two successive phosphorylations of mevalonate with ATP followed by an ATP-dependent dehydration and decarboxylation. The condensation of one IPP and one DMAPP in a head-to-tail reaction produces the 10-carbon geranyl pyrophosphate (GPP) and is catalyzed by the enzyme geranyl pyrophosphate synthase (GPS). The reaction is initiated by the loss of pyrophosphate from DMAPP, which generates a resonance-stabilized allylic carbonium ion. Attack of the π -bond of IPP generates a tertiary carbonium ion. Stereospecific loss of a proton generates the product GPP (Chappell, 1995). These polymerization reactions, unlike the more general nucleophilic condensations in other biosynthetic pathways, are electrophilic in nature.

The addition of another IPP molecule to GPP generates farnesyl pyrophosphate (FPP). Additional polymerization reactions give rise to geranylgeranyl pyrophosphate (GGPP) and other isoprenoids. FPP and GGPP are two intermediate isoprenoid compounds known to be transferred to proteins.

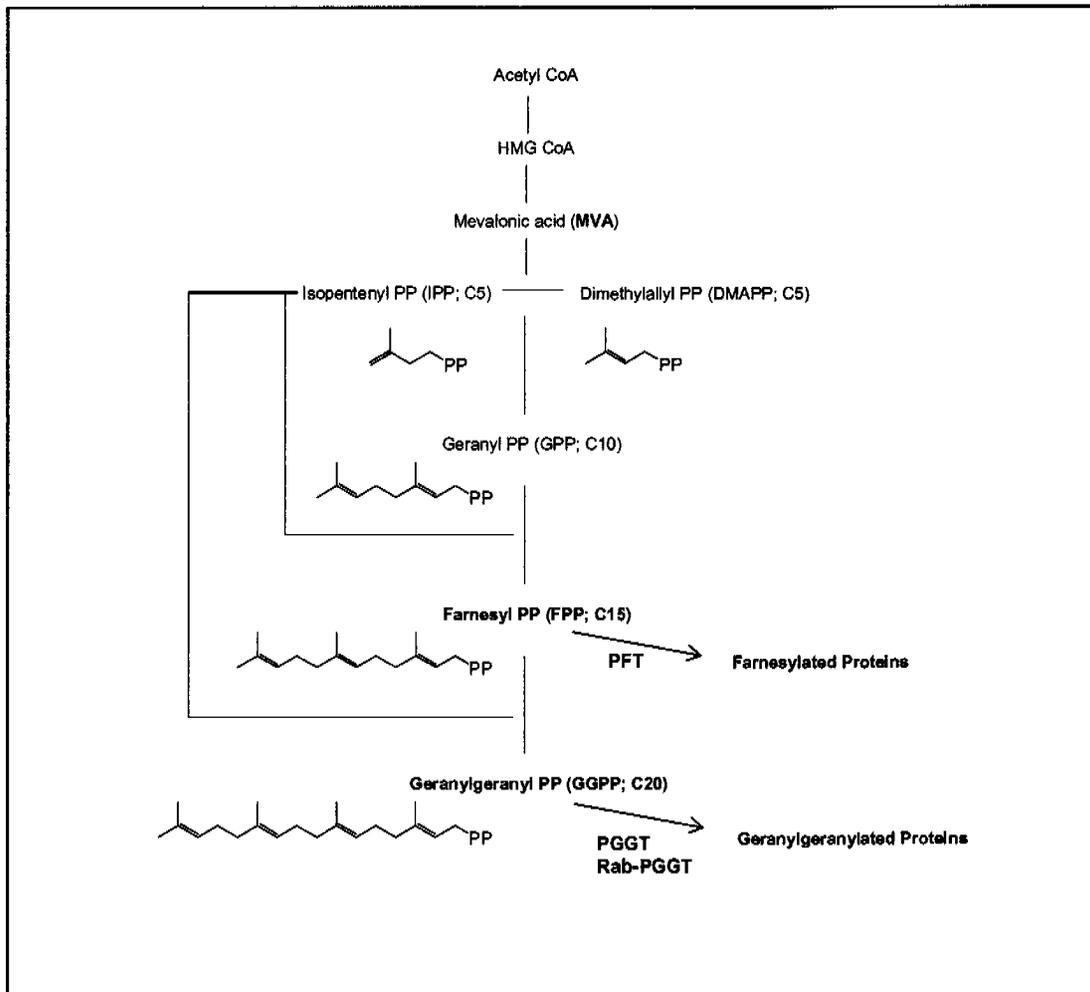


Figure 1: The mevalonate dependent isoprenoid pathway (MVA-pathway)

PFT: protein farnesyltransferase; PGGT: protein geranylgeranyltransferase; Rab-PGGT: Rab-protein geranylgeranyl transferase.

1.2 The discovery of prenylation

The first evidence that isoprenoids can modify polypeptides came from studies in the late 1970s and early 1980s on the structures of certain fungal mating factors. These mating factors were found to consist of a short peptide terminating in a cysteine residue to which a farnesyl group or, in one instance, an oxidized farnesyl group, was covalently linked (Kamiya *et al.*, 1978; Ishibashi *et al.*, 1984).

The discovery that mammalian proteins could be subject to prenylation took much longer. The elucidation of the cholesterol biosynthesis pathway laid the

groundwork for this discovery. The synthesis of mevalonate by the enzyme HMG-CoA reductase (HMGR) is the committed step in cholesterol formation. Identification of compactin, a specific inhibitor of HMGR, allowed experiments to study the metabolism of mevalonate in cells (Endo, 1992). In a series of studies in the early 1980s, a product of mevalonate metabolism other than cholesterol was found to be required for entry of cells into the S phase of the cell cycle (Schmidt *et al.*, 1982; Fairbanks *et al.*, 1984). The search for the required metabolite of mevalonate, which involved following the fate of exogenously added [³H]mevalonate to compactin treated cells, revealed that metabolites of mevalonic acid are incorporated into proteins (Schmidt *et al.*, 1984; Sinensky and Logel, 1985; Maltese and Sheridan, 1987). Subsequently, one of the proteins was identified as the nuclear envelope protein lamin B (Beck *et al.*, 1988; Wolda and Glomset, 1988), and the modifying species was shown to be a farnesyl group (Farnsworth *et al.*, 1989). At about the same time, independent studies demonstrated that Ras proteins are farnesylated (Hancock *et al.*, 1989; Schafer *et al.*, 1989; Casey *et al.*, 1989). Ras proteins are small GTP-binding proteins that play crucial roles in signaling pathways controlling cell growth and differentiation (Burgering and Bos, 1995; Boguski and McCormick, 1993). Yeast genetics provided important clues for the discovery that Ras is modified by an isoprenoid lipid. Genes required for the posttranslational maturation of both Ras and the peptide-mating pheromone a-factor in *S. cerevisiae* was discovered, indicating that both polypeptides are processed via a common route (Schafer and Rine, 1992; Powers *et al.*, 1986). The finding that the a-factor contained a farnesylated cysteine at its carboxyl terminus (Anderegg *et al.*, 1988) prompted investigators to examine whether Ras proteins were also subject to this modification. This search led to the discovery that Ras proteins are farnesylated and that the modification is required for oncogenic forms of Ras to transform cells (Hancock *et al.*, 1989; Schafer *et al.*, 1989; Casey *et al.*, 1989). These studies dramatically stimulated interest in the field of protein prenylation (Gibbs, 1991).

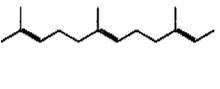
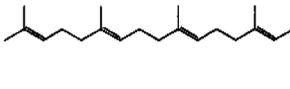
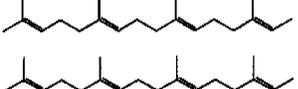
The coding sequences of lamin B, Ras proteins and a-factors indicated that the C-termini of all of these proteins contain a cysteine residue three amino acids

proximal to the carboxyl-terminus (Farnsworth *et al.*, 1989; Powers *et al.*, 1986; Stimmel *et al.*, 1990), the CaaX motif (C: cysteine, a: aliphatic amino acid, X: any amino acid). The identification of the CaaX motif as the site of prenylation united the field from lower eukaryotes to mammalian systems (Clarke, 1992; Schafer and Rine, 1992). Searches of sequence databases revealed that a variety of proteins contain the CaaX motif (Clarke, 1992; Cox and Der, 1992). Comparison of the CaaX motif with the C-termini of mature proteins also indicated that the processing of CaaX-containing proteins consists of at least three steps because, in addition to being prenylated, the mature proteins lacked the three C-terminal residues (-aaX) and contained a carboxymethyl group on the prenylcysteine (Clarke, 1992).

1.3 Prenyltransferases

Prenylated proteins can be grouped into two major classes: those containing the CaaX motif and those containing the CC- or CxC motif (Glomset and Farnsworth, 1994). The former class contains a diverse group of proteins (Clarke, 1992; Cox and Der, 1992), whereas the latter is almost exclusively composed of members of the Rab family of small GTP-binding proteins that participate in intracellular membrane trafficking (Novick and Brennwald, 1993). Three known enzymes catalyze isoprenoid addition to proteins (Table 1): the protein farnesyltransferase (PFT), the protein geranylgeranyltransferase type I (PGGT) and the protein geranylgeranyltransferase type II (Ras-PGGT).

PFT and PGGT are collectively known as CaaX prenyltransferases, which attach their respective isoprenoid to the cysteine residue of a C-terminal CaaX motif. Rab-PGGT attaches the geranylgeranyl group to two C-terminal cysteine residues in the Rab family of Ras-related GTPases. All three enzymes are heterodimers, but only PFT and PGGT share a common α subunit.

Protein prenyltransferases			
	Protein farnesyltransferase (PFT)	Protein geranylgeranyltransferase (PGGT1)	Protein geranylgeranyltransferase 2 (Ras-PGGT)
Gene name (α -subunit)	FTA	FTA (GGT1A)	GGT2A
Gene name (β -subunit)	FTB	GGT1B	GGT2B
Substrate motif	Carboxy-terminal CaaX-box	Carboxy-terminal CaaX-box	Carboxy-terminal -CC, -CXC, -CCX, -CCXX, -CCXXX, -CXXX
Lipid anchor	 Farnesyl	 Geranylgeranyl	 Geranylgeranyl or two moieties per substrate

1.3.1 Reaction path of the protein farnesyltransferase PFT

PFT is a Zn^{2+} -dependent prenyltransferase containing an α and β subunit, which catalyzes the farnesylation on a C-terminal CaaX motif of a protein. In a first step, FPP binds to the active site of PFT followed by the CaaX peptide substrate. This leads to the formation of the ternary reaction complex. The two reactants are brought into juxtaposition not by major protein conformational changes, but by a rotation involving the first and the second isoprenoid units (atoms C1 to C10) of the farnesyl moiety, resulting in a movement of the C1 atom toward the cysteine residue. The Zn^{2+} -activated thiolate of the cysteine orients a free pair of electrons for in-line nucleophilic attack on the C1. It is still unknown, whether the reaction follows a nucleophilic substitution mechanism S_N1 or S_N2 (S_N1 : unimolecular reaction, 1st order kinetics; S_N2 : bimolecular reaction; 2nd order kinetics). Mg^{2+} , which is also required for optimal enzymatic activity, may coordinate the diphosphate moiety of FPP. The three C-terminal amino acids of the farnesylated peptide product make extensive van der Waals contacts with the FPP molecule, suggesting that the amino-acid sequence of the CaaX motif may modulate product release. Furthermore, the PFT releases the farnesylated protein product only when a new molecule of FPP binds to its active site. Since the latter may be substantially membrane bound under intracellular conditions, release of the farnesylated precursor protein from the PFT may occur at a membrane surface rather than in the cytoplasm. Following

initial farnesylation, the C-termini of the CaaX-box proteins are successively processed by two additional reactions. The terminal –aaX residues are first removed by an endopeptidase, an integral membrane protein localized to the endoplasmatic reticulum. Following removal of their terminal –aaX sequence, proteins are carboxyl-methylated on their farnesylated cysteine residue (Figure 2).

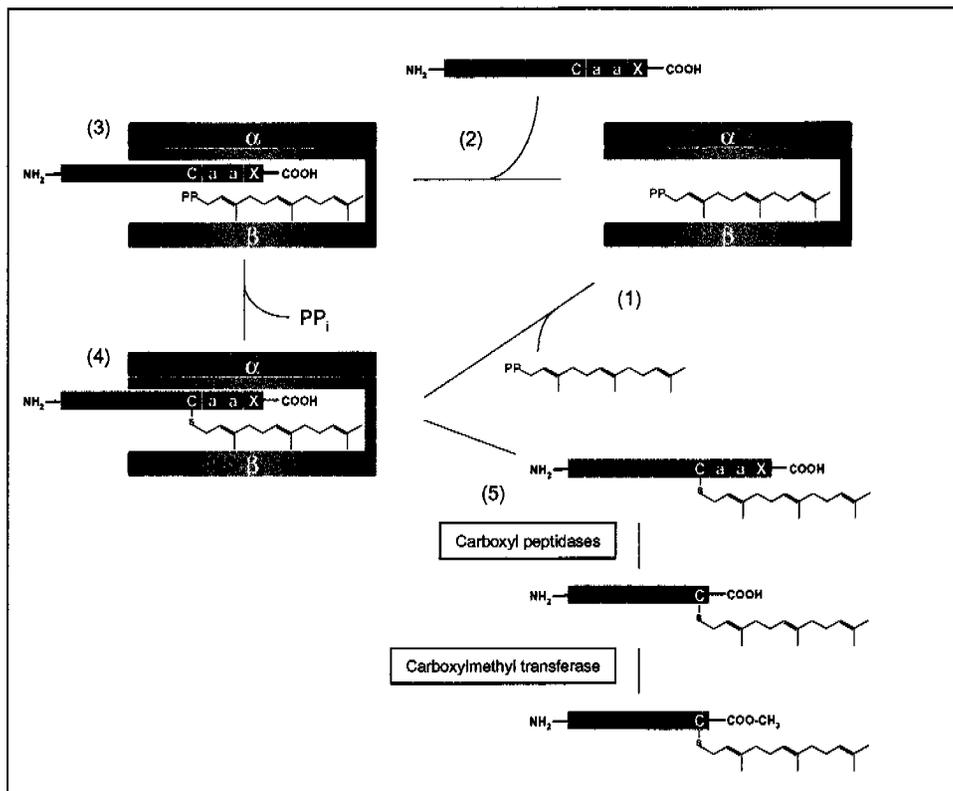


Figure 2: Mechanism of farnesylation

FPP binds to PFT (1) followed by the binding of the CaaX peptide substrate (2) leading in the ternary reaction complex (3). Farnesyl will be attached to the peptide substrate by releasing the pyrophosphate (4). After the attachment of a new FPP, the farnesylated peptide product will dissociate (5) and finally be processed (Long *et al.*, 2002).

1.3.2 Sequence dependence of protein farnesylation

As already described, the C-terminal amino acid sequence of a protein that can be farnesylated is referred to a Ca₁a₂X box: C: cysteine, a₁a₂: aliphatic amino acids X: any amino acid. The acronym is hardly a rigid rule; non-aliphatic residues, for example, are often found within the CaaX sequence of proteins that are isoprenylated, and proteins having CaaX-like sequences are not always

isoprenylated. More than 10 years ago, Moores *et al.* (1991) evaluated the determinants of the CaaX box by introducing mutations at the final three positions of the Ras-CaaX protein. The CaaX variants were scored for their ability to act as protein substrate in prenyltransferase assays. The transferase activity was normalized to that of a control protein substrate (Ras-CVLS for farnesylation).

Basic and aromatic side chains were found to be tolerated at a_1 , but not at a_2 . Proline was tolerated at both a_1 and a_2 but appeared to have a more deleterious effect at a_1 . Changes in the X position altered both, prenylation efficiency and specificity. The effect of amino acid substitutions at the X position can be divided into two groups: substrates with a high normalized activity (glycine, alanine, serine, threonine, histidine, asparagine, glutamine, cysteine, valine, isoleucine, phenylalanine and methionine) and substrates with very low activity (proline, tyrosine, tryptophan, lysine, arginine, aspartic acid, glutamic acid). Detailed kinetic determinations with purified protein farnesyltransferases combined with *in vivo* isoprenoid identification would be required to clarify whether the different CaaX sequences have absolute specificity for isoprenylation under physiological conditions and substrate concentrations. Tetrapeptide competition experiments combined with a prenylation assay indicated that the CaaX box residues by themselves appear to contain all of the critical determinants for effective interaction between enzyme and protein substrate, with the amino acid in the last position of CaaX being an important recognition site for isoprenylation. However, Moores *et al.* (1991) could not rule out that amino acids outside the CaaX may also influence isoprenoid reactions.

1.3.3 Evolutionary history

Protein prenyltransferases are currently known only in eukaryotes, but they are widespread, being found in vertebrates, insects, nematodes, protozoa, fungi and plants. The chromosomal locations and number of exons of prenyltransferase genes in the major eukaryotic model organisms are shown in Table 2. All organisms show a similar structure in gene organization. The α subunits of PFT and PGGT are encoded by the same gene (FTA, GGT1A). All the other subunits (FTB, GGT1B, GGT2A and GGT2B) are each

Organism	Subunit	Chromosome	Number of exons
<i>Mus musculus</i>	FTA (GGT1A)	8	9
	FTB	12	6
	GGT1B	18	9
	GGT2A	14	15
<i>Drosophila melanogaster</i>	GGT2B	3	8
	FTA (GGT1A)	2	3
	FTB	3	4
	GGT1B	2	4
<i>Caenorhabditis elegans</i>	GGT2A	3	1
	GGT2B	2	2
	FTA (GGT1A)	4	4
	FTB	5	6
<i>Arabidopsis thaliana</i>	GGT1B	2	6
	GGT2A	4	9
	GGT2B	3	6
	FTA (GGT1A)	3	5
<i>Saccharomyces cerevisiae</i>	FTB	5	14
	GGT1B	2	11
	GGT2A	4	8
	GGT2B	3	9
<i>Saccharomyces cerevisiae</i>	FTA (GGT1A)	11	1
	FTB	4	1
	GGT1B	7	1
	GGT2A	10	1
	GGT2B	16	1

encoded by a single gene. The α and β subunits have different folds, thus are unlikely to have arisen from a common ancestor. Molecular phylogenetic analysis has shown that orthologous proteins in different species are more closely related to each other than to their paralogs in the same species; the relationships between the β subunits are not fully clear (Andres *et al.*, 1993; Dhawan *et al.*, 1998).

1.4 Protein prenylation in human: Ras proteins and cancer

In human cancers, mutated Ras genes are one of the most frequently identified classes of oncogenes. Mutated Ras proteins have been found in many human cancer tissues including bladder, breast, stomach, colon, liver, lung, kidney, ovary, and pancreas (Barbacid, 1987). The Ras and Ras-related proteins are ubiquitous in eukaryotes. They are small GTP binding proteins (20-35 kD) termed the Ras super family. Most known prenylated proteins are members of this super family. On the basis of amino acid sequence similarity, over 80 Ras-

related proteins have been divided into four subfamilies: Ras, Rho (Rho and Rac), Rab/Ypt, and Ran (Feig, 1994; Sano and Ohashi, 1995). They are involved in the control of a wide variety of cellular functions such as cytoskeletal organization of polymerized actin to produce stress fibers (Rho) (Ridley and Hall, 1992), regulation of intracellular vesicular transport (Rab/Ypt) (Novick and Brennwald, 1993), mediation of protein import into the nucleus (Ran) (Moore and Blobel, 1993), and control of the cell cycle (Ras) and cell polarity (Rho) (Khosravi-Far and Der, 1994; Khosravi-Far *et al.*, 1991).

Although geranylgeranylation is the predominant form of prenylation for the Ras-related proteins including the Rho and Rab/Ypt families, farnesylation is required for the functional Ras proteins. To date, Ras proteins are the best characterized class of prenylated proteins and serve as model systems in studying protein prenylation.

The finding that farnesylation is required for the function of Ras led to the awareness that, as first proposed by Schafer *et al.* (1989), inhibition of Ras farnesylation could be a novel anti-cancer approach. Ever since then, the development of protein farnesylation inhibitors as potential anti-cancer drugs has been a focus for scientists' world-wide (Khosravi-Far and Der, 1994; Kohl *et al.*, 1995; Gibbs *et al.*, 1994). Ras proteins are key elements in signal transduction pathways associated with cell proliferation and differentiation. Oncogenic Ras proteins constitutively induce unrestricted cell growth and division. However, the oncogenic Ras loses its ability to transform cells if it is not farnesylated (Powers *et al.*, 1986). A large number of inhibitors that suppress the activity of PFT have been chemically synthesized. Based on the design template, they can be divided into 3 groups: FPP analogs, CaaX peptide analogs, and bisubstrate analogs (reviewed in Zhang and Casey, 1996). In addition, several microbial products have been identified that block farnesylation, such as manumycin and pepticinnamin from *Streptomyces sp.* and chaetomelic acid from *Chaetomela acutisetata* (Hara *et al.*, 1993; Tamanoi, 1993).

1.5 Protein prenylation in plants and the *Arabidopsis* mutant *era1-2*

Plant protein isoprenylation was first suggested to occur in etiolated radish seedlings (Bach, 1987) and spinach (Epstein *et al.*, 1991). Subsequently, protein isoprenylation in suspension-cultured tobacco cells was described (Randall *et al.*, 1993). Numerous proteins with covalent isoprenoid modifications were shown to exist in cultured tobacco cells by [¹⁴C]mevalonate-labeling. Most of these proteins were associated with the membrane fraction. [³H]Mevalonate-labeling of spinach leaves also provided evidence for plant protein isoprenylation (Swiezewska *et al.*, 1993; Shipton *et al.*, 1995). Many studies established the existence of isoprenylated plant proteins (Lorraine *et al.*, 1996) and provided evidence that protein-linked isoprenoids are derived from the mevalonate pathway (MVA-pathway). Currently, there is no evidence that the recently discovered alternative isoprenoid pathway, the non-mevalonate pathway (MEP-pathway), contributes to protein isoprenylation in plants (Lichtenthaler *et al.*, 1997a, 1997b).

Isoprenylated plant proteins are generally localized to intracellular membranes, although some appear to reside in the cytosolic fraction. Interestingly, several reports suggest that isoprenylated plant proteins might also be found in nuclei, mitochondria and plastids (Schultz *et al.*, 1981; Zhu *et al.*, 1993; Parmryd *et al.*, 1997; Rodriguez-Concepcion *et al.*, 1999).

In plants, as in mammals, protein isoprenylation is involved in cell cycle control and therefore in the control of developmental processes (see review Galichet and Gruissem, 2003). Current evidence for the involvement of isoprenylated protein(s) in plant cell cycle regulation is based on observations using pharmacological inhibitors of protein isoprenylation. For example, inhibitors of protein isoprenylation (e.g. lovastatin and perillyl alcohol) were shown to arrest the growth of tobacco BY-2 cultures, but only when added prior to or during the cell division phase of the culture, which was determined by [³H]thymidine-labeling (Morehead *et al.*, 1995). Interestingly, protein farnesyltransferase activity was found to be maximal during this important phase of culture growth (Morehead *et al.*, 1995). Subsequently, Qian *et al.* (1996) demonstrated that

protein farnesylation is required for cell cycle progression in synchronized tobacco BY-2 cells using manumycin, a highly specific fungal inhibitor of protein farnesyltransferase. Transcripts encoding PFT α - and β -subunits were also shown to be expressed predominantly in pea tissues undergoing active cell division, such as root tips, nodules, and floral buds (Qian *et al.*, 1996).

Protein isoprenylation is also required for normal ABA signaling and seed dormancy in *Arabidopsis thaliana* (Cutler *et al.*, 1996; Pei *et al.*, 1998). *Arabidopsis* mutants with an enhanced response to ABA were identified by screening for mutagenized seeds that failed to germinate in the presence of ABA at a concentration that does not inhibit germination of wild type seeds (Cutler *et al.*, 1996). Ultimately, five recessive mutants representing three complementation groups (*era1*, *era2*, and *era3*) were identified. One complementation group (*era1*) consisted of a T-DNA tagged mutant (*era1-1*; ecotype Wassilewskija) and two fast neutron mutants (*era1-2* and *era1-3*; ecotype Columbia). The T-DNA insert of *era1-1* allowed the cloning of genomic DNA sequences flanking the insert, which were used to identify wild type cDNA and genomic clones of *ERA1* (Cutler *et al.*, 1996). Southern blot analysis using an *ERA1* cDNA probe confirmed the disruption of the *ERA1* gene in all three *era1* mutants. Interestingly, the deduced amino acid sequence of the *ERA1* protein product exhibited significant relatedness to plant, mammalian, and yeast protein farnesyltransferase β -subunits and, unlike wild type plants, floral extracts from *era1* mutant plants lacked protein farnesyltransferase activity. Consistent with a role in seed development, *ERA1* mRNA was detected in wild type floral buds. All three *era1* mutants exhibited poor germination in the absence of a germination stimulus, but were capable of breaking dormancy following a germination stimulus (i.e. a vernalization treatment), demonstrating that the *era1* mutation increases seed dormancy. Furthermore, double mutants defective at the *era1* locus and the *aba1-1* (ABA biosynthesis), *abi1-1* (ABA sensitivity), or *abi3-6* (ABA sensitivity) locus exhibited decreased dormancy relative to *era1*, suggesting that the extreme dormancy of *era1* mutants is dependent on ABA biosynthesis and perception. The ABA hypersensitivity of

era1 null mutants strongly hints to one or more farnesylated proteins in negative regulation of ABA signaling.

Stomatal function in *Arabidopsis* is also regulated by protein farnesylation (Pei *et al.*, 1998). The *era1*-associated ABA hypersensitivity also profoundly affects stomatal function (Pei *et al.*, 1998). ABA normally regulates the activity of guard cell ion channels and, consequently, governs transpirational water loss by promoting stomatal closure. Pei *et al.* (1998) demonstrated a marked increase in ABA-stimulated anion currents in guard cells from *era1* versus wild type plants using whole cell patch-clamp recordings. However, basal anion channel activity in the absence of ABA was not affected by the *era1* mutation. This effect on guard cell anion channel activation correlated with the effects of the *era1* mutation on ABA-induced stomatal closure: specifically, apertures of *era1* stomates were significantly smaller than those of wild type stomates following ABA treatment. In addition, *era1* plants were better at retaining water under drought conditions than wild type plants, suggesting that protein farnesyltransferase may represent an excellent target for introducing drought tolerance into genetically modified plants. These effects on ABA-stimulated anion currents and stomatal function were confirmed using wild type guard cells treated with competitive inhibitors of protein farnesyltransferase (α -hydroxyfarnesylphosphonic acid and manumycin). In addition, the *era1* lesion was shown to restore ABA-stimulated anion currents and stomatal function to *abi1* and *abi2* mutants, which are ABA insensitive in the absence of *era1* due to defects in related type 2C protein phosphatases (Leung *et al.*, 1994; Leung *et al.*, 1997). Thus, *era1* mutations are epistatic to *abi1* and *abi2*, providing evidence that *era1* functions in the ABA signal transduction pathway in concert with, or downstream of, *abi1* and *abi2* (Nambara and McCourt, 1999; Pei *et al.*, 1998). Current models predict that one or more farnesylated proteins negatively regulate ABA signaling and that protein dephosphorylation inhibits regulator function (Nambara and McCourt, 1999; Cutler *et al.*, 1996; Pei *et al.*, 1998). Protein isoprenylation is a regulated process in plants. For example, PFT activity has been linked to cell division activity in tobacco BY2 cultures (Qian *et al.*, 1996; Morehead *et al.*, 1995). In addition, Calmodulin 53 (CaM53)

localization has been shown to be related to the metabolic status of the plant cell (Yalovsky *et al.*, 1999; Rodríguez-Concepción *et al.*, 1999). While CaM53 was found to be plasma membrane-localized in leaf explants incubated for 3 days in the light, the protein was nuclear-localized in explants incubated for the same time in the dark. Dark-induced nuclear localization of CaM53 was reversed by sucrose, suggesting that the isoprenylation, localization, and function of CaM53 are controlled by the metabolic status of the cell (i.e. light and/or sugar promote the isoprenylation and plasma membrane localization of CaM53). Given the data described above that indicate negative control of protein farnesyltransferase β -subunit expression by light and sugar (Yang *et al.*, 1993; Zhou *et al.*, 1997), regulation of plant protein isoprenylation by environmental and developmental stimuli is certain to become a complex story. Regulation of protein isoprenyltransferase gene expression, protein isoprenyltransferase activity (e.g. possible activation of PFT activity by α -subunit phosphorylation (Goalstone *et al.*, 1998)), and availability of protein and isoprenoid substrates (e.g. regulation of HMGR by SNF-1 kinases, etc.) are likely to influence the isoprenylation, targeting, and function of many plant proteins.

In addition, there is increasing evidence that PFT is required for the regulation of plant growth and development (Yalovsky *et al.*, 2000). The developmental and growth phenotype observed in *era1-2* suggest that PFT acts positively or negatively on proteins that control specific processes such as meristem cell division activity and size, organ size and numbers and time to flowering, although most of the proteins that are probably farnesylated by PFT are unknown (Rodríguez-Concepción *et al.*, 1999). The most obviously phenotypical alterations are the larger rosette leaves, more flowers and more sepals, petals and carpels per flower. The average size of *era1-2* rosette leaves is 2.5-fold that of comparable wild-type leaves. To understand the role of PFT during induction of flower development, the time needed for transition from vegetative to reproductive growth was measured. Under long-day growth conditions, *era1-2* plants bolted after 25 to 26 days, whereas wild-type plants did so after 20 to 21 days. These data led to the conclusion that PFT acts on

positive regulators during the normal transition to flowering in *Arabidopsis* (Yalovsky *et al.*, 2000).

The *Arabidopsis* mutant *PLURIPETALA* (*plp*) which contains lesions in the alpha-subunit of both the PFT and PGGT has been characterized recently (Running *et al.*, 2004). *plp* shows some characteristics similar to, but more severe than *era1-2*, as well as additional phenotypical features, as a consequence of its function in both farnesyltransferase and geranylgeranyltransferase-I. The complete loss of protein-prenylation in *plp* strongly decreases growth, as judged by the extremely small plant size in contrast to *era1-2*. *plp* is used as a unique *in vivo* system to study the role of prenylation in eukaryotes.

1.6 Leaf development in *Arabidopsis*

As a consequence of the enlarged leaf size described for *era1-2* (Rodríguez-Concepción *et al.*, 1999), the study of leaf development became interesting to our work.

Leaf organogenesis is usually explained from the perspective of cell theory. According to this theory, the cell is the basic unit of a multicellular organism; therefore, the unit of organogenesis or morphogenesis should be the cell.

In leaf morphogenesis, the control of cell proliferation seems to be related to the control of cell size (Meijer and Murray, 2001; Tsukaya, 2002). There may be a relationship between the control of leaf shape and the control of leaf-cell behavior. In addition to the coordination of leaf cell behavior at the organ level, leaf size also appears to be regulated by coordination at the whole-plant level. Several studies suggest that leaves do not reach their full growth potential, that is, that leaves have the potential to become larger than typically seen under standard conditions. Moreover, leaf development is complex in dicotyledonous plants, in which cell division and cell elongation occur simultaneously in the same region throughout leaf expansion (Maksymowych, 1963; Poethig and Sussex, 1985).

The details in leaf development still remain unclear, due to the complex processes in which cell proliferation and elongation as well as actions of plant hormones is involved.

Prior to a description of heterophylly by Röbbelen (1957), studies of leaf development in *Arabidopsis* had been very limited. Since then many mutants have been isolated with alterations in leaf morphology (e.g., Rédei, 1962; Lee-Chen and Steinitz-Sears, 1967; Barabas and Rédei, 1971). Early anatomical analyses of leaf development were reported for several other species, such as tobacco, *Xanthium*, and *Phaseolus* (for reviews, see Maksymowych, 1963; Marx, 1983; Dale, 1988) and, until the mid 1990s, the genetic and molecular analysis of leaf development was restricted to analysis of the function of *Knotted1 (kn1)* gene in maize (*Zea mays* L.; Smith and Hake, 1992, 1993).

The *angustifolia (an)* mutant of *Arabidopsis* was isolated as a mutant with narrower and thicker leaves than the wild type (Rédei, 1962). The phenotype of the *an* mutant is specific to leaves and floral organs (modified leaves), and it is caused not by a reduction in cell number but rather by a specific defect in the elongation of cells in the transverse (leaf-width) direction of the leaf (Tsukaya *et al.*, 1994; Tsuge *et al.*, 1996). This polar defect in cell elongation in the *an* mutant was observed in all leaf cells examined, including epidermal cells, trichomes, and parenchymatous cells of the leaf. The *AN* gene is thought to be a key gene in regulation of the polar elongation of leaf cells in the leaf-width direction specifically (Tsuge *et al.*, 1996). The interpretation of the *an* mutation led to the postulate of the existence of another kind of mutant in leaf morphology, namely, a mutant with a polar defect in cell elongation in the longitudinal (leaf-length) direction of the leaf. Such a mutant was identified as the *rotundifolia3-1 (rot3-1)* mutant (Tsuge *et al.*, 1996), which had a defect in the elongation of leaf cells in the leaf-length direction, without any change in the normal number of cells. The phenotype was apparent only in leaves and floral organs. Thus, the *ROT3* gene appears to be the key gene that regulates the elongation of leaf cells in the leaf-length direction. The *an rot3-1* double mutant had an additive phenotype, suggesting that the two genes act independently (Tsuge *et al.*, 1996).

All known mutants of *Arabidopsis* with mutations related to the actions of brassinosteroids (BR) develop smaller leaves than the wild type. It is believed that the main role of BR is to stimulate the elongation of cells (Altmann, 1998). The *brassinolide insensitive 1 (bri1)* mutant (Li and Chory, 1997) has a defect in the perception of brassinosteroid and the *cabbage1 (cbb1 = dwf1-6; dim)*, *cbb2*, *cbb3*, *constitutive photomorphogenesis and dwarfism (cpd)*, *de-etiolated 2 (det2)*, *dwarf4 (dwf4)*, and *dwf5* mutants respectively, all have defects in the biosynthesis of brassinosteroids (Feldmann *et al.*, 1989; Takahashi *et al.*, 1995; Kauschmann *et al.*, 1996; Szekeres *et al.*, 1996; Fujioka *et al.*, 1997; Azpiroz *et al.*, 1998; Cho *et al.*, 2000). Several steps in the biosynthesis of BR are catalyzed by cytochrome P450-dependent enzymes, which are homologous to ROT3 in terms of amino acid sequence. The *DET2* gene is thought to regulate the level of expression of the *KORRIGAN (KOR)* gene, which encodes a plasma-membrane-bound endo- 1,4- β -D-glucanase that is essential for the initiation of cell expansion in *Arabidopsis* (Nicol *et al.*, 1998). The *kor* mutant develops smaller leaves than the wild type. Evidence is also beginning to accumulate suggesting that brassinosteroid might, in addition, be involved in the proliferation of cells (Hu *et al.*, 2000).

Two types of histological defects in leaves have been reported in *auxin-resistant (axr)* mutants of *Arabidopsis*. The small leaves and short inflorescence stems of *axr1* plants, which have a mutation in a gene for ubiquitin activating enzyme E1 (Leyser *et al.*, 1993), are attributed to decreases in the numbers of cells that make up these organs (Lincoln *et al.*, 1994). By contrast, in the *axr2* mutant, there is a dramatic decrease in the lengths of cells in stems, with a less conspicuous decrease in cell number. However, no anatomical studies of *axr2* leaves have been reported to date (Timpote *et al.*, 1992). The *AXR2* gene encodes a member of the Aux/IAA family that is thought to play a role in auxin signaling (Nagpal *et al.*, 2000).

Developmental plasticity in response to environmental and physiological conditions is a unique feature of plant development and is one of the most important current targets of studies of mechanisms that control plant development (Sultan, 2000). Among the environmental factors that influence the

developmental plasticity of plants, light has a particularly significant effect on leaf morphology, since leaves should receive photons as much as possible for photosynthesis in adaptation to the light environment. The PHYTOCHROME (PHY) gene controls the expansion of leaf blades and the elongation of petioles in *Arabidopsis* (Tsukaya *et al.*, 2002). The *far-red elongated1* (*fre1 = phyA*) mutant was reported to exhibit no changes in leaf morphology (Nagatani *et al.* 1993), while the long-hypocotyl *hy3* (= *phyB*) mutant does have defects in leaf morphology. All studies of this mutation agree that *hy3* mutant plants have longer petioles than those of wild-type plants (Nagatani *et al.*, 1991; Nagatani *et al.*, 1993; Reed *et al.*, 1993).

The proliferation of leaf cells is affected by many factors but most such factors influence cell proliferation ubiquitously in all plant organs. Kim *et al.* (1998) analyzed a mutant with a defect in the *CURLY LEAF* (*CLF*) gene, which encodes a member of the family of polycomb proteins that is required for stable repression, in vegetative shoots, of a member of the *MADS* gene family, the *AGAMOUS* (*AG*) gene (Yanofsky *et al.*, 1990). The *clf* mutant develops normal roots, hypocotyls and cotyledons, but the foliage leaves are significantly smaller and the inflorescence stems are shorter than those of the wild type. Both the extent of cell elongation and the number of cells are reduced in the *clf* mutant (Kim *et al.*, 1998). The ectopic expression of *AG* results in small leaves, whereas the *clf ag* double mutant has large leaves (Goodrich *et al.*, 1997).

1.7 Gene-expression profiling by using microarray technology

The complete sequencing of several genomes, including that of *Arabidopsis*, has signaled the beginning of a post-genomic era in which scientists are becoming increasingly interested in functional genomics; that is, uncovering the functional roles of different genes, and how these genes may interact with and/or influence one another. However, this question no longer needs to be answered by examining individual genes or proteins, but rather by simultaneously studying hundreds to thousands of unique genetic elements at a

time. Already, the post-genomic era is beginning to subdivide into distinct -omic domains, such as transcriptomics, proteomics, and metabolomics. This enables researchers to examine not only genetic elements, but also the corresponding proteins and metabolites.

To date, the most widely studied of these -omic domains are that of transcriptomics, which is able to reveal subtle differences in the levels of thousands of mRNA between experimental samples. Although proteins are the end product derived from a gene, the transcription of a gene is both critical and highly regulated, thereby providing an ideal point of investigation (Brazma and Vilo, 2000).

At present, two technologies dominate the field of high-density microarrays: the cDNA printed array and the oligonucleotide array. The cDNA array has a long history of development (Ekins, 1998) stemming from immunodiagnostic work done in the 1980s; however, it has been most widely developed in recent years by Stanford researchers, using the technique of depositing cDNA tags onto a glass slide with precise robotic printers (De Risi *et al.*, 1997). Labeled cDNA fragments are then hybridized to the cDNA probes on a chip and differences in the mRNA between samples can be identified and visualized using an arbitrary coloring scheme. The GeneChip[®] oligonucleotide array developed by Affymetrix, Inc (Santa Clara, CA, USA) (Chee *et al.*, 1996), involves synthesizing short (25-mers) probes directly onto a glass slide using photolithographic masks (Pease *et al.*, 1994; Barone *et al.*, 2001). Sample processing includes the production of labeled cRNA, which is then hybridized on the array, washing and laser scanning, resulting in a corresponding signal for each cRNA level. Regardless of the array used, the output can be readily transferred into various data analysis programs, where the selection and clustering of genes with different expression profiles can be examined.

1.8 Interactive Metabolic Pathway Map

The Interactive Metabolic Pathway Map was a joint project with my colleague Oliver Laule. The construction of databases as well as the scientific work necessary for visualization of gene expression data in a metabolic context was done in equal shares. Thus, chapter (2.3) is identical in his thesis: "Coordination of MVA and MEP Isoprenoid Pathways in *Arabidopsis thaliana*". The programming of the graphical user interface was performed by Matthias Hirsch-Hoffmann.

Technologies like whole-genome expression arrays (Celis *et al.*, 2000; De Risi *et al.*, 1997; Michaut *et al.*, 2003; Wang *et al.*, 2003) generate huge multiparameter data sets, which would have been unimaginable a few years ago. Their exploitation is limited by our ability to interpret them. Many studies just use them to earmark candidate genes. To realize their potential in providing a comprehensive analysis of system responses, it is necessary to combine them with a portfolio of interpretational tools. While many tools are available to analyze data sets by clustering and supervised machine learning, relatively few allow the data to be organized and displayed by the user in the context of pre-existing biological knowledge.

Example of tools that allow data sets to be viewed in the context of biological pathways, gene regulatory networks, or protein-protein interactions includes GenMAPP (<http://www.GenMAPP.org>), PATHWAYASSIST (<http://www.ariadne-genomics.com>), PATHWAY Processor (Grosu *et al.*, 2002) and BIOMINER (<http://voyager.bioinf.uni-sb.de/HPL/Projects/BioMiner>). Their usefulness for plant data sets, however, is restricted. First, they were developed for microbial or animal systems, so irrelevant categories are imported and plant-specific pathways and processes are absent. Second, their flexibility is restricted; for example, they often do not display family members individually. Plants have small- to medium-sized families for enzymes in central metabolism and very large families for many classes of enzymes involved in biosynthetic and secondary metabolism (Arabidopsis Genome Initiative, 2000). Tools that do not

resolve them will not realize the full potential of genome arrays. Third, an incomplete knowledge base hampers approaches that depend on bottom-up reconstruction of pathways.

1.9 The aim of this study

Part 1:

A detailed analysis of the *era1-2* mutant phenotype had previously been performed in our group. The aim of the first part of this thesis was to analyze the described alterations in leaf morphology in *era1-2* on a cellular level. For this a kinematic approach was used to get insights into the pattern of cell proliferation and cell expansion during leaf development.

Part 2:

In a second part of the work the focus was directed to the behavior of global gene expression caused by the lack of farnesylation in *era1-2*, using the Affymetrix GeneChip® system. The objective was to identify genes coding for biochemical- or signal transduction pathways related to the observed phenotype in *era1-2*.

Part 3:

The development of new methods of data analysis for microarray experiments, with the goal of establishing new standards to objectively process the massive datasets produced from functional genomic experiments, represents the third part of my work.

Several databases containing biological knowledge were established. For metabolic pathways, the database information was displayed on an interactive map, including all currently described biochemical pathways in Arabidopsis. In parallel, bioinformatics tools were developed for pre-analysis of gene expression data as well as automatic extraction of database knowledge out of the expression datasets.

2. Materials and Methods

Cultivation of plants as well as the isolation of total RNA and cRNA preparation was done by Arnaud Galichet (Gruissem Laboratory). The array hybridization, staining and scanning, and handling of the Affymetrix GeneChip[®] system, was performed together with Arnaud.

The characterization of cell cycle and cell elongation parameters using a kinematics approach was done by Gerrit Beemster (Flanders University, Ghent) and is described by De Veylder *et al.* (2001).

2.1 Plant material

Seeds from *Arabidopsis thaliana* (Columbia) wild type and *era1-2* mutant plants, respectively, were sown in 1x MS medium, supplemented with 1g/l sucrose and 0.5g/l MES pH 6.0, and stored for 4 days at 4°C before transfer to the growth chamber. Plants were grown at 20°C under long-day (16-hr-light/8-hr-dark) conditions at a photon fluence rate of 80 μ mol m⁻² s⁻¹ PAR supplied by “cool white” fluorescent tubes.

For RNA extraction, the first and second leaves were harvested 9, 15 and 21 days (8 hours after the switch from dark to light) after sowing for WT and after 11, 17 and 23 days for *era1-2*. The delay in development of 2 days is observed in parallel with the hypersensitivity of *era1-2* to ABA.

2.2 Gene-expression profiling

2.2.1 Sample preparation

2.2.1.1 Isolation of total RNA

RNA from plant tissue from the first and second leaf from *era1-2* and WT were extracted using Qiagen RNeasy mini kits (Basel, Switzerland). After extraction,

equal amounts of material were pooled to achieve 15µg of total RNA per sample. The absorbance of all RNA samples was checked at 260 and 280nm for determination of sample concentration and purity, and samples were additionally analyzed by agarose gel electrophoresis for intact 18s and 28s rRNA.

2.2.1.2 cRNA preparation

Five µg total RNA was used as starting material for all samples. The first and second strand cDNA synthesis was performed using the SuperScript Choice System (Life Technologies) according to the manufacturer's instructions, but using the oligo-dT primer containing a T7 RNA polymerase binding site (5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-(dT)₂₄-3').

Labeled cRNA was prepared using the MEGAscript. In Vitro Transcription kit IVT (Ambion). Biotin labeled CTP and UTP (Enzo) was used together with unlabeled NTP's in the reaction. Following the IVT reaction, the unincorporated nucleotides were removed by filtration through the RNeasy columns (Qiagen).

2.2.1.3 Array hybridization, staining and scanning

cRNA (15µg) was fragmented at 94°C for 35min in buffer containing 40mM Tris-acetate pH 8.1, 100mM KOAc, 30mM MgOAc. Prior to hybridization, the fragmented cRNA diluted in the hybridization cocktail (Eukaryotic Hybridization control mix, 0.1mg/ml herring sperm DNA, 0.5mg/ml acetylated BSA and 1x hybridization buffer) was heated to 99°C for 5min, subsequently cooled to 45°C for 5min and loaded onto the Affymetrix probe array cartridge. The probe array was incubated for 16h at 45°C at constant rotation (60rpm). After the hybridization, the probe array was exposed to 10 washes with non-stringent (user protocol) wash buffer at 25°C followed by 4 washes with stringent (user protocol) wash buffer at 50°C. The biotinylated cRNA was stained with a

streptavidin-phycoerythrin conjugate (SAPE solution) for 10min at 25°C followed by 10 washes with non-stringent wash buffer at 25°C. The second stain was done for 10min at 25°C in the antibody solution followed by the third stain for another 10min at 25°C with the SAPE solution. The final 15 washes were done at 25°C with non-stringent wash buffer. The probe arrays were scanned at 560nm using a confocal laser scanning microscope (made for Affymetrix by Agilent). The readings from the quantitative scanning were analyzed by the Affymetrix gene expression analysis software, Microarray Suite 5.0 (MAS 5.0).

2.2.2 Pre-analysis of the expression data

Gene expression array technology is rapidly becoming standard in many areas of biological research. High-density oligonucleotide GeneChip[®] (Affymetrix, Santa Clara, CA) provide a convenient medium on near genomic scale for plant research. Each gene is represented on the GeneChip[®] by an array of diverse probe pairs (typically 11). Each probe pair consists of a perfect match (PM) probe and the corresponding mismatch (MM) probe. Each PM probe contains thousands of identical short (25-mer) sequences from the transcribed sequence of the gene represented by that probe set. The MM probe on the other hand contains sequences identical to the PM probe except for a single nucleotide at the center of the sequence. Upon hybridization with the fluorescent-dye-labeled cRNA, each probe in the probe set captures a certain amount of cRNA, whose fluorescent intensity is subsequently measured. The measured intensities of all probes in the probe set reflect in a unique way the expression level of that particular gene.

2.2.2.1 Detection call: absent, marginal and present

The detection call answers the question if the transcript of a particular probe set is reliably detected by the probe array. The answer is given by the calls “absent” or “present”. In this context, absent means the expression level is below the threshold of detection. A “marginal” call is used in the case of uncertainty.

The signal represents the amount of transcript in solution. For each PM intensity, a matching MM probe provides reference background hybridization intensity. If the MM value is less than the PM value, the PM value is used directly. However, if the MM value is larger than the PM value, the algorithm creates an adjustment MM value based on the average difference intensity between \log_2 PM and \log_2 MM, or if that measure is too small, some fraction of PM. The adjusted MM values are used to calculate the \log_2 PM – \log_2 adjusted MM for each probe pair. The signal for a probe set is calculated as the one-step biweight estimate of the combined differences of all the probe pairs in the probe set.

The difference between PM and MM is the basis to compute the corresponding call. A one-side Wilcoxon’s signed rank test is used to calculate the p-value that reflects the significance of the differences between PM and MM. The p-value or statistical significance of a result is the probability that the observed change in a sample occurred by chance. The lower the p-value, the greater the probability the results are significant. To make a call, the p-value for a probe set is examined on an axis with two pre-defined thresholds, α_1 and α_2 ($\alpha_1 = 0.04$ and $\alpha_2 = 0.06$). For p-values between zero and 0.5, α_1 and α_2 define the threshold for the calls. If the p-value $< \alpha_1$, the detection call will be present, if $\alpha_1 \leq$ p-value $< \alpha_2$, it is called marginal (at the limit of detection) and if the p-value $> \alpha_2$, it is called absent. For further analysis, I included only datasets with a present call through all experiments.

2.2.2.2 Data normalization

Typically, the first transformation applied to expression data, referred to as normalization, adjusts the individual hybridization intensities to balance them appropriately so that meaningful biological comparisons can be made. There are a number of reasons why data must be normalized, including unequal quantities of starting RNA, differences in labeling or detection efficiencies and systematic biases in the measured expression levels. Conceptually, normalization is similar to adjusting expression levels measured by Northern analysis or quantitative reverse transcription PCR (RT-PCR) relative to the expression of one or more reference genes whose levels are assumed to be constant between samples.

In a first step, the data were scaled to a predefined target signal (500, according to instructions from the Affymetrix field specialist). Scaling is a mathematical technique applied to the data from several different probe arrays to minimize discrepancies. The scaling factor is also applied to the noise value. I used two independent experiments for comparative analysis only if the difference between the scaling factors is less than 5.

$$\text{Scaling Factor} = \frac{\text{Target Signal}}{\text{Trimmed Mean Signal}_{\text{probe array}}}$$

Target Signal = 500 (pre-defined)

Trimmed Mean Signal_{probe array} = $\sum \text{probe signal}_i$ (i=1 up to N) / N

N = number of total probe signals

2.2.2.3 Fold change and log transformation

Selecting interesting genes can be done in various ways. However, two selection methods are used very widely. The first such selection method involves simply comparing the expression levels in the experiment versus control. The genes having expression values very different in the experiment versus control are selected. A difference is considered as significant if it is at least 2-fold. Sometimes, this selection method is used in parallel on expression estimates provided by several techniques. The second widely used selection method involves selecting the genes for which the ratio of the experiment and control values is a certain distance from the mean experiment/control ratio. A number of other *ad hoc* threshold and selection procedures have also been used.

In this work, I selected genes if they show at least once a 2-fold change between the sample and the respective control over all experiments.

There are three common types of values that can be associated with the fold change of a gene. The first, a signed fold change (such as a +1.8 or -3.1 fold change, corresponding to induction or repression, respectively) has the most intuitive appeal but has a discontinuity spanning from -1.0 to +1.0 that can be problematic. The second type, an unsigned change (such as 1.8 or 0.35 fold change, again corresponding to induction or repression, respectively) has no such discontinuity but is bounded in the lower range by zero and is unbounded in the upper range. This asymmetry about unity hinders analysis. The third type, a logarithm (base 2, base 10 or natural) of the unsigned fold change, is undoubtedly the most tractable. No fold change in expression is represented by zero, induction is positive and repression is negative. Most importantly, there are no discontinuities or asymmetries. In addition, effects on the intensity of GeneChip[®] signals tend to be multiplicative; for example, doubling the amount of RNA should double the signal over a wide range of absolute intensities. The logarithm transformation converts these multiplicative effects (ratios) into additive effects (differences), which are easier to model. For the global analysis the datasets were further log transformed to the basis of 2.

2.2.3 Data mining: Introduction to bioinformatics

To get more insight into the tendencies within the expression profiles, the data sets were clustered using the Eisen software (Eisen *et al.*, 1998). After log-transformation, I grouped the data based on the self organizing maps algorithm. All genes and arrays were normalized to 1. Afterwards the sets were clustered in a hierarchical manner (complete linkage clustering) in two dimensions (genes and arrays). For visualization the results were colored using the TreeView software (part of the Eisen software). The principles behind this strategy are explained in the following.

2.2.3.1 Information extraction by data clustering

GeneChips[®] facilitate the monitoring of changes in the expression patterns of large collections of genes. The analysis of expression array data has become a computationally-intensive task that requires the development of bioinformatics technology for a number of key stages in the process, such as image analysis, database storage, gene clustering and information extraction.

In general, there are two types of experiments: those involving the comparison of two conditions (typically the condition of interest versus a reference) and those involving the study of many conditions (e.g. time courses, dosage series, series of patients, tissues, etc.). Typically, multi-condition experiments are represented by a matrix of gene expression values, with genes in rows and conditions in columns. Depending on the experiment, the values of gene expression can be used to classify conditions (columns) or gene expression profiles (rows). Both cases involve an initial grouping step, either to obtain sets of conditions with similar gene expression values or to obtain sets of genes with similar expression profiles for the conditions studied. The grouping is usually performed using clustering methods.

Data can be clustered in two different ways: in a hierarchical or non-hierarchical manner. Hierarchical clustering allows detection of higher-order relationships

between clusters and profiles, whereas the majority of non-hierarchical classification techniques work by allocating expression profiles to a pre-defined number of clusters, with no assumptions as to the inter-cluster relationships.

2.2.3.1.1 Hierarchical clustering

Aggregative hierarchical clustering in its different variants (average-linkage, single-linkage, complete-linkage) (Sneath and Sokal, 1973) is one of the most popular choices for the analysis of pattern of gene expression. This is in part due to the availability of software for running these methods, either within standard statistical packages or those designed specifically for gene expression data (Eisen *et al.*, 1998). Standard aggregative hierarchical clustering produces a representation of the data in the shape of a binary tree, where the most similar patterns are clustered in a hierarchy of nested subsets (Sneath and Sokal, 1973). This method has been used to analyze several of different datasets, from yeast (Eisen *et al.*, 1998) to human cells (Wen *et al.*, 1998; Scherf *et al.*, 2000).

The basic idea behind hierarchical clustering is to assemble a set of items (genes or arrays) into a tree, where items are joined by very short branches if they are very similar to each other, and by increasingly longer branches as their similarity decreases. First of all the similarity has to be defined. The most commonly used similarity metrics are based on Pearson correlation. The Pearson correlation coefficient r between any two series of number $X=\{X_1, X_2, X_3, \dots, X_N\}$ and $Y=\{Y_1, Y_2, Y_3, \dots, Y_N\}$ is defined as:

$$r = \frac{1}{N} \sum_{i=1, N} \left(\frac{X_i - \bar{X}}{\sigma_x} \right) \left(\frac{Y_i - \bar{Y}}{\sigma_y} \right)$$

where \bar{X} is the average of values in X , and σ_x is the standard deviation of these values. If X and Y are plotted as curves, the Pearson correlation coefficient r will tell how similar the shapes of the two curves are. r is always between -1 and 1, with 1 meaning that the two series are identical, 0 meaning they are completely

independent and -1 meaning they are perfect opposites. The correlation coefficient is invariant under scalar transformation of the data. Thus, two curves that have an identical shape, but different magnitude, will still have a correlation of 1. To avoid this, the following modified equation for *un-centered* correlation was used:

$$r = \frac{1}{N} \sum_{i=1,N} \left(\frac{X_i}{\sqrt{\frac{1}{N} \sum_{i=1,N} (X_i)^2}} \right) \left(\frac{Y_i}{\sqrt{\frac{1}{N} \sum_{i=1,N} (Y_i)^2}} \right)$$

Taken together, if N is the number of data observations, the first iteration merges the two most similar clusters to form N_1 clusters. The Pearson correlation coefficient was used to calculate the similarity. On the next iteration, N_2 clusters are formed, by merging the two most similar of N_1 clusters. The process continues until only a single cluster remains, containing all N data points.

2.2.3.1.2 Self organizing map

Standard hierarchical clustering works very well for clustering conditions (represented by columns, i.e. a small number of items), but several authors (Tamayo *et al.*, 1999) have noted that this clustering method is not very robust when applied to clustering thousands of gene expression profiles. In addition, typical runtimes of standard methods based on distance matrices can range from N^2 to N^4 , which makes them very slow when thousands of items are to be analyzed. Neural networks have been proposed as an alternative for overcoming some of the above mentioned problems (Tamayo *et al.*, 1999; Törönen *et al.*, 1999; Herrero *et al.*, 2001). Unsupervised neural networks, such as self organizing maps (SOM) (Kohonen, 1990) or the self organizing tree algorithm (SOTA) (Dopazo and Carazo, 1997), provide a more robust framework, appropriate for clustering large amounts of noisy data. Because of their properties, neural networks are suitable for the analysis of gene

expression patterns. They can deal with data sets containing noisy, ill-defined items with irrelevant variables and outliers and whose statistical distribution does not need to be parametric. SOM is a method of clustering that are

SOM algorithm:

Initialize the N cluster centers: c_1, c_2, \dots, c_N

Repeat steps 1 through 4 until convergence is reached.

For iterations n:

1. Select a gene g from the normalized data.
2. Calculate the distance d_i from the gene to each cluster center c_j .
 $d_i = D(g \| c_j), i = 1, 2, \dots, N$
3. Identify the cluster k closest to g
 $k = \arg \min \{d_j\}$
4. Update the weights of the k th cluster and its immediate neighbors
using the following learning rule:
 - a) $c_k(n+1) = c_k(n) + (n)(g - c_k(n))$
 - b) $c_j(n+1) = c_j(n) + (n)(g - c_j(n)); j = k-1, k+1$
 - c) The learning rates are given by (learning rates are adjusted every 10 epoches for better stability during training):

$$\eta_1(n+1) = \eta_1(n)/(1+0.0005n^{0.01})$$

$$\eta_2(n+1) = \eta_2(n)/(1+0.0005n^{0.01})$$

5. Check the conditions: If $(\|c(n+1) - c(n)\|_2) / (\|c(n+1)\|_2) < \epsilon$ then stop, else continue.

somewhat related to k-means clustering. The design of SOM starts with defining a geometric configuration for the partitions in one- or two-dimensional grid. Then, the random weight vectors are assigned to each partition. During training, a gene expression profile is picked randomly.

The weight vector closest to

the expression profile is identified. The identified weight vector and its neighbors are adjusted to look similar to the expression profile. This process is repeated until the weight vectors converge. During operation, SOM maps gene expression profiles to the relevant partitions based on the weight vectors to which they are most similar. However, in circumstances where the expected number of partitions (clusters) available in gene expression data is unknown, the validation of SOM's clustering result becomes a critical issue. One may heuristically validate the clustering results to identify the approximate number of clusters.

Many validation techniques have been implemented to evaluate clustering results: figure of merit (FOM), gap statistic, Hubert and Jaccard index, and adjusted rand index, and entropy measure. All these algorithms have been proven valuable by some experiments. However, given the large volume and high dimension of gene expression data, the use of such heuristic evaluation technique makes the clustering process extremely time-consuming and complicated.

2.2.3.1.3 Combination of SOM and hierarchical clustering

A number of methods have been proposed in the literature to accomplish data partitioning and cluster validation/visualization either simultaneously or independently. Herrero and Dopazo (2002) proposed a combination of SOM and hierarchical method for clustering gene expression data. The use of a previous step of pre-clustering with SOM (dimensionality reduction), followed by the application of hierarchical clustering methods, constitutes a fast and reasonably accurate method for exploratory analysis of large amounts of data. Given the speed of the procedure, it can be used interactively even for several thousands of expression profiles.

The used version varies slightly from that of Tamayo *et al.* (1999), in that it restricts the analysis one-dimensional SOMs along each axis. The result is similar to the result of k-means clustering, except that, unlike k-means, the nodes in a SOM are ordered. This tends to result in a relatively smooth transition between groups. The final picture is a hierarchical tree in which the branch lengths are proportional to the differences of the average expression patterns of the clusters under the bifurcations of the tree. Our datasets were first clustered according to SOM and afterwards grouped using the hierarchical clustering algorithm (complete linkage clustering). All genes as well as all experiments were equally weighted to 1.

2.3 Interactive Metabolic Pathway Map

2.3.1 Overall design of the metabolic pathway map

The Metabolic Pathway Map consists of an *Arabidopsis thaliana* specific database for biochemical pathways combined with a visualization tool to display global gene expression data in a functional and compartment specific context. Expression data is displayed within the biochemical network and can then be followed up to the level of individual catalytic reaction for which a particular gene encodes the respective enzyme.

2.3.1.1 Database construction for biochemical pathways in *A. thaliana*

To build the database for plant-specific pathways, initially 10 main categories corresponding to different areas of functions (proteinogenic amino acids; carbohydrates; lipids; vitamins and cofactors; complex carbohydrates; complex lipids; non- proteinogenic amino acids, other (secondary) substances, nucleotides, plant hormones) were defined. Each category was then extended into several sub-categories in a hierarchical manner (Fig. 2.1). Each sub-category consists of an individual pathway or represents a metabolic base unit. Pathway and gene annotations were accomplished by manual literature search. Sequences of genes shown to encode pathway enzymes were extracted from the NCBI and TAIR database and were then blasted against the Affymetrix Arabidopsis ATH1-121501 GeneChip[®], utilizing the Affymetrix Probe Match Tool (https://www.affymetrix.com/analysis/probe_match.prma). Only genes showing a perfect match with the query sequence (Probe set of 11 oligonucleotides) were selected and linked to our database via their unique Affymetrix gene identifiers (ChipID).

The data base encompasses currently 1250 genes encoding pathway enzymes and 468 enzymatic reactions in 59 pathways or metabolic units containing 609 different compounds. Since one gene can encode proteins possessing multiple

enzymatic functions and one enzyme function can require the protein products of several genes, ChipIDs are linked to proteins and vice versa in a “many-to-many” mode. Metabolic compounds are represented as their chemical structures created by using the ChemDraw Pro 7.0 software, imported and stored into the database as bitmap (BMP) files.

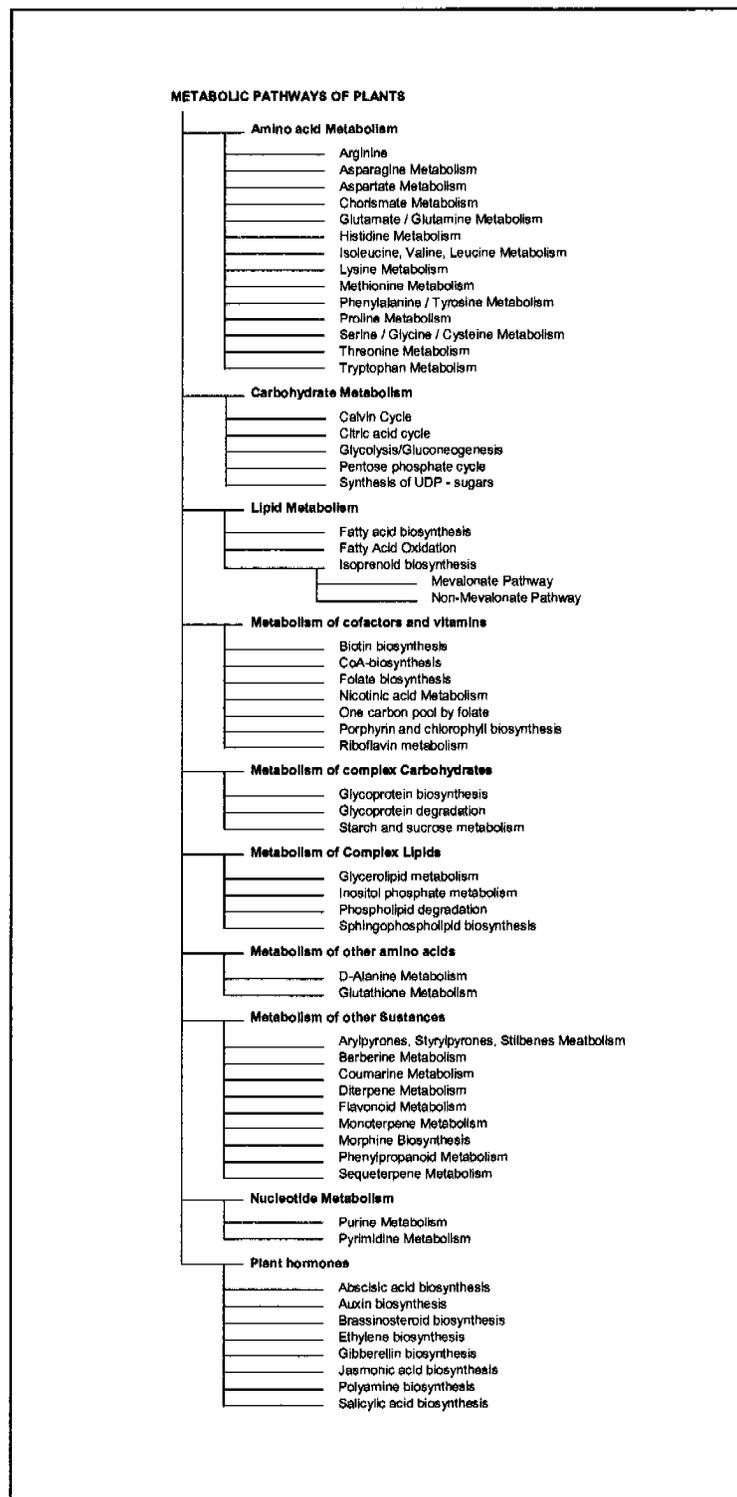


Figure 3: Hierarchical order defined for *A. thaliana* metabolic database. The metabolism is organized in 10 main categories and 59 sub-categories.

2.3.1.2 Integration of compartment specific gene annotation

Information about compartment specificity in *A. thaliana* was obtained by using TargetP, a method for predicting the subcellular location of proteins by identifying the presence of N-terminal pre-sequences, such as chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP), or secretory pathway signal peptide (SP) (<http://www.cbs.dtu.dk/services/TargetP/>). The genes encoding proteins assigned to a certain compartment were connected to the unique Affymetrix identifiers via their AgiCodes, and the information was integrated in the biochemical pathway database.

2.3.2 Visualization Software

The visualization software was developed to map genome-wide gene expression data onto biochemical pathways contained in the underlying database. The tool depicts pathways of different topologies using different layouts (linear, circular, and tree structured). The user can include or suppress information such as the chemical structures of substrates.

The map is classified hierarchically at three levels of detail, i.e., three levels of resolutions (Fig. 2.2, 2.3, 2.4). Resolution 1 is an overview of the network of Arabidopsis metabolic pathways shown in a single screen. Resolution 2 provides medium grade resolution in terms of functionality like carbohydrate metabolism, nucleotide metabolism, etc. and contains pathways related to that function. Resolution 3 shows the reaction network as well as the compounds and the enzymes involved in individual pathways. The connections between the different pathways are displayed in all levels.

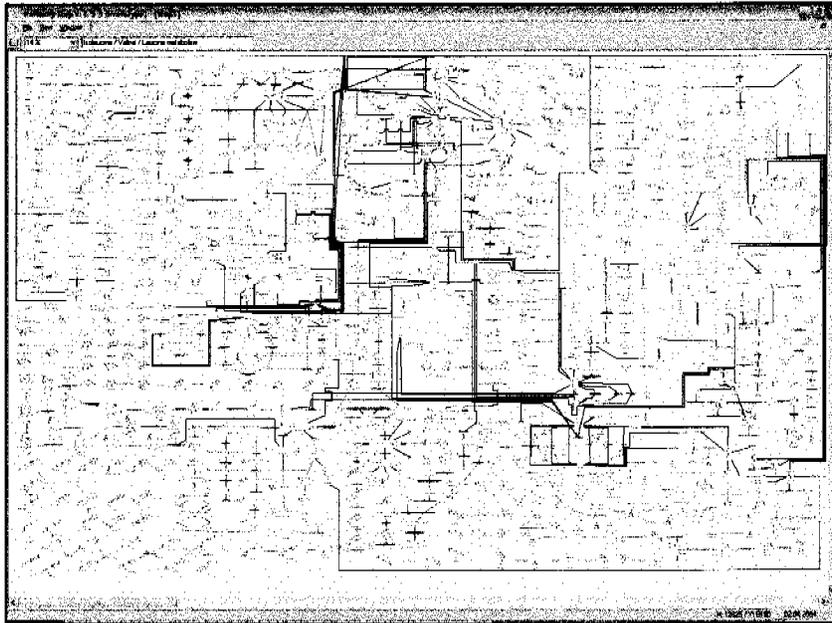


Figure 4: General overview of Arabidopsis metabolic pathways. Enzymes belonging to an individual pathway or a functional unit are symbolized by a common color.

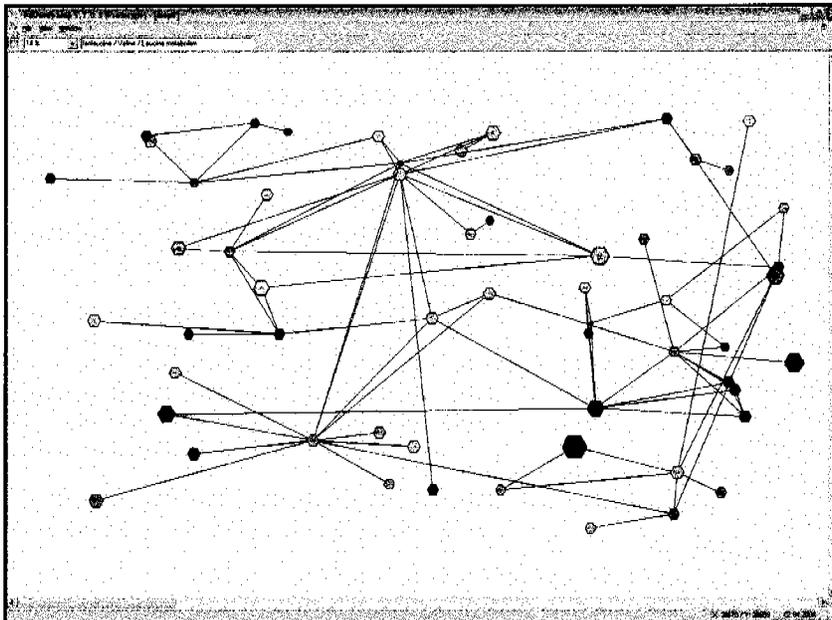


Figure 5: Interconnection view of metabolic pathways within the biochemical network. Each hexagon represents an individual pathway or a functional unit.

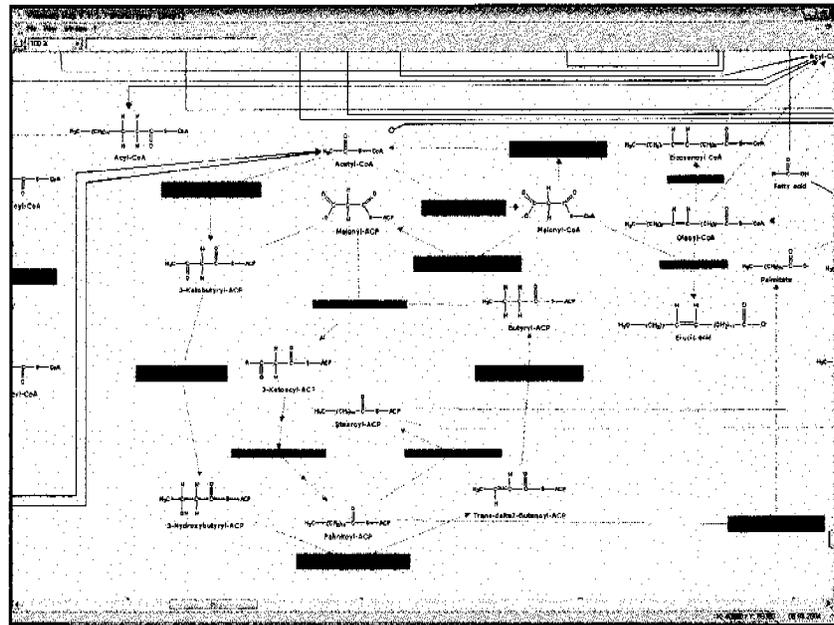


Figure 6: Individual pathway view. As an example, fatty acid biosynthesis is depicted.

To view gene expression data on the biochemical pathway map, the data can be imported in a column-separated excel format and saved as a defined experiment. Loaded experiments can be assigned individually as samples or controls to compare the respective gene expression levels.

The log ratio of gene expression data sample/control is calculated automatically after data sets are imported into the map and is then visualized by color-coding (Fig. 2.5) genes based on user presetting. The color code for differences in gene expression (red: up-; green: down-regulated) is based on a modified algorithm developed by Eisen *et al.* (1998). For the study of gene expression data in the context of subcellular compartments, only compartment-specific changes in gene expression are displayed by color code. The biochemical pathway map includes a sample data set and an extensive help tutorial.

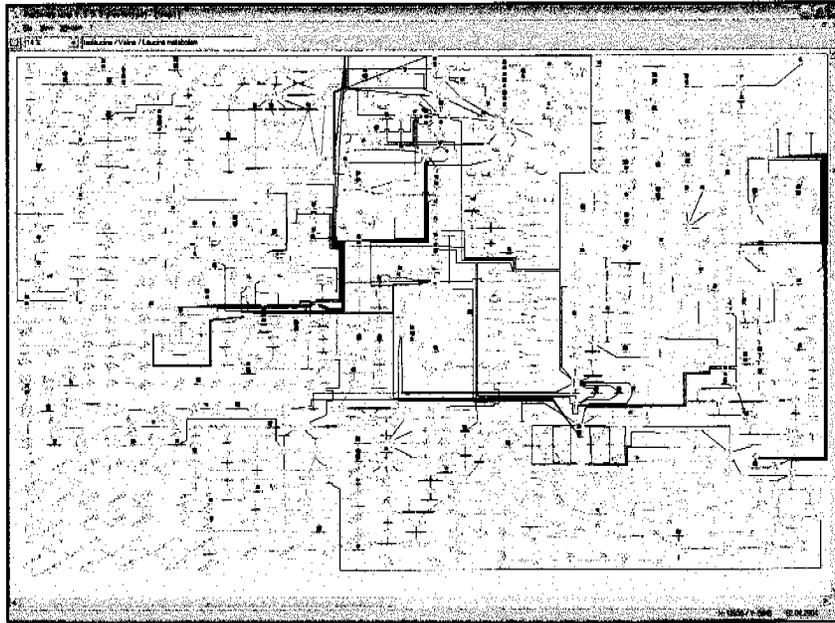


Figure 7: Overview of *A. thaliana* metabolic pathways with gene expression data (fold-changes) visualized by color code. Red represents an increase in relative expression whereas green represents a decrease.

2.3.2.1 Outlook

The metabolic pathway map requires further development and extension of plant- specific biochemical pathways. It is further planned to include new features such as the possibility to display metabolic as well as proteomic data to compare changes on multiple levels in order to get better insights in biological processes. The pathway map will be available on the web, from which instructions on how to introduce experimental data sets can be downloaded. This will allow users to view and analyze their data in a biochemical context.

2.4 Additional databases

To expedite the proceedings of global gene expression analysis, additional databases which allow a functional data analysis were constructed.

2.4.1 *A. thaliana* database for regulatory elements

The *Arabidopsis thaliana* database for regulatory elements is based on keyword searches, performed on Affymetrix gene annotations for the ATH1-12150 GeneChip®. To implement keyword searches, a visual basic macro was developed. The number of genes obtained for each keyword, e.g., 1150 kinases, is given in Fig. 2.6

Keywords:	
Calcium	(92)
Calmodulins	(77)
Carriers	(68)
CDK cdc cyclins	(81)
Channels	(66)
CLAVATA	(19)
Cyclophilins	(13)
Cytochromes	(270)
Farnesyl- geranyltransferase	(8)
G-proteins	(80)
Histones	(67)
Importin exportin	(15)
Inositols	(62)
Kinases	(1150)
Peroxidases	(87)
Phosphatases	(213)
Phospholipases	(42)
Receptors	(423)
Senescence associated	(23)
Thioredoxins	(40)
Transcription factors	(926)
Transporters	(305)

Figure 8: Keyword database for *A. thaliana*. Keywords used to construct the *A. thaliana* database of regulatory elements by screening the ATH1-12150 GeneChip®. Numbers of genes encoding a respective element are shown in parentheses.

2.4.2 *A. thaliana* database for sub-cellular genomes

A list of AgiCodes for genes which constitute the plastome or the mitochondrial genome, respectively, was downloaded from the Munich Information Center for Protein Sequences (MIPS). The genes encoding proteins assigned to either plastome (82) or mitochondrial genome (106) were then linked to the unique Affymetrix identifiers via AgiCodes.

2.5 Database information extraction

User friendly tools were developed to establish new standards to objectively process the massive datasets of gene expression profiles. This includes raw data selection based on the detection calls or p-values, ratio calculation between two experiments, and the selection of genes, that are significantly altered in their gene expression levels. The information stored in the different databases (described above) could be easily extracted out of the expression data by creating functional annotated lists or by visualization in a simplified metabolic cell.

2.5.1 Java applications

The expression profile analyzer1.0.0 for plants, the functional group analyzer and expression profile analyzer for the Human GeneChip are written in JAVA 1.3 and run on different operating systems (Windows, Mac OS X, Linux, UNIX, etc.). This software package can run on any computer for which there exists a JAVA virtual machine (JVM).

Requires for PC: Windows 98/NT/2000/XP__Pentium III or later; 256 RAM*; 50 MB of free disk space* __1400 x 1024 display

Requires for Mac: Mac OS 9.1/9.2/OS X v10.2.1 or later __256 RAM*; 50 MB of free disk space* __1400 x 1024 display

*Large data sets may require significantly more RAM and storage for optimum performance.

All the software can be installed from a CD.

2.5.1.1 Expression profile analyzer 1.0.0 for plants

The expression profile analyzer 1.0.0 for plants consists of four independent modules (fig. 9). The first module can be used to compare two different sets of expression data. The second module is designed for the raw data analysis. Genes could be selected if they have a present-call through all experiments of interest or according to the p-value, which represents the liability of the signal detection from the 11 independent probe sets. Changes in gene expression can be calculated by the ratio function between a sample and the corresponding control and genes could be selected if they have a significant fold change at least in one experiment, which can be defined by the user. The third module is developed to extract the database information (pathways, compartments, genomes, keywords) out of the expression profiles. A keyword search represents the fourth module. Any keyword (gene name, chipID, expression level etc.) could be used to screen through any list of interest.

Every result is displayed on the screen within the software and automatically stored on the respective folder under the name of the entry list, linked with a short description of the analysis which was done. The lists for the analysis have to be loaded as a text delimited file (.txt) consisting of columns and rows. It is important that the first column consist the ChipID and the second the GeneChip descriptions. The following columns are dependent on the analysis which would be performed.

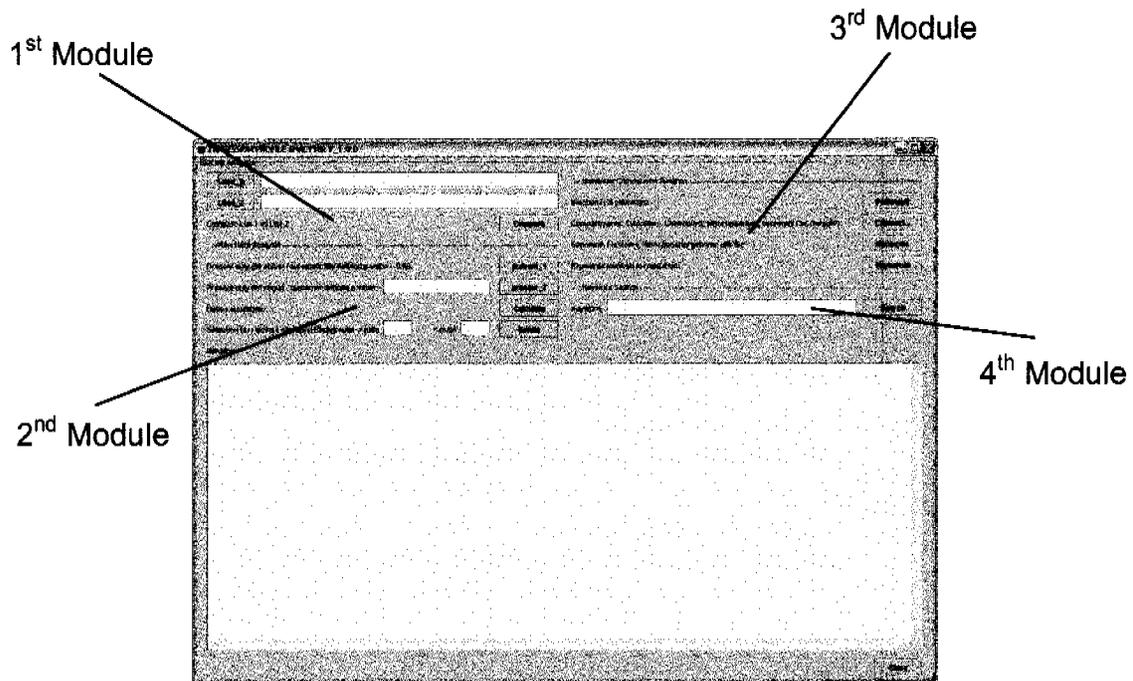


Figure 9: Expression profile analyzer for plants

Screen of the expression profile analyzer 1.0.0 for plants (Java application). The software contains four different functional modules. 1st Module: comparative analysis; 2nd Module: raw data analysis; 3rd Module: database supported analysis; 4th Module: key word search.

2.5.1.2 Functional group analyzer

The functional group analyzer displays all the biochemical pathways from the Arabidopsis database in a simplified cell, containing four distinct separated compartments: cytoplasm, chloroplast, mitochondrion and secretory pathways (fig. 10). Each metabolic unit is represented by a grey square. Single pathways are grouped to a metabolic unit. After loading the gene expression data, all the genes for one pathway are in linear order of squares colored according to the previously explained color code (green for repressed genes and red for induced genes). This kind of analyses gives a first overview on the level of gene expression for biochemical pathways and their compartment predictions. The next step will include a detailed analysis with the Metabolic Map with all the chemical structures and the links within the cellular network.

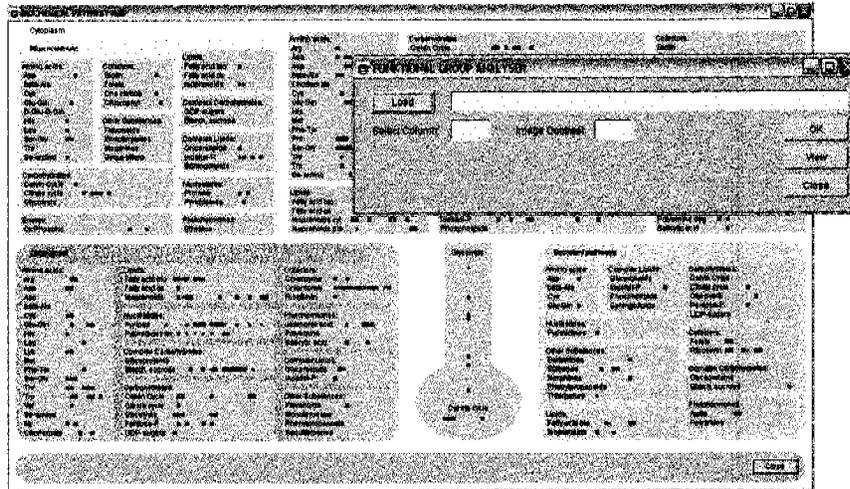


Figure 10: Functional group analyzer for plants

Screen of the functional group analyzer. The small screen can be used for the data transformation, whereas the large screen is used for the visualization of the different compartment and the involved biochemical pathways.

2.6 Biocomputational approach to screen the *Arabidopsis* protein-sequence database for putative PFT targets

To find putative PFT (protein farnesyltransferase) targets, the *Arabidopsis* protein database was screened for proteins containing a Ca_1a_2X box. According to Moores *et al.* (1991) only a selected group of amino acids at the position a_1 , a_2 or X were proven to be functional in context of PFT targets. Nevertheless, each candidate has to be experimentally analyzed for its ability to be targeted by FPP. As a consequence, the biocomputational approach was done in three steps with increasing stringency. In a first loop the protein database was screened for proteins, which have a cysteine (C) at the fourth position from the C-terminal end ($Cxxx$). In a second step, the group obtained in the first step was analyzed for proteins with either a G, A, I, F, M, C, S, T, H, N or a Q at the C-terminal end ($CxxX$). In the last step the database was screened for a theoretical functional $CaaX$ group according to Moores *et al.* (1991). For the first "aliphatic" position a_1 the amino acids G, A, V, L, I, P, F, Y, W, M, C, S, T, K, R and H and for the second "aliphatic" position a_2 , which has experimentally been proven to be much more specific, the amino acids G, A, V, L, I, P, M, C, S and T were used for the screen ($CaaX$). For the third step, each amino acid, within the

-aaX box, got an efficiency score (table 3). The score is based on the farnesylation efficiency compared to a control (Moore *et al.*, 1991).

	aa	a		a		X	
			e coeff		e coeff		e coeff
aliphatic	G	x	1.00	x	1.00	x	0.30
	A	x	1.00	x	1.00	x	0.80
	V	x	1.00	x	1.00	x	0.40
	L	x	1.00	x	1.00		
	I	x	1.00	x	1.00	x	0.25
	P	x	0.25	x	0.30		
aromatic	F	x	0.55			x	0.30
	Y	x	0.55				
	W	x	0.55				
sulfurous	M	x	1.00	x	1.00	x	0.60
	C	x	1.00	x	1.00	x	1.40
alcoholic	S	x	1.00	x	1.00	x	1.20
	T	x	1.00	x	1.00	x	0.45
alkaline	K	x	0.50				
	R	x	0.50				
	H	x	0.50			x	0.35
acidic	N					x	0.40
	Q					x	0.85
	D						
	E						

Table 3: Table for putative PFT target prediction according to Moore *et al.*, 1991

All amino acids are described if they are functional at a certain position within the CaaX box and with the corresponding efficiency coefficient.

By multiplication of the three scores, the resulting value is the probability of an efficient farnesylation. An efficiency score of 1 indicates that the amino acid combination has the same efficiency as the control combination used in the approach of Moore *et al.* (1991). The calculation was done for all proteins obtained by the third step. According to *Apetala 2*, which is experimentally proven to be a target for farnesylation, with a score of 0.3 a cut off of 0.25 was used to determine putative PFT target.

Pseudo-code for the putative PFT target search:

1. Loop

Cxxx

for C: if sequence length = 3="C"
select and write in a new file

2. Loop

CxxX

for C: if sequence length = 3="C"
and for X: if X= "C", "M", "S", "A", "Q", "G", "H", "T", "N", "V", "I", "F"
select and write in a new file

3. Loop

```

C a1 a2 X
  for a1: if a1= "D", "E", "Q", "N"
    'Exclude
  else
    Case "K", "R", "H"
      eff coeff A = 0.5
    Case "F", "Y", "W"
      eff coeff A = 0.55
    Case "P"
      eff coeff A = 0.25
    Case Else
      eff coeff A = 1

C a1 a2 X
  for a2: if a2= "D", "E", "Q", "N", "K", "R", "H", "F", "Y", "W"
    'Exclude
  else
    Case "P"
      eff coeff B = 0.3
    Case Else
      eff coeff B = 1

C a1 a2 X
  for X: if X= "C", "M", "S", "A", "Q", "G", "H", "T", "N", "V", "I", "F"
    Case "C"
      eff coeff C = 1.4
    Case "M"
      eff coeff C = 0.6
    Case "S"
      eff coeff C = 1.2
    Case "A"
      eff coeff C = 0.8
    Case "Q"
      eff coeff C = 0.85
    Case "G"
      eff coeff C = 0.3
    Case "H"
      eff coeff C = 0.35
    Case "T"
      eff coeff C = 0.45
    Case "N"
      eff coeff C = 0.4
    Case "V"
      eff coeff C = 0.4
    Case "I"
      eff coeff C = 0.25
    Case "F"
      eff coeff C = 0.3

Calculation_CaaX = eff coeff A * eff coeff B * eff coeff C

if Calculation_CaaX >= 0.25 → putative PFT target

```

2.7 Electronic Northern as a control for expression status of putative PFT targets

To define whether a certain gene is expressed or not, or at least once during the entire time course, a cut off value for the expression level was calculated. A typical expression profile was used to count the genes with a present- or an absent call, respectively, and graphically displayed (fig. 11). On the x axis are the logarithmical transformed values of the expression levels and on the y axis the amount of genes for each expression level. The genes with a present call are represented in the upper half, whereas the genes with an absent call are in the lower half of the figure. We observed a significant correlation between the expression level and the detection call: the higher a gene is expressed, the higher is the probability that this gene has a present call and vice versa. As a consequence, either the chip is not able to detect low expression levels, or the levels of the transcripts are not high enough to be translated and therefore not functional. The further analysis was based on the second consequence.

A differential curve between the “present genes” and the “absent genes” in respect to expression levels was calculated. The curve cuts the x axis at 2.09 which correspond to an expression level of 125. On an expression level of 125 the chip technology is able to detect the same amount of genes with a present call as well as with an absent call. As a consequence, up to 125 transcripts were defined as functional and therefore they could serve as targets for translation.

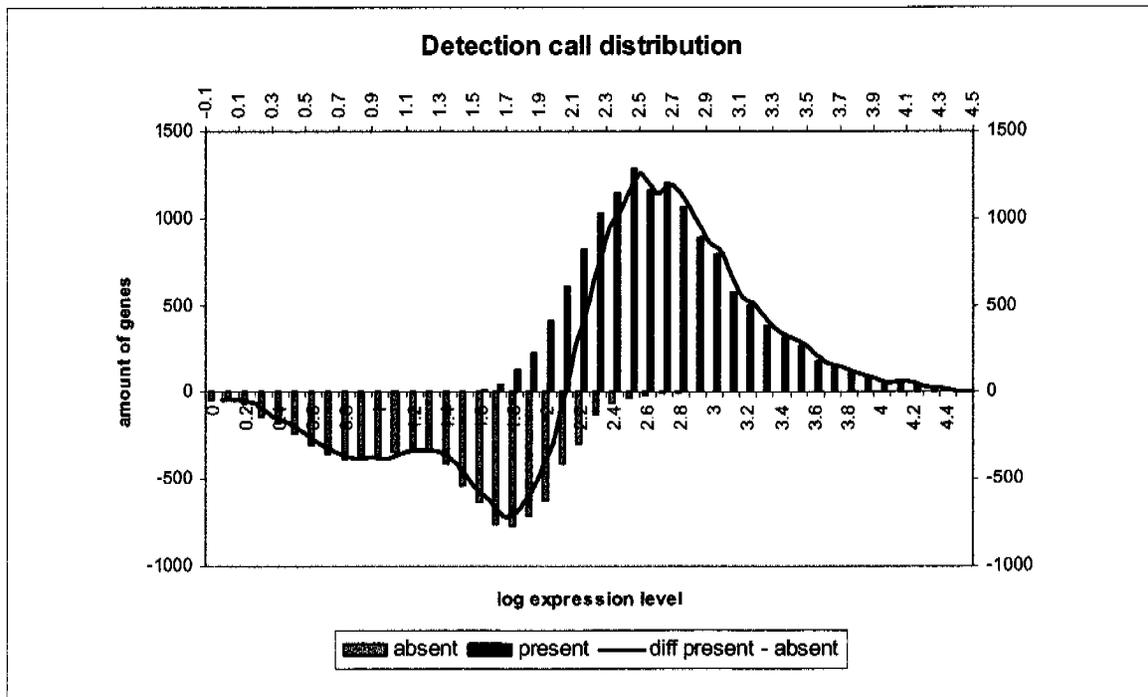


Figure 11: Detection call distribution within a WT gene expression profile

Graphical view of the gene expression values (log transformed) vs the corresponding amount of genes. Red symbolizes the genes with a present call and blue the genes with an absent call. The black line defines the differential spectrum present call minus absent calls.

All genes, which were found to be putative PFT targets, were analyzed in the context of a functional active expression level.

3. Results

3.1 Kinematics

A kinematics analysis was performed in collaboration with Gerrit Beemster (Flanders University, Ghent). In Ghent we measured the different parameters (leaf size, cell number, cell size and DNA content) that underlie this kinematic study. Data related to leaf development and cell cycle regulation in WT *Arabidopsis thaliana* will be published separately. In Zurich, the data were used for the comparative analysis of leaf development between *era1-2* and WT. Arnaud Galichet analysed the data in the context of cell cycle and cell expansion, and I analysed the data in the context of biochemical pathways and their regulation.

3.1.1 Kinematics analysis

Among other macroscopic alterations (described by Yalovsky *et al.*, 2000), an increase in mature leaf size of about 70% (Fig. 12) was observed for *era1-2* compared to WT *Arabidopsis*.

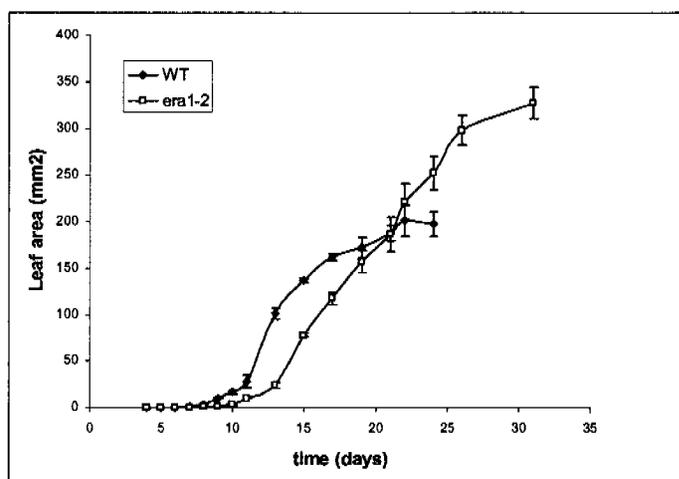


Figure 12: Kinematics I: Leaf area

Increase in leaf size in WT and *era1-2* measured over a time period of 32 days for *era1-2* and over a period of 25 days for WT.

Although plant development is influenced by environmental factors, it appears that the intrinsic size of plant organs like leaves is controlled by cell number and cell size. To investigate the difference in mature leaf size between *era1-2* and WT on a cellular basis, the progression of the growth process was analyzed over time by using a kinematics approach. For this purpose, the first and the second leaf of WT and *era1-2* was analyzed for leaf area, cell size, cell number and DNA content. For the cell size and cell number measurements, only the epidermal cell layer was used. The measurements were conducted over a period of 32 days for *era1-2* and over a period of 25 days for WT, starting with day 4 after sowing.

3.1.1.1 Leaf development in WT

Leaf development in WT plants is divided into three different phases. A slowly increasing initial growth phase is followed by an exponential growth until day 19, after which leaf size remains approximately constant (Fig. 12). The first phase comprises mainly cell proliferation, whereas exponential growth in the second phase is due to cell expansion. After 19 days, the WT leaf reaches its mature size, which represents the third phase (Fig. 13).

Cell proliferation phase:

From day 6 to day 13 we observed a strong increase in cell number from 270 ± 32 cells to 14538 ± 811 cells, whereas cell size was not significantly affected. During this stage the leaf reached a size of $101.7 \pm 5.9 \text{ mm}^2$.

Cell expansion phase:

From day 13 to day 19, the increase in cell number stagnated. The cell size increased from 707.7 ± 54.9 to $1042.0 \pm 26.9 \text{ }\mu\text{m}^2$ during this time, and in parallel we observed an increase in leaf size from 101.7 ± 5.9 to $172.8 \pm 10.8 \text{ mm}^2$.

Mature phase:

In the mature phase, the final leaf size increased from 172.8 ± 10.8 to $197.7 \pm 13.2 \text{ mm}^2$ (14.4%) in WT plants, due to small changes in cell number and cell size.

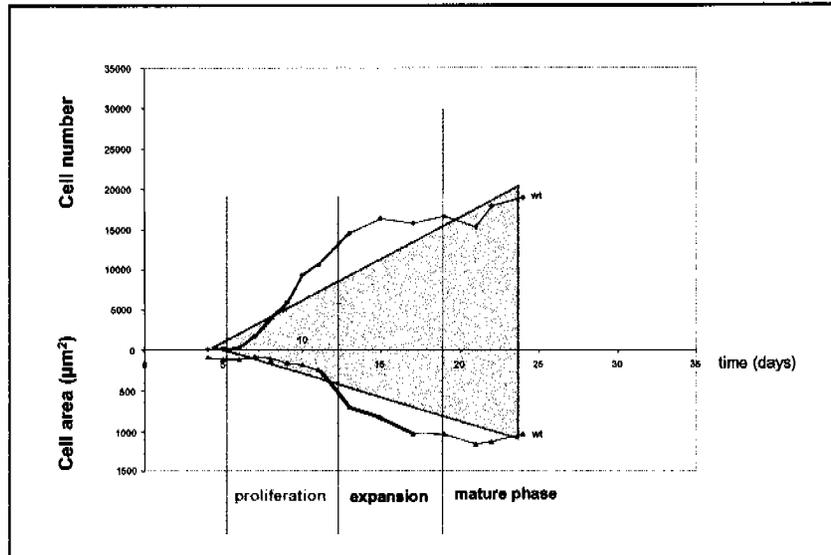


Figure 13: Kinematics II: Overview of leaf development in WT

A kinematics analysis of WT was performed over a time period of 25 days. The triangle in brown represents the leaf area. The development of a leaf is divided into three different phases: the proliferation phase, the expansion phase and the mature phase. The progression of the cell numbers are represented in the upper part, whereas the progression of the cell area is in the lower part.

3.1.1.2 Leaf development in *era1-2*

In principle, *era1-2* followed the same kinetics, and leaf development in the mutant can also be subdivided into three different phases as described for WT (Fig. 14).

Cell proliferation phase:

Cell proliferation in *era1-2* leaves occurred with a two day delay and can be subdivided into 2 distinct phases that were separated by the cell expansion phase. We observed a first proliferation phase from day 8 to day 15 and an additional and therefore second proliferation phase from day 21 to day 26,

whereas cell size was not significantly changed during these phases. From day 8 to day 15 the leaf reached a size of $77.9 \pm 1.3 \text{ mm}^2$ and after 26 days the leaf size was found to be $297.8 \pm 15.9 \text{ mm}^2$. Compared to WT, the final cell number after 26 days was found to be ~57% higher in *era1-2* leaves (18918 ± 703 cells in WT; 29784 ± 2455 cells in *era1-2*).

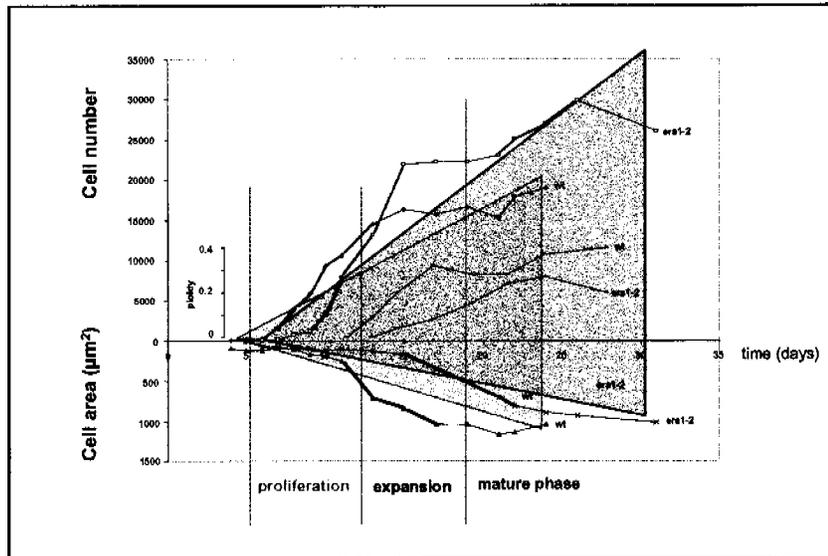


Figure 14: Kinematics II: Overview of leaf development in WT and *era1-2*

A kinematics analysis of WT and *era1-2* leaves was performed over a time period of 32 days for *era1-2* and over a period of 25 days for WT. The triangle in brown represents the leaf area in WT; the triangle in green represents the leaf area in *era1-2*. The endoreplication curve is represented in black and depicted only for 8C cells.

Cell expansion phase:

From day 15 to day 21 we observed an increase in cell size from 358.5 ± 16.0 to $809.3 \pm 38.4 \text{ µm}^2$ in *era1-2*, whereas the cell number remained constant during this time. The final cell size of $1262.0 \pm 33.0 \text{ µm}^2$ resembled the size of $1167.6 \pm 81.3 \text{ µm}^2$ in WT leaves. Compared to WT, the cell expansion phase in *era1-2* was initiated with fewer cells having undergone endoreplication cycles: in WT we observed 30% 2C cells, 61% 4C cells and 9% 8C cells, whereas in *era1-2*, 55% 2C cells, 41% 4C cells and 4% 8C cells.

Mature phase:

era1-2 leaves reached the mature phase with a delay of 7 days compared to WT. The final leaf size was $327.2 \pm 16.4 \text{ mm}^2$ after which cell size and cell number remained constant.

3.2 Gene expression analysis using Affymetrix GeneChips®

To establish a genomic scale analysis of gene function in Arabidopsis leaves, we performed expression profiling experiments using Affymetrix GeneChips® (ATH1-12150), representing 22746 genes of the Arabidopsis genome. Experiments were performed over a time course of 23 days. The RNA was extracted from the first and second leaf during the proliferation phase at day 9 for WT and day 11 for *era1-2*; during the expansion phase at day 15 for WT and day 17 for *era1-2*, and during the mature phase at day 21 for WT and during the second proliferation phase at day 23 for *era1-2* (for a total of 6 microarray hybridisations). In addition, for each individual hybridisation, RNA samples from three different leaves of three independent plants were extracted and pooled (a total of 18 independent RNA extractions).

For each microarray, overall intensity normalization was performed as described in Materials and Methods. For further analysis, only genes were used showing a present call over all time points, which indicates that the transcripts were reliably detected.

3.2.1 Global analysis of gene expression data

Among the 22746 genes represented on the Affymetrix ATH1-12150 GeneChip®, 9724 were detected with a present call over all experiments. Based on the two-day delay in germination between *era1-2* and WT plants, *era1-2* at day 11, 17 and 23 was compared with WT at day 9, 15 and 21, respectively. Within the comparisons, 4771 genes show at least once a significant 2-fold

change in gene expression. Among these 4771 genes, the proteins encoded by 1118 genes were predicted to be targeted to the chloroplast, 2519 to the cytoplasm, 480 to the mitochondrion and 654 to secretory pathways (Fig. 15).

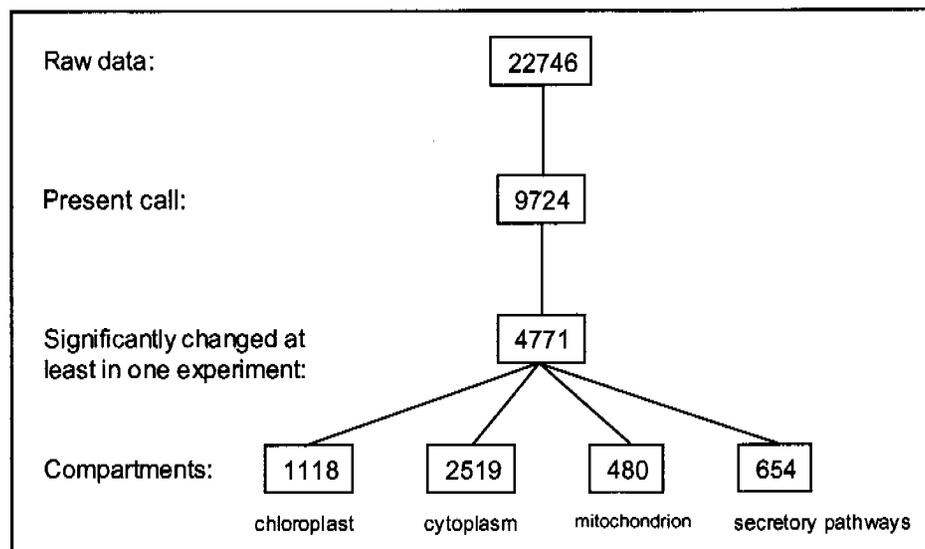


Figure 15: Raw data analysis in four steps

This analysis includes raw data collection, detection call analysis (present call), significant changes (at least 2-fold), compartment predictions of the corresponding proteins. Boxes represent the number of genes, obtained by the respective analysis.

3.2.1.1 Cluster analysis of co-regulated genes

To obtain insight into the overall behaviour of gene expression levels in *era1-2* and WT, a cluster analysis according to Eisen *et al.* (1998) was performed. A hierarchical clustering was used to structure the pre-clustered dataset obtained by using self organizing maps (SOM). The analysis was done by a pair wise comparison of growth stages within WT and *era1-2*, and additionally between *era1-2* and the respective WT growth stages, as shown in Table 4. The results are shown in a matrix (Fig. 16), in which each column represents a comparison between two experiments and each row represents a single gene. The binary grouping of genes within the clusters is symbolized by the cluster

wt 15d vs wt 9d	wt 21d vs wt 15d	era1-2 17d vs era1-2 11d	era1-2 23d vs era1-2 17d	era1-2 11d vs wt 9d	era1-2 17d vs wt 15d	era1-2 23d vs wt 21d
1	2	3	4	5	6	7

Table 4:
Pairs of combined experiments

tree. The changes in gene expression are represented by a colour-code according to Eisen *et al.* (1998). Red is used for genes with a higher level in expression compared to the respective control, whereas green is used for genes with a reduced level in gene expression compared to the control.

The cluster analysis was done with 4771 genes, which showed at least in one comparison a 2-fold change in gene expression (Fig. 16).

3.2.1.1.1 Global gene expression analysis during leaf development in WT

To investigate in changes in gene expression during WT leaf development, leaves at day 15 were compared to leaves at day 9, which represents the transition from the proliferation to the expansion phase. Leaves at day 21 were compared to leaves at day 15, which represents the transition to maturation (Fig. 16). Interestingly, only about 20% of all genes followed the same pattern in gene expression during the entire leaf development, which is indicated by the same colouring in both columns. As a consequence, 80% of the selected genes were found to be differentially regulated during leaf maturation when compared to the transition from the proliferation phase to the expansion phase. The genes were found either to be up-regulated during the first transition and down-regulated during the second transition, or vice versa.

3.2.1.1.2 Global gene expression analysis during leaf development in *era1-2*

Changes in gene expression in *era1-2* were analyzed in the same way as WT: *era1-2* leaves at day 17 were compared to leaves at day 11 and leaves at day 23 were compared to leaves at day 17. In general, gene expression patterns during *era1-2* leaf development were similar to WT. Thus, at a global scale gene expression patterns follow a similar trend in WT and *era1-2* leaves. The following analysis will therefore focus on the comparison of changes in gene

expression between WT and *era1-2* to reveal possible differences in gene expression to explain the difference in leaf size.

3.2.1.1.3 Differences between *era1-2* and WT gene expression profiles

To detect changes in gene expression, *era1-2* leaf development stages were compared to the respective WT leaf development stages. During the proliferation phase as well as the expansion phase, several genes were found to be significantly changed in their expression between WT and the mutant. Between *era1-2* leaves at day 11 and WT leaves at day 9, 82 genes were down-regulated and 36 genes were up-regulated, whereas between *era1-2* leaves at day 17 and WT leaves at day 15, 98 genes were down-regulated and 70 genes were up-regulated.

During maturation, however many more genes were detected with changes in gene expression between WT and the mutant. Between *era1-2* leaves at day 23 and WT leaves at day 21, 704 genes were down-regulated and 1118 were up-regulated (Fig. 16). This apparent induction in gene expression is the result of a stronger down-regulation in the expression of the corresponding genes during maturation in WT leaf development than in *era1-2* leaf development. A reduction in gene expression during maturation between *era1-2* and WT is consequently a result of a stronger up-regulation of the corresponding genes in WT than in *era1-2*.

Genes, which show changes in gene expression during maturation were also analysed in terms of the predicted localization of the corresponding proteins. Five hundred seventeen nuclear-encoded genes code for proteins predicted for chloroplast localization. Within this group of genes 420 were up-regulated and 97 were down-regulated. The corresponding proteins of 892 genes were predicted to localize to cytoplasm, 163 to the mitochondrion and 251 to the secretory pathway (Fig. 16).

Together, the loss of protein-farnesylation in *era1-2* resulted mainly in an increased expression of nuclear genes coding for proteins predicted to localize to chloroplasts.

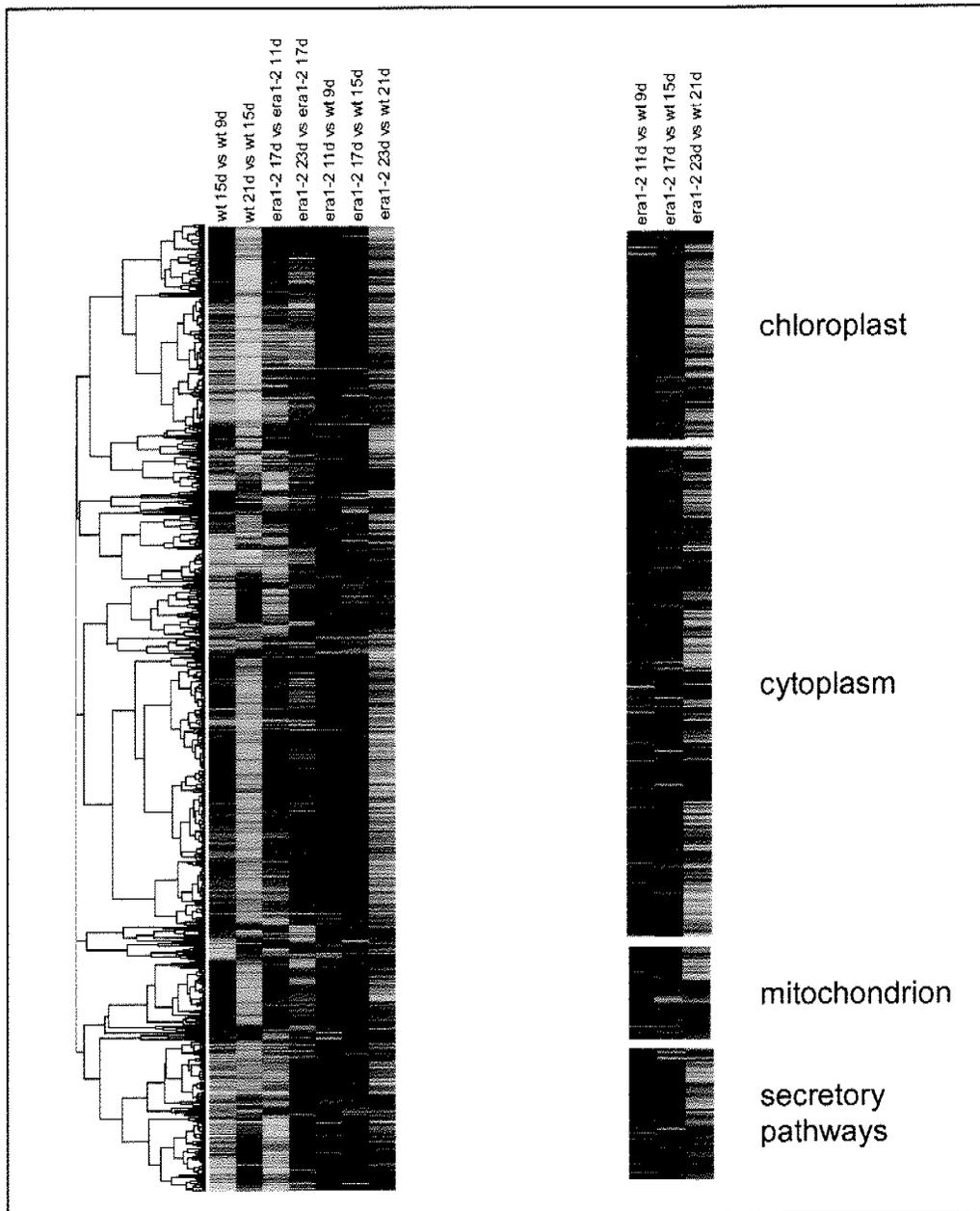


Figure 16: Cluster gram of gene expression during leaf development

Cluster gram of 4771 genes during leaf development in *era1-2* and WT (left) and separated for each compartment prediction of the corresponding protein (right). Each column represents the comparison of two experiments. Gene expression profiles are in rows: red, transcriptional up-regulation; green: down-regulated compared to the respective control. The cluster tree illustrates the nodes of co-regulation within gene expression over all comparisons.

3.2.1.2 Analysis of genes involved in leaf development

During maturation, after 20 days in leaf development, an additional cell proliferation phase is responsible for the enlarged leaves in *era1-2*. In addition to the kinematics results, most of the changes in gene expression between WT and *era1-2* were also observed during this time period. Therefore we focused our analysis on genes with changes in gene expression during maturation and that have been reported to function in leaf development, especially in leaf expansion, in cell cycle and cell expansion.

3.2.1.2.1 Genes involved in leaf expansion

Four genes have been described to control leaf expansion: *ANGUSTIFOLIA* (*AN*), *ROTUNDIFOLIA* (*ROT3-1*), *CURLY LEAF* (*CLF*) and *AGAMOUS* (*AG*) (Rédei, 1962; Tsuge *et al.*, 1996; Kim *et al.* 1998; Yanofsky *et al.*, 1990). For each of them, a knockout mutant has been characterized (Rédei, 1962; Tsuge *et al.*, 1996; Kim *et al.* 1998). The *an* mutant has a specific defect in the elongation of cells in the transverse (leaf-width) direction of the leaf (Tsukaya *et al.*, 1994; Tsuge *et al.*, 1996). The *rot3-1* has a defect in the elongation of leaf cells in the leaf-length direction (Tsuge *et al.*, 1996), whereas the double mutant *clf ag* has larger leaves than the WT (Goodrich *et al.*, 1997). Expression of the *AN* gene is not significantly changed in *era1-2*. Expression calls for *ROT3-1*, *CLF* and *AG* were absent and therefore the raw expression data were not considered in the analysis (Table 5).

Table 5: Genes involved in leaf expansion

Grey boxes represent genes, for which expression calls were absent. The raw expression data were therefore and not considered in the analysis.

ChIPID	Description	<i>era1-2</i> 11d vs WT 6d	<i>era1-2</i> 17d vs WT 16d	<i>era1-2</i> 23d vs WT 21d
259438_at	ANGUSTIFOLIA	0.94	1.15	1.10
246219_at	ROTUNDIFOLIA	4.46	1.34	0.73
287129_at	CURLY LEAF	0.97	1.43	2.19
254669_at	AGAMOUS	0.24	1.59	0.48

3.2.1.2.2 Genes involved in cell proliferation and cell expansion

Genes involved in cell cycle and cell elongation were studied in terms of changes in their expression during leaf maturation. In *era1-2* leaves all genes shown in Table 6 were found to be up-regulated at day 23 compared to WT at day 21. Among this group are genes encoding cyclin D3;1 and D3;3, CAK1At (cyclin-dependent kinase activator 1) and AINTEGUMENTA (*ANT*), which are known to be involved in cell proliferation. The expression of *CYCD3;1* in Arabidopsis is associated with proliferating tissues and is induced by brassinosteroids (BRs). CAK1At is a CDK activating kinase that promotes cell cycle progression, and ANT supports cell proliferation by increasing *CYCD3;1* expression (Dewitte *et al.*, 2003; Umeda *et al.*, 1998; Klucher *et al.*, 1996). With regard to the control of cell expansion, we found the cytokinin-responsive *STUNTED PLANT 1 (STP1)* (Baskin *et al.*, 1995) gene and the steroid biosynthesis *DIMINUTO/DWARF1 (DIM/DWF1)* to be up-regulated. Both genes are known to be involved in the control of cell expansion (Takahashi *et al.*, 1995; Klahre *et al.*, 1998). A basic requirement for cell expansion is the controlled remodelling of cell wall components. Therefore we examined the expression of genes involved in cellulose biosynthesis and cell wall matrix modification. We found one gene encoding a cellulose synthase and four pectinesterase genes that were up-regulated at day 23 in *era1-2*. We also found four genes that encode xyloglucan modifying proteins to be up-regulated. Among the cell matrix modifying proteins, four genes encoding α -expansins and a gene encoding a β -expansin were up-regulated. Genetic studies have led to the identification of proteins that play a critical role in determining cell expansion by regulating microtubules and actin dynamics (Azimzadeh *et al.*, 2001). The Arabidopsis gene *KATANIN* is involved in cortical cytoskeleton organization (McNally and Vale, 1993) and is among the group of up-regulated genes. Furthermore, three genes coding for members of the tubulin family and a gene encoding the actin depolymerizing factor 4 were up-regulated.

Together, these results suggest that at the time when WT leaves enter the maturation phase, leaf one and two in *era1-2* appear to reinitiate a juvenile program of cell proliferation.

ChpID	gene name	<i>era1-2 11d vs WT 9d</i>	<i>era1-2 17d vs WT 15d</i>	<i>era1-2 23d vs WT 21d</i>
253270_at	cyclin D3.1	1.60	1.69	3.13
252189_at	cyclin D3.3	1.26	1.85	2.00
253756_at	Cdk-activating kinase CAK1A1	1.13	1.09	2.13
263010_at	aintegumenta (ANT)	1.15	0.97	6.79
262456_at	STUNTED plant	0.55	0.42	2.59
267938_at	cell elongation protein, Dwarf1	1.04	1.40	2.09
253428_at	cellulose synthase catalytic subunit (RSW1)	0.83	0.93	2.09
268760_at	putative peclnacetylsterase	0.97	0.99	2.10
266735_at	putative peclnesterase	0.90	1.31	2.13
247239_at	pectin methylsterase-like protein	1.02	0.77	2.00
251181_at	putative protein peclnesterase homolog	0.88	0.91	2.03
252907_at	xyloglucan endo-transglycosylase	0.47	0.90	2.34
263841_at	putative xyloglucan endo-transglycosylase	1.08	0.86	3.98
247162_at	xyloglucan endo-transglycosylase-like protein	0.76	0.95	4.81
253040_at	endo-xyloglucan transferase	0.60	0.91	2.13
266790_at	expansin AtEx6	0.84	1.12	2.27
267158_at	putative expansin	0.55	0.98	2.56
268003_at	expansin At-EXP5	0.86	1.19	3.68
255822_at	putative expansin	0.38	0.79	3.54
253815_at	putative beta-expansin (cim1)	0.68	1.43	2.32
246056_at	katanin p80 subunit	1.04	1.09	2.25
261639_at	tubulin alpha-2/alpha-4 chain	1.02	0.97	2.18
261230_at	beta tubulin 1	0.93	0.91	2.81
260317_at	tubulin beta-6 chain (sp P29514)	1.01	1.06	2.76
247856_at	actin depolymerizing factor 4	0.77	0.90	2.03

Table 6: Genes involved in developmental processes found to be up-regulated in *era1-2* compared to the WT control (red: up-regulated).

3.2.2 Gene-specific analysis of gene expression data

By using the kinematics approach we observed that the enlarged leaf sizes in *era1-2* are caused by an additional proliferation phase after 20 days in leaf development. This additional proliferation phase agrees with changes in gene expression during maturation in *era1-2* at day 23 compared to WT at day 21, during which genes involved in cell cycle and cell elongation were up-regulated in *era1-2*.

The loss of protein-farnesylation therefore results in an up-regulation of genes involved in cell cycle and cell elongation during maturation in *era1-2* leaf development. As a consequence, the further analysis focused on changes in gene expression during the time period between day 11 and day 23. To integrate all three time points for *era1-2* and WT we compared transition stages in addition to developmental phases. The transition from the proliferation phase

to the expansion phase in *era1-2* (day 11 compared to day 17) was compared to the same transition in WT (day 9 compared to day 15); and the transition between the expansion phase and the mature phase in *era1-2* (day 17 compared to day 23) was compared to the corresponding transition in WT (day 15 compared to day 21). The resulting data set contains three different groups of genes. A first group of genes shows changes in expression only at the first transition, a second shows changes at both transition stages, and a third group of genes that shows only changes in expression at the second transition from the expansion phase to the mature phase (Fig. 17).

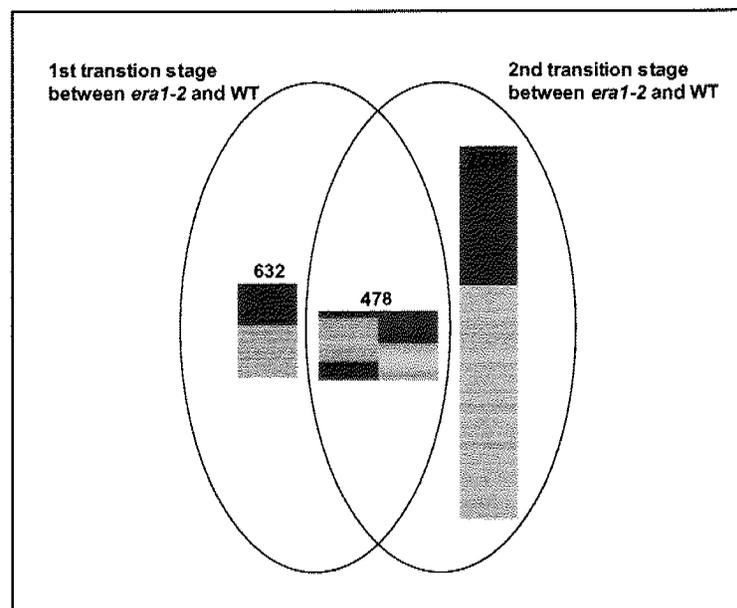


Figure 17: Combination of a Venn-diagram cluster gram analysis

Numbers represent the total number of genes for a specific transition stage in leaf development.

Approximately 4 times more genes (2510) were altered in their expression levels at the second transition compared to the first transition stage (632). Only 478 genes showed changes in gene expression at both transition stages, and among them, 125 genes were up- and 59 genes were down-regulated at both transitions and 294 genes were either up-regulated during the first transition and down-regulated at the second or vice versa.

The three groups of genes shown in Fig. 17 were used to find specific genes for biochemical and signal transduction pathways using different databases in combination with the Metabolic Pathway Map for Arabidopsis (see Materials

and Methods). Furthermore, the three groups of genes were studied in terms of localization prediction for the encoded proteins.

3.2.2.1 Specific analysis of genes for biochemical pathways

The three groups of genes (Fig. 17) were analysed for genes encoding proteins involved in biochemical pathways. Thirty-four genes for biochemical pathways were found to be regulated at the transition from the proliferation phase to the expansion phase and 160 genes were found with changes in their expression at the transition from the expansion phase to the mature phase. A total number of 38 genes were identified with significantly altered levels in gene expression at both transition stages (Fig. 18).

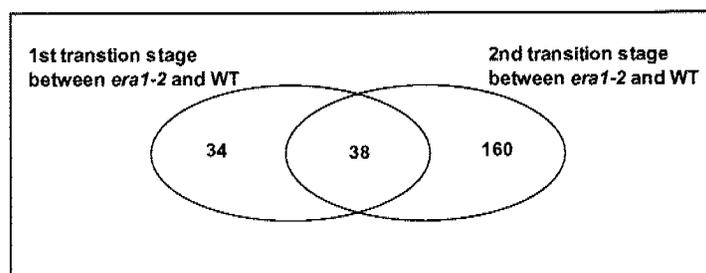


Figure 18: Venn-diagram of genes involved in metabolic pathways during leaf development

Genes involved in metabolic pathways can be divided into groups based on their changes in expression at the corresponding transition stage in leaf development.

Protein-farnesyltransferase (PFT) uses the isoprenoid farnesylpyrophosphate (FPP) to modify target proteins. Therefore, genes for pathways involved in the synthesis of the common isoprenoid precursor isopentenylpyrophosphate (IPP) or pathways downstream of the IPP biosynthesis were analysed in more detail.

3.2.2.1.1 Changes in gene expression at the first transition stage in leaf development

Changes expression at the transition from the proliferation to the expansion phase were analysed for genes coding for protein involved in biochemical

pathways. Among 34 genes, 27 were up-regulated and only 7 genes were down-regulated (Table 7). Interestingly, all nuclear encoded genes coding for proteins predicted for chloroplast localization were up-regulated. Within this group are genes coding for enzymes involved in the Calvin cycle, glycolysis and fatty acid biosynthesis. The gene encoding the mitochondrial-localized farnesylpyrophosphate synthase 1 (FPS1) was also up-regulated.

	ChIPID	atg#	Enzyme	1st transition stage	Pathway	Metabolism	
chloroplast	249615_at	At5g38530	Tryptophan synthase beta	2.27	Tryptophan Metabolism	Amino acid Metabolism	
	255720_at	At1g32060	Phosphonobulokinase	2.08	Calvin Cycle	Carbohydrate Metabolism	
	248283_at	At5g52920	Pyruvate kinase	2.09	Glycolysis/Gluconeogenesis	Carbohydrate Metabolism	
	265026_at	At1g24360	3-Ketoacyl-ACP reductase	2.11	Fatty acid biosynthesis	Lipid Metabolism	
	266035_at	At2g05990	2,3-trans-Enoyl-ACP reductase	3.04			
	267308_at	At2g03200	Malonyl-CoA:ACP transacylase	2.35	Purine Metabolism	Nucleotide Metabolism	
	253252_at	At4g34740	Amidophosphoribosyltransferase 2 precursor	2.32			
	246490_at	At5g15950	S-adenosylmethionine decarboxylase (adoMetDC2)	2.30	Polyamine Metabolism		
	254740_s_at	At4g13890	glycine hydroxymethyltransferase - like protein	2.01	Serine / Glycine / Cysteine	Amino acid Metabolism	
	255946_at	At1g22020	glycine hydroxymethyltransferase, putative similar to	0.45			
259222_at	At3g03980	putative phosphoribosylanthranilate transferase	2.34	Tryptophan Metabolism			
cytoplasm	252022_at	At3g52830	Fructose-bisphosphate aldolase	2.01	Calvin Cycle		
	253135_at	At4g35830	Aconitase	2.16	Citric acid cycle	Carbohydrate Metabolism	
	253054_at	At4g26970	Aconitase	2.03			
	263988_at	At2g42790	Citrate synthase	0.44	Glycolysis/Gluconeogenesis		
	252022_at	At3g52890	Fructose-bisphosphate aldolase	2.01			
	264045_at	At2g22450	putative GTP cyclohydrolase	0.43	Riboflavin metabolism	Metabolism of cofactors and vitamins	
	250007_at	At5g18670	alpha-Amylase	2.00	Starch and sucrose metabolism	Metabolism of complex Carbohydrates	
	236681_at	At3g52340	Sucrose-phosphatase	2.21	Inositol phosphate metabolism	Metabolism of Complex Lipids	
	253002_at	At4g38530	Phospholipase C	2.48			
	247275_at	At5g34370	Beta-ureidoptopionase	0.49	beta-Alanine Metabolism	Metabolism of other amino acids	
	254926_at	At4g11280	ACC synthase (AIACS-6)	0.21	Ethylene biosynthesis	Plant hormones	
	mito	263706_s_at	At1g31180	isopropylmalate dehydrogenase	2.09	Leucine Metabolism	Amino acid Metabolism
		263563_at	At2g17130	isocitrate dehydrogenase	2.09	Citric acid cycle	Carbohydrate Metabolism
		248718_at	At5g17770	Geranyltransferase (FPP-synthase)	2.12	isoprenoid biosynthesis cytoplasmic pathway	Lipid Metabolism
	secretory pathways	263428_at	At5g22250	Aspartate aminotransferase	2.08	Aspartate Metabolism	Amino acid Metabolism
249266_at		At5g11670	beta-Phosphogluconate dehydrogenase	2.12	Pentose Phosphate cycle	Carbohydrate Metabolism	
247520_at		At5g11310	Cytochrome-c oxidase	2.21	Oxidative phosphorylation	Energy Metabolism	
250891_at		At5g04530	(3-ketoacyl-CoA synthase)	3.58	Fatty acid biosynthesis	Lipid Metabolism	
253285_at		At4g34250	(3-ketoacyl-CoA synthase)	3.05			
266938_at		At2g18950	Phytyl/prenyl transferase	2.17	isoprenoid biosynthesis plastidic pathway		
263083_at		At2g27190	purple acid phosphatase precursor	0.44	Riboflavin metabolism	Metabolism of cofactors and vitamins	
263563_at		At2g16430	putative purple acid phosphatase precursor	0.42			
263595_at		At2g01890	putative purple acid phosphatase	3.05			

Table 7: Pathway genes regulated at the first transition stage

Genes, which are found to be regulated at the transition from the proliferation to the expansion phase only.

3.2.2.1.2 Changes in gene expression at the second transition stage in leaf development

Many more genes (160) were identified whose expression changed at the transition from the expansion to the mature phase (Table 8a and 8b). Hundred and twelve genes were up-regulated and 48 were down-regulated. Sixty-nine genes were found coding for proteins predicted to localize to chloroplasts. Among this group of genes are genes encoding proteins involved in the Calvin

cycle, glycolysis, pentose phosphate cycle, fatty acid biosynthesis, isoprenoid biosynthesis, porphyrin and chlorophyll biosynthesis as well as starch and sucrose metabolism. The genes encoding those proteins were up-regulated. Three genes encoding proteins involved in the chloroplast-localized isoprenoid pathway were up-regulated: 1-D-deoxyxylulose 5-phosphate synthase, 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase and 2C-methyl-D-erythritol-phosphate-CMP kinase. Two genes coding for enzymes involved in phytosterol synthesis, squalene synthetase and cycloartenol synthase, were also up-regulated (Fig. 8a).

	ChpID	atg#	Enzyme	2nd transition stage	Pathway	Metabolism
chloroplast	264134_at	A14g24630	Argininosuccinate synthase	0.35	Arginine	
	256046_at	A13g21220	Asparagine synthetase	0.24	Aspartate Metabolism	
	258281_at	A13g26900	putative strikimate kinase	3.50	Chorismate Metabolism	
	262177_at	A11g74710	isochorismate mutase / isochorismate synthase	0.06	Cysteine biosynthesis	
	260566_at	A12g43750	Cysteine synthase	3.13	Glutamate / Glutamine Metabolism	
	245701_at	A15g04140	ferredoxin-dependent glutamate synthase	2.16	Homidine Metabolism	
	245287_at	A14g14910	imidazoleglycerol-phosphate dehydratase	2.43	Isoleucine Metabolism	
	256882_at	A13g23940	Dihydroxy-acid dehydratase	0.24	Lysine Metabolism	Amino acid Metabolism
	251392_at	A13g06060	Dihydrodipicolinate synthase	0.30	Serine / Glycine / Cysteine	
	253192_at	A14g35630	Phosphoserine transaminase	0.22	Aspartate kinase	
	253274_at	A14g34200	Phosphoglycerate dehydrogenase	0.27	Aspartate kinase	
	250291_at	A15g13290	Aspartate kinase	0.30	Aspartate semialdehyde dehydrogenase	
	262841_at	A11g14810	Aspartate semialdehyde dehydrogenase	0.49	anthranilate synthase beta chain	
	247604_s_at	A15g57890	anthranilate synthase beta chain	0.42	Indole-3-glycerol-phosphate synthase (IGPS)	
	246688_at	A15g48220	Indole-3-glycerol-phosphate synthase (IGPS)	2.02	putative indole-3-glycerol phosphate synthase	
	263807_at	A12g04400	putative indole-3-glycerol phosphate synthase	0.38	Dihydroxy-acid dehydratase	
	256882_at	A13g23940	Dihydroxy-acid dehydratase	0.24	3-Phosphoglycerat kinase	
	256226_at	A11g56190	3-Phosphoglycerat kinase	4.24	3-Phosphoglycerat kinase	
	257699_at	A13g12790	3-Phosphoglycerat kinase	2.87	Ribose 5-phosphate isomerase	
	259096_at	A13g04730	Ribose 5-phosphate isomerase	2.41	glucose-6-phosphate isomerase	
	254111_at	A14g24620	glucose-6-phosphate isomerase	3.16	Pyruvate dehydrogenase	
	256190_at	A11g30120	Pyruvate dehydrogenase	2.47	3-Phosphoglycerat kinase	
	256226_at	A11g56190	3-Phosphoglycerat kinase	4.24	Fructokinase	
	257699_at	A13g12790	3-Phosphoglycerat kinase	2.87	Pyruvate dehydrogenase	
	260107_at	A11g56430	Fructokinase	2.28	Glucose 6-phosphate dehydrogenase	
	266904_at	A12g34590	Pyruvate dehydrogenase	7.99	glucose-6-phosphate isomerase	
	249694_at	A15g35790	Glucose 6-phosphate dehydrogenase	8.06	Ribose 5-phosphate isomerase	
	254141_at	A14g24620	glucose-6-phosphate isomerase	2.14	Ribose 5-phosphate isomerase	
	258096_at	A13g04790	Ribose 5-phosphate isomerase	2.41	Transaldolase	
	260967_at	A11g12230	Transaldolase	3.63	UDP-glucuronate decarboxylase	
	247720_at	A15g59290	UDP-glucuronate decarboxylase	3.65	ferredoxin-NADP+ reductase	
	247131_at	A15g08190	ferredoxin-NADP+ reductase	3.02	ferredoxin-NADP+ reductase	
	259230_at	A14g05390	ferredoxin-NADP+ reductase	0.44	Acetyl-CoA carboxylase	
	246613_at	A15g53360	Acetyl-CoA carboxylase	2.72	3-Hydroxyacyl-CoA dehydratase	
	250470_at	A15g10160	3-Hydroxyacyl-CoA dehydratase	2.84	Acetyl-CoA carboxylase	
	258289_at	A13g15690	Acetyl-CoA carboxylase	2.50	Farnesyltransferase (FGPP synthase)	
	252696_s_at	A14g38460	Farnesyltransferase (FGPP synthase)	3.22	1-D-deoxyxylulose 5-phosphate synthase	
	245281_at	A14g15660	1-D-deoxyxylulose 5-phosphate synthase	2.80	Lycopene epsilon-cyclase + Lycopene beta-cyclase	
	247936_at	A15g57030	Lycopene epsilon-cyclase + Lycopene beta-cyclase	3.66	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	
	260324_at	A11g36970	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	2.30	MEP-CMP kinase	
	266883_at	A12g26930	MEP-CMP kinase	2.07	Dihydroxy-acid dehydratase	
	256882_at	A13g23940	Dihydroxy-acid dehydratase	0.24	Acetolactate synthase	
	263490_at	A12g31810	Acetolactate synthase	3.94	Chlorophyll synthetase / Chlorophyllase	
	249308_at	A13g51820	Chlorophyll synthetase / Chlorophyllase	33.76	Farnesolase	
	246870_at	A15g78930	Farnesolase	0.16	Mg-protoporphyrin IX methyltransferase	
	254105_at	A14g25060	Mg-protoporphyrin IX methyltransferase	6.86	ALA dehydratase	
	250370_at	A11g69740	ALA dehydratase	3.39	Protochlorophyllide oxidoreductase	
	264830_at	A11g03630	Protochlorophyllide oxidoreductase	3.53	Farnesolase	
	264717_at	A12g30390	Farnesolase	2.47	GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone-4-phosphate synthase	
	247894_at	A15g59750	GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone-4-phosphate synthase	3.19	Alpha-dextrin endo-1,6-alpha-glucosidase	
	245712_at	A15g04390	Alpha-dextrin endo-1,6-alpha-glucosidase	3.68	Cellulose synthase (DesA)	
	246460_at	A15g16910	Cellulose synthase (DesA)	0.38	4-alpha-glucanotransferase (Glucosyltransferase (D-enzyme))	
	247216_at	A15g64860	4-alpha-glucanotransferase (Glucosyltransferase (D-enzyme))	8.89	Starch synthase	
	249785_at	A15g24300	Starch synthase	2.63	Glucose-1-phosphate adenylyltransferase	
	249927_at	A16g13220	Glucose-1-phosphate adenylyltransferase	4.66	1,4-alpha-glucan branching enzyme	
	259896_at	A15g03690	1,4-alpha-glucan branching enzyme	2.37	alpha Amylase	
	260412_at	A11g69830	alpha Amylase	2.39	Glycerol-3-phosphate dehydrogenase	
	245790_at	A11g32290	Glycerol-3-phosphate dehydrogenase	3.21	Phosphatidylserine decarboxylase	
	254035_at	A14g25970	Phosphatidylserine decarboxylase	3.09	Glutathione synthetase	
	246785_at	A15g27390	Glutathione synthetase	2.12	Chalcone isomerase	
	251827_at	A13g55120	Chalcone isomerase	0.48	Gusnylate kinase	
	256401_at	A13g06200	Gusnylate kinase	0.36	Adenylosuccinate lyase like protein	
	256481_s_at	A14g18440	Adenylosuccinate lyase like protein	2.04	SAICAR synthetase	
	256978_at	A13g21110	SAICAR synthetase	0.38	UMP kinase	
	254052_at	A14g28280	UMP kinase	3.07	lipoxygenase A1, OXG	
	252618_at	A13g48140	lipoxygenase A1, OXG	0.43	allene oxide cyclase	
	259396_at	A11g13280	allene oxide cyclase	3.48	12-oxophytoylsuccinate reductase (OPR3)	
	265530_at	A12g06090	12-oxophytoylsuccinate reductase (OPR3)	0.44	isochorismate synthase	
	262177_at	A11g74710	isochorismate synthase	0.08		

Table 8a: Pathway genes regulated at the second transition stage
Regulated nuclear encoded genes coding for proteins predicted to localize to chloroplasts.

	ChiplD	atg#	Enzyme	2nd transition stage	Pathway	Metabolism	
cytoplasm	245901_at	AT5g19550	Aspartate aminotransferase	0.37	Aspartate Metabolism	Amino acid Metabolism	
	247216_at	AT5g85010	Asparagine synthetase	3.28			
	252415_at	AT3g47340	Asparagine synthetase	30.19			
	245951_at	AT5g19550	aspartate aminotransferase Asp2	0.37			
	262933_at	AT1g65940	Polyamine oxidase	0.42			
	248248_at	AT5g33120	sermidine synthase	3.27			
	246584_at	AT5g14800	P5C reductase	0.22			
	253438_at	AT4g32520	glycine hydroxymethyltransferase (EC 2.1.2.1) - like protein	0.26			
	256403_at	AT1g17745	Phosphoglycerate dehydrogenase	0.10			
	245745_at	AT1g51115	similarity to anthranilate synthase large subunit	2.86			
	255487_at	AT4g02510	Tryptophan synthase alpha	2.50			
	253971_at	AT4g26530	Fructose-bisphosphate aldolase	3.23	Calvin Cycle	Carbohydrate Metabolism	
	265742_at	AT2g01290	Ribose 5-phosphate isomerase	0.42			
	246461_s_at	AT5g50590	Fumerase	2.48			
	247969_at	AT5g36350	Pyruvate kinase	0.35			
	253836_at	AT4g27800	6-phosphofruktokinase	3.14			
	253971_t_at	AT4g26530	Fructose-bisphosphate aldolase	3.23	Glycolysis/Gluconeogenesis	Carbohydrate Metabolism	
	256672_at	AT3g06590	Phosphoglycerate mutase	0.37			
	265742_at	AT2g01290	Ribose 5-phosphate isomerase	0.42	Penicose Phosphate cycle	Energy Metabolism	
	262174_at	AT1g74910	GDP-mannose pyrophosphorylase	4.74			
	250235_s_at	AT5g13430	Ubiquinol-cytochrome C reductase	0.49	Oxidative phosphorylation	Energy Metabolism	
	256966_s_at	AT2g28715	ATP synthase	2.23			
	261570_at	AT1g01120	(3-ketoacyl-CoA synthase)	2.27	Fatty acid biosynthesis	Lipid Metabolism	
	263192_at	AT1g36160	Acetyl-CoA carboxylase	2.04			
	263432_at	AT2g22230	3-Hydroxyacyl-ACP dehydratase	5.07	Fatty Acid Oxidation	Lipid Metabolism	
	246304_at	AT3g51840	Acyl-CoA oxidase	3.19			
	246778_at	AT5g27450	Mevalonate kinase	2.07	Isoprenoid biosynthesis cytoplasmic pathway	Lipid Metabolism	
	248511_at	AT5g50375	Cycloartenol cyclisomerase	4.98			
	246990_at	AT5g49230	Acetyl-CoA C-acyltransferase (ThiOase)	3.76			
	253206_at	AT3g34640	Squalene synthase	2.35			
	268485_at	AT2g07050	Cycloartenol synthase	3.08			
	259959_at	AT1g52340	Xanthoxin oxidase	2.23	Isoprenoid biosynthesis plastidic pathway	Lipid Metabolism	
	251759_at	AT3g55630	tetrahydrofolylpolyglutamate synthase precursor - like protein	2.53			
	256762_at	AT0g11750	dihydroorotidin aldolase-like protein	2.17	Folate biosynthesis	Metabolism of cofactors and vitamins	
	261664_s_at	AT1g00480	Fomyltetrahydrofolate synthetase	2.85			
	264820_at	AT1g00475	Coproporphyrinogen III oxidase	2.81	Porphyrin and chlorophyll biosynthesis	Metabolism of cofactors and vitamins	
	257067_at	AT3g20500	purple acid phosphatase, putative	2.25			
	245094_at	AT2g40640	4-alpha-glucanotransferase (Glucosyltransferase (D-enzyme))	2.83	Starch and sucrose metabolism	Metabolism of complex Carbohydrates	
	248425_at	AT5g17420	Cellulose synthase (CesA)	2.20			
	247251_at	AT5g64740	Cellulose synthase (CesA)	2.59			
	253627_at	AT5g05130	Cellulose synthase (CesA)	2.33			
	253586_at	AT4g39330	Cellulose synthase (CesA)	2.45			
	253428_at	AT4g32410	Cellulose synthase (CesA)	2.83			
	267115_s_at	AT2g32840	Cellulose synthase (CesA)	0.45			
	261309_at	AT1g48500	Phosphoethanolamine N-methyltransferase	2.89			Glycerolipid metabolism
262367_s_at	AT1g49340	Phosphatidylinositol 4-kinase	2.33				
248248_at	AT5g51120	sermidine synthase	5.27	beta-Alanine Metabolism			Metabolism of other amino acids
249078_at	AT5g44070	Phytochelatin synthase	0.46				
262838_at	AT2g36880	putative s-adenosylmethionine synthetase	3.05	Selenocysteine acid Metabolism	Metabolism of other Substances		
245690_at	AT5g04230	PAL 3	2.99				
252943_at	AT4g39330	Cinnamoyl alcohol dehydrogenase (NADPH)	2.89	Phenylpropanoid Metabolism	Metabolism of other Substances		
246651_at	AT5g35170	putative adenylate kinase	3.08				
256092_at	AT4g09320	Nucleoside-diphosphate kinase	0.43	Purine Metabolism	Nucleotide Metabolism		
267537_at	AT2g11860	Guanylate kinase	0.42				
252079_at	AT4g04950	Dihydropyrimidine	0.50	Pyrimidine Metabolism	Nucleotide Metabolism		
256875_s_at	AT1g79660	12-oxophylloquinone reductase (OPR2)	0.15				
245690_at	AT5g04230	PAL3	2.99	Oxalate/oxonol Pathway / Amino acid Biosynthesis	Pathogen metabolism		
252827_at	AT4g39950	cytochrome P450 - like protein : CYP79B	0.33				
260637_at	AT1g67380	ACC oxidase, putative	2.25	Ethylene biosynthesis	Plant hormones		
263838_at	AT2g36880	putative s-adenosylmethionine synthetase	2.85				
265194_at	AT1g09010	1-aminoacylpropane-1-carboxylate oxidase	0.46	Terpene Biosynthesis	Terpenoid Metabolism		
253206_at	AT4g34640	farnesyl-diphosphate farnesyltransferase (squalene synthase 1)	2.30				
266495_at	AT2g07050	Cycloartenol synthase, (S)-2,3-epoxy-squalene mutase	3.08	Aspartate Metabolism Serine / Glycine / Cysteine Calvin Cycle	Amino acid Metabolism		
259710_at	AT1g77670	Aspartate aminotransferase	0.43				
246800_at	AT5g28700	glycine hydroxymethyltransferase - like protein	3.18				
247278_at	AT2g48360	Fructose-1,6-bisphosphatase	4.08				
253950_at	AT4g26910	Dihydrodipicolinate succinyltransferase	0.37				
256439_at	AT3g17240	Lipoamide dehydrogenase	0.48				
247278_at	AT5g64380	Fructose-1,6-bisphosphatase	4.08				
258439_at	AT3g17240	Lipoamide dehydrogenase	0.45				
258774_at	AT3g15640	Cytochrome c oxidase	0.47				
260948_at	AT2g43360	Biotin synthase	2.95				
264394_at	AT1g11660	aminomethyltransferase-like precursor/protein	4.38	One carbon pool by folate	Metabolism of cofactors and vitamins		
250221_at	AT5g14220	Protocorphylinogen IX oxidase	2.04				
245441_at	AT4g16700	Phosphatidylserine decarboxylase	0.47	Inositol phosphate metabolism	Metabolism of Complex Lipids		
252459_s_at	AT3g47220	Phospholipase C	2.07				
256156_at	AT3g08100	Phospholipase C	3.56	Purine Metabolism	Nucleotide Metabolism		
254191_at	AT2g33900	Nucleoside-diphosphate kinase	2.32				
265956_at	AT2g37250	putative adenylate kinase	0.22	Citric acid cycle	Carbohydrate Metabolism		
251541_at	AT3g58750	Citrate synthase	2.88				
247924_at	AT5g57655	Xylose isomerase	2.78	Glycolysis/Gluconeogenesis	Carbohydrate Metabolism		
249099_at	AT5g18640	Triacylglycerol lipase	2.91				
250008_at	AT5g18630	Triacylglycerol lipase	3.08	Fatty acid biosynthesis	Lipid Metabolism		
245809_at	AT1g56440	Squalene oxidase	2.21				
256335_at	AT2g32440	ent-Kaurenic acid hydroxylase	2.27	Isoprenoid biosynthesis cytoplasmic pathway Isoprenoid biosynthesis plastidic pathway	Lipid Metabolism		
264250_at	AT1g78680	gamma-glutamyl hydrolase, putative	2.49				
246636_at	AT5g24830	acid phosphatase-like protein	2.31	Folate biosynthesis	Metabolism of cofactors and vitamins		
250336_at	AT3g11720	beta-Glucosidase	3.22				
252952_at	AT4g38570	Phosphatidylinositol synthase	3.23	Starch and sucrose metabolism	Metabolism of complex Carbohydrates		
251679_at	AT3g27030	putative oxoamino strictosidine synthase	2.26				
246880_s_at	AT5g25960	myrosinase TSG2	2.33	Monoterpenoid indole alkaloid biosynthesis	Metabolism of other Substances		
						Auxin biosynthesis	Plant hormones

Table 8b: Pathway genes regulated at the second transition stage
 Regulated nuclear encoded genes coding for proteins predicted to localize to the cytoplasm, the mitochondrion, or predicted to be involved in secretory pathways.

3.2.2.1.3 Changes in gene expression at both the first and the second transition stage in leaf development

Only 38 genes were found with changes in expression at both transition stages, during all developmental phases. Within this group of genes, 16 genes were up-regulated and 4 genes were down-regulated (Table 9). These 20 genes are probably regulated independently of the developmental stage of the leaf and their changes in expression might result only from the loss of farnesylation. A group of 18 genes was up-regulated at the first transition stage and down-regulated at the second transition stage or vice versa. The expression of these genes might be affected by the loss of farnesylation as well as the developmental stage of the leaf.

	ChiplD	atg#	Enzyme	1st transition stage	2nd transition stage	Pathway	Metabolism
chloroplast	249710_at	At5g35630	Glutamine synthetase GS	2.01	2.74	Glutamate / Glutamine Metabolism	Amino acid Metabolism
	266694_at	At3g10050	Threonine deaminase	2.25	0.35	Isoleucine Metabolism	
	245145_at	At2g45440	putative dihydrodipicolinate synthase	3.38	0.25	Lysine Metabolism	
	266671_at	At2g29690	anthranilate synthase, alpha subunit	2.22	2.68	Tryptophan Metabolism	
	251685_at	At3g54050	Fructose-1,6-bisphosphatase	2.02	2.77	Glycolysis/Gluconeogenesis	Carbohydrate Metabolism
	266086_at	At3g25660	Dihydrodiposimide S-Acetyltransferase	2.05	3.78	Oxidative phosphorylation	Energy Metabolism
	255046_at	At4g06850	H ⁺ transporting ATPase	2.37	3.49	Oxidative phosphorylation	Energy Metabolism
	250096_at	At5g17230	Phytoene synthase	2.29	2.84	Isoprenoid biosynthesis plastidic pathway	Lipid Metabolism
	262675_at	At1g04970	Tocopherol O-methyltransferase	0.49	2.02	Isoprenoid biosynthesis plastidic pathway	Lipid Metabolism
	247392_at	At5g36370	Glutamate-1-semialdehyd de aminotransferase	2.38	0.49	Porphyrin and chlorophyll biosynthesis	Metabolism of cofactors and vitamins
cytoplasm	267421_at	At2g36040	Phosphoribosylaminimidazolecarboxamide formyltransferase	0.49	0.27	Purine Metabolism	Nucleotide Metabolism
	256002_at	At1g29900	Carbamoylphosphate synthetase	2.38	0.30	Pyrimidine Metabolism	
	263637_at	At1g04970	Fructose-1,6-bisphosphatase	2.18	2.31	Glycolysis/Gluconeogenesis	Carbohydrate Metabolism
	263924_at	At2g36530	Enolase	2.43	0.27	Glycolysis/Gluconeogenesis	
	249732_at	At5g24420	6-Phosphoglucono lactonase	2.99	4.27	Pentose Phosphate cycle	
	249733_at	At5g24400	6-Phosphoglucono lactonase	2.02	2.19	Pentose Phosphate cycle	
	269517_at	At1g02630	Catalase	0.41	2.50	Fatty Acid Oxidation	Lipid Metabolism
	254445_at	At4g11820	Hydroxymethylglutaryl-CoA synthase	0.46	2.67	Isoprenoid biosynthesis cytoplasmic pathway	
	264799_at	At1g08550	Violaxanthin de-epoxidase	0.39	6.86	Isoprenoid biosynthesis cytoplasmic pathway	
	245242_at	At1g44440	Chlorophyll a oxygenase	2.18	2.36	Porphyrin and chloroplast biosynthesis	Metabolism of cofactors and vitamins
256109_at	At3g23910	α-Glucosidase	0.49	3.70	Starch and sucrose metabolism	Metabolism of complex Carbohydrates	
260699_at	At1g12240	Invertase (β-fructofuranosidase)	0.43	2.66	Starch and sucrose metabolism		
236613_at	At3g02870	myo-inositol-1-phosphatase	2.13	3.29	Inositol phosphate metabolism	Metabolism of Complex Lipids	
265725_at	At2g32010	Inositol polyphosphate 3-phosphatase	2.22	3.29	Inositol phosphate metabolism	Metabolism of Complex Lipids	
251643_at	At3g07350	Quercetinase	0.49	0.46	Purine Metabolism	Nucleotide Metabolism	
247735_at	At5g29440	Thymidylate kinase	0.50	0.50	Pyrimidine Metabolism		
246259_at	At1g23160	Phosphoribosyl-AMP cyclohydrolase	2.22	0.30	Histidine Metabolism	Amino acid Metabolism	
mito	246038_at	At5g56070	Dihydrodiposimide succinyltransferase	2.16	0.23	Citric acid cycle	Carbohydrate Metabolism
	249828_at	At5g23250	Succinyl-CoA synthetase	2.02	0.41	Citric acid cycle	
	254737_at	At4g13840	Fatty Acid Elongase	2.02	2.18	Fatty acid biosynthesis	Lipid Metabolism
	260655_at	At1g21920	Phosphatidylinositol-4-phosphate 5-kinase	0.42	2.39	Inositol phosphate metabolism	Metabolism of Complex Lipids
secretory pathway	253039_at	At4g37760	Squalene epoxidase	0.44	2.39	Isoprenoid biosynthesis cytoplasmic pathway	Lipid Metabolism
	245637_at	At1g25230	putative purple acid phosphatase precursor	2.13	2.22	Inositol phosphate metabolism	Metabolism of Complex Lipids
	262630_at	At1g14700	purple acid phosphatase, putative	2.23	3.94	Inositol phosphate metabolism	Lipid Metabolism
	249046_at	At5g44400	berberine bridge enzyme	2.22	2.24	Riboflavin metabolism	Metabolism of cofactors and vitamins
	260391_at	At1g74020	putative striclosidine synthase	0.45	0.42	Biosynthesis of Berberine	Metabolism of other Substances
	257702_at	At3g12870	GTP synthetase	2.24	0.25	Monoterpenoid indole alkaloid biosynthesis	Metabolism of other Substances
	257702_at	At3g12870	GTP synthetase	2.24	0.25	Pyrimidine Metabolism	Nucleotide Metabolism
	263534_at	At4g51500	cytochrome P450 monooxygenase / CYP763B	2.18	0.43	Auxin biosynthesis	Plant hormones

Table 9: Pathway genes regulated at the first and the second transition stage
Genes coding for enzymes involved in biochemical pathways were found to be regulated from day 9 to 23 during leaf development.

Phytoene synthase was up-regulated at both transition stages. This enzyme has shown to be a rate limiting step in the synthesis of carotenoids (Lindgren *et al.*, 2003). The seed-specific overexpression of an endogenous Arabidopsis phytoene synthase gene resulted in delayed germination and increased levels

of carotenoids, chlorophyll, and abscisic acid (Lindgren *et al.*, 2003). The gene for squalene epoxidase was down-regulated at the first transition and up-regulated at the second transition stage.

Together, the loss of farnesylation resulted in an up-regulation of genes encoding enzymes involved in the Calvin cycle, MEP-pathway and the biosynthesis of porphyrin and chlorophyll at the second transition stage in leaf development in *era1-2* (Table 10). This suggests that farnesylation is involved in the transcriptional regulation of these chloroplast localized pathways. For these three pathways, it has been shown that certain of their enzymes are putative targets for the posttranslational activation by ferredoxin-dependent thioredoxins f and m (Balmer *et al.*, 2002). The genes for the enzymes in one of the three pathways shown in Table 10 are either up-regulated or not changed in their expression levels. Eleven of those proteins, whose genes do not show changes in expression, serve as putative targets for the regulation by thioredoxins f and m.

	Chip ID	Atg #	Name	2nd transition stage	Thioredoxin f, m targets
Calvin Cycle	256228_at	A11g58190	Ribulose-1,5-bisphosphate carboxylase/oxygenase	2.24	
	257099_at	A13g12780	3-Phosphoglycerat kinase	2.67	
			Glycerinaldehyde-3-phosphate dehydrogenase	n.c.	
			Triose-phosphate isomerase	n.c.	
			Fructose-bisphosphate aldolase	n.c.	
			Fructose-1,6-bisphosphatase	n.c.	
			Transketolase	n.c.	
			Aldolase	n.c.	
			Sedoheptulose-1,7-bisphosphatase	n.c.	
			Transketolase	n.c.	
	259068_at	A13g04790	Ribose 5-phosphate isomerase	2.41	
			Phosphoributokinase	n.c.	
			Phosphopentose epimerase	n.c.	
MEP-pathway	245281_at	A14g15560	1-D-deoxyxylulose 5-phosphate synthase	2.8	
			1-D-deoxyxylulose 5-phosphate reductoisomerase	n.c.	
			MEP-CMP transferase	n.c.	
	266863_at	A12g26930	MEP-CMP kinase	2.07	
	260324_at	A11g53970	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	2.35	
			GcpE protein	n.c.	
		Isopentenyl-diphosphate delta-isomerase	n.c.		
Porphyrin and chlorophyll biosynthesis			Glu-tRNA synthetase	0.49	
			Glutamate-t-semialdehyde aminotransferase	n.c.	
	260370_at	A11g09740	ALA dehydratase	3.06	
			Porphobilinogen deaminase	n.c.	
			Uroporphyrinogen III decarboxylase	n.c.	
	264820_at	A11g03475	Coproporphyrinogen III oxidase	2.61	
	250221_at	A15g14220	Protoporphyrinogen IX oxidase	2.94	
			Magnesium chelatase	n.c.	
	254105_at	A14g25080	Mg-protoporphyrin IX methyltransferase	6.68	
	264839_at	A11g03630	Protochlorophyllide oxidoreductase	3.53	
246308_at	A13g51820	Chlorophyll synthetase / Chlorophyllase	3.78		
245242_at	A11g44446	Chlorophyll a oxygenase	2.56		

Table 10: The Calvin cycle, the MEP-pathway and the biosynthesis of porphyrin and chlorophyll are shown with all enzymatic steps. Red represents enzymes whose corresponding genes were up-regulated at the second transition stage in *era1-2*. Blue represents putative targets for ferredoxin-dependent thioredoxin activation.

3.2.2.2 Specific analysis of genes for signal transduction pathways

The above results suggest that protein-farnesylation appears to be important for the transcriptional regulation of the Calvin cycle, the MEP-pathway and the biosynthesis of porphyrin and chlorophyll. Therefore, the three groups of genes shown in Fig. 17 were analysed to identify genes coding for proteins involved in ferredoxin-dependent redox regulation. At the second transition stage, genes for thioredoxin f1, f2, m and the ferredoxin-thioredoxin reductase were up-regulated (Table 11).

ChiplD	atg#	Name	2nd transition stage
263624_at	At2g04700	ferredoxin-thioredoxin reductase	3.21
245118_at	At2g41680	thioredoxin reductase	2.37
256607_at	At3g02730	thioredoxin f1	2.29
250133_at	At5g16400	thioredoxin f2	11.33
265476_at	At2g15570	thioredoxin M	2.17
256398_at	At3g15360	thioredoxin m4	2.16
255379_at	At4g03520	M-type thioredoxin	2.09
251815_at	At3g54900	thioredoxin-like protein	2.30

Table 11: Nuclear encoded genes coding for thioredoxins and thioredoxin-reductase predicted to localize into the chloroplasts. All genes were up-regulated during the second transition stage in *era1-2* leaves.

The up-regulation of genes for thioredoxins f1, f2 and m as well as the up-regulation of the gene for thioredoxin reductase at the second transition stage suggest that the ferredoxin-dependent redox regulation might be affected by the loss of farnesylation at later leaf development stages. Furthermore, the regulation of the Calvin cycle, the MEP-pathway and the synthesis of porphyrin and chlorophyll appears to be affected by the loss of farnesylation at the transcriptional level.

3.2.2.3 Genome specific analysis

Genes encoded by the plastid and mitochondrial genomes were analysed at both transition stages during leaf development between *era1-2* and WT. No genes encoded by the mitochondrial genome showed altered expression levels. In contrast, expression of several genes encoded by the plastid genome was increased at least at one transition stage (Table 12). Interestingly, all genes coding for proteins involved in the chloroplast translational machinery were up-regulated at the transition from the expansion phase to the mature phase, indicating that the loss of protein-farnesylation might have an inductive effect on protein synthesis in the chloroplast. Genes coding for proteins involved in photosynthetic processes are either up-regulated at one transition or up-regulated at both transition stages.

ChipID	Name	1st transition stage	2nd transition stage
244964_at	PSII cytochrome b559	2.02	n.c.
244963_at	PSII cytochrome b559	2.15	n.c.
245016_at	carboxyltransferase beta subunit	2.02	2.33
244966_at	cytochrome b6-f complex, subunit V	2.98	2.58
244965_at	hypothetical protein	2.12	2.64
245050_at	PSII K protein	n.c.	7.15
245025_at	ATPase I subunit	n.c.	3.04
245023_at	PSII I protein	n.c.	6.12
245018_at	hypothetical protein	n.c.	2.30
244995_at	ATPase a subunit	n.c.	3.30
244988_s_at	ribosomal protein L23	n.c.	2.34
244986_at	ribosomal protein S19	n.c.	2.41
244985_at	ribosomal protein L22	n.c.	2.28
244981_at	ribosomal protein S6	n.c.	2.26
244980_at	ribosomal protein L36	n.c.	2.26
244969_at	ribosomal protein S18	n.c.	2.29
244939_at	ribosomal protein S12 (trans-splice part 1 of 2)	n.c.	2.05

Table 12: Plastid encoded genes

All genes were either not changed (n.c.) or up-regulated in their gene expression levels.

Furthermore, a keyword search was performed for nuclear genes coding for proteins predicted to localize to the chloroplast and that are involved in the translational machinery, i.e. ribosomes and initiation factors. At the second transition stage 26 of such genes were up-regulated in their expression (Table 13).

ChiplD	atg#	Name	2nd transition stage	Keyword	
245852_at	At5g13610	ribosomal protein L10-like	4.86	ribosomal	
246294_at	At3g56910	ribosomal protein, chloroplast ribosomal protein PsCL18 precursor	3.49		
247201_at	At5g66220	50S ribosomal protein L29	3.96		
248174_at	At5g54600	50S ribosomal protein L24	3.85		
249331_at	At5g40950	50S ribosomal protein L27	2.13		
250190_at	At5g14320	ribosomal protein precursor - like ribosomal protein S13 precursor	4.05		
251120_at	At3g63490	chloroplast ribosomal L1	3.10		
251883_at	At3g54210	ribosomal protein L17	3.00		
251941_at	At3g53470	putative protein ribosomal protein S25	3.33		
255623_at	At4g01310	putative L5 ribosomal protein	3.66		
255850_at	At2g33450	putative chloroplast 50S ribosomal protein L28	4.16		
256855_at	At3g15190	30S ribosomal protein S20	5.39		
257190_at	At3g13120	30S ribosomal protein S10	3.03		
257225_s_at	At3g27850	50S ribosomal protein L12-C	2.02		
260898_at	At1g29070	plastid ribosomal protein L34 precursor	4.15		
261119_at	At1g75350	chloroplast 50S ribosomal protein L31	3.91		
261190_at	At1g32990	ribosomal protein L11	5.33		
261954_at	At1g64510	plastid ribosomal protein S6	5.22		
263131_at	At1g78630	ribosomal protein L13	4.38		
264575_at	At1g05190	putative chloroplast 50S ribosomal protein	3.89		
265247_at	At2g43030	50S ribosomal protein L3	2.70		
266570_at	At2g24090	putative chloroplast ribosomal protein L35	3.81		
267088_at	At2g38140	30S ribosomal protein S31	3.55		
267435_at	At2g33800	30S ribosomal protein S5	4.10		
263597_at	At4g30690	putative protein translation initiation factor, IF3	11.41		initiation factor
262483_at	At1g17220	putative translation initiation factor IF2	2.14		

Table 13: Keyword search for protein involved in the plastidic translational machinery

3.3 Biocomputational approach to identify putative PFT targets

Farnesylation is a posttranslational modification of proteins. The loss of protein-farnesylation resulted in changes in gene expression in *era1-2* leaf development. Nuclear genes coding for proteins predicted to localize to the chloroplast and involved in biochemical and signal transduction pathways were up-regulated. The relationship between a farnesylated protein and the altered transcriptional regulation of these genes, however, was still unclear. Therefore the Arabidopsis protein database was used to identify putative PFT targets by screening for proteins with a CaaX-motif. In total, 654 proteins were found with a cysteine at the fourth position from the C-terminal end (Cxxx). This group of proteins contains putative targets for the protein-farnesyltransferase (PFT) and the two protein-geranylgeranyltransferases (PGGT I and II). Within this group, 123 were identified to match the criteria of a PFT target according to Moores *et al.* (1991). Proteins were further selected if the corresponding gene was expressed at least in one experiment during the three different developmental phases in *era1-2* and WT leaves. The proteins were then analysed in the context of compartment prediction. Because protein-farnesylation occurs in the cytoplasm, the focus was on proteins predicted to localize to the cytoplasm or to

secretory pathways. Finally, a keyword search in combination with a search for genes involved in metabolic pathways was performed to select well known proteins (Table 14).

After an intensive literature search, two proteins were selected which were promising in the context of our analysis: the transcription factor Hap5a and cytochrome b561.

ChipID	atg#	CaaX - box	Name	era1 vs wt1	era2 vs wt2	Keyword
255936_at	A11g12880	CRAF	calcium dependent protein kinase	n.c.	n.c.	Calcium
249089_at	A15g44240	CKDN	ATPase, calcium-transporting	n.c.	n.c.	
266782_at	A12g29120	CEGE	putative ligand-gated ion channel protein	6.40	2.36	Channel
262206_at	A12g01090	CVSV	putative ubiquinol-cytochrome c reductase	n.c.	n.c.	
245818_at	A11g26140	CL68	cytochrome b561	n.c.	n.c.	Cytochrome
256598_at	A13g30180	CSPY	cytochrome P450 homolog	2.30	n.c.	
262728_at	A11g75840	CVFL	rac-like GTP binding protein (ARAC5)	n.c.	0.48	G-proteins
245777_at	A11g73540	CQTG	diphosphoinositol polyphosphate phosphohydrolase	0.39	0.44	Inositol metabolism
263076_at	A12g17520	CDII	putative protein kinase/endoribonuclease	n.c.	n.c.	
247338_at	A15g63680	CVVK	pyruvate kinase	n.c.	n.c.	Kinase
250526_at	A15g08670	CVVK	pyruvate kinase	n.c.	n.c.	
249879_at	A15g23170	CRVW	serine/threonine protein kinase	n.c.	n.c.	
253507_at	A14g32000	CLCV	serine threonine protein kinase	n.c.	n.c.	Receptor
263278_at	A14g34220	CVLV	receptor protein kinase - TMKL1	n.c.	0.39	
252468_at	A13g46970	CPVP	starch phosphorylase H (cytosolic form)	n.c.	n.c.	Starch and sucrose metabolism
265078_at	A11g55460	CKLA	zinc finger protein rts2	n.c.	n.c.	
261376_at	A11g18660	CRMQ	zinc finger protein	n.c.	n.c.	
251763_at	A13g55730	CARS	R2R3-MYB transcription factor	n.c.	n.c.	Transcription factor
250582_at	A15g07580	CDAF	ethylene responsive element binding factor 5	n.c.	n.c.	
249246_at	A15g28120	CDL3	putative protein transcription factor Hap5a	n.c.	n.c.	
261959_at	A11g64500	CFVN	peptide transporter	2.76	3.27	Transporter

Table 14: Putative PFT-targets found by a biocomputational analysis

Proteins were selected only, if they were predicted for the cytoplasm or for the secretory pathways. The selected proteins were used for an intensive literature analysis.

4. Discussion

The *Arabidopsis thaliana* mutant *era1-2* was analyzed at a cellular and at a gene expression level to study the effects of farnesylation on leaf development and plant growth. For this purpose, a kinematic study was combined with extensive microarray experiments. In addition, a biocomputational approach was used to identify putative protein-farnesyltransferase (PFT) targets which might not be farnesylated in *era1-2* resulting in the observed phenotypical as well as gene expression alterations.

4.1 An additional proliferation phase is responsible for the increase in mature leaf size in era1-2

Among other phenotypical attributes described by Yalovsky *et al.* (2000), the loss of farnesylation in *era1-2* causes an increase in mature leaf size. A kinematic approach was used to examine leaf development in *era1-2* and in wild type (WT) at a cellular level. Using this approach, a second proliferation phase was discovered in *era1-2* after 20 days, whereas only a single proliferation phase was found in WT. It appears that the additional proliferation phase is the main reason for the extended leaf size in *era1-2*, because similar patterns were obtained in the kinetic studies for cell number and cell size during the first 20 days of leaf development in both WT and *era1-2*.

In animals, Ras proteins, which are targets for farnesylation, play fundamental roles in cell signal transduction pathways that regulate cell growth, cell differentiation and cell proliferation. *ras* gene mutations are highly prevalent in many human cancers (Crul *et al.*, 2001) and have been implicated in tumorigenesis (Adjei, 2001). Because of its key role as a molecular switch in signal transduction processes and its frequent mutation in human cancer, Ras has become an important target for antineoplastic therapy (Adjei, 2001). Selective blockage of Ras farnesylation, which essentially blocks membrane attachment and interrupts Ras effector pathways, has emerged as an approach to develop novel anti-cancer therapies that stop cell cycle progression (Zang

and Casey, 1996; Sebti and Hamilton, 2000). These data are in contrast with our results since the loss of farnesylation leads to a re-entry into the cell cycle during maturation in leaf development.

4.2 Transcriptional control of genes involved in cell cycle regulation in *era1-2*

A global gene expression analysis was performed to obtain insights into the overall behavior of changes in gene expression during leaf development. In general, global gene expression in WT and *era1-2* followed the same pattern during leaf development. This pattern was found to be nearly identical at the transition from the cell proliferation phase to the cell expansion phase (1st transition). At the transition from the cell expansion phase to the mature phase (2nd transition), differentially expressed genes behaved similar, however the extent of changes in gene expression was more pronounced in WT than in *era1-2*. As a consequence, most changes in gene expression between *era1-2* and WT were observed during maturation in leaf development.

A gene specific analysis was performed to study the expression pattern of genes coding for proteins involved in the regulation of leaf growth. In *Arabidopsis*, leaf growth is the result of both cell expansion and cell proliferation (Tsukaya *et al.*, 1993; Tsukaya *et al.*, 1994; Neff and Van Volkenburgh 1994; Kauschmann *et al.*, 1996; Szekeres *et al.*, 1996; Tsuge *et al.*, 1996; Kim *et al.*, 1998; Kim *et al.*, 1999; Nicol *et al.*, 1998; Van Volkenburgh 1999; Donnelly *et al.*, 1999). Following our observation that a second cell proliferation phase was initiated in *era1-2*, we analyzed the gene expression pattern in the context of cell cycle control. Several genes coding for proteins involved in cell cycle regulation were up-regulated in *era1-2* at day 23 relative to WT leaves at day 21. Thus, the re-entry of *era1-2* into a cell proliferation phase was confirmed at the gene expression level. Taken together, the data strongly suggest that farnesylation of proteins is required to control cell cycle regulation during maturation in leaf development.

4.3 Farnesylation is involved in the transcriptional regulation of chloroplast metabolism

The loss of farnesylation resulted in changes in gene expression during maturation in leaf development. Around 80% of genes with changes in expression coding for proteins predicted to localize to the chloroplast were up-regulated between *era1-2* and WT. This group of genes was analyzed in the context of functional categories.

Genes coding for enzymes involved in the Calvin cycle, fatty acid biosynthesis, starch and sucrose metabolism, porphyrin and chlorophyll biosynthesis and in the MEP-pathway for isoprenoid biosynthesis, were up-regulated in *era1-2* leaves. In addition to these pathways, genes coding for proteins involved in the chloroplast translational machinery and localized in the nuclear as well as the chloroplast genome were found to be up-regulated. This led to the conclusion that farnesylation is involved in the transcriptional control of chloroplast localized metabolic pathways and protein synthesis within the chloroplast. Therefore, farnesylation appears to have a critical role in the developmental and functional control of the chloroplast metabolism.

4.4 Farnesylation leads to redox-mediated control of chloroplast metabolism

Redox regulation involving thioredoxins is a basic regulatory mechanism during the transition from night to day (Scheibe, 1991; Danon and Mayfield, 1994). The ferredoxin-dependent thioredoxin f and m proteins play a key role in the regulation of enzymes involved in the Calvin cycle, fatty acid biosynthesis and ATP synthase localized in the chloroplast (Schürmann and Jacquot, 2000). Using a proteomics approach, Balmer *et al.* (2003) identified new putative thioredoxin-interacting proteins. Among this group of proteins, two are involved in the MEP-pathway (1-D-deoxyxylulose 5-phosphate reductoisomerase, GcpE-protein), three in chlorophyll biosynthesis (glutamate-1-semialdehyde

aminotransferase, uroporphyrinogen III decarboxylase, magnesium chelatase) and three new target proteins in the Calvin cycle (triose-phosphate isomerase, transketolase, phosphopentose epimerase). The genes coding for these proteins were not changed in their expression levels between *era1-2* and WT during leaf development. The genes coding for chloroplast-targeted thioredoxin f and m and for the thioredoxin reductase, however, were found to be up-regulated in *era1-2* leaves. Thus, the majority of enzymes involved in the Calvin cycle, the chlorophyll biosynthesis and the MEP-pathway were either up-regulated on the gene expression level or might be activated by chloroplast thioredoxins. This led to the conclusion that farnesylation is required for the correct transcriptional as well as the possible posttranslational regulation of enzymes in these three pathways.

4.5 Regulation of isoprenoid biosynthesis controls cell cycle progression

As discussed above, the loss of farnesylation positively affects plastidic isoprenoid biosynthesis at both the transcriptional and possibly the posttranslational level, which would consequently lead in an induced level of isoprenoid precursor (IPP) formation in the chloroplast. A unidirectional exchange of a MEP-pathway intermediate (most likely IPP) from the chloroplast to the cytoplasm was proposed by Laule *et al.* (2003). As a result, in *era1-2* leaves an accumulation of IPP in the cytosol can be expected during maturation (Fig. 20). The up-regulation of genes coding for squalene synthetase, squalene epoxidase and cycloartenol synthase, which catalyze the first three steps in the phytosterol biosynthesis, is consistent with a possible channelling of cytosolic IPPs into the sterol forming branch of cytosolic isoprenoid biosynthesis and subsequently into the formation of brassinosteroids (BRs). BRs have been shown to be involved in promoting cell division in *Arabidopsis* suspension culture systems by the induction of *CycD3* (Hu *et al.*, 2000). The gene expression levels of *CycD3.1* and *CycD3.3* were up-regulated in *era1-2* during leaf maturation. Furthermore BRs were described to control the proliferation of

leaf cells in *Arabidopsis* (Nakaya *et al.*, 2002). Therefore a putative scenario would be that the loss of farnesylation results in an increase of BR levels and consequently in the promotion of cell cycle progression in mature leaves of *era1-2* (Fig. 20).

4.6 Biocomputational approach to identify putative PFT targets

To identify farnesylated candidate proteins that could mediate the communication between the cytosol and the chloroplasts, a computational analysis was performed to screen for putative PFT target proteins. The CaaX-motif was used as an identifier according to Moores *et al.* (1991). It has been shown that protein-farnesylation takes place in the cytosol (Lange and Ghassemian, 2003). Therefore, only proteins which are predicted to the cytosol or to the secretory pathway and which possess a putative farnesylation site were selected for this analysis. By this analysis, twenty-one proteins were obtained and a potential role of these proteins in leaf development was verified by extensive literature search. The literature search revealed two promising putative PFT targets: the cytochrome b561 and the transcription factor Hap5a.

4.6.1 Cytochrome b561

In plant and animal cells, cytochromes b561 belongs to a class of membrane proteins that catalyze transmembrane electron transfer with ascorbate (ASC) as the electron donor, thereby contributing to ASC-mediated redox metabolism (Njus and Kelley, 1993; Asard *et al.*, 2001). Members of this protein family use monodehydroascorbate as an electron acceptor to regenerate fully reduced ascorbate. Four putative Cyt b561 isoforms have been identified in *Arabidopsis* and only one of them has the typical farnesylation motif at the C-terminus. Little is known about the subcellular localization and physiological function of these proteins however biochemical evidence suggests the presence of at least one

ASC-reducible Cyt b561 in the plasma-membrane fraction from plants, including *Arabidopsis* (Asard *et al.*, 1989; Bérczi *et al.*, 2001).

In a hypothetical model one could assign the Cyt b561-CaaX to the chloroplast membrane, where it would control the redox status in the plastids by the reduction of oxidized ascorbate (Fig. 19). Ascorbate is synthesized at the inner membrane of the mitochondria (Siendones *et al.*, 1999; Bartoli *et al.*, 2000) from where it is distributed to all intracellular compartments. Compartments, such as the chloroplast, that can not synthesize ascorbate require mechanisms to regenerate reduced ascorbate molecules. This can be achieved by either *trans*-membrane transport of the molecule (Horemans *et al.*, 2000a and 2000b), or by *trans*-membrane transport of electrons using monodehydroascorbate as an electron acceptor (Asard *et al.*, 1992; Horemans *et al.*, 1994), or by a combination of both mechanisms. In the following model (Fig. 19) we suggest that the regeneration of reduced ascorbate in the chloroplast is accomplished by a combination of both mechanisms.

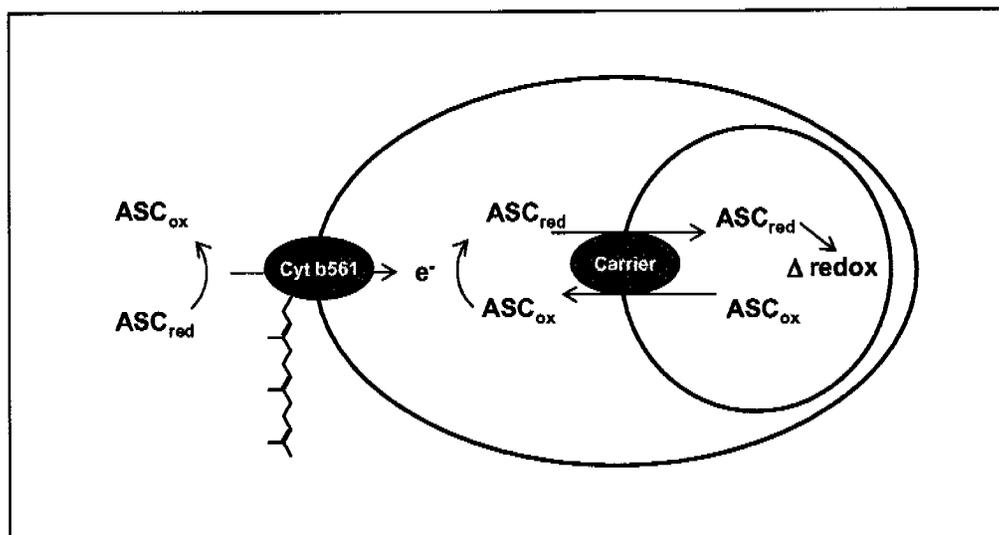


Figure 19: Model for the regeneration of reduced ascorbate in the chloroplast

Electrons would be transferred from reduced ascorbate by Cyt b561 through the outer chloroplast membrane into the intermembrane lumen, where they would be used for the reduction of oxidized ascorbate. This transfer mechanism might be under the control of farnesylation. Reduced ascorbate would then be transported through the inner chloroplast membrane by an ascorbate carrier. An accumulation of reduced ascorbate would change the redox status in the chloroplast.

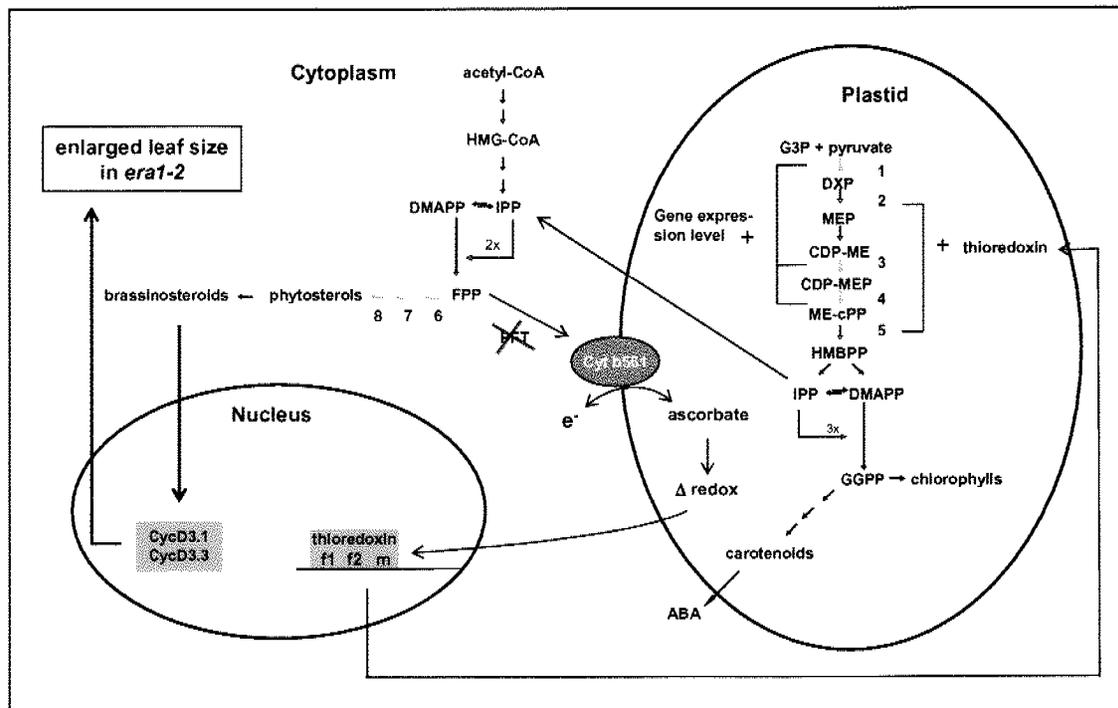
Electrons might be transferred from reduced ascorbate by Cyt b561 through the outer chloroplast membrane to reduce oxidized ascorbate in the intermembrane matrix. Reduced ascorbate would then be transferred through the inner chloroplast membrane by an ascorbate carrier as proposed for pea-leaf protoplasts (Foyer and Lelandes, 1996). This would lead in an altered redox status in the chloroplast, which could influence the transcription of nuclear-encoded thioredoxins which are targeted to the chloroplast and finally activate the thioredoxin-regulated metabolic pathways (Fig. 19).

In figure 20, such a model is shown for the activation of the plastidic isoprenoid pathway. The loss of farnesylation could result in the mistargeting or altered activity of Cyt b561. It is possible that this could result in the accumulation of reduced ascorbate, resulting in an altered redox status in plastids. The redox status has been shown to act as a signal for chloroplast-nucleus communication (Pfannschmidt *et al.*, 2003) and could therefore be responsible for the induced transcription of genes coding for 1-D-deoxyxylulose 5-phosphate synthase, the 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase and the 2C-methyl-D-erythritol-phosphate-CMP kinase and the thioredoxin f1, f2 and m which were all up-regulated in *era1-2* leaves compared to WT.

Figure 20 (see page 80): Hypothetical model which shows the relation ship between loss of farnesylation and enlarged leaf size observed in *era1-2*

Red arrows and red boxes represent genes, which were up-regulated in *era1-2*. Numbers 1 – 5 represent enzymes involved in the plastidic isoprenoid pathway: (1) 1-D-deoxyxylulose 5-phosphate synthase; (2) 1-D-deoxyxylulose 5-phosphate reductoisomerase; (3) 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; (4) 2C-methyl-D-erythritol-phosphate-CMP kinase; (5) GcpE-protein. Numbers 6 – 8 represent enzymes involved in the phytosterol biosynthesis: (6) squalene synthetase; (7) squalene epoxidase; (8) cycloartenol synthase.

The loss of farnesylation would lead in an up-regulation of the plastidic isoprenoid pathway at both the transcriptional and the posttranscriptional level. The activation at the posttranslational level would be triggered by a mistargeted or altered activity of Cyt b561, resulting in a possibly altered redox status. As a consequence of the induced MEP-pathway, an accumulation of chloroplast IPP would be expected. The unidirectional exchange of chloroplast IPP to the cytoplasm would lead to an increase of IPP in the cytosol. The up-regulation of the genes 6 -8 would favour the channelling of IPPs into the branch of sterol synthesis and might be in an induction of BR biosynthesis. BRs have shown to promote cell cycle progression. A promotion of cell cycle would result in an additional proliferation phase and finally lead in the enlarged leaf size observed in *era1-2*.



4.6.2 The transcription factor Hap5a

The CCAAT-box is one of the most common promoter elements found in a large number of genes in animals, fungi and plants (Maity and de Crombrughe, 1998; Mantovani, 1999). The Hap complex, which is also called NF-Y or CBF in animals, binds the CCAAT-box and regulates gene expression. This complex consists of three distinct subunits: HAP2/NF-YA/CBF-B, HAP3/NF-YB/CBF-A and HAP5/NF-YC/CBF-C, all of which are necessary for DNA binding. Each subunit contains an evolutionary conserved domain that is responsible for DNA binding and protein-protein interaction. Arabidopsis has 9 HAP2 subunit genes, 10 HAP3 subunit genes and 10 HAP5 subunit genes (Edwards *et al.*, 1998; Gusmaroli *et al.*, 2001, 2002). Among them, only *LEAFY COTYLEDON1* (*LEC1*), a member of the HAP3 family, has been characterized in detail. Miyoshi *et al.* (2003) have isolated three genes that potentially encode a HAP3/ nuclear factor-YB (NF-YB)/CCAAT binding factor-A (CBF-A) subunit of a CCAAT-box binding complex in rice. The genes were expressed in various organs, including leaves. Transgenic rice plants with antisense or RNAi constructs have pale green leaves in which the amount of chlorophyll was reduced and chloroplasts

were degenerated. Chloroplast degeneration was accompanied by reduced expression of nuclear-encoded genes such as *RIBULOSE BIPHOSPHATE CARBOXYLASE* and *CHLOROPHYLL A-B BINDING PROTEIN*, which were found to be up-regulated in *era1-2* (data not shown). Only one HAP gene containing the typical farnesylation box was expressed in *era1-2* leaves. This observation suggests a potential function of farnesylation within the chloroplast biogenesis as well as in the expression of genes coding for important photosynthetic proteins. Beside APETALA1 (AP1) (Madel *et al.*, 1992), HAP5a would be the second transcription factor in plants with a farnesylation motif. AP1 is involved in the transition from the vegetative to the reproductive growth (Yalovsky *et al.*, 2000). The gene expression level of *AP1* was below the detection threshold in *era1-2* and WT leaves.

Taken together, the loss of farnesylation caused an increase in mature leaf size in *era1-2*. The kinematics study combined with the gene expression analysis showed an additional proliferation phase in *era1-2* after 20 days in leaf development. Most of the changes in gene expression including genes coding for proteins involved in cell cycle control were observed during leaf maturation between *era1-2* and WT. Our results suggest that farnesylation might play an important role during later stages in leaf development.

A hypothetical model inferred from our data suggests that a farnesylated protein has a role in communication between the cytoplasm and the chloroplast, and is therefore involved in the transcriptional and possibly posttranslational control of chloroplast metabolic pathways. A possible scenario would be that the loss of farnesylation in *era1-2* leads to an increased synthesis of isoprenoid precursors. The induction of brassinosteroid synthesis, which could be assumed from our data, would be consistent with this scenario, and would finally result in an unscheduled activation of the cell cycle. The activated cell cycle would cause the additional proliferation phase, which appears to be the main reason for the enlarged leaf size observed in *era1-2*.

5. References

Adjei, A.A. (2001). Blocking oncogenic Ras signaling for cancer therapy. *J. Natl. Cancer Inst.* **93**:1062–74.

Altmann, T. (1998). Recent advances in brassinosteroid molecular genetics. *Curr. Opin. Plant Biol.* **1**:378-83.

Anderegg, R.J., Betz, R., Carr, S.A., Crabb, J.W., Duntze, W. (1988). Structure of *Saccharomyces cerevisiae* mating hormone a-factor. *J. Biol. Chem.* **263**:18236-40.

Andres, D.A., Milatovich, A., Ozcelik, T., Wenzlau, J.M., Brown, M.S., Goldstein, J.L., Francke, U. (1993). cDNA cloning of the two subunits of human CAAX farnesyltransferase and chromosomal mapping of FNTA and FNTB loci and related sequences. *Genomics* **18**:105-12.

Arigoni, D., Sagner, S., Latzel, C., Eisenreich, W., Bacher, A., Zenk, M.H. (1997). Terpenoid biosynthesis from 1-deoxy-D-xylulose in higher plants by intramolecular skeletal rearrangement. *Proc. Natl. Acad. Sci. U S A* **94**:10600-5.

Asard, H., Horemans, N., Caubergs, R.J. (1992). Transmembrane electron transport in ascorbate-loaded plasma membrane vesicles from higher plants involves a *b*-type cytochrome. *FEBS Lett.* **306**:143-6.

Asard, H., Kapila, J., Verelst, W., Bérczi, A. (2001). Higher-plant plasma membrane cytochrome b561: a protein in search for a function. *Protoplasma* **217**:77-93.

Asard, H., Venken, M., Caubergs, R., Reijnders, W., Oltmann, F.L., De Greef, J.A. (1989). *b*-Type cytochromes in higher plant plasma membranes. *Plant Physiol.* **90**:1077-83.

-
- Azimzadeh, J., Traas, J., Pastuglia, M.** (2001). Molecular aspects of microtubule dynamics in plants. *Curr. Opin. Plant Biol.* **4**:513-9.
- Azpiroz, R., Wu, Y., LoCascio, J.C., and Feldmann, K.A.** (1998). An Arabidopsis brassinosteroid-dependent mutant is blocked in cell elongation. *Plant Cell* **10**:219-30.
- Bach T.J.** (1987). Synthesis and metabolism of mevalonic acid in plants. *Plant Physiol. Biochem.* **25**:163-78.
- Balmer, Y., Koller, A., del Val, G., Manieri, W., Schürmann, P., Buchanan, B.B.** (2002). Proteomics gives insight into the regulatory function of chloroplast thioredoxins. *Proc. Natl. Acad. Sci. U S A* **100**:370-5.
- Barabas, Z., Rédei, G.P.** (1971). Facilitation of crossing by the use of appropriate parental stocks. *Arabidopsis Inf. Ser.* **8**: 7-8.
- Barbacid, M.** (1987). Ras genes. *Annu. Rev. Biochem.* **56**:779-827.
- Barnes, D., Mayfield, S.P.** (2003). Redox control of posttranscriptional processes in the chloroplast. *Antioxid Redox Signal* **5**:89-94.
- Barone, A.D., Beecher, J.E., Bury, P.A., Chen, C., Doede, T., Fidanza, J.A., McGall, G.H.** (2001). Photolytographic synthesis of high-density oligonucleotide probe arrays. *Nucleosides Nucleotides Nucleic Acids* **20**:525-31.
- Bartoli, C.G., Pastori, G.M., Foyer, C.H.** (2000). Ascorbate biosynthesis in mitochondria is linked to the electron transport chain between complexes III and IV. *Plant Physiol.* **123**:335-44.

- Baskin, T.I., Cork, A., Williamson, R.E., Gorst, J.R.** (1995). STUNTED PLANT 1, a gene required for expansion in rapidly elongating but not in dividing cells and mediating root growth responses to applied cytokinin. *Plant Physiol.* **107**:233-43.
- Beck, L.A., Hosick, T.J., Sinensky, M.** (1988). Incorporation of a product of mevalonic acid metabolism into proteins of Chinese hamster ovary cell nuclei. *J Cell. Biol.* **107**:1307-16.
- Bérczi, A., Lüthje, S., Asard, H.** (2001). B-Type cytochromes in plasma membranes of *Phaseolus vulgaris* hypocotyls, *Arabidopsis thaliana* leaves, and *Zea mays* roots. *Protoplasma* **217**:50-5.
- Boguski, M.S., McCormick, F.** (1993). Proteins regulating Ras and its relatives. *Nature* **366**:643-54.
- Brazma, A., Vilo, J.** (2000). Gene expression data analysis. *FEBS Lett.* **480**:17-24.
- Burgering, B.M., Bos, J.L.** (1995). Regulation of Ras-mediated signalling: more than one way to skin a cat. *Trends Biochem. Sci.* **20**:18-22.
- Casey, P.J., Solski, P.A., Der, C.J., Buss, J.E.** (1989). p21ras is modified by a farnesyl isoprenoid. *Proc. Natl. Acad. Sci. U S A* **86**:8323-7.
- Celis, J.E., Kruhoffer, M., Gromov, I., Frederiksen, C., Ostergaard, M., Thykjaer, T., Gromov, P., Yu, J., Palsdottir, H., Magnusson, N., Orntoft, T.F.** (2000). Gene expression profiling: monitoring transcription and translation products using DNA microarrays and proteomics. *FEBS Lett.* **480**:2-16.
- Chappell, J.** (1995). Biochemistry and molecular biology of the isoprenoid biosynthesis pathway. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**:521-47.

- Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X.C., Stern, D., Winkler, J., Lockhard, D.J., Morris, M.S., Fodor, S.P.** (1996). Accessing genetic information with high-density DNA arrays. *Science* **274**:610-14.
- Cho, H.-T., Cosgrove, D.** (2000). Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **97**:9783-8.
- Clarke, S.** (1992). Protein isoprenylation and methylation at carboxyl-terminal cysteine residues. *Annu. Rev. Biochem.* **61**:355-86.
- Cox, A.D., Der, C.J.** (1992). The ras/cholesterol connection: implications for ras oncogenicity. *Crit. Rev. Oncog.* **3**:365-400.
- Crul, M., de Klerk, G.J., Beijnen, J.H., Schellens, J.H.** (2001). Ras biochemistry and farnesyl transferase inhibitors: a literature survey. *Anticancer Drugs* **12**:163-84.
- Cutler, S., Ghassemian, M., Bonetta, D., Cooney, S., McCourt, P.** (1996). A protein farnesyl transferase involved in abscisic acid signal transduction in *Arabidopsis*. *Science* **273**:1239-41.
- Dale, J.E.** (1988). The control of leaf expansion. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **39**:267-95.
- Danon, A., Mayfield, S.P.** (1994). Light-regulated translation of chloroplast messenger RNAs through redox potential. *Science* **266**:1717-9.
- De Risi, J.L., Iyer, V.R., Brown, P.O.** (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**:680-6.

De Veylder L, Beeckman T, Beemster GT, Krols L, Terras F, Landrieu I, van der Schueren E, Maes S, Naudts M, Inze D. (2001). Functional analysis of cyclin-dependent kinase inhibitors of *Arabidopsis*. *Plant Cell* **13**:1653-68.

Dewitte, W., Riou-Khamlichi, C., Scofield, S., Healy, J.M., Jacquard, A., Kilby, N.J., Murray, J.A. (2003). Altered cell cycle distribution, hyperplasia, and inhibited differentiation in *Arabidopsis* caused by the D-type cyclin CYCD3. *Plant Cell* **15**:79-92.

Dhawan, P., Yang, E., Kumar, A., Mehta, K.D. (1998). Genetic complexity of the human geranylgeranyltransferase I beta-subunit gene: a multigene family of pseudogenes derived from mis-spliced transcripts. *Gene* **210**:9-15.

Donnelly, P.M., Bonetta, D., Tsukaya, H., Dengler, R., Dengler, N.G. (1999). Cell cycling and cell enlargement in developing leaves of *Arabidopsis*. *Dev. Biol.* **215**:407-19.

Dopazo, J., Carazo, J.M. (1997). Phylogenetic reconstruction using an unsupervised growing neural network that adopts the topology of a phylogenetic tree. *J. Mol. Evol.* **44**:226-33.

Draghici, S., Kulaeva, O., Hoff B., Petrov, A., Shams, S., Tainsky, M.A. (2003). Noise sampling method: an ANOVA approach allowing robust selection of differentially regulated genes measured by DNA microarrays. *Bioinformatics* **19**:1348-59.

Edwards, D., Murray, J.A.H., Smith, A.G. (1998). Multiple genes encoding the conserved CCAAT-box transcription factor complex are expressed in *Arabidopsis*. *Plant Physiol.* **117**:1015-22.

-
- Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D.** (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U S A* **95**:14863-8.
- Ekins, R.P.** (1998). Ligand assays: from electrophoresis to miniaturized microarrays. *Clin. Chem.* **44**:2015-30.
- Endo, A.** (1992). The discovery and development of HMG-CoA reductase inhibitors. *J. Lipid* **33**:1569-82.
- Epstein, W.W., Lever, D., Leining, L.M., Bruenger, E., Rilling, H.C.** (1991). Quantitation of prenylcysteines by a selective cleavage reaction. *Proc. Natl. Acad. Sci. U S A* **88**:9668-70.
- Fairbanks, K.P., Witte, L.D., Goodman, D.S.** (1984). Relationship between mevalonate and mitogenesis in human fibroblasts stimulated with platelet-derived growth factor. *J. Biol. Chem.* **259**:1546-51.
- Farnsworth, C.C., Wolda, S.L., Gelb, M.H., Glomset, J.A.** (1989). Human lamin B contains a farnesylated cysteine residue. *J. Biol. Chem.* **264**:20422-9.
- Feig, L.A.** (1994). Guanine-nucleotide exchange factors: a family of positive regulators of Ras and related GTPases. *Curr. Opin. Cell Biol.* **6**:204-11.
- Feldmann, K.A., Marks, M.D., Christianson, M.L., Quatrano, R.S.** (1989). A dwarf mutant of *Arabidopsis* generated by T-DNA insertion mutagenesis. *Science* **243**:1351-4.
- Foyer, C.H., Lelandais, M.** (1996). A comparison of the relative rates of transport of ascorbate and glucose across the thylakoid, chloroplast and plasmalemma membranes of pea leaf mesophyll cells. *J. Plant Physiol.* **148**:391-8.

- Fujioka, S., Li, J., Choi, Y.H., Seto, H., Takatsuto, S., Noguchi, T., Watanabe, T., Kuriyama, H., Yokota, T., Chory, J., Sakurai, A. (1997).** The *Arabidopsis deetiolated2* mutant is blocked early in brassinosteroid biosynthesis. *Plant Cell* **9**:1951-62.
- Galichet, A., Gruissem, W. (2003).** Protein farnesylation in plants – conserved mechanisms but different targets. *Curr. Opin. Plant Biol.* **6**:530-5.
- Gibbs, J.B. (1991).** Ras C-terminal processing enzymes--new drug targets? *Cell* **65**:1-4.
- Gibbs, J.B., Oliff, A., Kohl, N.E. (1994).** Farnesyltransferase inhibitors: Ras research yields a potential cancer therapeutic. *Cell* **77**:175-8.
- Glomset, J.A., Farnsworth, C.C. (1994).** Role of protein modification reactions in programming interactions between ras-related GTPases and cell membranes. *Annu. Rev. Cell. Biol.* **10**:181-205.
- Goalstone, M.L., Leitner, J.W., Wall, K., Dolgonos, L., Rother, K.I., Accilli, D., Draznin, B. (1998).** Effect of insulin on farnesyltransferase. Specificity of insulin action and potentiation of nuclear effects of insulin-like growth factor-1, epidermal growth factor, and platelet-derived growth factor. *J. Biol. Chem.* **273**:23892-6.
- Goldstein, J.L., Brown, M.S. (1990).** Regulation of the mevalonate pathway. *Nature* **343**:425-30.
- Goldstein, J.L., Brown, M.S., Stradley, S.J., Reiss, Y., and Gierasch, L.M. (1991).** Nonfarnesylated tetrapeptide inhibitors of protein farnesyltransferase. *J. Biol. Chem.* **266**:15575-8.

- Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., Coupland, G. (1997).** A Polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature* **386**:44-51.
- Grosu, P., Townsend, J.P., Hartl, D.L., Cavalieri, D. (2002).** Pathway Processor: a tool for integrating whole-genome expression results into metabolic networks. *Genome Res.* **12**:1121-6.
- Gusmaroli, G., Tonelli, C., Mantovani, R. (2001).** Regulation of the CCAAT-binding NF-Y subunits in *Arabidopsis thaliana*. *Gene* **264**:173-85.
- Gusmaroli, G., Tonelli, C., Mantovani, R. (2002).** Regulation of novel members of *Arabidopsis thaliana* CCAAT-binding nuclear factor Y subunits. *Gene* **283**:41-8.
- Hancock, J.F., Magee, A.I., Childs, J.E., Marshall, C.J. (1989).** All ras proteins are polyisoprenylated but only some are palmitoylated. *Cell* **57**:1167-77.
- Hara, M., Akasaka, K., Akinaga, S., Okabe, M., Nakano, H., Gomez, R., Wood, D., Uh, M., Tamanoi, F. (1993).** Identification of Ras farnesyltransferase inhibitors by microbial screening. *Proc. Natl. Acad. Sci. U S A* **90**:2281-5.
- Herrero, J., Dopazo, J. (2002).** Combining hierarchical clustering and self-organizing maps for exploratory analysis of gene expression patterns. *J. Proteome Res.* **1**:467-70.
- Herrero, J., Valencia, A., Dopazo, J. (2001).** A hierarchical unsupervised growing neural network for clustering gene expression patterns. *Bioinformatics* **17**:126-36.

- Horemans, N., Asard, H., Caubergs, R.J.** (1994). The role of ascorbate free radical as an electron acceptor to cytochrome *b*-mediated trans-plasma membrane electron transport in higher plants. *Plant Physiol.* **104**:1455-8.
- Horemans, N., Foyer, C., Asard, H.** (2000a). Transport and action of ascorbate at the plant plasma membrane. *Trends Plant Sci.* **5**:263-7.
- Horemans, N., Foyer, C., Potters, G., Asard, H.** (2000b). Ascorbate function and associated transport systems. *Plant Physiol. Biochem.* **38**:531-40.
- Hu, Y., Bao, F., Li, J.** (2000). Promotive effect of brassinosteroids on cell division involves a distinct *CycD3*-induction pathway in *Arabidopsis*. *Plant J.* **24**:693-701.
- Ishibashi, Y., Sakagami, Y., Isogai, A., Suzuki, A.** (1984). Structures of tremmerogens A9291-I and A-9291-VII: peptidyl sex hormones of *Tremella brasiliensis*. *Biochemistry* **23**:1399-1404.
- Kamiya, Y., Sakurai, A., Tamura, S., Takahashi, N.** (1978). Structure of rhodotorucine A, a novel lipopeptide, inducing mating tube formation in *Rhodospidium toruloides*. *Biochem. Biophys. Res. Commun.* **83**:1077-83.
- Kauschmann, A., Jessop, A., Koncz, C., Szekeres, M., Willmitzer, L., Altmann, T.** (1996). Genetic evidence for an essential role of brassinosteroids in plant development. *Plant J.* **9**:701-13.
- Khosravi-Far, R., Lutz, R. J., Cox, A.D., Conroy, L., Bourne, J.R., Sinensky, M., Balch, W.E., Buss, J.E., Der, C.J.** (1991). Isoprenoid modification of rab proteins terminating in CC or CXC motifs. *Proc. Natl. Acad. Sci. USA* **88**:6264-8.

Khosravi-Far, R., Der, C.J. (1994). The Ras signal transduction pathway. *Canc. Metastasis Rev.* **13**:67-89.

Kim, G.-T., Tsukaya, H., Saito, Y., Uchimiya, H. (1999). Changes in the shapes of leaves and flowers upon overexpression of cytochrome P450 in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **96**:9433-7.

Kim, G.-T., Tsukaya, H., Uchimiya, H. (1998). The *CURLY LEAF* gene controls both division and elongation of cells during the expansion of the leaf blade in *Arabidopsis thaliana*. *Planta* **206**:175-83.

Klahre, U., Noguchi, T., Fujioka, S., Takatsuto, S., Yokota, T., Nomura, T., Yoshida, S., Chua, N.H. (1998). The *Arabidopsis* DIMINUTO/DWARF1 gene encodes a protein involved in steroid synthesis. *Plant Cell* **10**:1677-90.

Klucher, K.M., Chow, H., Reiser, L., Fischer, R.L. (1996). The AINTEGUMENTA gene of *Arabidopsis* required for ovule and female gametophyte development is related to the floral homeotic gene APETALA2. *Plant Cell* **8**:137-53.

Kohl, N., Conner, M., Gibbs, J., Graham, S., Hartman, G., Oliff, A. (1995). Development of inhibitors of protein farnesylation as potential chemotherapeutic agents. *J. Cell Biochem. Suppl.* **22**:145-50.

Kohonen, T. (1990). The self-organizing map. *Proc. IEEE* **78**:1464-80.

Lange, B.M., Ghassemian, M. (2003). Genome organization in *Arabidopsis thaliana*: a survey for genes involved in isoprenoid and chlorophyll metabolism. *Plant Mol. Biol.* **51**:925-48.

- Laule, O., Fürholz, A., Chang, H.S., Zhu, T., Wang, X., Heifetz, P.B., Grisse, W., Lange, M.** (2003). Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **100**:6866-71.
- Lee-Chen, S., Steinitz-Sears, L.M.** (1967). The location of linkage groups in *Arabidopsis thaliana*. *Can. J. Genet. Cytol.* **9**:381-4.
- Leung, J., Bouvier-Durand, M., Morris, P.C., Guerrier, D., Cheddor, F., Giraudat, J.** (1994). Arabidopsis ABA response gene ABI1: features of a calcium-modulated protein phosphatase. *Science* **264**:1448-52.
- Leung, J., Merlot, S., Giraudat, J.** (1997). The Arabidopsis ABSCISIC ACID-INSENSITIVE2 (ABI2) and ABI1 genes encode homologous protein phosphatases 2C involved in abscisic acid signal transduction. *Plant Cell* **9**:759-71.
- Leyser, H.M.O., Lincoln, C.A., Timpfe, C., Lammer, D., Turner, J., Estelle, M.** (1993). *Arabidopsis* auxin-resistance gene *AXR1* encodes a protein related to ubiquitin-activating enzyme E1. *Nature* **364**:161-4.
- Li, J., Chory, J.** (1997). A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. *Cell* **90**:929-38.
- Lichtenthaler, H.K., Rohmer, M. & Schwender, J.** (1997a). Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. *Physiol. Plant.* **101**:643-52.
- Lichtenthaler, H.K., Schwender, J., Disch, A., Rohmer, M.** (1997b). Biosynthesis of isoprenoids in higher plant chloroplasts proceeds via a mevalonate-independent pathway. *FEBS Lett.* **400**:271-4.

- Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K., Hake, S.** (1994). A *Knotted1*-like homeobox gene in *Arabidopsis* is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. *Plant Cell* **6**:1859-76.
- Lindgren, L.O., Stahlberg, K.G., Hoglund, A.S.** (2003). Seed-specific overexpression of an endogenous *Arabidopsis* phytoene synthase gene results in delayed germination and increased levels of carotenoids, chlorophyll, and abscisic acid. *Plant Physiol.* **132**:779-85.
- Long, S.B., Casey, P.J., Beese, L.S.** (2002). Reaction path of protein farnesyltransferase at atomic resolution. *Nature* **419**:645-50.
- Loraine, A.E., Yalovsky, S., Fabry, S., Gruissem, W.** (1996). Tomato Rab1A homologs as molecular tools for studying Rab geranylgeranyl transferase in plant cells. *Plant Physiol.* **110**:1337-47.
- Maity, S.N., de Crombrughe, B.** (1998). Role of the CCAAT-binding protein CBF/NF-Y in transcription. *Trends Biol. Sci.* **23**:174-8.
- Maksymowych, R.** (1963). Cell division and cell elongation in leaf development of *Xanthium pennsylvanicum*. *Amer. J. Bot.* **50**:891-901.
- Maltese, W.A., Sheridan, K.M.** (1987). Isoprenylated proteins in cultured cells: subcellular distribution and changes related to altered morphology and growth arrest induced by mevalonate deprivation. *J. Cell Physiol.* **133**:471-81.
- Mandel, M.A., Gustafson-Brown, C., Savidge, B., Yanofsky, M.F.** (1992). Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1*. *Nature* **360**:273-7.
- Mantovani, R.** (1999). The molecular biology of the CCAAT-binding factor NF-Y. *Gene* **239**:15-27.

- Marx, G.A.** (1983). Developmental mutants in some annual seed plants. *Annu. Rev. Plant Physiol.* **34**:389-417.
- McNally, F.J., Vale, R.D.** (1993). Identification of katanin, an ATPase that severs and disassembles stable microtubules. *Cell* **75**:419-29.
- Meijer, M., Murray, J.A.H.** (2001). Cell cycle controls and the development of plant form. *Curr. Opin. Plant Biol.* **4**:44-9.
- Michaut, L., Flister, S., Neeb, M., White, K.P., Certa, U., Gehring, W.J.** (2003). Analysis of the eye developmental pathway in *Drosophila* using DNA microarrays. *Proc. Natl. Acad. Sci. USA* **100**:4024-9.
- Millar, A.A., Kunst, L.** (1997). Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. *Plant J.* **12**:121-31.
- Miyoshi K, Ito Y, Serizawa A, Kurata N.** (2003). OsHAP3 genes regulate chloroplast biogenesis in rice. *Plant J.* **36**:532-40.
- Moore, M., Blobel, G.** (1993). The GTP binding protein ran/TC4 is required for protein import into the nucleus. *Nature* **365**:661-3.
- Moores, S.L., Schaber, M.D., Mosser, S.D., Rands, E., O'Hara, M.B., Garsky, V.M., Marshall, M.S., Pompliano, D.L., Gibbs, J.B.** (1991). Sequence dependence of protein isoprenylation. *J. Biol. Chem.* **266**:14603-10.
- Morehead, T.A., Biermann, B.J., Crowell, D.N., Randall, S.K.** (1995). Changes in protein isoprenylation during the growth of suspension-cultured tobacco cells. *Plant Physiol.* **109**:277-84.

- Nagatani, A., Chory, J., Furuya, M.** (1991). Phytochrome B is not detectable in the *hy3* mutant of *Arabidopsis*, which is deficient in responding to end-of-day farred light treatments. *Plant Cell Physiol.* **32**:1119-22.
- Nagatani, A., Reed, J.W., Chory, J.** (1993). Isolation and initial characterization of *Arabidopsis* mutants that are deficient in phytochrome A. *Plant Physiol.* **102**:269-77.
- Nagpal, P., Walker, L.M., Young, J.C., Sonawala, A., Timpte, C., Estelle, M., Reed, J.W.** (2000). *AXR2* encodes a member of the Aux/IAA protein family. *Plant Physiol.* **123**:563-73.
- Nakaya, M., Tsukaya, H., Murakami, N., Kato, M.** (2002). Brassinosteroids control the proliferation of leaf cells of *Arabidopsis thaliana*. *Plant Cell Physiol.* **43**:239-44.
- Nambara, E., McCourt, P.** (1999). Protein farnesylation in plants: a greasy tale. *Curr. Opin. Plant Biol.* **2**:388-92.
- Neff, M.M., Van Volkenburgh, E.** (1994). Light-stimulated cotyledon expansion in *Arabidopsis* seedlings. The role of phytochrome B. *Plant Physiol.* **104**:1027-32.
- Nicol, F., His, I., Jauneau, A., Vernhettes, S., Canut, H., Hofte, H.** (1998). A plasma membrane-bound putative endo-1,4- β -Dglucanase is required for normal wall assembly and cell elongation in *Arabidopsis*. *EMBO J.* **17**:5563 - 76.
- Njus, D., Kelley, P.M.** (1993). The secretory-vesicle ascorbate-regenerating system: a chain of concerted H⁺/e⁻-transfer reactions. *Biochim. Biophys. Acta* **1144**:235-48.

- Novick, P., Brennwald, P.** (1993). Friends and family: the role of the Rab GTPases in vesicular traffic. *Cell* **75**:597-601.
- Parmryd, I., Shipton, C.A., Swiezewska, E., Dallner, G., Andersson, B.** (1997). Chloroplastic prenylated proteins. *FEBS Lett.* **414**:527-31.
- Pease, A.C., Solas, D., Sullivan, E.J., Cronin, M.T., Holmes, C.P., Fodor, S.P.** (1994). Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci. USA* **91**:5022-6.
- Pei, Z.M., Ghassemian, M., Kwak, C.M., McCourt, P., Schroeder, J.I.** (1998). Role of farnesyltransferase in ABA regulation of guard cell anion channels and plant water loss. *Science* **282**:287-90.
- Pfannschmidt, T., Schutze, K., Fey, V., Sherameti, I., Oelmüller, R.** (2003). Chloroplast redox control of nuclear gene expression—a new class of plastid signals in interorganellar communication. *Antioxid Redox Signal.* **5**:95-101.
- Poethig, R.S., Sussex, I.M.** (1985). The developmental morphology and growth dynamics of the tobacco leaf. *Planta* **165**:158–69.
- Powers, S., Michaelis, S., Broek, D., Santa Anna, S., Field, J., Herskowitz, I., Wigler, M.** (1986). RAM, a gene of yeast required for a functional modification of RAS proteins and for production of mating pheromone a-factor. *Cell* **47**:413-22.
- Qian, D., Zhou, D., Ju, R., Cramer, C.L., Yang, Z.** (1996). Protein farnesyltransferase in plants: molecular characterization and involvement in cell cycle control. *Plant Cell* **8**:2381-94.
- Randall, S.K., Marshall, M.S., Crowell, D.N.** (1993). Protein isoprenylation in suspension-cultured tobacco cells. *Plant Cell* **5**:433-42.

- Rédei, G.P.** (1962). Single locus heterosis. *Z. Vererbungs.* **93**:164-70.
- Reed, J.W., Nagpal, P., Poole, D.S., Furuya, M., Chory, J.** (1993). Mutations in the gene for the red/far-red light receptor phytochrome B alter cell elongation and physiological responses throughout *Arabidopsis* development. *Plant Cell* **5**:147-57.
- Ridley, A.J., Hall, A.** (1992). The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* **70**:389-99.
- Röbbelen, G.** (1957). Über Heterophyllie bei *Arabidopsis thaliana* (L.) Heynh. *Ber. Dt. Bot. Ges.* **70**:39-44.
- Rodriguez-Concepcion, M., Yalovsky, S., Zik, M., Fromm, H., Grisse, W.** (1999). The prenylation status of a novel plant calmodulin directs plasma membrane or nuclear localization of the protein. *EMBO J.* **18**:1996-2007.
- Running, M.P., Lavy, M., Sternberg, H., Galichet, A., Grisse, W., Hake, S., Ori, N., Yalovsky, S.** (2004). Enlarged meristems and delayed growth in *plp* mutants result from lack of CaaX prenyltransferases. *Proc. Natl. Acad. Sci. USA* **20**:7815-20.
- Sano, H., Ohashi, Y.** (1995). Involvement of small GTP-binding proteins in defense signal transduction pathways of higher plants. *Proc. Natl. Acad. Sci. USA* **92**:4138-44.
- Schafer, W.R., Kim, R., Sterne, R., Thorner, J., Kim, S.H., Rine, J.** (1989). Genetic and pharmacological suppression of oncogenic mutations in ras genes of yeast and humans. *Science* **245**:379-85.

- Schafer, W.R., Rine, J.** (1992). Protein prenylation: genes, enzymes, targets, and functions. *Ann. Rev. Genet.* **26**:209-37.
- Scheibe, R., Kampfenkel, K., Wessels, R., Tripier, D.** (1991). Primary structure and analysis of the location of the regulatory disulfide bond of pea chloroplast NADP-malate dehydrogenase. *Biochim. Biophys. Acta.* **1076**:1-8.
- Scherf, U., Ross, D.T., Waltham, M., Smith, L.H., Lee, J.K., Tanabe, L., Kohn, K.W., Reinhold, W.C., Myers, T.G., Andrews, D.T., Scudiero, D.A., Eisen, M.B., Sausville, E.A., Pommier, Y., Botstein, D., Brown, P.O., Weinstein, J.N.** (2000). A gene expression database for the molecular pharmacology of cancer. *Nat. Genet.* **24**:236-44.
- Schmidt, R.A., Glomset, J.A., Wight, T.N., Habenicht, A.J., Ross, R.** (1982). A study of the Influence of mevalonic acid and its metabolites on the morphology of swiss 3T3 cells. *J. Cell Biol.* **95**:144-53.
- Schmidt, R.A., Schneider, C.J., Glomset, J.A.** (1984). Evidence for post-translational incorporation of a product of mevalonic acid into Swiss 3T3 cell proteins. *J. Biol. Chem.* **259**:10175-80.
- Schultz, G., Ellerbrock, B.H., Soll, J.** (1981). Site of prenylation reaction in synthesis of phyloquinone (vitamin K1) by spinach chloroplasts. *Eur. J. Biochem.* **117**:329-32.
- Schürmann, P., Jacquot, J.P.** (2000). Plant thioredoxin system revisited. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **51**:371-400.
- Sebti, S.M., Hamilton, A.D.** (2000). Farnesyltransferase and geranylgeranyltransferase I inhibitors and cancer therapy: lessons from mechanism and bench-to-bedside translational studies. *Oncogene* **19**:6584-93.

- Siendones, E., Gonzalez-Reyes, J.A., Santos-Ocana, C., Navas, P., Córdoba, F.** (1999). Biosynthesis of ascorbic acid in kidney bean. L-galactono-gamma-lactone dehydrogenase is an intrinsic protein located at the mitochondrial inner membrane. *Plant Physiol.* **120**:907-12.
- Shipton, C.A., Parmryd, I., Swiezewska, E., Andersson, B., Dallner, G.** (1995). Isoprenylation of plant proteins in vivo. Isoprenylated proteins are abundant in the mitochondria and nuclei of spinach. *J. Biol. Chem.* **270**:566-72.
- Sinensky, M., Logel, J.** (1985). Defective macromolecule biosynthesis and cell-cycle progression in a mammalian cell starved for mevalonate. *Proc. Natl. Acad. Sci. U S A* **82**:3257-61.
- Smith, L.G., and Hake, S.** (1992). The initiation and determination of leaves. *Plant Cell* **4**:1017-27.
- Smith, L.G., and Hake, S.** (1993). Molecular genetic approaches to leaf development: *Knotted* and beyond. *Can. J. Bot.* **72**:617-25.
- Sneath, P.H.A., Sokal, R.R.** (1973). *Numerical Taxonomy*. Freeman, San Francisco.
- Stimmel, J.B., Deschenes, R.J., Volker, C., Stock, J., Clarke, S.** (1990). Evidence for an S-farnesylcysteine methyl ester at the carboxyl terminus of the *Saccharomyces cerevisiae* RAS2 protein. *Biochemistry* **29**:9651-9.
- Sultan, S.E.** (2000). Phenotypic plasticity for plant development, function and life history. *Trends Plant Sci.* **5**:537-42.
- Swiezewska, E., Thelin, A., Dallner, G., Andersson, B., Ernster, L.** (1993). Occurrence of prenylated proteins in plant cells. *Biochem. Biophys. Res. Commun.* **192**:161-6.

- Szekeres, M., Németh, K., Koncz-Kálmán, Z., Mathur, J., Kauschmann, A., Altmann, T., Rédei, G.P., Nagy, F., Schell, J., Koncz, C. (1996).** Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in *Arabidopsis*. *Cell* **85**:171–82.
- Takahashi, T., Gasch, A., Nishizawa, N., Chua, N.-H. (1995).** The *DIMINUTO* gene of *Arabidopsis* is involved in regulating cell elongation. *Genes Dev.* **9**:97-107.
- Tamanoi, F. (1993).** Inhibitors of Ras farnesyltransferase. *Trends Biochem. Sci.* **18**:349-53.
- Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., Golub, T.R. (1999).** Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. U S A* **96**:2907-12.
- Thompson Jr, G.A., Okuyama, H. (2000).** Lipid-linked proteins in plants. *Progress in Lipid Research* **39**:19-39.
- Timpte, C.C., Wilson, A.K., Estelle, M. (1992).** Effects of the *axr2* mutation of *Arabidopsis* on cell shape in hypocotyl and inflorescence. *Planta* **188**:271-78.
- Törönen, P., Kolehmainen, M., Wong, G., Castren, E. (1999).** Analysis of gene expression data using self-organizing maps. *FEBS Lett.* **451**:142-6.
- Tsuge, T., Tsukaya, H., Uchimiya, H. (1996).** Two independent and polarized processes of cell elongation regulate leaf blade expansion in *Arabidopsis thaliana* (L.) Heynh. *Development* **122**:1589-600.

- Tsukaya, H.** (2002). Interpretation of mutants in leaf morphology: genetic evidence for a compensatory system in leaf morphogenesis that provides a new link between cell and organismal theory. *Int. Rev. Cytol.* **217**:1–39.
- Tsukaya H, Kozuka T, Kim GT.** (2002). Genetic control of petiole length in *Arabidopsis thaliana*. *Plant Cell Physiol.* **43**:1221-8.
- Tsukaya, H., Naito, S., Rédei, G.P., Komeda, Y.** (1993). A new class of mutations in *Arabidopsis thaliana*, *acaulis1*, affecting the development of both inflorescence and leaves. *Development* **118**:751–64.
- Tsukaya, H., Tsuge, T., Uchimiya, H.** (1994). The cotyledon: a superior system for studies of leaf development. *Planta* **195**:309-12.
- Umeda, M., Bhalerao, R.P., Schell, J., Uchimiya, H., Koncz, C.** (1998). A distinct cyclin-dependent kinase-activating kinase of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U S A* **95**:5021-6.
- Van Volkenburgh, E.** (1999). Leaf expansion – an integrating plant behavior. *Plant Cell Environ.* **22**:1463–73.
- Wang, R., Okamoto, M., Xing, X., Crawford, N.M.** (2003). Microarray analysis of the nitrate response in *Arabidopsis* roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. *Plant Physiol.* **132**:556-67.
- Wen, X., Fuhrman, S., Michaels, G.S., Carr, D.B., Smith, S., Barker, J.L., Somogyi, R.** (1998). Large-scale temporal gene expression mapping of central nervous system development. *Proc. Natl. Acad. Sci. U S A* **95**:334-9.
- Wolda, S.L., Glomset, J.A.** (1988). Evidence for modification of lamin B by a product of mevalonic acid. *J. Biol. Chem.* **263**:5997-6000.

- Yalovsky, S., Kulukian, A., Rodriguez-Concepcion, M., Young, C.A., Grisse, W.** (2000). Functional requirement of plant farnesyltransferase during development in *Arabidopsis*. *Plant Cell* **12**:1267-78.
- Yalovsky, S., Rodriguez-Concepcion, M., Grisse, W.** (1999). Lipid modifications of proteins - slipping in and out of membranes. *Trends Plant Sci.* **4**:439-45.
- Yang, Z., Cramer, C.L., Watson, J.C.** (1993). Protein farnesyltransferase in plants. Molecular cloning and expression of a homolog of the beta subunit from the garden pea. *Plant Physiol.* **101**:667-74.
- Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A., Meyerowitz, E.M.** (1990). The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcriptional factors. *Nature* **346**:35-9.
- Zhang, F.L., Casey, P.J.** (1996). Protein prenylation: molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* **65**:241-69.
- Zhang, N., Portis, A.R. Jr.** (1999). Mechanism of light regulation of Rubisco: a specific role for the larger Rubisco activase isoform involving reductive activation by thioredoxin-f. *Proc. Natl. Acad. Sci. U S A* **96**:9438-43.
- Zhou, D., Qian, D., Cramer, C.L., Yang, Z.** (1997). Developmental and environmental regulation of tissue- and cell-specific expression for a pea protein farnesyltransferase gene in transgenic plants. *Plant J.* **12**:921-30.
- Zhu, J.K., Bressan, R.A., Hasegawa, P.M.** (1993). Isoprenylation of the plant molecular chaperone ANJ1 facilitates membrane association and function at high temperature. *Proc. Natl. Acad. Sci. U S A* **90**:8557-61.

Acknowledgements

First and foremost, I would like to thank Prof. Wilhelm Gruissem for giving me the opportunity to work in his group on this exciting project and for his professional supervision, guidance and encouragement throughout my graduate studies.

I would like to extend my appreciation also to Prof. N. Amrhein and Prof. M. Aebi for having accepted the role of co-referee in the examination of this doctoral thesis.

Special thanks go to my colleague Oliver Laule for a relaxed, familiar and cooperative atmosphere within our co-joint project.

I am also grateful to my colleagues in the lab: Arnaud Galichet for his support within the project and all the others for happy hours, cosy lunch time and very good atmosphere.

In a special way, I owe my deepest gratitude to my wife Julia for her patience, advice and emotional support during the whole thesis.

Last but not least I express my deepest thank to my parents and my friends for their patience and never ending support during my education.

Curriculum vitae

Name: Andreas Fürholz

Date of birth: 29.08.1967

Civil status: married



- 1974 - 1985 Primar- and Bezirksschule in Biberist
Kantonsschule in Solothurn
- 1985 – 1990 Teacher on different education levels in Biberist and Zuchwil
- 1990 - 1993 Eidgenössische Maturität Typus C (AKAD)
- 1993 - 1995 Teacher on different education levels
- 1995 - 1999 Studies in biology at the Swiss Federal Institute of Technology ETHZ
- 1999 Diploma thesis at the Institute of Plant Sciences ETHZ
Biochemistry and Physiology of Plants (Prof. N. Amrhein, Dr. A.
Schaller)
Overexpression, purification and characterisation of a flavin
dependent reductase from *Lycopersicon esculentum* Mill.
- 2000 Practical course at the department of clinical research, University
hospital Berne
- 2001 - 2004 PhD position at the Institute of Plant Sciences ETHZ
Biotechnology of Plants (Prof. W. Gruissem)