This is to certify that the

dissertation entitled

VACUOLAR TRANSPORT OF ctVSS-BEARING PROTEINS: A GENETIC APPROACH

presented by

Sridhar Venkataraman

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Biochemistry and Molecular Biology

N. Raikhel
Major professor

Date 2/12/01

VACUOLAR TRANSPORT OF ctVSS-BEARING PROTEINS: A GENETIC APPROACH

By

Sridhar Venkataraman

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry and Molecular Biology

2001

ABSTRACT

VACUOLAR TRANSPORT OF ctVSS-BEARING PROTEINS: A GENETIC APPROACH

By

Sridhar Venkataraman

Plants respond to stress and pathogen attack by synthesizing and storing defense-related proteins in the vacuole. Many of these proteins are targeted to the vacuole by virtue of sorting information residing in a C-terminal vacuolar sorting signal (ctVSS). While many studies have been conducted to identify the nature of these sorting signals, very little is known about the factors they interact with or the factors involved in the transport mechanism that delivers these proteins to the vacuole. A genetic approach has been used in *Arabidopsis thaliana* (L.) Heynh, to identify loci involved in the ctVSS-bearing protein transport pathway. Mutants were screened from a collection of EMS-treated *Arabidopsis thaliana* ecotype Columbia seeds expressing rat-preputial β-D glucuronidase with a ctVSS from tobacco chitinase and barley lectin. Mutants *cvs1* and *cvs2* (for ctVSS vacuolar sorting) were identified and characterized by immunoelectron microscopy. In these mutants, barley lectin was found to be partially mislocalized in the apoplast along the cell wall and in intercellular spaces of root sections. Such results were not observed in leaf sections, indicating the identification of root specific mutants. The mutant locus of *cvs1* was mapped to a contig of BACs between AtSO191 and DFR on chromosome V. Further analyses will have to be carried out to clone the *cvs1* mutant locus. However, the research presented describes identification of a factor potentially involved in transporting ctVSS bearing proteins to the vacuole.

To my parents Vasantha and R. Venkataraman and my wife Meera

ACKNOWLEDGMENTS

My graduate studies have involved many colleagues and faculty members who have been part of my development as a student. First and foremost, I would like to thank my advisor Natasha for having supported me all these years and helped me conduct my research. She has been instrumental in teaching me how to be critical in reviewing scientific work. I would also like to thank Dr Kenneth Keegstra for helping me join the PRL. I owe a lot to him. My committee members, Dr Kroos, Dr Preiss and Dr Smith, have also helped me through these years as a PhD student.

A number of colleagues have taught me a lot in these years: Ahmed Faik, John Froehlich, Brett Brzobohaty, Philipp Kapranov, Krzysztof Szczyglowski and Sergei Mekhedov. Coming to the Raikhel Lab: Past – Jim "Jim is always right" Dombrowski, Maor "Bapeti" Bar-Peled, Sharif "Jim II" Ahmed; and present – Diane Bassham, Enrique Rojo, Valya Kovaleva, Rodrigo Sarria-Millan, Haiyan Zheng, Ivan Delgado, Emily Avila-Teegarden, Curtis Wilkerson and Sybil Myers.

I am especially grateful to Jim Dombrowski to have initiated my project. These past few months would have been impossible without the help of Tony Sanderfoot (the Universal Thesis Master) and John Scott-Craig (Seminario Oratorio Major).

Finally my family members: my wife Meera, my parents Vasantha and R. Venkataraman, Latha, Zoltan and now Milan. I must not forget some of my friends who have been around when I needed them: Mundath Narayanan Menon, Penmetsa Ramachandra Varma, Hemant Varma and Satyashayee Misra.

TABLE OF CONTENTS

CHA

REFE

APPE

Mapp

APPE

Bradf

APPE

Tables
and the

LIST

LIST OF TABLES

LIS

Figu

Figu

Figu

Figu

Figu

Figu

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

Figure

LIST OF FIGURES

ABI

AP-
BAC
BCI
BL
BP
BSA
CAP
CCV
cDN
cM
Col
COP
COP
Con
CTAB
ctVSS
CVS
CW
Da
DEPC
DMSO
DNA
dNTPs
E coli
EDTA
EMS
ER
F_1
F_2
FV
GTP
Het
ICS
Ler
LV
M_1
M_2

M_3

MES
MOPS
mRNA
MU

ABBREVIATIONS

AP-1	Clathrin-associated adaptor complex 1
BAC	Bacterial artificial chromosome
BCIP	5-Bromo-4-chloro-3-indolyl phosphate p-Toluidine Salt
BL	Barley lectin
BP80	Binding Protein 80
BSA	Bovine serum albumin (fraction V)
CAPS	Cleaved amplified polymorphic sequences
CCV	Clathrin-coated vesicles
cDNA	copy deoxyribo nucleic acid
cM	centi Morgan
Col	*Arabidopsis thaliana* ecotype Columbia
COPI	Coat-protein (I) coated vesicles
COPII	Coat-protein (II) coated vesicles
Cotyl	Cotyledon
CTAB	Cetyltrimethylammonium bromide
ctVSS	C-terminal vacuolar sorting signal
CVS	ctVSS vacuolar sorting (gene or mutant)
CW	Cell wall
Da	Dalton
DEPC	Diethyl pyrocarbonate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribo nucleic acid
dNTPs	Deoxynucleotide triphosphate
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid disodium salt
EMS	Ethyl methane sulfonate
ER	Endoplasmic reticulum
F_1	First filial generation
F_2	Second filial generation (usually self-fertilized progeny of F_1)
FV	Fused vacuole
GTP	Guanosine 5'-triphosphate Sodium salt
Het	Heterozygous
ICS	Itercellular spaces
L*er*	*Arabidopsis thaliana* ecotype landsberg *erecta*
LV	Lytic vacuole
M_1	Primary mutagenized generation
M_2	Secondary mutagenized generation (usually self-fertilized progeny of M_1)
M_3	Tertiary mutagenized generation (usually self-fertilized progeny of M_2)
MES	2-(N-Morpholino)ethanesulfonic acid
MOPS	3-(N-Morpholino)propanesulfonic acid
mRNA	Messenger ribonucleic acid
MU	Methyl umbelliferone

ORF
PAC
PBS
PCR
PH
PV
PS
Ps
PT
PV
Ra
Ra

Ra
RF
RF
RN
RR
SDS
SDS
SDV
SNA
SP
SSL
SV
SV
T
T
T
T
T
TBE
TC
T-DN
TE
TGN
TIP
V
VV
VPA
VPS
WV
WGA
WT
X-Gl

ORF	Open reading frame
PAC	Precursor accumulating vesicles
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PHA	Phytohemagglutinin
PMSF	Phenylmethylsulfonyl fluoride
PSV	Protein storage vacuole
PsVSS	Physical structure vacuolar sorting signal
PT	Potato inhibitor terminator
PVC	Prevacuolar compartment
Rat-GUS	Rat preputial β-glucuronidase
Rat-GUS-ctVSS	Rat preputial β-glucuronidase fused to tobacco basic chitinase ctVSS
Rat-GUS-Delta	Secreted Rat preputial β-glucuronidase
RF	Recombination frequency
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RR Buffer	RNA resuspension buffer
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate poly acrylamide gel electrophoresis
SDV	Smooth dense vesicles
SNARE	soluble N-ethyl maleimide sensitive factor adaptor protein receptor
SP	Signal peptide
SSLP	Simple sequence length polymorphism
SsVSS	Sequence specific vacuolar sorting signal
SV	Secretory vesicle
T_1	Primary transformants
T_2	Secondary transformants
T_3	Tertiary transformants
T_4	Quaternary transformants
T_5	Fifth generation transformants
TBE	Tris Borate EDTA
TC	Tobacco basic chitinase
T-DNA	Transfer DNA
TE	Tris EDTA
TGN	*Trans*-Golgi network
TIP	Tonoplast intrinsic protein
V	Vacuole
v/v	Volume/volume
VPA	Vacuolar protein aggregates
VPS	Vacuolar protein sorting
w/v	weight/volume
WGA	Wheat germ agglutinin
WT	Wild type
X-GlcU	5-Bromo-4-chloro-3-indolyl beta-D-glucuronide Sodium salt

YAC	Yeast artificial chromosome

Chapter 1

Protein targeting to the plant vacuole: the ctVSS story

Plant

and

depe

repo

com

encl

barl

199

exist

199

vacu

enzy

199

PSV

Mart

Storag

TIP,

inabil

indica

INTRODUCTION

Plant vacuoles

Plant vacuoles were first observed to be large spaces devoid of cytoplasmic matter and hence termed vacuole. This large organelle occupies 30 to 90 % of the plant cell depending on the type of cell under observation. Functionally the vacuole is a large repository of inorganic ions, organic acids, sugars, secondary metabolites, plant defense compounds, stress-related solutes, detoxified substances, enzymes and storage proteins enclosed by the vacuole membrane (tonoplast) (Marty, 1999). Recent evidence from barley root cells, barley aleurone cells and tobacco cell (Paris et al, 1996, Swanson et al, 1998 and Di Sansebastiano et al, 1998) indicates that two or more types of vacuoles may exist in a plant cell, each characterized by different pH optima (Di Sansebastiano et al, 1998) and different aquaporins in the respective tonoplasts (Jauh et al, 1998). The low pH vacuole is termed the lytic vacuole (LV) due to the presence of low pH optima hydrolytic enzymes, whereas the near-neutral vacuole, the protein storage vacuole (PSV) (Marty, 1999). The LV is analogous to the yeast vacuole and the mammalian lysosome while the PSV is unique to plants (Marty, 1999). The lytic vacuole has been reviewed in detail by Marty (1999) and will not be discussed.

Storage proteins and the PSV

The protein storage vacuole has been found to have a tonoplast decorated with α-TIP, an aquaporin found in seeds and root tips. It differs from the lytic vacuole in its inability to be stained with neutral red, a dye that accumulates in acidic compartments, indicating a neutral pH (Di Sansebastiano et al, 1998). This compartment accumulates

storage proteins such as globulins and prolamins, proteins characterized by their relative solubility in aqueous or ethanolic buffers. Globulins are found in all dicots and many monocots. Globulins are of the vicilin type (7S globulins) or the legumin type (11S globulins). Common 7S globulins include vicillin (from *Pisum sativum*), phaseolin (from *Phaseolus vulgaris*) and conglycinin (from *Glycine max*). The legumin category of proteins includes glycinin from *Glycine max*. Prolamins are predominantly found in the endosperm of monocots such as wheat and rice. Apart from storage proteins, the PSV also contains many ancillary storage proteins such as barley lectin and bean phytohemagglutinin as well as hydrolase inhibitors such as α-amylase inhibitors (Marty, 1999). These proteins may serve a defense role in plants against herbivory, insect and pathogen attack (reviewed by Marty, 1999; Herman and Larkins, 1999).

The ability of plants to synthesize and store proteins has made them very important as a primary source of protein for animal species. In modern times, these aspects of plants have been utilized in developing engineered proteins in plants to improve the nutritive value of plants and their products (Lawrence et al, 1994). These aspects have also been exploited in developing plants with therapeutic properties. The *Escherichia coli* enterotoxin B protein has been expressed in potatoes as an oral vaccine against *E. coli*-induced diarrhea (Mason et al, 1998). Similarly, the hepatitis B surface antigen has been expressed in potatoes as an oral vaccine against hepatitis B (Richter et al, 2000) and the Guy 13 monoclonal antibody against *Streptococcus mutans* (agent of dental caries) was expressed in tobacco (Cabanes-Macheteau et al, 1999). While most of these developments are still in their infancy, the prospects of plant based therapies for common ailments are exciting.

Protein targeting to the vacuole in plants

Plant proteins destined for the vacuole are synthesized as precursors by ribosomes associated with the endoplasmic reticulum (ER), and are co-translationally inserted into the ER lumen (see Figure 1.1). In the ER lumen, they undergo several modifications such as the proteolytic processing of the signal peptide, N-linked glycosylation, disulfide bond formation, and protein folding. These steps are regulated by a "quality control" step following which these proteins are thought to be packaged into COPII vesicles bound for the Golgi complex (see Pimpl et al, 2000). Certain proteins such as cereal prolamins and pumpkin 2S albumin are packaged into protein bodies or precursor accumulating (PAC) vesicles bound for the protein storage vacuole. After the COPII vesicles deliver their cargo to the *cis*-Golgi, proteins further undergo maturation processes such as complex mannosylation of the glycan as they move through the medial and *trans*-Golgi via a COPI vesicle mediated manner (see Pimpl et al, 2000). Eventually proteins arrive at the *trans*-Golgi network (TGN) where they are sorted according to localization signals residing on the protein. Proteins with vacuolar targeting signals are sent to the vacuole via the prevacuolar compartment (PVC) in clathrin-coated vesicles or smooth dense vesicles (depending on the targeting signal). Proteins, which lack a vacuolar targeting signal or a Golgi retention signal, are delivered to the extracellular space via the default pathway of exocytosis.

Figure
Solubl
Once i
contai
Vesicl
the Go
(dark
(PSV).

1) ssV
the pre
2) Pro
packag
and L
3) Pro
secret

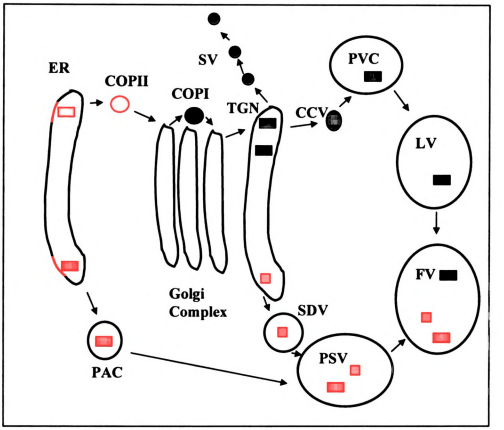

Figure 1.1 Model of the plant secretory pathway

Soluble proteins are synthesized by ribosomes associated with the Endoplasmic Reticulum (ER). Once in the lumen of the ER they are processed and eventually packaged into a) COPII vesicles containing cargo such as barley lectin or sporamin (yellow) or b) Precursor Accumulating Vesicles (PAC) as in the case of pumpkin 2S albumin (red). COPII vesicles deliver their cargo to the Golgi complex where they are processed as they proceed by a COPI vesicle mediated process (dark green). PAC vesicles, in contrast, deliver their cargo (red) to the Protein Storage Vacuole (PSV). Proteins processed in the Golgi network reach the TGN where they are sorted.

1) ssVSS bearing proteins (blue) are packaged into Clathrin-Coated Vesicles (CCV) en route to the prevacuolar compartment (PVC). These proteins ultimately reach the Lytic Vacuole (LV).
2) Proteins bearing a ctVSS such as barley lectin or a psVSS such as legumin (brown) are packaged into Smooth Dense Vesicles (SDV) en route to the PSV. In vegetative cells, the PSV and LV often fuse to form one large vacuole (FV).
3) Proteins which lack a Golgi retention signal or a vacuolar targeting signal are packaged into secretory vesicles and their cargo (green) are secreted into the apoplast.

The Endoplasmic reticulum

Storage proteins are synthesized on the rough endoplasmic reticulum (ER) and are co-translationally inserted into the ER lumen. This entry into the ER is specified by a signal peptide residing at the N-terminus of the nascent polypeptide and consists of 16 to 30 residues organized in three domains: a positively charged amino terminus, followed by a hydrophobic region and a polar region terminated by a cleavage site (von Heijne, 1985). The signal peptide is cleaved upon entry into the ER. Once in the ER lumen, proteins undergo various modifications such as N-linked glycosylation, disulphide bond formation, oxidation of prolines to hydroxyprolines, trimming of the glycan and protein folding (reviewed in Galili et al, 1998). There is significant evidence to indicate the existence of a quality control step in the ER resulting in the reverse translocation of misfolded proteins to the cytosol (reviewed in Vitale and Denecke, 1999). Proteins such as pumpkin 2S albumin and monocot prolamins exit the ER via PAC vesicles or protein bodies respectively and are transported directly to the vacuole (Okita and Rogers, 1996; Galili et al, 1998; Hara-Nishimura et al, 1998; Choi et al, 2000).

The transport of soluble proteins from the ER to the Golgi has been found to occur via COPII vesicles in yeast (for review see Barlowe, 1998). This mechanism involves Sec12p (GTP-exchange protein), Sar1p (a GTPase), Sec23p, Sec24p, Sec13p and Sec31p (for review see Barlowe, 1998). Similarly, in plants the ER to Golgi traffic is attributed to a homologous COPII system. This hypothesis was supported by the identification of plant homologues of the yeast proteins. The homologues identified are AtSAR1 (Bar-Peled and Raikhel, 1997; Takeuchi et al, 2000), AtSEC12 (Bar-Peled and Raikhel, 1997), and AtSEC23 (Movafeghi et al, 1999).

Golgi and the plant secretory pathway

The plant Golgi body can be discerned as a stack of membrane-bound compartments (cisternae) associated with a large network of tubules and vesicle buds called the *trans*-Golgi network or TGN. The number of Golgi stacks varies depending on the plant species and can range from one stack per *Chlamydomonas* cell to thousands in giant fiber cells in cotton (for review Andreeva et al, 1998). Plant Golgi bodies carry out a number of functions including protein glycosylation, lipid biosynthesis and polysaccharide synthesis. Based on the localization of Golgi enzymes, polysaccharides, and various modifications that proteins undergo in the Golgi stacks, the Golgi cisternae are classified into *cis*-, medial- and *trans*-Golgi (Dupree and Sherrier, 1998).

Typically, proteins are delivered from the ER to the *cis*-Golgi, proceed through the medial- and *trans*-Golgi undergoing various modifications. The modifications include protein glycan α1,3-fucosylation (Lerouge et al, 1998), protein glycan β1,2-xylosylation (Lerouge et al, 1998), proteolytic processing (Jiang and Rogers, 1999) and complex mannosylations of the glycan (Lerouge et al, 1998). Intra-Golgi traffic in plants has been proposed to occur via a COPI vesicle-mediated manner (Andreeva et al, 2000, Pimpl et al, 2000) or via cisternal maturation as in the case of scales in algae (Brown, 1969) and procollagen in mammalian cells (Bonfanti et al, 1998).

Eventually, proteins arrive at the TGN en route for two broad destinations: the extracellular space via the default pathway of exocytosis, or the vacuole via a sorting determinant-mediated pathway. It can be said that in the absence of a positive sorting determinant, soluble proteins are secreted via the default pathway. Currently, very little is known about the vesicular machinery of exocytosis. However it is clear that the process is

Figure
SP = S
terminal

the case

from the

lectins, s

mediated

mediated by calcium ions and annexins and that these vesicles are delivered to the plasma membrane by microfilaments (Carroll et al, 1998, Battey et al, 1999, Sutter et al. 2000). The presence of a positive sorting determinant directs proteins to the vacuole (Vitale and Raikhel, 1999).

Sorting determinants

Sorting determinants can be broadly classified into two categories, propeptide mediated sorting and mature protein domain-mediated (psVSS) sorting (Figure 1.2). In

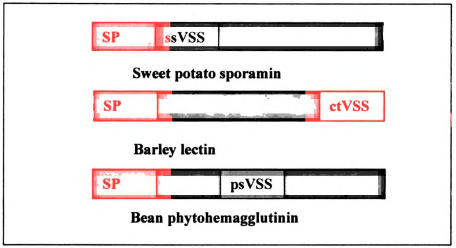

Figure 1.2 Vacuolar soluble proteins and their targeting signals
SP = Signal peptide, ssVSS = sequence specific vacuolar sorting signal, ctVSS = C-terminal vacuolar sorting signal, psVSS = protein structure vacuolar sorting signal.

the case of propeptide-bearing proteins, the propeptide is eventually cleaved and is absent from the mature protein in the vacuole. The study of a number of vacuolar hydrolases, lectins, seed storage proteins and protease inhibitors indicates that vacuolar sorting is mediated by the presence of a propeptide signal at the N-terminus (ssVSS) of the protein

(adjacent to the signal peptide) as in the case of sweet potato sporamin and barley aleurain (Figure 1.2) or at the C-terminus (ctVSS) of the protein (Figure 1.2) as in barley lectin (BL), tobacco chitinase (TC), kiwi actinidin (Paul et al, 1995), bean phaseolin and brazil nut 2S albumin (reviewed by Venkataraman and Raikhel, 1998; Matsuoka and Neuhaus, 1999; Vitale and Raikhel, 1999). The analysis of the psVSS has been reviewed in detail by Neuhaus and Rogers (1998) and will not be covered in this review.

ssVSS

Recently Matsuoka (2000) summarized the properties of the ssVSS as to consist of an X_1-X_2-I/L-X_3-X_4 motif where X_1 lacks a small hydrophobic side chain and Asn is preferred, X_2 is a non acidic amino acid, X_3 is any amino acid and X_4 is a large and preferably hydrophobic amino acid (Table 1.1). Although these signals were originally

Table 1.1 List of representative vacuolar sorting signals

Protein	**ssVSS**
Sweet potato sporamin[1]	. . RF**NPIRL**PTTHEPA . .
Barley aleurain[1]	. . SSSSFADS**NPIRP**VDTDRAAST . .
Arabidopsis aleurain[2]	. . ANIGFDES**NPIRM**VSDGLREVC . .

Protein	**ctVSS**
Barley lectin[1]	. . VFAEAIAANSTLVAE
Tobacco chitinase[1]	. . GLLVDTM
Bean phaseollin[1]	. . AFVY
Tobacco osmotin[1]	. . CPYGSAHNETTNFPLEMPTSSTHEVAK
Kiwi actinidin[3]	. . QNHPKPYSSLINPPAFSMSKDGPVGVDDGQRYSA

ssVSS = sequence specific vacuolar sorting signal; ctVSS = C-terminal vacuolar sorting signal. Letters in bold face represent amino acids forming the conserved motif. 1 = reviewed by Matsuoka and Neuhaus (1999). 2 = Ahmed et al (2000). 3 = Paul et al (1995)

identified at the N-terminus of cargo proteins, Koide et al (1997) demonstrated that the signal is also functional if artificially placed at the C-terminus of a cargo protein. This finding may explain the properties of the C-terminally located targeting signals of Brazil nut 2S albumin (Kirsh et al, 1996), the pumpkin 2S albumin (Shimada et al, 1997) and NaPI (Miller et al, 1999). These C-terminally located signals have a X_1X_2I/LX_3X_4 motif and bind the ssVSS-specific vacuolar sorting receptor BP80 or its orthologs (Matsuoka, 2000). Jiang and Rogers (1998) have demonstrated that the cleavage of the ssVSS occurs in the PVC and that the mature protein is then transported to the vacuole.

To differentiate between proteins with a vacuolar sorting determinant and those that lack a sorting determinant, the existence of receptors at the *trans*-Golgi or *trans*-Golgi network was postulated. Using synthetic peptides corresponding to the ssVSS of barley aleurain and the non-functional mutant ssVSS, Kirsch et al (1994) purified from pea clathrin coated vesicles, an 80 kDa protein (BP80). This protein was found to interact with the aleurain ssVSS with a k_d of 37 nM in a pH dependent manner. Later the Raikhel group cloned AtELP (Ahmed et al, 1997), the *Arabidopsis* ssVSS receptor, and demonstrated that this protein is localized in the TGN and interacts with synthetic peptides representing ssVSSs *in vitro* (Ahmed et al, 2000). The role of AtELP as an ssVSS receptor was further strengthened by observing the co-localization of AtELP and sporamin in *Arabidopsis* cells. However AtELP did not co-localize with BL in *Arabidopsis* cells indicating that the receptor was specific for ssVSS bearing proteins *in vivo*. Synthetic peptides representing the cytosolic domain of AtELP were found to interact with mammalian AP-1 adaptor complexes in-vitro confirming the role of AtELP in sorting ssVSS-bearing proteins and delivering them into clathrin-coated vesicles

(Sanderfoot et al, 1998). AtELP was also co-localized with AtSYP21 (a syntaxin found on the PVC) indicating that this receptor shuttles between the TGN and the PVC (Sanderfoot et al, 1998).

ctVSS

While a consensus motif has been identified for the ssVSS sorting determinants, ctVSS signals do not possess a common motif (Table 1.1) at the level of their primary structure. This type of sorting determinant is found in a number of lectins, storage proteins, hydrolases and defense related proteins. To date ctVSSs from barley lectin (Dombrowski et al, 1993), tobacco chitinase (Neuhaus et al, 1994), kiwi actinidin (Paul et al, 1995), bean phaseolin (Frigerio et al, 1998) and tobacco osmotin (Sato et al, 1995) have been characterized. The deletion of this determinant from the protein results in the missorting to the apoplast, and conversely the addition of this signal to heterologous secreted proteins such as cucumber chitinase (Neuhaus et al, 1991), hen egg white lysozyme (Neuhaus et al, 1995), rat preputial β-glucuronidase (Neuhaus et al, 1995), Green Fluorescent Protein (Di Sansebastiano et al, 1998) and sporamin with a non-functional ssVSS (Matsuoka et al, 1995) resulted in the vacuolar localization of the respective chimeric proteins. It has been postulated that the ctVSS may be cleaved in the vacuole by an aspartic proteinase (Amidon et al, 1999). Extensive site directed mutation analyses have been performed on the tobacco chitinase ctVSS: 'GLLVDTM' (Neuhaus et al, 1994) and the BL ctVSS 'VFAEAIAANSTLVAE' (Dombrowski et al, 1993). The terminal methionine of the TC-ctVSS could be deleted without effect, while all other residues were required. Replacing ^{2}LLV^{4} with SSS or ^{4}VDTM^{7} with LLLL did not reduce

the sorting efficiency while replacing the terminal M with a G reduced vacuolar sorting by 50%. In the case of the BL-ctVSS as little as ^{1}VFA^{3} was shown to be functional in vacuolar targeting, while the addition of two G residues at the C-terminus of the motif abrogated vacuolar targeting. Although glycosylation is not necessary for vacuolar targeting in plants, translocating the glycosylation site from a -7 position to a -1 position in the ctVSS of BL resulted in secretion of the protein (for review Dombrowski and Raikhel, 1996). This is attributed to a steric hindrance rather than a role for the glycan. The lack of a consensus motif within the ctVSSs indicates that a structural feature of the ctVSS might be responsible for the selectivity. It has been suggested that the ctVSS of barley lectin adopts an amphipathic alpha helix structure (Dombrowski and Raikhel, 1996). However, such a structural feature has not been demonstrated to be necessary for ctVSS-mediated vacuolar sorting.

Currently, no receptor has been identified for the ctVSS pathway. It has been shown that if ctVSS proteins are expressed to high levels, then the proteins can be secreted (Neuhaus et al, 1994; Frigerio et al, 1998) indicating that the pathway is saturable. Frigerio et al (1998) suggest the existence of a specific machinery involved in sorting ctVSS bearing proteins onward to the vacuole. However, such proteins have not been identified.

Sorting machinery

While a large body of evidence has been accumulated regarding the transport of ssVSS-bearing proteins, very little is known about the pathway transporting ctVSS-bearing proteins. However, there are several differences between the two pathways.

Matsuoka et al (1995) demonstrated that both barley lectin and tobacco chitinase are secreted from tobacco protoplasts when treated with 33 μM of wortmannin, whereas sporamin is correctly targeted to the vacuole. Though wortmannin is an inhibitor of phosphatidyl inositol-3 kinase in mammalian cells, its specific role in the plant secretory pathway has not been elucidated. It is also known that ctVSS-bearing proteins have not been found in clathrin-coated vesicles that contain aleurain (an ssVSS-bearing protein). Finally, the final destinations for these proteins are different as the ssVSS-bearing proteins are delivered to the lytic vacuole (decorated with γ-TIP, an aquaporin), while ctVSS-bearing proteins are delivered to the PSV (decorated with α-TIP). These evidences demonstrate that the ctVSS and ssVSS-mediated pathways are different. The lack of a consensus motif in the ctVSS and the variation in the length of the ctVSS among the various known ctVSS sequences has hampered the understanding of this part of the secretory pathway.

In summary, the study of vacuolar protein transport has come a long way since observing spaces in the cell that were devoid of cytoplasmic contents. A number of vacuolar proteins have been identified and their targeting motifs dissected. The characterization of these sorting signals led to the identification of sorting receptors and other proteins required for this machinery. However, the lack of a consensus sequence has hampered the understanding of the machinery involved in sorting ctVSS-bearing soluble proteins to the vacuole.

"Statement of purpose"

The research presented aims at identifying proteins involved in ctVSS-mediated protein transport. So far it is known that a class of soluble proteins reach the vacuole by virtue of their ctVSSs. In view of the general features of the various secretory pathway models proposed, it is conceivable to expect a variety of proteins involved in the pathway. Some of these might be receptors recognizing ctVSS-bearing cargo and sorting them into vesicles bound for the vacuole. This pathway may also involve a specialized class of vesicles decorated with a complement of proteins. The fate of these vesicles may be regulated by certain factors.

To characterize the pathway, a genetic approach has been undertaken in *Arabdopsis thaliana* to identify factors involved in the pathway, described in Chapters 2 to 6. Through a mutant screen *cvs1* and *cvs2* were identified as ctVSS vacuolar sorting mutants. *cvs1* was found to be a single gene recessive mutation causing a leaky sorting defect in root cells. This mutation was mapped to a contig of bacterial artificial chromosomes between AtSO191 and DFR on chromosome V.

Chapter 2

Development of a reporter line for use in a screen to identify mutants impaired in ctVSS-sorting in *Arabidopsis*

INTRODUCTION

The study of the secretory pathway in yeast has advanced due to the large collection of *vps* mutants impaired in vacuolar protein transport (Raymond et al, 1992). More than 50 genes have been identified whose products play important roles in the pathway. The identification of the plant homologues of yeast and mammalian secretory pathway proteins has advanced the understanding of this mechanism in plants. This approach was initiated by complementing the yeast *pep12* mutant with an *Arabidopsis* cDNA library resulting in the identification of AtPEP12 (later renamed SYP21) (Bassham et al, 1995), the first plant syntaxin to be identified. Since then AtSYP41 and AtSYP42 (Bassham et al, 2000), AtVPS45 (Bassham et al, 2000), AtVTI1a and b (Zheng et al, 1999), have been identified and characterized. It is likely that these proteins are involved in the ssVSS pathway. On the other hand, wortmannin sensitivity studies (Matsuoka et al, 1995) and ctVSS mutation studies (Dombrowski et al, 1993, Neuhaus et al, 1994) have shown that the ctVSS sorting pathway may be different from that of the ssVSS-sorting pathway. The ctVSS mediated pathway appears to be unique to plants as Gal and Raikhel (1994) showed that the yeast vacuolar targeting machinery does not transport barley lectin to the yeast vacuole. Taken together the ctVSS-mediated pathway may be unique to plants and may involve proteins hitherto unknown in other systems.

To identify proteins involved in targeting ctVSS-bearing proteins to the vacuole, a plant mutant screen was envisaged. A number of candidate ctVSS-bearing proteins have been identified in *Arabidopsis* such as AtOsmotin (Capelli et al, 1997), myrosinase-binding protein (Takechi et al, 1999), vacuolar invertases such as AtFruct3 and AtFruct4 (Haouazine-Takvorian et al, 1997), and a vacuolar chitinase, AtChib (Samac et al, 1990).

However, the role of the putative ctVSS in these proteins has not been characterized. Barley lectin (BL), a root cap and developing embryo protein has been expressed in *Arabidopsis* and extensively characterized (Dombrowski, 1995; Ahmed et al, 2000). In *Arabidopsis thaliana* ecotype RLD, constitutively-expressed BL was localized in the vacuole by immunoelectron microscopy and the precursor protein was shown by pulse-chase analysis to be proteolytically cleaved to the mature size. It was also demonstrated that BL lacking a ctVSS was secreted and localized in the apoplast. Another reporter active in plant vacuoles is rat preputial β-glucuronidase (Rat-GUS, Nishimura et al, 1986) when fused to the ctVSS of tobacco chitinase (Neuhaus et al, 1995). Rat-GUS is a homotetrameric protein consisting of 75 kDa (648 amino acids) subunits, synthesized in rat cells as a precursor with a 15 amino acid C-terminal extension. This extension bears a glycosylation site for the phosphomannosyl glycan required for targeting to the lysosome via the mannose-6 phosphate receptor in animal cells (Johnson et al, 1990). The deletion of the C-terminal extension (Rat-GUS-Delta) resulted in the secretion of the protein in tobacco cell suspension protoplasts and a reduction in activity. The level of expression, however, was not affected. With the addition of the ctVSS from tobacco chitinase (Rat-GUS-ctVSS), the protein was efficiently directed to tobacco protoplast vacuoles (Neuhaus et al, 1995).

To identify mutants impaired in the ctVSS pathway in the absence of known endogenous reporter, a transgenic approach was used whereby mutants would be identified which mislocalized heterologous ctVSS-bearing proteins. As the mutant screen involved chemical mutagenesis (see Chapter 3 for details), two reporters were used, one as a primary reporter to screen mutants visually and the other as a secondary reporter to

verify the mutants by electron microscopy. Such an approach would ensure that *cis*-mutants (defective only in the targeting signal of one reporter) would be eliminated by the secondary screen. To screen a large collection of mutants, a visual screen was devised with Rat-GUS-ctVSS as the primary reporter as its enzymatic activity is easily detected *in vivo*. To perform the secondary screen, BL was chosen as a reporter. The localization of BL could be performed by immunoelectron microscopy. Since this screen required a stable line expressing both reporters, it was necessary to develop a line with a single insertion locus that was homozygous for the reporter construct. This would facilitate mapping and positional cloning at a later stage. The development of a transgenic homozygous single locus seed line for EMS mutagenesis is presented.

RESULTS AND DISCUSSION

While important results have been obtained describing protein sequences necessary for sorting soluble proteins to the vacuole, little is known of the machinery that carries out the sorting in plants. The goal of this study was to develop a line of *Arabidopsis* plants stably transformed with a vacuolar reporter construct for creating ctVSS-sorting mutants. Mutants with aberrant sorting pathways would then be expected to mis-direct their cargo to the apoplast via the default pathway of exocytosis. In order to screen for mutants defective in the transport machinery, a reporter construct was developed containing cDNAs encoding two reporter proteins: Rat-GUS-ctVSS and BL. Rat-GUS-ctVSS served as the primary reporter in a visual screen and BL as a secondary reporter to verify the *trans* nature of the mutation. Two reporters were used in the screen to eliminate false positives that would arise due to chemically mutagenized vacuolar targeting signals (*cis*-mutations). As a control, a

transgenic line was developed (Rat-GUS-Delta) which lacks a vacuolar targeting signal and would be secreted into the apoplast.

Rat-GUS-ctVSS/BL Reporter construct and Rat-GUS-Delta construct

A binary vector construct was created in pMOG800 (Mogen International, The Netherlands), with Rat-GUS-ctVSS driven by a double 35S promoter and BL driven by a second double 35S promoter resulting in pMOGCRB. *Agrobacterium tumefaciens* (GV3101 pMP90) cells were transformed with this construct and used to infect *Arabidopsis thaliana* ecotype Columbia plants by vacuum infiltration (Bent et al, 1994). In parallel, Rat-GUS-Delta (Neuhaus et al, 1995) driven by a 35S promoter was cloned into a pMOG800 vector and used to similarly transform *Arabidopsis thaliana* ecotype Columbia plants. This control line helped establish the screening assay.

Primary transformants and evaluations

Eleven *Arabidopsis thaliana* ecotype Columbia plants were independently infiltrated with a suspension of *Agrobacterium* cells carrying the pMOGCRB construct according to Bent et al (1994). About 3000 seeds per parent were screened on kanamycin media and an average of 15 plants per parent were found resistant to the antibiotic after 2 weeks of treatment. Since the insertion of the T-DNA is random, the activity of the transgene is known to vary from plant to plant depending on the site of insertion. Also, within a single independently transformed plant, several transformation events could occur giving rise to different levels of expression of the transgene. Leaves of kanamycin resistant seedlings were assayed for Rat-GUS activity. Briefly, leaves were removed from three

seedlings per transformed parent and wounded to allow the entry of the Rat-GUS substrate, 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-GlcU), into the cell. The activity was evaluated visually. The results were compared with wild type plants (Figure 2.1a, b, c and d). Although wild type *Arabidopsis* has been reported to contain β-glucuronidase activity (Wozniak and Owens, 1994), wild type *Arabidopsis* plants did not pick up any stain after 2 hours of staining at 37 ^{0}C. In contrast, pMOGCRB plants stained very strongly. Activity was detected in trichomes and leaf veins. This confirmed that Rat-GUS was active when expressed in plants as previously shown by Neuhaus et al (1995). Among independent transformants, seedlings were selected visually for high ß-glucuronidase activity. These plants were subsequently transferred to soil in growth chambers. It is likely that some of the lines have multiple insertions of the T-DNA and therefore exhibit higher levels of activity. At this stage however, it was difficult to evaluate the number of insertions based solely on a visual comparison of the level of activity.

The histochemical assay was complemented by a quantitative fluorometric assay (adapted from Jefferson, 1987). β-Glucuronidase activity could be detected only in pMOGCRB plants (Figure 2.2 and Table 2.1) using 10 μg total protein per assay. The reaction appeared to proceed in a linear manner with respect to time. To confirm the presence of the secondary reporter, protein extracts were analyzed for the presence of BL. BL was detected by western blotting using anti-WGA sera (Ahmed et al, 2000). At a dilution of 1:1000, the antibody cross-reacted with a single product of the expected size (Figure 2.3).

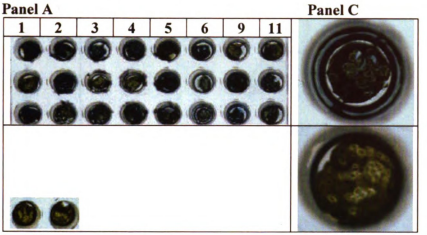

Figure 2.1 Verification of Rat-GUS activity in primary transformants

A single leaf was harvested from each of 3 kanamycin resistant seedlings per independent transformant and wounded with a sterile micropipet tip (to allow the dye to enter the cells) in a microtiter plate. The leaf was then incubated overnight in 100 ml of assay buffer (0.1 M sodium acetate pH 4.8, 1 mg/ml BSA, 2 mM X-GlcU, 0.5 mM potassium ferrocyanide and 0.5 mM potassium ferricyanide) at 37 ^{0}C (modified from Guerineau and Mullineaux, 1993). Panel A) Leaf samples from kanamycin resistant primary transformant seedlings of 8 independent Columbia plants transformed with pMOBCRB stained for Rat-GUS activity (blue). Panel B) Leaf samples from wild type Columbia seedlings stained for Rat-GUS activity. Panel C) Close up of one microtiter well with a pMOGCRB primary transformant leaf stained for Rat-GUS activity (blue). Panel D) Close up of one microtiter well with a wild type leaf stained for Rat-GUS activity. The comparison of panels C and D indicate the strong blue stain in pMOGCRB transformants unlike wild type leaf samples. Images in this dissertation are presented in color.

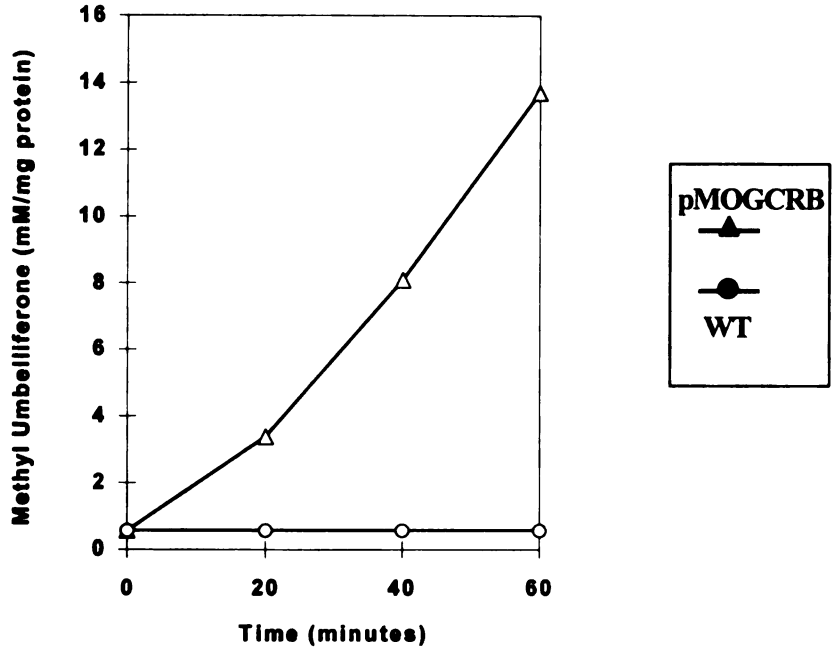

Figure 2.2 β-glucuronidase activity in protein extract from primary transformant
Specific Activity of 10 μg of leaf protein from a) pMOGCRB plants (blue) and b) WT plants (green) measured at 20-minute intervals by estimating the concentration of methyl umbelliferone released from methyl umbelliferyl β-D glucuronide. This assay demonstrated that pMOGCRB transformants expressed an active Rat-GUS as seen by the linear activity of a one-hour time course. This activity was absent in wild type tissue samples.

Tab

Cor
sar
Bu
Ar
exp
act

Figu
BL
whe
Col
prot
Thi
tiss

Table 2.1 β-glucuronidase activity in protein extract from primary transformant

Time (minutes)	Buffer	pMOGCRB	Wild Type
0	0.56	0.56	0.56
20	0.56	3.3	0.56
40	0.56	8.06	0.56
60	0.56	13.68	0.56

Concentration of Methyl Umbelliferone (MU) released by β-glucuronidase activity in samples measured in mM of MU released per mg total protein per minute. Samples include Buffer (Extraction buffer), pMOGCRB transformants and Wild Type (Wild type *Arabidopsis* Ecotype Columbia). This assay demonstrated that pMOGCRB transformants expressed an active Rat-GUS as seen by the linear activity of a one-hour time course. This activity was absent in wild type tissue samples.

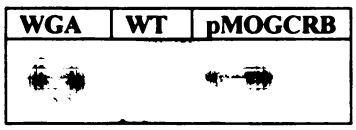

Figure 2.3 Western blot to verify the expression of the barley lectin (BL)
BL was detected using rabbit anti-wheat germ agglutinin (WGA), Lane WGA (100 ng wheat germ agglutinin) was used as a size marker, Lane WT (60 μg total wild type Columbia leaf protein), pMOGCRB (60 μg total pMOGCRB primary transformant leaf protein). Proteins were detected using rabbit anti-WGA (1:1000) as a primary antibody. This analysis demonstrated that pMOGCRB transformants expressed BL unlike wild type tissues.

Sec

tra

so

for

do

be

tra

sec

Tab

Lin

lev

the

ap

wi

ob

ar

th

Secondary transformants

Lines with a single reporter construct are expected to be hemizygous primary transformants and their progeny after self-fertilization (secondary transformants) segregate so that ¼ (25 %) of the progeny are homozygous for the reporter, ½ (50 %) are hemizygous for the reporter and ¼ (25 %) lack the reporter. Since resistance to kanamycin selection is a dominant trait, ¾ (75 %) of the secondary transformants with a single reporter locus should be resistant to kanamycin. To identify single insertion secondary transformants, secondary transformants were germinated on kanamycin media and the ratio of kanamycin resistant seedlings to the total number of seedlings was determined (see Table 2.2). To identify lines

Table 2.2 Frequency of T_2 progeny surviving kanamycin selection

Line	Resistant	Susceptible	Level of resistance (%)	χ^2	*p*-value
CRB T_2 –1	49	1	98	14.1	1×10^{-4}
CRB T_2 –2	60	0	100	20	7.7×10^{-6}
CRB T_2 –3	73	0	100	24.3	8.1×10^{-7}
CRB T_2 –5	86	0	100	28.6	8.5×10^{-8}
CRB T_2 –6	65	20	76.5*	0.098	0.75
CRB T_2 –9	64	12	84*	3.43	0.063
CRB T_2 -11	56	3	94.9	13.44	2×10^{-4}

Lines of transgenic lines and level of resistance to kanamycin are displayed. * At a 0.05 level of significance, the frequency fits a 3:1 ratio indicating a single gene insertion into the genome (using the chi-square test of goodness of fit). Lines CRB T_2-6 and CRB T_2-9 appeared to have a single insertion locus.

with a single reporter locus, a chi-square test of goodness of fit was performed on the ratios obtained against an expected ratio of 75 % (3:1) and the *p*-value was estimated for each line. These were tested against a critical *p*-value of 0.05. Lines, that had a *p*-value smaller than 0.05 were discarded, as their resistance ratio did not fit the required ratio of 75 %.

Lines $CRBT_2$-6 (p-value = 0.75) and $CRBT_2$-9 (p-value = 0.063) were selected, as their p-values were greater than 0.05 indicating a low deviation from the expected ratio of 75%. Data from the progeny of lines $CRBT_2$-4, 7, 8 and 10 were not consistent with an expected 3:1 ratio and these lines were not pursued.

Tertiary transformants

Once a single locus line is identified, it is necessary to identify homozygous progeny of this line. This was accomplished by scoring the progeny of twenty individual T_2 plants from each progeny pool of lines $CRBT_26$ and $CRBT_29$. Seeds from these lines (T_3) were germinated on kanamycin media and the ratio of kanamycin resistant seedling to the total number of seedlings was determined (see Tables 2.3 and 2.4). Ratios thus obtained were categorized in classes of 0 % (reporter-less segregants), 75 % (progeny of heterozygous secondary transformants) and 100 % (progeny of homozygous secondary transformants) by performing a chi-square goodness of fit test. As none of the T_3 pools of line $CRBT_2$-9 were completely susceptible to kanamycin, these were discarded. Among the T_3 pools of line $CRBT_2$-6, lines $CRBT_36$-18 and $CRBT_36$-24 were identified to be homozygous and were analyzed further.

Leaves were harvested from 10 individual T_3 plants from these lines and total protein extracts prepared. These extracts were analyzed for BL by western analysis and for Rat-GUS by fluorometric activity assays (Table 2.5). Significant activity and protein levels could not be detected from the progeny of line $CRBT_26.24$ and these lines were

Table 2.3 Reporter segregation in 20 T_3 progeny pools of line $CRBT_3$-6

$CRBT_3$-6	Total	% Resistance	Category
16	45	71	75*
17	119	92	100
18	86	99	100**
19	120	97	100
20	96	nd	nd
21	107	82	75*
22	127	98	100
23	129	89	100
24	90	100	100**
25	72	86	75*
26	72	100	100
27	126	97	100
28	66	100	100
29	52	92	100
30	84	76	75*
31	47	91	100
32	48	77	75*
33	36	0	0
34	101	92	100
35	32	7	0

The segregation of the reporter construct in the progeny of line $CRBT_3$-6 was evaluated by estimating the resistance to kanamycin in 20 progeny pools of this line. Lines of T_3 ($CRBT_3$-6), the total number of seedlings sampled, the percentage that survived the selection and the category they were classified in. * 3:1 segregation of reporter construct ** Lines selected for further studies. Lines $CRBT_3$-6.18 and $CRBT_3$-6.24 were selected for further studies.

Ta

Tha
by
(CR
sel
Th
tex

Table 2.4 Reporter segregation in 20 T_3 progeny pools of line $CRBT_3$-9

$CRBT_3$-9	Total	% Resistance	Category
1	69	96	100
2	48	73	75*
3	44	82	75*
4	104	98	100
5	196	99	100
6	69	78	75*
7	219	97	100
8	91	100	100
9	62	73	75*
10	55	96	100
11	73	95	100
12	44	77	75*
13	73	97	100
14	53	91	100
15	69	93	100
16	98	100	100
17	76	96	100
18	37	89	75*
19	47	100	100
20	109	89	100

The segregation of the reporter construct in the progeny of line $CRBT_3$-9 was evaluated by estimating the resistance to kanamycin in 20 progeny pools of this line. Lines of T_3 ($CRBT_3$-9), the total number of seedlings sampled, the percentage that survived the selection and the category they were classified in. * 3:1 segregation of reporter construct. This line was discarded as kanamycin susceptible progeny could not be detected (see text).

Table 2.5 Evaluation of reporter levels in tertiary transformants

Sample	Activity (mM MU /mg protein/min)	BL
Ext Buffer	0.00	-
Wt	0.31	-
$CRBT_36$-18-1	5.19	-
$CRBT_36$-18-2	0.92	+
$CRBT_36$-18-3	2.14	-
$CRBT_36$-18-4	8.25	+
$CRBT_36$-18-5	3.36	+
$CRBT_36$-18-6	2.75	+
$CRBT_36$-18-7	3.36	-
$CRBT_36$-18-8	1.83	-
$CRBT_36$-18-9	1.83	-
$CRBT_36$-18-10	0.92	-/+
$CRBT_36$-24-1	0.00	-
$CRBT_36$-24-2	0.00	-
$CRBT_36$-24-3	0.00	-
$CRBT_36$-24-4	0.00	-
$CRBT_36$-24-5	0.00	-
$CRBT_36$-24-6	0.00	-
$CRBT_36$-24-7	0.31	-
$CRBT_36$-24-8	0.00	-
$CRBT_36$-24-9	0.00	-
$CRBT_36$-24-10	0.00	-

The levels of reporters in tertiary transformants were evaluated by measuring Rat-GUS activity and BL by western analysis. β-Glucuronidase activity (mM methyl umbelliferone/mg protein/min). Samples include Buffer (Extraction buffer), leaf protein from each of ten plants selected from $CRBT_26$-18 and $CRBT_26$-24 (transgenic for a single insertion locus of pMOGCRB) and Wild Type (Wild type *Arabidopsis* ecotype Columbia). BL was detected by western analysis. + = High levels of BL, -/+ = moderate levels of protein and - = no protein. Lines $CRBT_36$-18-1, $CRBT_36$-18-4 and $CRBT_36$-18-5 were selected for further analysis.

discarded. Lines $CRBT_3$6.18.01, 6.18.04 and 6.18.05 (see Table 2.5) appeared to have significant GUS activities and BL levels. Seeds from three plants ($CRBT_3$6.18.01, 6.18.04 and 6.18.05) were harvested for further analysis.

Quaternary transformants

To confirm that homozygous tertiary transformants yielded stable progeny, the progeny of T_3 lines ($CRBT_3$6.18.01, 6.18.04 and 6.18.05) were analyzed by plating on kanamycin media and scoring for survival (Table 2.6). BL analyses by western blotting

Table 2.6 Kanamycin resistance in quaternary transformants

Genotype	Total # of Seeds	% Kanamycin Resistance
$CRBT_4$6.18.1	81	96.3
$CRBT_4$6.18.4	79	88.6
$CRBT_4$6.18.5	80	100*

Progeny of lines $CRBT_4$6.18.1, $CRBT_4$6.18.4 and $CRBT_4$6.18.5 were germinated on kanamycin media and the survival ratio was determined. Genotype = Lines of T_4 (progeny of $CRBT_4$6-18). Total # of Seeds = total number of seedlings sampled. % Kanamycin Resistance = the percentage of seedlings that survived the selection. * Complete resistance to kanamycin (indicating homozygosity). Line $CRBT_4$6.18.5 was found to have a 100 % kanamycin resistance indicating a stable homozygous reporter and was thus chosen for further studies.

(Figure 2.4) and β-glucuronidase activity assays were performed on leaf tissue (summarized in Table 2.7). Line $CRBT_4$6.18.05 was found to have a 100 % survival on kanamycin and detectable amounts of BL protein and Rat-GUS activity. This line was bulked for EMS mutagenesis (yielding line $CRBT_5$6.18.05$_X$). Qualitatively, it was observed that the levels of both reporters were lower in $CRBT_5$6.18.05$_X$ than in primary

Fig
To
imm
rab
ph
an
6.

Ta

Ev
Bl
me
bu
CR
we
CR

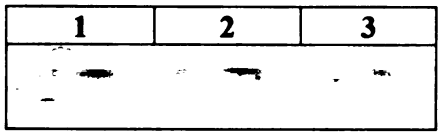

Figure 2.4 Verification of BL in $CRBT_4$ seedlings by western analysis
Total leaf protein from 3 $CRBT_4$ seedlings (40 µg) was resolved by SDS-PAGE and immunoblotted. Resolved proteins were visualized by a primary antibody treatment with rabbit anti-WGA sera, followed by treatment with goat anti-rabbit coupled to alkaline phosphatase. Blots were developed according to Harlow and Lane, 1988 (see Materials and Methods). Lane 1 = $CRBT_4$-6.18-1, Lane 2 = $CRBT_4$-6.18-4 and Lane 3 = $CRBT_4$-6.18-5. All 3 lines tested expressed BL.

Table 2.7 Reporter level estimation in quaternary transformants

Genotype	Activity mM MU/mg-min	BL
Ext. Buffer	0.00	-
WT	0.05	-
$CRBT_4$6.18.1	1.73	+
$CRBT_4$6.18.4	1.36	+
$CRBT_4$6.18.5	2.23	+

Evaluation of reporter levels in tertiary transformants by measuring Rat-GUS activity and BL by western analysis (see Figure 2.4 for western analysis). β-Glucuronidase activity is measured in mM methyl umbelliferone/mg protein/min. Samples include Buffer (Extraction buffer), leaf protein from quaternary transformant lines $CRBT_4$6.18-1, $CRBT_4$6.18-4 and $CRBT_4$6.18-5 and WT (Wild type *Arabidopsis* ecotype Columbia). BL was detected by western analysis (see Figure 2.4). + = High levels of BL and - = no protein. Line $CRBT_4$6.18.5 performed well with respect to BL and Rat-GUS.

transformants. In order to determine if this line was suitable for a mutant screen a "screen test" was performed.

Screen test

To screen for mutants missorting Rat-GUS and BL to the apoplast (see Chapter 3 for details), it is necessary to detect Rat-GUS activity specifically in the apoplast. To achieve this, an assay has been designed in which the substrate for Rat-GUS [5-bromo-4-chloro-3-indolyl â-D-glucuronide (X-GlcU)] is localized in the apoplast. X-GlcU is a stain that requires 2-5 % detergent to penetrate the plasma membrane. By applying X-GlcU to seedlings without any detergent, the dye diffuses into the apoplast but does not cross the plasma membrane (Trull and Deikman, 1998). However, upon the application of X-GlcU in the presence of 2 % (v/v) Triton X-100, the cell membranes are permeabilized and the substrate diffuses into the cell.

After obtaining seeds homozygous for a single reporter locus expressing BL and Rat-GUS (line $CRBT_5 6.18.05_X$), a control screen was performed to test the line for suitability in a mutant screen. As a control seeds expressing a secreted Rat-GUS (Rat-GUS-Delta, Rat-GUS lacking a vacuolar targeting signal) were used. In this assay it was expected that Rat-GUS-ctVSS would be localized in the vacuole in the absence of detergent. Seeds of wild type *Arabidopsis*, $CRBT_5 6.18.05_x$ and Rat-GUS-Delta were germinated on media plates vertically and stained with X-GlcU. While roots of Rat-GUS-Delta turned blue (indicating Rat-GUS activity in the apoplast) within 2 hours at room temperature, roots of wild type seedlings and $CRBT_5 6.18.05_x$ seedlings did not pick up any stain. To verify that the $CRBT_5 6.18.05_x$ seedlings were white due to a vacuole-localized Rat-GUS, seedlings of

wild type *Arabidopsis*, $CRBT_5 6.18.05_x$ and Rat-GUS-Delta were germinated on media plates vertically and stained with X-GlcU in the presence of 2 % (v/v) Triton X-100. It was observed that only the wild type seedlings remained white while both Rat-GUS-Delta roots and $CRBT_5 6.18.05_x$ root turned blue. This confirmed that $CRBT_5 6.18.05_x$ seedlings were expressing Rat-GUS but it was localized intracellularly, while Rat-GUS-Delta was secreted and hydrolyzed X-GlcU resulting in blue roots. Upon the application of Triton X-100, the cell membrane was permeabilized and the dye was accessible to the enzyme. Thus both Rat-GUS-Delta seedlings and $CRBT_5 6.18.05_x$ seedlings turned blue.

Since the secondary screen was based on the localization of BL, the expression of BL protein in $CRBT_5 6.18.05_x$ seedlings was investigated. Ten day-old $CRBT_5 6.18.05_x$ seedlings were analyzed for the presence of BL by germinating parental seeds on vertical media plates for 10 days and harvesting the tissues in 3 categories: root tips, lateral root zone and cotyledons. Total protein was extracted and analyzed by western blotting (Figure 2.5). BL was detected in the lateral root and in the root tip but not in cotyledonary leaves.

CONCLUSION

Using an *Agrobacterium*-mediated transformation approach pMOGCRB, a reporter construct encoding Rat-GUS-ctVSS and BL was introduced into *Arabidopsis thaliana* ecotype Columbia. Following each generation by kanamycin selection and measuring activities of Rat-GUS and protein levels of BL, a line ($CRBT_5 6.18.05_X$)

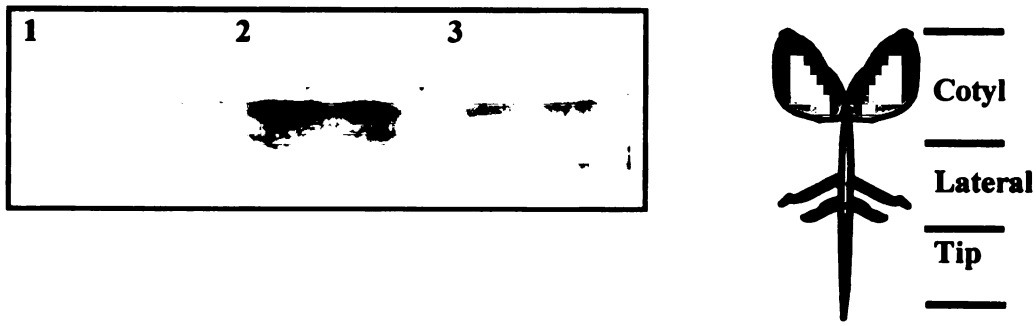

Figure 2.5 Western analysis of protein extracts from different seedling tissues
Total protein from cotyledonary leaves, lateral roots and root tips (~40 μg) was resolved by SDS-PAGE and immunoblotted. Resolved proteins were visualized by a primary antibody treatment with rabbit anti WGA, followed by treatment with goat anti-rabbit coupled to alkaline phosphatase. Blots were developed according to Harlow and Lane, 1988 (see Materials and Methods). Lane 1 = Cotyledonary leaf protein ("Cotyl" in diagram), Lane 2 = Lateral root protein ("Lateral" in diagram) and Lane 3 = Root tip protein ("Tip" in diagram). While BL was not detected in cotyledonary tissues, it was more abundant in lateral roots rather than root tips.

homozygous for the reporter construct was developed. Although both reporters could be detected in line $CRBT_5 6.18.05_X$, the activity levels were not very high possibly due to cossuppression of the reporters. By performing a screen test for Rat-GUS secretion and root tip protein analysis for BL, the line was found to be suitable for EMS mutagenesis.

MATERIALS AND METHODS

Development of the reporter construct pMOGCRB

All chemicals, unless otherwise indicated, were purchased from Sigma Chemical Co (St. Louis, MO, USA). The KpnI site of pMOG843 (Mogen International, The Netherlands) was removed by blunt-ending and religation. The plasmid was then cleaved with XbaI. The cauliflower mosaic virus 35S promoter (35S)/potato inhibitor terminator (PT) fragment was

introduced into the XbaI site of pBluescript (modified by eliminating the HindIII site). Two clones were isolated (in both orientations) designated 10pBS843 (orientation: SacII, 35S, PT and KpnI) and 14 pBS843 (orientation: KpnI, 35S, PT and SacII). The coding region of BL was excised from pTAW50 (Matsuoka et al, 1995) with EcoRI and introduced into the HindIII site of 10pBS843 by blunt end cloning, yielding p10BL4. The coding region of Rat-GUS-ctVSS was excised from pRatGUSTail (Dr J-M. Neuhaus, Universite de Neuchatel, Switzerland) with BamHI and PstI and inserted into the BglII-PstI sites of pIC19H (Marsh et al, 1984) yielding p19HRGT. The coding region of Rat-GUS-ctVSS was excised from p19HRGT with HindIII and inserted into the HindIII site of 14pBS843 yielding p14RGT2. The coding region of p10BL4 was excised as an XbaI fragment and inserted into the SpeI site of p14RGT2 yielding pBLRT9. A 2.1 kbp SacII fragment from pMOG800 (Mogen International, The Netherlands) containing the multiple cloning site of pMOG800 was inserted into the SacII site of pBLRT9, yielding pSM10. The BL-Rat-GUS-ctVSS coding region of pSM10 was excised as an XhoI fragment and inserted into the XhoI site of pMOG800, yielding pMOGCRB (Figure 2.6).

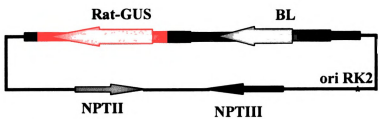

Figure 2.6 Diagram of plasmid pMOGCRB
Rat-GUS = Rat-GUS-ctVSS (driven by a double 35S promoter and terminated with a potato protease inhibitor terminator). BL = Barley lectin (driven by a double 35S promoter and terminated with a potato protease inhibitor terminator). Ori RK2 = origin of replication. NPTII = Neomycin phosphotransferase for kanamycin selection in plants. NPTIII = Neomycin phosphotransferase for kanamycin selection in bacteria.

De

Sw

De

site

2.

Fig

Ra

ter

ph

ph

Pla

tra

tra

tum

in

ger

Development of the control construct: pMOGRGD

Clone pJMNRGDelta was obtained from Dr J-M Neuhaus (Universite de Neuchatel, Switzerland) and digested with EcoR1. A 2722 bp EcoRI fragment consisting of Rat-GUS-Delta driven by a 35S promoter and with a 35S terminator was subcloned into the EcoRI site of pMOG800 (Mogen International, The Netherlands) yielding pMOGRGD (Figure 2.7).

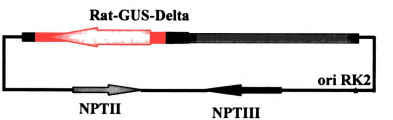

Figure 2.7 Diagram of plasmid pMOGRGD
Rat-GUS-Delta = Rat-GUS-Delta (secreted version, driven by a 35S promoter and terminated with a 35S terminator). Ori RK2 = origin of replication. NPTII = Neomycin phosphotransferase for kanamycin selection in plants. NPTIII = Neomycin phosphotransferase for kanamycin selection in bacteria.

Plants, Bacteria

All DNA constructions were carried out in *Escherichia coli* strain DH5á. All plant transformations were carried out in *Arabidopsis thaliana*, ecotype Columbia. All plant transformations were carried out by vacuum infiltrating plants with *Agrobacterium tumefaciens* GV3101 pMP90 electroporated with pMOGCRB or pMOGRGD. Vacuum infiltration was carried according to Bent et al, (1994). Seeds (T1) were harvested and germinated on media containing 40 mg/l kanamycin sulfate and 250 mg/l vancomycin.

Kanamycin resistant seedlings were transferred to 'hard selection media' containing 1.5 % (w/v) phytoagar for 1 week and then transplanted into soil.

Leaf activity assay

A single leaf was harvested from kanamycin resistant seedlings and wounded with a sterile micropipet tip in a microtiter plate. The leaf was then incubated overnight in 100 µl of assay buffer [0.1 M sodium acetate pH 4.8, 1 mg/ml bovine serum albumin (BSA), 2 mM 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-GlcU), 0.5 mM potassium ferrocyanide and 0.5 mM potassium ferricyanide] at 37 ^{0}C (modified from Guerineau and Mullineaux, 1993). The level of expression of Rat-GUS was evaluated visually.

Quantitative β-glucuronidase Assay

Tissue (0.2 g) was harvested from seedlings in soil and homogenized with liquid nitrogen, 0.5 g purified sand (J.T. Baker Chemical Co. Phillipsburg, NJ, USA), 100 mg polyvinyl polypyrrolidone using a mortar and a pestle. The powder was resuspended in 2.5 ml Rat-GUS extraction buffer [phosphate-buffered saline (PBS), 0.1% (v/v) Triton X-100, 0.1 % (w/v) sodium-N-lauryl sarcosine, 1 mM phenylmethylsulfonyl fluoride (PMSF)], and then centrifuged twice at 4 ^{0}C for 10 minutes at 15000 g. Protein contents were estimated by Bradford's assay (Bio-Rad Laboratories, Hercules, CA) using the Bradford analysis package, bradfordhsa.htm (available at http://www.msu.edu/~venkata1/bradfordhsa.htm, see Appendix B). Protein extracts were diluted to a concentration of 0.4 µg/µl with 0.1 M sodium acetate pH 4.8 containing 1 mg/ml BSA. The reaction was initiated by adding assay buffer (0.1 M sodium acetate pH 4.8, 1 mg/ml BSA, 2 mM methyl umbelliferone ß-D

glucuronide) to 25 µl of diluted protein extracts and incubated at 37 ^{0}C. Every 20 minutes 10 µl was sampled and the reaction was stopped with 1.5 ml 0.25 M sodium carbonate. Fluorescence was measured by exciting at 365 nm and measuring the emission at 455 nm (adapted from Jefferson, 1987).

"Roots, shoots and tips" Assay

Seeds of $CRBT_5$-6.18.5$_x$ were sterilized and germinated vertically on media plates [Gamborg's-B5 basal medium with minimal organics (Sigma) 0.32 % (w/v), 2-(N-morpholino) ethanesulfonic acid (MES) 0.05% (w/v) and sucrose 2% (w/v), pH 5.7] for 10 days. This was achieved by setting the plates upright to allow the roots to germinate on the surface of the medium. Tissue was harvested from seedlings in three categories: root tips, lateral roots and cotyledonary leaves and homogenized with liquid nitrogen, using a Kontes motorized pestle. The powder was resuspended in protein extraction buffer [PBS, 0.1% (v/v) Triton X-100, 0.1 % (w/v) sodium-N-lauryl sarcosine, 1 mM PMSF], and then centrifuged twice at 4 ^{0}C for 10 minutes at 15000 g. Protein contents were estimated by Bradford's assay (Bio-Rad Laboratories, Hercules, CA) using the Bradford analysis package, bradfordhsa.htm (see Appendix B).

SDS-PAGE and Western blotting

Protein extracts were diluted with half a volume of 3X Laemmli buffer (Harlow and Lane, 1988). The diluted samples (25 µl) were boiled, treated with 5 µl acetamide solution (5 µl 1X Laemmli buffer, 2 mg Tris base, 1.33 mg iodoacetamide) and the mixture was incubated at 37 ^{0}C for 30 minutes prior to SDS-PAGE. Proteins were immunoblotted onto

nitrocellulose filters, transfer-verified by staining with Ponceau S, blocked and incubated with primary antibodies (rabbit anti-WGA, Ahmed et al, 2000) at 1:1000 dilution and with secondary antibodies (goat anti-rabbit coupled to alkaline phosphatase) at 1:1000 dilution according to Harlow and Lane (1988). Protein bands were visualized using 50 ml of alkaline phosphatase buffer (100 mM sodium chloride, 5 mM magnesium chloride, 100 mM Tris, pH 9.5) containing 0.16 mg/ml 5-bromo-4-chloro-3-indolyl phosphate and 0.35 mg/ml nitro blue tetrazolium chloride. After the development of bands, the blots were rinsed with distilled water and photographed.

Visual test Screen

Seeds were sterilized and germinated vertically on media plates [Gamborg's-B5 basal medium with minimal organics (Sigma Chemical Company, St Louis MO, USA) 0.32 % (w/v), MES 0.05% (w/v) and sucrose 2% (w/v), pH 5.7) for one week. This was achieved by setting the plates upright to allow the roots to germinate on the surface of the medium. Staining was carried out with X-GlcU solution [X-GlcU (0.1 mg/ml of final suspension), dissolved in DMSO (0.1 % of the final suspension) and the solution dispersed in 0.05 % sterile agarose]. Seedlings were overlaid with 10 ml of X-GlcU solution and scored after 5 days. Detergent test were carried as above except that staining was carried out with X-GlcU (0.1 mg/ml of final suspension), dissolved in DMSO [0.1 % (v/v) of the final suspension] and the solution dispersed in 0.05 % (v/v) sterile agarose containing 2 % (v/v) Triton X-100.

Chapter 3

EMS mutagenesis and characterization of ctVSS-missorting plants

INTRODUCTION

The current understanding of the various proteins involved in the secretory pathway has been largely aided by the development of yeast mutants, aberrant for a variety of morphological characters or in the transport of cargo proteins such as carboxypeptidase Y (for review Klionsky et al, 1990). Mutants can be developed in plants using several approaches such as insertion mutagenesis, physical mutagenesis and chemical mutagenesis. Insertion mutagens include transposons (reviewed by Coupland, 1992) and T-DNA (reviewed by Koncz et al, 1992). Both of these mutagens are inserted into the genome at random and this aspect as been used to generate large collections of insertion mutants. This technique makes cloning the locus of interest relatively easy as the sequence of the insertion is known and several molecular techniques are available to clone and sequence regions flanking the insertion. Physical mutagens such as fast neutron bombardment, X-ray and γ-ray irradiation cause deletions in the genome. Both insertion mutagenesis and physical mutagenesis result in null mutants as the gene of interest is either interrupted by an insertion or deleted by a physical mutagen. Chemical mutagenesis using ethyl methane sulphonate (EMS) has been a common approach in *Arabidopsis thaliana* due to the generation of a large number of point mutations in a single step (Redei and Koncz, 1992). Point mutations provide a convenient approach in studying essential events as they do not necessarily result in the loss of the gene function but often result in alterations in the gene product. Many of these mutations are prominent enough to result in a measurable phenotype but not deleterious enough to be lethal to the plant. A mutagenesis scheme resulting in null mutants would thus not be appropriate to study protein transport to the vacuole, as the vacuole is an essential organelle and a complete

loss of a particular gene product may cause lethality. Point mutants however would be more appropriate to study vacuolar protein sorting, as the mutants could be leaky. In this chapter the identification of mutants aberrant in vacuolar protein transport is presented.

RESULTS AND DISCUSSION

Seeds of line $CRBT_5 6.18.05_X$ (see Chapter 2), homozygous for BL and Rat-GUS-ctVSS were mutagenized with 0.3 % EMS and germinated in soil (according to Haughn and Somerville, 1986). They were allowed to self-fertilize and set seeds (M_2 seeds). Each pool thus represented a large family of mutants (M_2).

The screen consisted of identifying seedlings whose roots acquired a blue color due to the hydrolysis of 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-GlcU) by mislocalized Rat-GUS in the apoplast. To screen mutants, M_2 seeds were germinated vertically on media for one week and then treated with X-GlcU (without any detergent). Within 48 hours, roots of seedlings could be seen to stain clearly (Figure 3.1). As a standard control, wild type Columbia seeds, Rat-GUS-Delta seeds (a control construct lacking the vacuolar targeting signal, see Chapter 2) and $CRBT_5 6.18.05_X$ seeds were treated the same way and analyzed. Trull and Deikman (1998) have demonstrated that secreted acid phosphatase can be detected by staining seedlings with 5-bromo-4-chloro-3-indolyl phosphate (BCIP), indicating that BCIP is localized in the apoplast and not membrane permeable. This screen was based on similar properties of X-GlcU. X-GlcU is a stain that requires 2-5 % detergent to penetrate the plasma membrane.

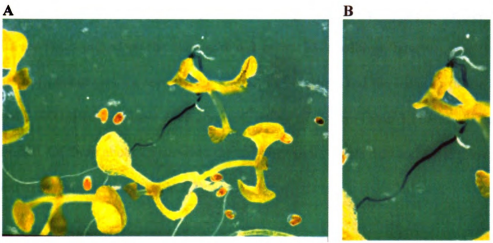

Figure 3.1 Mutant Screen on vertical plates
Seedlings were germinated on media plates (see Materials and methods) for five days. To obtain roots developing on the surface of the medium, the plates were set upright. To visualize the secreted Rat-GUS activity, seedlings were stained with 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-GlcU) in 0.1 % agarose. Subsequently plates were maintained horizontally for the stain solution to spread on the plate. Seedlings that turned blue due to secreted Rat-GUS activity were subsequently transferred to media plates lacking X-GlcU for one week and then transferred to soil. Panel A: one blue-root seedling among several white seedlings. Panel B: close up of blue seedling.

In all cases, wild type and parental seedlings did not stain blue, while Rat-GUS-Delta seedlings picked up the stain within 2 hours. M_2 seedlings, however, displayed a variety of phenotypes, as seen by the intensity of staining and the location of the stain. About 25000 seedlings have been screened out of 47 pools (Table 3.1) and 86 mutants identified. Of these about 25 survived the transfer to soil (Table 3.2) and 14 of these were selfed to obtain M_3 seeds (Table 3.3).

Analysis of M_3 seeds

Often a number of candidate mutants are identified in a preliminary screen of M_2 seeds. Some of these may not be determined by a genetic defect. This can be verified by evaluating the M_3 generation. In order to determine if the phenotype was heritable, the M_3 seeds were plated and stained (Table 3.3). It was observed that certain M_3 lines such as *e5a, e10a, e10b, e11c, e12-1, e13-1, e19-1, e4-1* and *e21a* secreted Rat-GUS activity as seen by the visual assay. In contrast, M_3 lines *e6a* and *e19-10* did not secrete Rat-GUS activity. This indicated that the defects in lines *e5a, e10a, e10b, e11c, e12-1, e13-1, e19-1, e4-1* and *e21a* were heritable and passed on to the next generation while lines *e6a* and *e19-10* did not have a heritable defect.

To investigate whether the secretion of Rat-GUS was due to a *cis*-mutation or a *trans*-mutation, the localization of BL in lines found to secrete Rat-GUS in a heritable manner, was investigated by immunoelectron microscopy. Blue seedlings were sectioned and immunoelectron microscopy was carried out using antisera specific for BL. All

Table 3.1 Pool-wise summary of number of mutants identified

Pool-wise summary	# of mutants	# screened
1	1	600
2	2	600
3	1	550
4	1	600
5	2	400
6	5	200
7	3	350
8	0	500
9	6	950
10	6	300
11	7	400
12	6	600
13	1	250
14	1	400
15	0	300
16	1	200
17	2	300
18	1	450
19	1	200
20	0	200
21	1	650
25	0	150
Total	86	24400
Controls		
$CRBT_5$-6.18-5_x	1	800
Rat-GUS-Delta	500	500
Wild Type	0	900

Pool-wise summary	# of mutants	# screened
26	3	400
28	1	200
29	0	250
30	5	950
31	0	300
32	0	600
33	0	800
34	0	1500
35	7	900
36	0	1350
37	6	750
38	3	700
39	2	400
40	3	1000
41	4	1100
42	0	950
43	3	350
44	0	900
45	0	500
46	1	400
47	0	950

To screen the mutagenized $CRBT_5$-6.18-5_x M_2 seeds, samples from each pool were germinated and set upright for roots to develop on the surface of the medium. Seedlings were stained for secreted Rat-GUS activity. This table shows a pool-wise summary of the number of mutants obtained and the number of seedlings screened. Control assays were performed with $CRBT_5$-6.18-5_x, Rat-GUS-Delta lines and Wild Type Columbia lines. A total of 24400 M_2 seedlings were scored out of 47 pools of M_2 families.

Table 3.2 List of mutants identified and the pools they were identified from

	Mutants	Pool
	e01-1	1
	e04-1	4
	e05a	5
	e06a	6
	e09a	9
	e10-1	10
	e10-3	10
	e10a	10
	e10b	10
	e10c	10
	e11a	11
	e11b	11
	e11c	11
	e12-1	12
	e12-2	12
	e13-1	13
	e14a	14
	e19-01	19
	e19-02	19
	e19-10	19
	e21a	21
	e37a	37
	e40a	40
	e40b	40
	e41a	41
Total	**25**	**15**

To screen the mutagenized $CRBT_5$-6.18-5_x M_2 seeds, samples from each pool were germinated and set upright for roots to develop on the surface of the medium. Seedlings were stained for secreted Rat-GUS activity. This lists 25 mutants belonging to 15 families (M_2 pools) that survived the transfer to soil.

Table 3.3 Evaluation of M_3 lines by Rat-GUS secretion assay and EM

M_3 Line	Blue: White	Secretion of BL by electron microscopy (EM)
e5a	41:0	Yes
e9a	ND	Yes
e10a	89:0	Yes
e10b	104:1	Yes
*E11c**	79:0	Yes
e12-1	37:15	Yes
*E13-1***	48:0	Yes
e19-1	3:2	No
e4-1	51:0	EM data not available
e6a	0:63	EM data not available
e10-3	ND	EM data not available
e10c	ND	EM data not available
E19-10	0:56	EM data not available
e21a	3:0	EM data not available

To determine the heritability of the secretion phenotype observed in the M2 seedlings, their progeny were scored for secretion of Rat-GUS by performing the root-staining assay. To investigate if the secretion of Rat-GUS was due to a *cis*-mutation, root and leaf sections were prepared from some of the secreting mutants. The localization of barley lectin (BL, the secondary reporter) was determined by immunoelectron microscopy (see Materials and Methods). This table summarizes the M_3 lines verified by listing the ratio of blue root-seedlings to white root seedlings and whether mis-localization (secretion) of BL in the apoplast was detected by electron microscopy. In certain cases exact counts could not be obtained (ND). In certain case electron microscopy was not performed and such situations have been described as "EM data not available". * Mutant *e11c* has been renamed *cvs1*. ** Mutant *e13-1* has been renamed *cvs2*

mutants analyzed (*e5a, e9a, e10a, e10b, e11c, e12-1, e13-1* and *e19-1*) showed the presence of BL in both the vacuole and the apoplast except for *e19-1* where BL was localized only in the vacuole. This suggested that the screen was robust and that the majority of mutants identified had defects in *trans* rather than a *cis*-mutation. It is possible that mutant *e19-1* had a mutation in the ctVSS of Rat-GUS causing it to be secreted while leaving BL unaffected and correctly localized in the vacuole.

In all the mutants that appeared to missort BL, BL was also found in the vacuole indicating that no "null" mutant was observed. Vacuolar BL was found in osmiophyllic vacuolar protein aggregates (VPA, Figure 3.2, 3.4 and 3.5). In contrast, apoplastic BL was either found as a "layer" along the cell wall (Figure 3.3 and 3.4) or in osmiophyllic protein aggregates found in the intercellular spaces (Figure 3.2 and 3.4) at the junction of two or more cells. Sections of leaves were also examined but no missorting was detected as all the BL was localized in the vacuole (Figure 3.5). As a control, sections were treated with non-immune sera (Figure 3.6) and no specific labeling was detected. Moreover the VPA in the non-immune treatment was not labeled with immunogold particles, indicating that the labeling was specific for BL. However, BL protein has not been detected by electron microscopy in the parental root tips.

Partial mislocalization of BL can be attributed to a leaky mutation in the secretory machinery causing some portion of the protein to be transported to the vacuole and the other portion secreted to the apoplast. Considering that the vacuole is an essential organelle of the plant cell, null mutants could have been lethal and therefore been eliminated from the screen. It was also interesting that the mutations only affected protein transport in the root and not in the leaf. This bias may have occurred as the visual

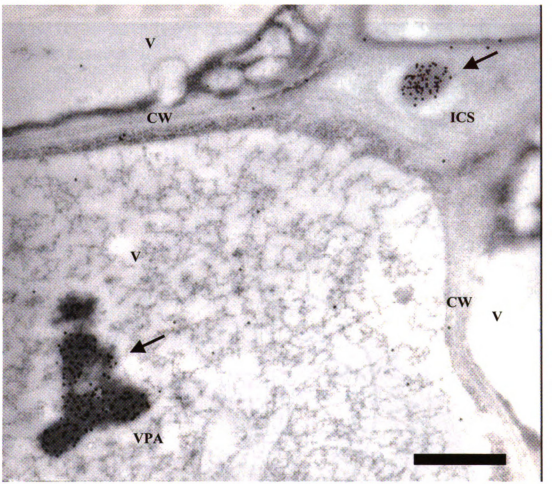

Figure 3.2 Immunolocalization of BL in section of root of mutant *e11c*
Blue roots were sectioned and prepared for immunoelectron microscopy (see Material and Methods) by Dr. V. Kovaleva. Immunodetection of BL was performed using rabbit anti-WGA antibodies. Arrow points to protein aggregates labeled with immunogold (10 nm) detecting the presence of BL. V = vacuole, VPA = vacuolar protein aggregate, CW = cell wall and ICS = intercellular space. Size bar = 0.5 μm.

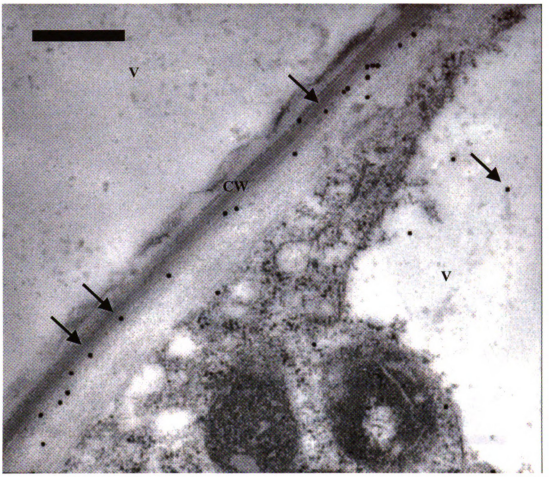

Figure 3.3 Immunolocalization of BL in section of root of mutant *e11c*
Blue roots were sectioned and prepared for immunoelectron microscopy (see Material and Methods) by Dr. V. Kovaleva. Immunodetection of BL was performed using rabbit anti-WGA antibodies. Arrow points to protein aggregates labeled with immunogold (10 nm) detecting the presence of BL. V = vacuole and CW = cell wall. Size bar = 0.25 μm.

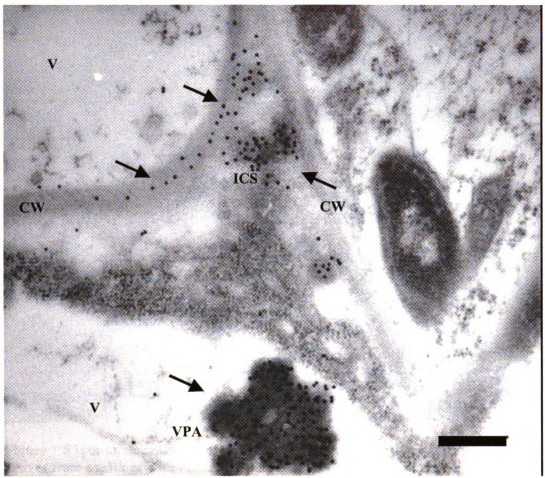

Figure 3.4 Immunolocalization of BL in section of root of mutant *e1Ic*
Blue roots were sectioned and prepared for immunoelectron microscopy (see Material and Methods) by Dr. V. Kovaleva. Immunodetection of BL was performed using rabbit anti-WGA antibodies. Arrow points to protein aggregates labeled with immunogold (10 nm) detecting the presence of BL. V = vacuole, VPA = vacuolar protein aggregate, CW = cell wall and ICS = intercellular space. Size bar = 0.25 μm.

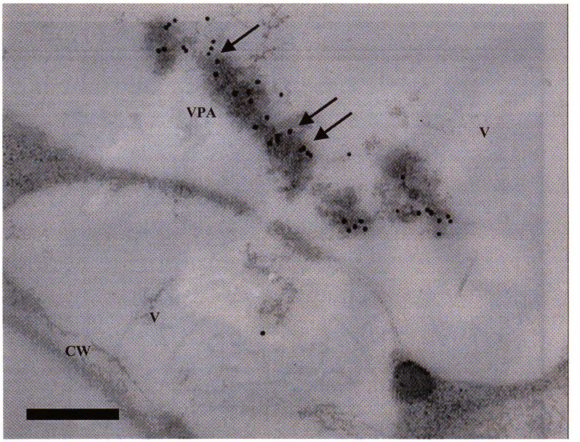

Figure 3.5 Immunolocalization of BL in section of leaf of mutant *e11c*
Leaves from seedlings with blue roots were sectioned and prepared for immunoelectron microscopy (see Material and Methods) by Dr. V. Kovaleva. Immunodetection of BL was performed using rabbit anti-WGA antibodies. Arrow points to protein aggregates labeled with immunogold (10 nm) detecting the presence of BL. V = vacuole, VPA = vacuolar protein aggregate and CW = cell wall. Size bar = 0.25 μm.

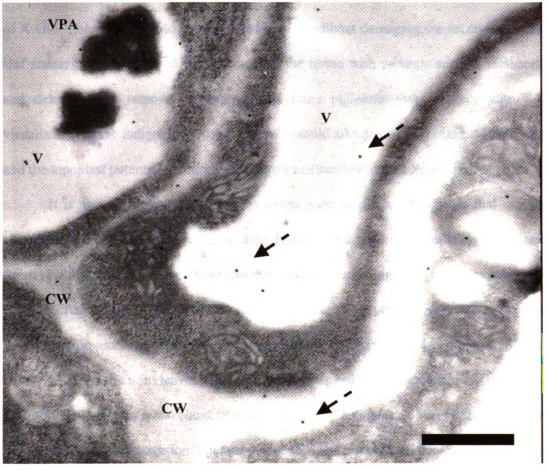

Figure 3.6 Non-immune control: section of root of mutant *e11c*
Blue roots were sectioned and prepared for immunoelectron microscopy (see Material and Methods) by Dr. V. Kovaleva. Control micrography was performed using non-immune sera. Broken arrow points to immunogold particles (10 nm). V = vacuole, VPA = vacuolar protein aggregate and CW = cell wall. Size bar = 0.5 μm.

screen was performed on roots. Since roots of *Arabidopsis* are white, upon the hydrolysis of X-GlcU, blue color could be visualized in vivo without damaging the seedling, unlike leaf staining which required the treatment of the tissue with solvents such as methanol and detergents to remove chlorophyll and other pigments that interfere with the visualization of the indigo dye. Such treatments would also damage the plasma membrane and the tonoplast potentially causing artefactual mislocalizations of Rat-GUS.

It is now apparent that higher organisms such as plants have genes that can be classified into gene families based on their function and sequence similarity. Clark and Roux (1999) have shown that the expression pattern of homologous genes or members of the same family can vary in a tissue specific manner. This phenomenon would explain why a specific mutation affecting the transport of ctVSS-bearing soluble proteins in roots does not affect that in leaves. Previous models based on components of the yeast secretory pathway did not account for tissue specificity and gene families as a unicellular model was under investigation. Kjemtrup et al (1995) demonstrated the transport of bean phytohemagglutinin (PHA) can vary in various tissues. In the meristem of the primary root, PHA is localized in the vacuole, whereas in elongated root cells, PHA is secreted. This indicates a tissue- and cell type-dependent variation in the properties of the secretory pathway. It is difficult, at this stage to determine the reason for obtaining root specific mutants. The positional cloning of the mutant loci will eventually shed some light on this phenomenon.

CONCLUSION

By screening a mutagenized a collection of seeds expressing Rat-GUS-ctVSS and BL, several mutants have been identified secreting Rat-GUS to the apoplast. Electron-microscopy revealed that the mutant phenotype is leaky as BL is observed both in the vacuole and in the apoplast. The lack of a mutant phenotype in leaf sections indicates the identification of root specific mutants that appear impaired in protein transport to the vacuole.

MATERIAL AND METHODS

EMS Mutagenesis

EMS mutagenesis was performed according to Haughn and Sommerville (1986). Rat-GUS-BL ($CRBT_5$ $6.18.5_X$) *A. thaliana* seeds (1 g) were soaked for 16 hours in 100 ml of 0.3 % (v/v) ethyl methane sulfonate (Sigma Chemical Company, St Louis MO, USA). They were subsequently washed with water 20 times, and sown in 12 flats (30 cm x 55 cm x 5 cm). They were grown under conditions of 22 ^{0}C, 110 μEinsteins of light at a 24-hour light duration for 2 months. Seeds were harvested in 47 pools, threshed and stored at 4 ^{o}C.

Visual Screen

Seeds (~5 mg) were sterilized and germinated vertically on media plates [Gamborg's-B5 basal medium with minimal organics (Sigma) 0.32 % (w/v), 2-(N-morpholino)ethanesulfonic acid (MES) 0.05% (w/v) and sucrose 2% (w/v), pH 5.7] for one week. Staining was carried out with X-GlcU (0.1 mg/ml of final suspension), dissolved in DMSO (0.1 % (v/v) of the final suspension) and the solution dispersed in 0.05 % (w/v)

sterile agarose. Seedlings were overlaid with 10 ml of X-GlcU solution and scored after 5 days. Seedlings with blue roots were identified and transferred to media plates (lacking X-GlcU) and grown until lateral roots developed and transferred to soil for recovery.

Immunoelectron microscopy

The root or leaf segments of mutant *Arabidopsis* seedlings were fixed in a mixture of 1.5% (v/v) formaldehyde, 0.5% (v/v) glutaraldehyde in 0.01 M sodium phosphate, pH 7.4, for 2.5 hours at room temperature on a rotator. The specimens were rinsed in the same buffer and postfixed in 0.5% (w/v) OsO_4 for 1 hour at room temperature. Dehydrated specimens were embedded in London Resin White (Polysciences, Warrington, PA). Ultrathin sections were made with an Ultracut S microtome (Reichert-Jung, Vienna, Austria) with a diamond knife and collected on nickel grids pre-coated with Formvar. Immunolabeling carried out according to Sanderfoot et al (1998) using primary rabbit polyclonal antibodies against wheat germ agglutinin (rabbit anti-WGA) and goat anti-rabbit IgG coupled directly to 10 nm colloidal gold particles. Control grids were treated with non-immune sera. The sections were examined with a Philips CM-10 transmission electron microscope. All labeling experiments were conducted several times each on independent sections. Tissue preparation and microscopy was carried out by Dr V. Kovaleva.

Chapter 4

Characterization of mutants and selection of *cvs1*

INTRODUCTION

Mutant screens often yield a large number of putative mutants that have to be characterized further prior to mapping the mutation. To identify mutations that affect the fate of reporter proteins, it is imperative to ensure that the reporter itself is not affected by the mutation. Transgenic reporters can also be affected by mutations that interfere with the expression of the transgene. Since a transgenic reporter construct was used to identify mutants impaired in the transport of ctVSS-bearing proteins, it was necessary to further analyze the performance of the reporters. Once a collection of mutants which display the desired phenotype is obtained, one can proceed with mapping and positional cloning.

RESULTS AND DISCUSSION

Identification and characterization of mutants

As the ctVSS-dependent sorting pathway is saturable (Frigerio et al, 1998 and Neuhaus et al, 1994), the possibility of secretion due to the over-expression of reporter proteins was investigated. The expression profile of Rat-GUS and BL in mutants *e5a, e11c, e12-1, e13-1*, $CRBT_5$-6.18.5_x (referred to in this chapter as "parental") was compared to that of a line of *Arabidopsis* plants expressing BL (referred to in this chapter as "control BL" line), in which BL is localized in the vacuole (Ahmed et al, 2000). Root tissue was harvested from plants grown in liquid culture. RNA was extracted from these tissues and analyzed by northern blotting. Filters were hybridized with randomly primed ^{32}P-labeled probes using fragments of cDNAs of Rat-GUS, BL or the 17S ribosomal RNA from rice as a loading control (Zarembinski and Theologis, 1993). Northern analysis (Figure 4.1) revealed that mutants could be categorized into two classes: low

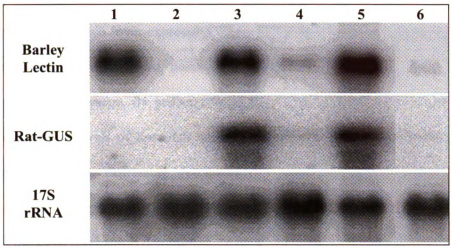

Figure 4.1 Northern analysis of root RNA from mature plants
Seedlings of mutant, $CRBT_5$-6.18.5_x and control barley lectin transformant were grown in liquid culture to obtain a large root mass (see Materials and Methods). RNA was extracted and 10 µg total RNA from each culture was resolved by MOPS-agarose electrophoresis. The RNA was then transferred to nitrocellulose and hybridized with randomly primed ^{32}P labeled probes. Probes used in the analysis were barley lectin, Rat-GUS and 17S ribosomal rDNA (see Materials and Methods) as a loading control.
Lane 1 = Barley lectin control, Lane 2 = $CRBT_5$-6.18.5_x, Lane 3 = mutant *e5a*, Lane 4 = mutant *e13-1*, Lane 5 = mutant *e12-1* and Lane 6 = mutant *e11c*.

expressers (*e11c* and *e13-1*) and over-expressers (*e5a* and *e12-1*). The levels of BL message in the two 'over-expressing' mutants (*e5a* and *e12-1*) matched with that of the control BL line. In contrast, mutants *e11c* and *e13-1* and parental lines had very low levels of BL message. By probing the northern blot with a probe for Rat-GUS, it was observed that levels of Rat-GUS message mimicked levels of BL message in all lines tested in that lines expressing high levels of BL also expressed high levels of Rat-GUS. Similarly, lines expressing low levels of BL had low levels of Rat-GUS. This indicated that the same mechanism determined the level of expression of both reporters.

The mRNA profile analysis was complemented with a western analysis using protein from the same tissues to determine the pattern of protein levels in mutants (*e5a, e11c, e12-1, e13-1*), parental and the control BL line. Total root protein was extracted from mature plants grown in liquid culture and the level of BL was detected by western analysis (Figure 4.2). It was observed that mutants could be categorized into two classes: low expressers (*e11c* and *e13-1*) and over-expressers (*e5a* and *e12-1*). The amount of BL in the low expresser was similar to that of the parental line indicating that the phenotype of secretion seen in these lines was not due to a saturation of the pathway but was perhaps a defect in the sorting machinery. This observation was further strengthened by results which demonstrated that the level of BL in the low expressing mutants was much lower than the control BL line, where BL is properly sorted to the vacuole. Taken together these results indicate that in mutants *e11c* and *e13-1*, the apoplastic localization of BL was not due to saturation of the transport machinery but may be due to a defect in the transport machinery caused by a *trans*-mutation.

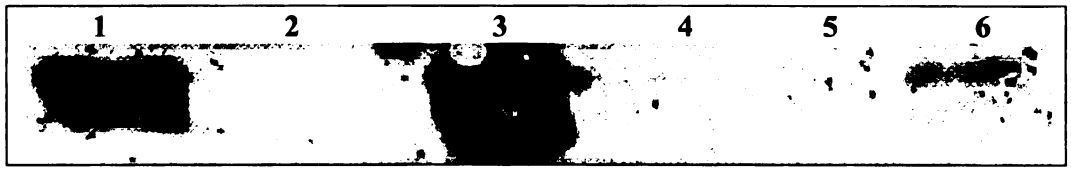

Figure 4.2 Western analysis of barley lectin (BL) in mutant and control tissues
Seedlings of mutant, $CRBT_5$-6.18.5$_x$ and control barley lectin transformant were grown in liquid culture to obtain a large root mass (see Materials and Methods). Protein was extracted and 10 μg total protein from each culture was resolved by SDS-PAGE and immunoblotted. Resolved proteins were visualized by a primary antibody treatment with rabbit anti-WGA, followed by treatment with goat anti-rabbit coupled to alkaline phosphatase. Blots were developed according to Harlow and Lane, 1988 (see Materials and Methods). Lane 1 = Barley lectin control, Lane 2 = $CRBT_5$-6.18.5$_x$, Lane 3 = mutant *e12-1*, Lane 4 = mutant *e13-1*, Lane 5 = mutant *e11c* and Lane 6 = mutant *e5a.*

A comparison of barley lectin protein levels in young parental tissues and mature tissues from liquid cultures revealed that while the protein can be detected in root tips (see Figure 2.5), it is absent in mature tissues. This was indicative of cosuppression, a phenomenon commonly seen in plants transformed with constitutive promoters such as the cauliflower mosaic 35S. These lines demonstrated a reduction of BL and Rat-GUS in mature tissues, such as those obtained from liquid cultures. This is very similar to results shown by Park et al (1996), where promoter homology was shown to cause the cossupression of transgenes driven by identical promoters. By using the same promoter (double 35S promoter) to drive BL and Rat-GUS, it is conceivable to suggest that promoter-homology may have induced silencing of the transgenes. Van Houdt et al (1997) showed that aging reinforced the silencing of transgenic neomycin phosphotransferase in tobacco plants. Palauqui et al (1996) showed that in tobacco lines, cosuppression can be induced during a period between 15 days after germination and flowering. Further Balandin and Castresana (1997) observed that silencing in tobacco occurs a few weeks after germination and is maintained throughout the vegetative phase

and flowering. However, silencing is reversed in developing seeds and fruits derived from meiotically divided cells (Balandin and Castresana, 1997). At this stage, it is difficult to determine if the silencing of BL and Rat-GUS are due to promoter homology or the presence of multiple insertions (tandem repeats) at the reporter locus.

The cossupressed condition in mutants *e11c* and *e13-1* would suggest that in these mutants, the ctVSS transport pathway may not be saturated and hence the partial mislocalization of BL in root tips as detected by electron microscopy, is due to a defect in the transport machinery. Unfortunately BL has not been detected in parental roots by electron microscopy although it is detectable in root tips by western analysis (see Figure 2.5). However, the identification of mutant *e19-1* (see Table 3.3) where Rat-GUS is secreted but BL is not mislocalized, indicates that transgenic BL is localized in the vacuole. This is also evident in leaf sections where BL was found exclusively in the vacuole. Taken together it can be implied that mutants *e11c* and *e13-1* partially missort BL to the apoplast due to a root-specific defect in the transport machinery and not saturation of the pathway.

It was also observed that in mutant *e11c*, the terminal inflorescence was "compressed" (Figure 4.3a and b) in comparison with the parental terminal inflorescence (Figure 4.3a and c), a phenotype that is consistent with a poor elongation of internodes in floral tissues. At this stage, it is not clear if the secretion phenotype and the inflorescence phenotype are linked. Mutant *e11c* has now been renamed *cvs1* for ctVSS vacuolar sorting defect and mutant *e13-1* has been renamed *cvs2*.

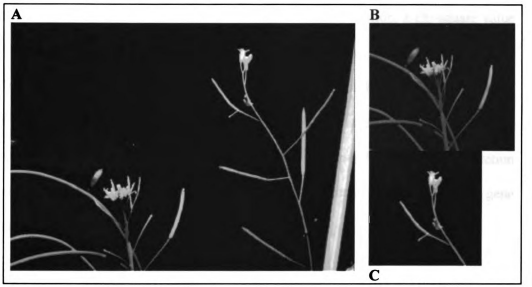

Figure 4.3 Examination of floral phenotype of *cvs1*
Mature inflorescences of $CRBT_5$-6.18.5_x and mutant *cvs1* were compared. Panel A = Inflorescence of mutant *cvs1* (left) and $CRBT_5$-6.18.5_x (right). Panel B = Close up of inflorescence of mutant *cvs1* with *"compressed"* phenotype. Panel C = Close up of inflorescence of $CRBT_5$-6.18.5_x

Genetics of mutant *cvs1*

In order to understand the nature of the mutation and eventually clone the gene whose mutation resulted in the phenotype, the mutant *cvs1* was backcrossed to the parental line (F_1 cross) (according to Raventos et al, 2000). This progeny was germinated in plates set vertically so that the roots develop on the surface of the medium. Upon staining with X-GlcU, it was found that none of the roots turned blue. This indicated a recessive phenotype, as mutant *cvs1* was homozygous (see Table 3.3). To confirm this finding, the F_1 seedlings were transferred to soil and their selfed progeny harvested. This progeny (F_2) was scored for segregation of the phenotype by performing the Rat-GUS secretion assay. Results indicated a 94 white seedlings (vacuolar) to 33 blue seedlings (secretion) ratio. Applying chi-square analysis to this information against an expected

segregation ratio of 3:1 (vacuolar: secreted) for a recessive mutation, a chi-square value of 0.066 and a *p*-value of 0.797 were obtained. (Note the smaller the chi-square the better and the higher the *p*-value the better so as to "fail to reject a null hypothesis of equality of ratios"). The large *p*-value strongly indicates the lack of a significant deviation in the sample from an expected ratio of 3:1. These findings confirmed the results of secretion assays performed on the F_1 generation indicating that mutant *cvs1* had a single gene recessive mutation.

CONCLUSION

By screening a population of EMS-mutagenized seeds expressing Rat-GUS-ctVSS and BL, several candidate mutants were identified using the Rat-GUS secretion assay (see Chapter 3). Immunoelectron microscopy revealed that in several mutants, BL was mislocalized in the apoplast in a "leaky" manner. The evaluation of reporter message levels and protein levels in these mutants allowed the identification of two mutants (*cvs1* and *cvs2*) potentially impaired in protein transport to the vacuole, resulting in the partial secretion of the reporter proteins. Based on the segregation ratio of the mutant phenotype after back crossing into the parental line, it appears that the phenotype of *cvs1* is determined by a single gene recessive mutation. It can be inferred that *cvs1* is a mutant with a single gene defect resulting in the partial secretion of ctVSS-bearing vacuolar reporter proteins.

MATERIALS AND METHODS

Arabidopsis Liquid Culture

Seeds (5-10) were surface-sterilized and cultured in sterile liquid culture medium [0.43 % Murashige and Skoog salts, 2-(N-morpholino)ethanesulfonic acid (MES) 0.05% (w/v), 1 % (w/v) sucrose at pH 5.7, supplemented with 0.1 mg/ml myo-inositol, 1 μg/ml nicotinic acid, 1 μg/ml pyridoxine and 1 μg/ml thiamine] with agitation at 250 rpm, 120 μEinstein light for 16 hours per day for 4 weeks. Root and shoot tissue was then frozen in liquid nitrogen and stored at –80 °C until needed.

Northern blotting

RNA for northern analysis was extracted using the approach of Houdebine and Puissant (1990) with certain modifications. Liquid cultured root tissue from mutant and wild type plants were harvested, frozen in liquid nitrogen and stored at –80 °C for extraction. One gram of tissue was homogenized with a mortar and pestle in liquid nitrogen, and the powder was extracted with 4 ml GTC solution [4 M guanidinium thiocyanate, 25 mM sodium citrate pH 7.0, 0.5 % (w/v) sodium lauryl sarkosinate and 10 mM β-mercaptoethanol]. The suspension was mixed with 650 μl of sodium acetate pH 4.0 by vortexing, followed by vortexing with 5 ml of phenol equilibrated with TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA). Finally 1 ml of chloroform-isoamyl alcohol (24:1) was added and the suspension centrifuged at 10000 g for 10 minutes at 4 °C. The aqueous phase (supernatant) was precipitated with 5 ml isopropyl alcohol by storing at 4 °C overnight. The suspension was then centrifuged at 10000 g for 10 minutes at 4 °C. The supernatant was discarded and the pellet of RNA dislodged with 1 ml of 4 M lithium

chloride. The precipitate was dispersed by vortexing and the RNA was precipitated by centrifugation. The lithium chloride step was repeated. The RNA pellet was then resuspended in 600 µl RR buffer [10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.5 % (w/v) SDS in DEPC-treated water]. The suspension was re-extracted with phenol-chloroform-isoamyl alcohol and finally RNA was precipitated with isopropanol. The yield of RNA was measured by absorption at 260 nm.

RNA (10 µg) was resolved by MOPS agarose gel electrophoresis as per Bar-Peled and Raikhel (1997) and transferred to nitrocellulose membranes. Membranes were hybridized with ^{32}P-labeled probes as per Bar-Peled and Raikhel (1997) synthesized from cDNA fragments encoding Rat-GUS (EcoRV fragment comprising of nucleotides 1516 to 1881) or barley lectin (ApaI-BamHI fragment comprising of nucleotides 40 to 546) or 17S ribosomal RNA from rice (see Zarembinski and Theologis, 1993). Hybridized blots were visualized using a Molecular Dynamics phophorimager.

Western Analysis of barley lectin

Liquid cultured root tissue from mutants, parental and control BL plants were harvested, frozen in liquid nitrogen and stored at –80 °C for extraction. One gram of tissue was homogenized with a mortar and pestle in liquid nitrogen, and the powder was resuspended in protein extraction buffer [PBS, 0.1% (v/v) Triton X-100, 0.1 % (w/v) sodium-N-lauryl sarcosine, 1 mM PMSF], and then centrifuged twice at 4 ^{0}C for 10 minutes at 15000 g. Protein contents were estimated by Bradford's assay (Bio-Rad Laboratories, Hercules, CA) using the Bradford analysis package, bradfordhsa.htm (see Appendix B).

Protein extracts were diluted with half a volume of 3X Laemmli buffer (Harlow and Lane, 1988). The diluted samples (25 µl) were boiled, treated with 5 µl acetamide solution (5 µl 1X Laemmli buffer, 2 mg Tris base, 1.33 mg iodoacetamide) and the mixture was incubated at 37 ^{0}C for 30 minutes prior to resolving by SDS-PAGE. Proteins were immunoblotted onto nitrocellulose filters, transfer confirmed by staining with Ponceau S, blocked and incubated with primary antibodies (rabbit anti-WGA) at 1:1000 dilution and with secondary antibodies (goat anti-rabbit linked to alkaline phosphatase) at 1:1000 dilution according to Harlow and Lane (1988). Protein bands were visualized using 50 ml of alkaline phosphatase buffer (100 mM sodium chloride, 5 mM magnesium chloride, 100 mM Tris, pH 9.5) containing 0.16 mg/ml 5-bromo-4-chloro-3-indolyl phosphate and 0.35 mg/ml nitro blue tetrazolium chloride. After the development of bands, the blots were rinsed with distilled water and photographed.

Crosses

Seedlings were grown under controlled conditions of 22 °C, 110 µEinsteins of light at an 18-hour light duration for 4 weeks until 3-4 siliques had fruited. Stamens from pollen donor flowers were dissected and used to pollinate emasculated flowers. The pollinated flowers were covered with plastic film for 4 days until the silique started developing and then the film was removed. Seeds were harvested and stored at 4 °C for further use.

Chapter 5

Mapping *cvs1* to a contig of BACs

INTRODUCTION

Upon the identification and characterization of a mutant, it is necessary to identify the mutant locus and map it as well as possible to identify the gene encoding the protein involved in the transport pathway. In *Arabidopsis* and other self-fertile species, this can be achieved by crossing the mutant into a different ecotype that offers a number of polymorphic characters for mapping. It is then possible to measure the linkage between the mutant locus and defined polymorphic loci in a segregating mapping population. This approach assumes that the phenotype is not ecotype-dependent. In *Arabidopsis*, a large number of polymorphic loci have been identified on all five chromosomes between ecotypes Columbia (Col) and Landsberg *erecta* (L*er*) enabling a fine mapping analysis.

RESULTS AND DISCUSSION

The backcrossed and self-fertilized line of *cvs1* (hereafter called ACWS) was crossed into Landsberg *erecta (Ler)*. The progeny from this cross was allowed to self-fertilize to generate a segregating population of mutants (11BC2F1L*er*F2). This population was scored for secretion of Rat-GUS by germination of seedlings on vertical media plates and staining the seedlings with 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-GlcU). The identification of seedlings secreting Rat-GUS activity confirmed that the phenotype under investigation was not affected by crossing into the mapping ecotype. All the individuals with blue roots were transferred to soil and genomic DNA was extracted from leaves. Two types of PCR-based codominant markers were used in the mapping analyses: Cleaved Amplified Polymorphic Sequences (CAPS) markers and Simple Sequence Length Polymorphisms (SSLP) markers. CAPS amplify fragments of the same

size in both Landsberg *erecta* and in Columbia (Col) but have a different restriction pattern (Konieczny and Ausubel 1993) while SSLP markers generate PCR products of different sizes without any further treatment (Bell and Ecker, 1994).

The analysis was initially performed on 10 DNA (Table 5.1) samples from segregating lines expressing the mutant phenotype, with markers representing various loci on the five chromosomes. Linkage analysis was carried out by evaluating the recombination frequency (RF) and calculating a chi-square for deviation from an expected ratio of 1:2:1 (Homozygous Col: Heterozygous: Homozygous L*er*). An unlinked locus would then have a recombination frequency of 0.5 (50 %) and a low chi-square value (high *p*-value), indicating a low deviation of the sample from the expected ratio of 1:2:1. Conversely a large deviation from a 1:2:1 ratio as indicated by a large chi-square value and a low *p*-value would indicate that the marker locus and the mutant locus are linked. These calculations were performed by developing an HTML-based code available at http://www.msu.edu/~venkata1/sslpfull.htm (see Appendix A). Recombination frequencies on chromosomes I to IV ranged between 0.4 and 0.6 indicating a lack of linkage between the markers and the mutant locus. However markers on chromosome V showed a linkage to the mutation as the RF ranged around 0.37 at marker ATSO262. Though this was not a close linkage, it was indicative of linkage considering the small sample size (N= 10). It was observed that the chi-square value was not very high in this small sample.

Table 5.1 10-Sample mapping analysis

Chromosome	Position	Marker	Total	RF	Linkage (cM)	chi-square	*p*-value
I	9.4	nga63	10	0.50	No Linkage	0.40	0.82
I	40.1	nga248	7	0.57	No Linkage	0.43	0.81
I	81.4	nga280	10	0.45	115.13	0.20	0.90
I	111.4	nga111	10	0.60	No Linkage	4.40	0.11
I	118.3	nga692	8	0.63	No Linkage	1.00	0.61
II	62	nga361	10	0.55	No Linkage	1.80	0.41
II	73	nga168	10	0.50	No Linkage	0.40	0.82
III	21	nga162	20	0.38	69.31	2.70	0.26
III	85.3	nga6	10	0.70	No Linkage	3.60	0.17
IV	19.8	nga12	6	0.42	89.59	3.00	0.22
IV	24.2	nga8	19	0.58	No Linkage	1.00	0.61
IV	102	nga1107	10	0.55	No Linkage	1.80	0.41
V	23.1	nga249	10	0.40	80.47	1.20	0.55
V	33	nga106	10	0.35	60.20	3.40	0.18
V	51.1	nga139	26	0.37	65.61	3.77	0.15
V	65.2	AthSO262	17	0.38	72.35	2.41	0.30
V	71.13	AthPHY-C	16	0.44	103.97	0.50	0.78
V	79	AthSO191	16	0.44	103.97	1.50	0.47

Summary of a 10-sample mapping analysis covering 5 chromosomes with various SSLP and CAPS markers. Position = Position of marker locus on chromosome (cM). Total = Number of segregants analyzed. RF = Recombination Frequency. Linkage = Distance (cM) between marker locus and *cvs1* estimated from the RF using Haldane's mapping function (for review Koorneef and Stam, 1992). Chi-square = chi-square calculated to estimate deviation of observed frequency from expected frequency in case of no linkage (Col : Het : L*er* = 1 : 2 : 1)

A similar analysis was performed on a larger sample (n = 45) shown in Table 5.2. Previously observed results were confirmed and a linkage to markers on chromosome V was evident with recombination frequencies of 0.28 at marker ATSO191 and higher RFs on either side of the marker. This is also evident from the analysis of chi-squares where a large deviation from the expected 1:2:1 ratio for an unlinked scenario was observed. As observed in the small sample analysis, the RF values were not as small as expected when approaching the mutant locus. This is likely due to the mis-identification of secreting mutants in the segregating population. Such a result is often obtained when phenotypes have to be evaluated subjectively (personal communication from Dr. Maarten Koornneef, Wageningen Agricultural University, Wageningen, The Netherlands). Further optimization of the RF by incorporating the secondary reporter, BL, could not be carried out as secreted BL in roots can only be detected by electron microscopy. The inclusion of non-mutants in the mapping analysis would lead to a larger recombination frequency than expected had there been no such individuals in the sample. To circumvent this problem, a reduction in the recombination frequency and a corresponding increase in chi-square value were used as indicators of linkage. A large-scale analysis using 135 DNA samples from secreting segregants was carried out using markers AthPHYC, ATSO191, DFR and LFY3 (Table 5.3). In this analysis, the RF was lowest around DFR. The trend of chi-square values reflected the trend of the RF values with marker DFR having the largest chi-square value (80.18). This suggested that mutation *cvs1* lay between AtSO191 and DFR.

As commercially available markers did not cover this region, new markers were developed. The region between AtSO191 and DFR is spanned by a set of 32 overlapping

Table 5.2 46-Sample mapping analysis

Chromosome	Position	Marker	Total	RF	Linkage (cM)	chi-square	*p*-value
I	40.6	nga248	46	0.58	No Linkage	3.05	0.22
I	81.71	nga280	46	0.50	No Linkage	0.22	0.90
I	111.4	nga111	46	0.46	122.84	5.34	0.07
II	50.56	nga1126	46	0.52	No Linkage	0.21	0.90
III	16.2	nga126	46	0.51	No Linkage	0.07	0.96
III	83.5	nga6	46	0.54	No Linkage	1.48	0.48
IV	24.1	nga8	46	0.63	No Linkage	10.52	0.01
IV	60.35	Aga	46	0.41	87.15	3.35	0.19
IV	102	nga1107	46	0.43	101.84	2.35	0.31
V	33.3	nga106	45	0.41	86.36	3.04	0.22
V	50.48	nga139	46	0.38	71.54	5.26	0.07
V	65.2	AthSO262	46	0.33	54.93	10.02	0.01
V	71.13	AthPHYC	45	0.30	45.81	16.40	0.00
V	79	AthSO191	46	0.28	39.93	17.80	0.00
V	87.63	DFR	46	0.33	52.49	9.80	0.01
V	115.01	LFY3	46	0.36	62.09	7.53	0.02

Summary of a 46-sample mapping analysis covering 5 chromosomes with various SSLP and CAPS markers. Position = Position of marker locus on chromosome (cM). Total = Number of segregants analyzed. RF = Recombination Frequency. Linkage = Distance (cM) between marker locus and *cvs1* estimated from the RF using Haldane's mapping function (for review Koorneef and Stam, 1992). Chi-square = Chi-square calculated to estimate deviation of observed frequency from expected frequency in case of no linkage (Col : Het : L*er* = 1 : 2 : 1)

Table 5.3 Fine map between AthPHYC and LFY3

Marker	AthPHYC	AthSO191	DFR	LFY3
Chromosome	V	V	V	V
Position	71.13	79.00	87.63	115.01
Total	135	135	135	135
RF	0.27	0.25	0.23	0.37
Linkage (cM)	39.04	34.66	31.72	67.06
chi-square	59.73	73.76	80.18	17.82
***p*-value**	0.00	0.00	0.00	0.00

Fine map of *cvs1* between AthPHYC and LFY3 on chromosome V. Position = Position of marker locus on chromosome V (cM). Total = Number of segregants analyzed. RF = Recombination Frequency. Linkage = Distance (cM) between marker locus and *cvs1*. Chi-square = chi-square calculated to estimate deviation of observed frequency from expected frequency in case of no linkage (Col: Het: Ler = 1:2:1)

bacterial artificial chromosomes (BAC) shown in Figure 5.1. New markers were developed for BACs MBB18, MKM21, K15E6 and MUL8. These were developed by performing multiple FASTA analyses (Pearson and Lipman, 1988) using do_Fasta.pl (Dr. Joe White, Department of Biochemistry, Michigan State University, East Lansing, MI, USA). The database of L*er* genomic sequences (available at http://www.tigr.org/tdb/at/atgenome/Ler.html) was searched for identities using the sequence of the 32 BACs. Within high identity regions, diverging domains were identified for developing polymorphism markers. Thus markers for BACs MBB18 (Figure 5.2), MKM21 (Figure 5.3), K15E6 (Figure 5.4) and MUL8 (Figure 5.5) were developed. Markers for BACs MBB18 and MUL8 were developed as SSLPs (Bell and Ecker, 1994), while markers for BACs MKM21 and K15E6 were developed as CAPS (Konieczny and Ausubel, 1993) markers (see Table 5.4). In addition CAPS markers for

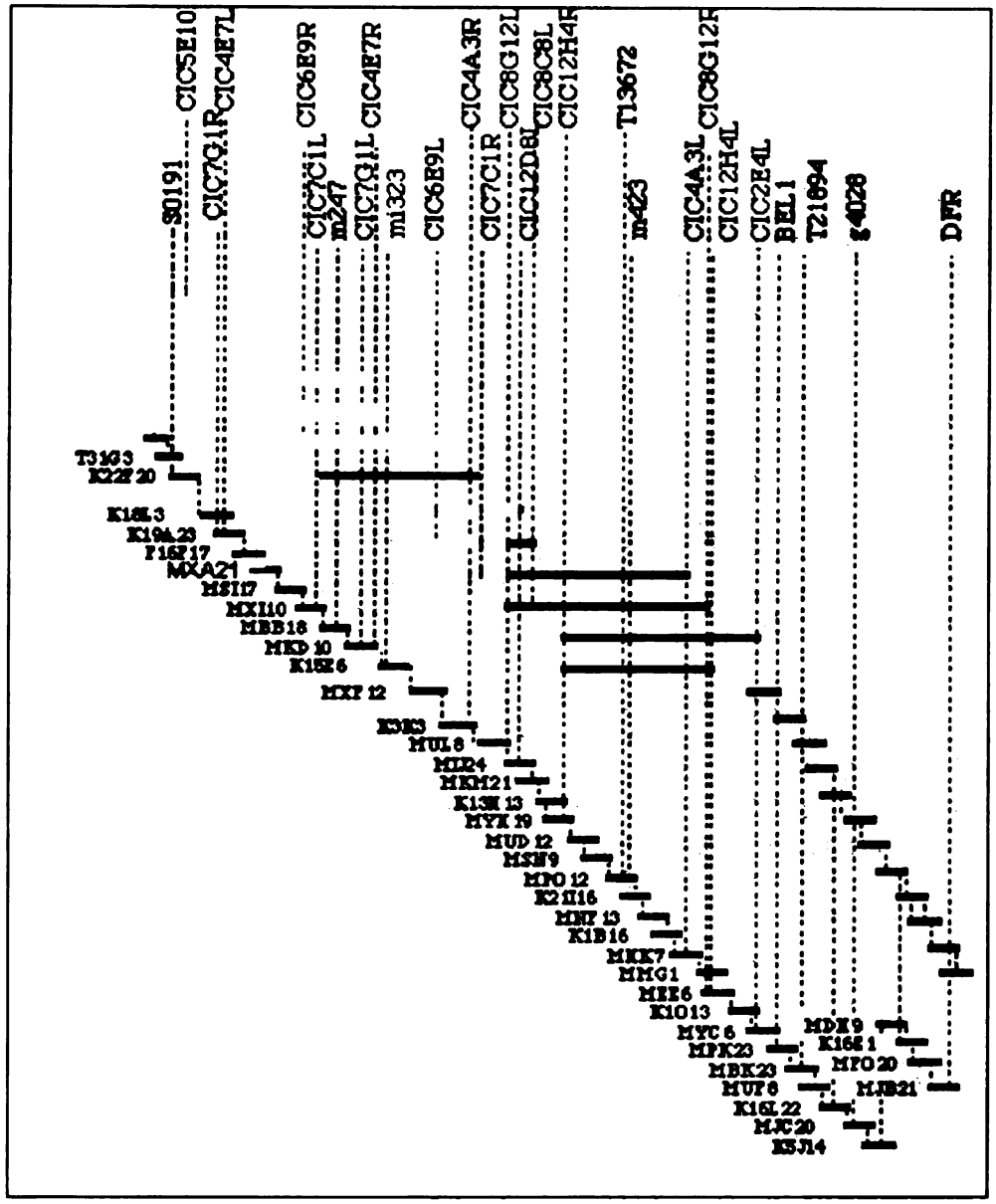

Figure 5.1 Alignment of BAC contig between AtSO191 and DFR on chromosome V
BACs are represented by blue bands and labeled. Black solid lines represent yeast artificial chromosomes (YACs) spanning the same region. Available from the Kazusa DNA Research Institute, Chiba, Japan

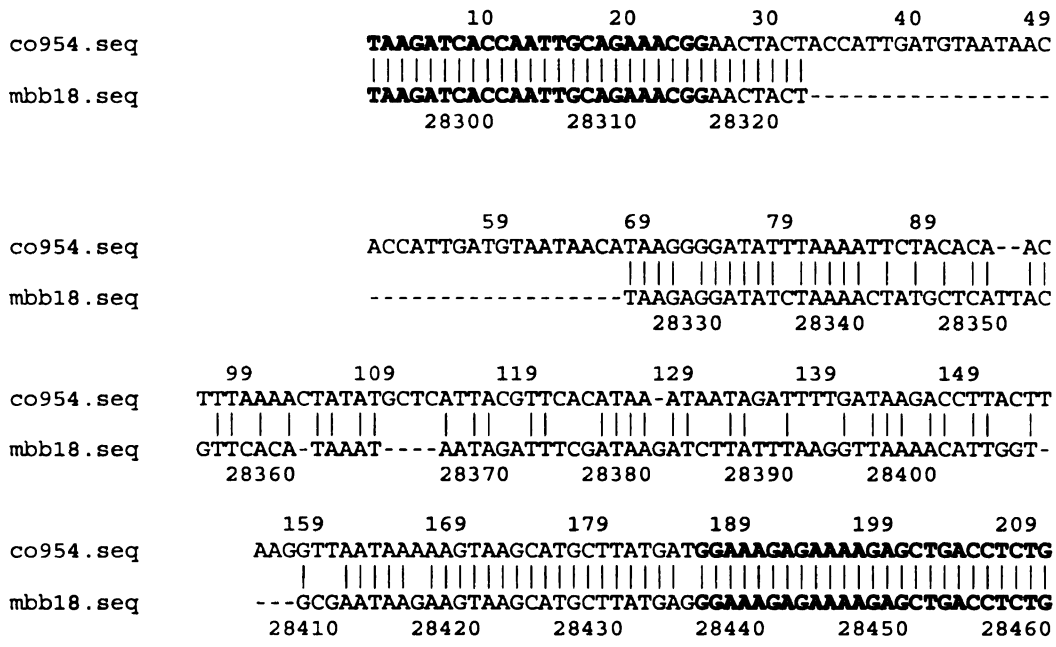

Figure 5.2 Alignment of BAC MBB18 (Col, 28295 – 28460) with Clone co954 (L*er*)
This alignment was used to identify polymorphic regions. Letters in bold indicate forward and reverse primers. Polymorphic stretch between 33 and 68 L*er* was utilized to generate an SSLP. Columbia DNA sequence was obtained from the Kazusa DNA Research Institute, Chiba, Japan and Landsberg *erecta* sequence from The Institute for Genomic Research, Rockville, MD, USA (see Materials and Methods).

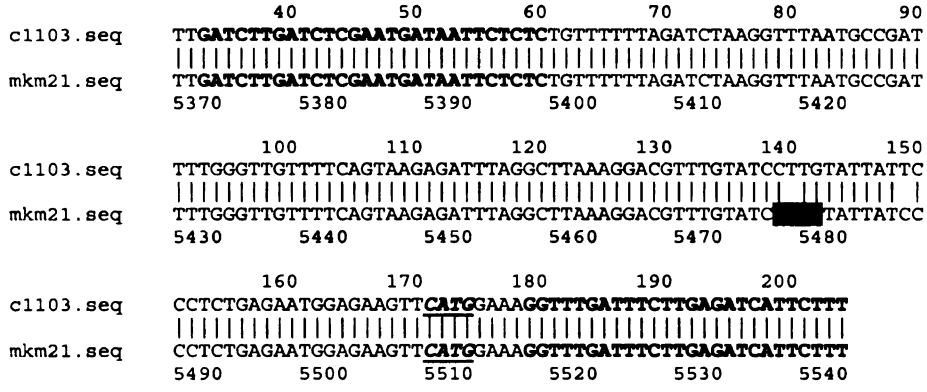

Figure 5.3 Sequence alignment of a section of BAC MKM21 (Col) and c1103 (L*er*)
This alignment was used to identify polymorphic regions. Letters in bold indicate forward and reverse primers. Shaded box on Col sequence at 5475-5478 = polymorphic NlaIII site Bold and underlined sequence indicates common NlaIII site in both sequences. Columbia DNA sequence was obtained from the Kazusa DNA Research Institute, Chiba, Japan and Landsberg *erecta* sequence from The Institute for Genomic Research, Rockville, MD, USA (see Materials and Methods).

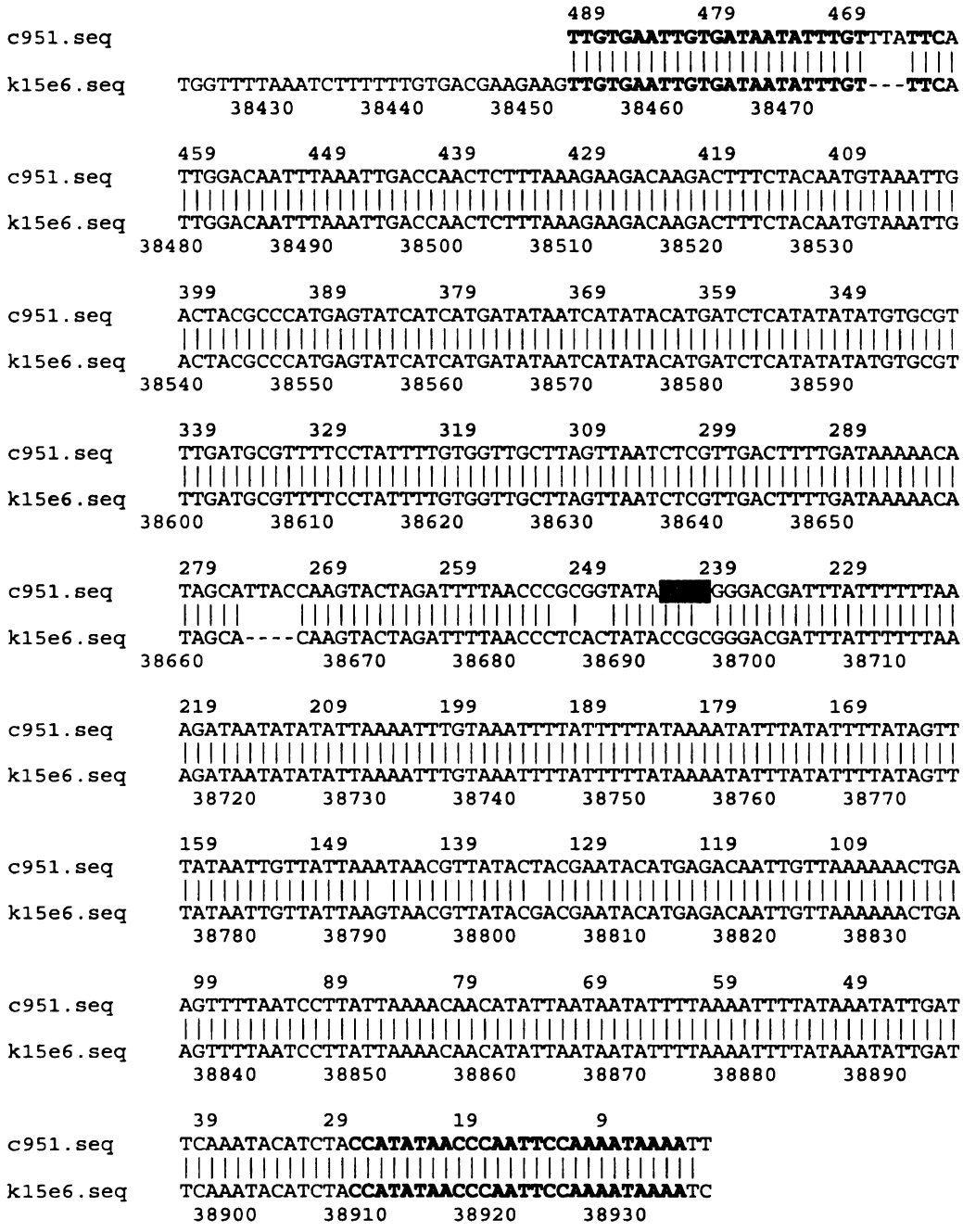

Figure 5.4 Sequence alignment of a section of BAC K15E6 (Col) and c951 (L*er*)
This alignment was used to identify polymorphic regions. Letters in bold indicate forward and reverse primers. Shaded box on L*er* sequence at 239-242 = HpaII site of polymorphism. Columbia DNA sequence was obtained from the Kazusa DNA Research Institute, Chiba, Japan and Landsberg *erecta* sequence from The Institute for Genomic Research, Rockville, MD, USA (see Materials and Methods).

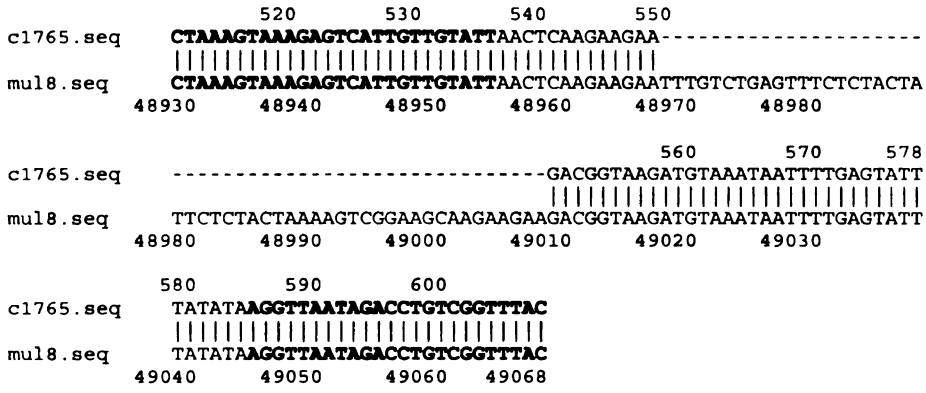

Figure 5.5 Alignment of BAC MUL8 (Col, 48930 – 49068) with Clone c1765 (L*er*)
This alignment was used to identify polymorphic regions. Letters in bold indicate forward and reverse primers. Polymorphic stretch between 48968 and 49008 in Col was utilized to generate an SSLP. Columbia DNA sequence was obtained from the Kazusa DNA Research Institute, Chiba, Japan and Landsberg *erecta* sequence from The Institute for Genomic Research, Rockville, MD, USA (see Materials and Methods).

Table 5.4 Markers used to map *cvs1* within between ATSO191 and DFR

Marker	Primer	Primer Sequence	Type	Polymorphism
MBB18[&]	Forward	TAAGATCACCAATTGCAGAAACGG	SSLP	Col = 169
	Reverse	CAGAGGTCAGCTCTTTTCTCTTTCC		Ler = 192
K15E6[&]	Forward	TTGTGAATTGTGATAATATTTGTTTC	CAPS	Col = 500
	Reverse	TTTTATTTTGGAATTGGGTTATATGG	*HpaII*	Ler= 250, 250
MUL8[&]	Forward	CTAAAGTAAAGAGTCATTGTTGTATT	SSLP	Col=140
	Reverse	GTAAACCGACAGGTCTATTAACCT		Ler= 99
MKM21[&]	Forward	GATCTTGATCTTGATCTCGAATGAT	CAPS	Col = 117, 32, 29
	Reverse	AAAGAATGATCTCAAGAAATCAAACC	*NlaIII*	Ler = 149, 29
MYH19*	Forward	ACTATTTATGGCTTAGCACTT	CAPS	Col=195
	Reverse	ATCTGTAAATCTTGTATATTACC	*DdeI*	Ler = 175, 20
MSN9R*	Forward	TGGCAGTAATTTAATGATAAACTA	CAPS	Col=200
	Reverse	CGTATTGAGCCTACATAACAA	*SpeI*	Ler = 180, 20
MNF13R*	Forward	CAGAGTCATGGTAAATGCCTG	CAPS	Col=190
	Reverse	GTTTCTTTGAGTATTAGAGTCC	*BstNI*	Ler = 170, 20

Markers for each BAC are listed along with the sequence of the amplification primers and the restriction enzymes (where appropriate). & = Primer sequence and polymorphism profile developed by multiple FASTA analyses using *A.t.* ecotype Columbia BAC sequences and *A.t.* ecotype Landsberg *erecta* random genomic clone sequences (The Institute for Genomic Research, Rockville, MD, USA). CAPS = Cleaved Amplified Polymorphic Sequences. SSLP = Simple Sequence Length Polymorphisms. Restriction enzymes used in developing the polymorphism are italicized. * = Ziegelhoffer et al (2000).

BACs MYH19, MSN9R and MNF13R were also used (Ziegelhoffer et al, 2000). These markers were used to map the meiotic recombination break points between AtSO191 and DFR (Table 5.5) indicating a recombination event between L*er* and Columbia chromosomal DNA in the neighborhood of *cvs1*.

Since the mutation was developed in a Columbia background, the region flanking the mutant locus is expected to be homozygous for Columbia markers (labeled blue in Table 5.5). The detection of a chromosomal break point on BACs MBB18 and MUL8 indicated that mutation *cvs1* lies "south" of the polymorphism marker for BAC MBB18 (25231 bp on BAC MBB18) and "north" of the marker for BAC MUL8 (48929 bp on BAC MUL8). The length of this region is about 417 kbp and is spanned by a contig of 6 overlapping BACs: MBB18, MKD10, K15E6, MXF12, K3K3 and MUL8 (Figure 5.6). As this region includes the CAPS marker for BAC K15E6, the mapping analysis performed with this marker revealed a Columbia pattern in all samples tested. This confirmed that the mutation lay in this region.

CONCLUSION

Using a combination of CAPS and SSLP markers the mutation in *cvs1* was initially localized between ATSO191 and DFR on chromosome V. By developing markers specific for the BACs spanning this region along with other markers, *cvs1* was mapped to a 6-BAC contig including MBB18, MKD10, K15E6, MXF12, K3K3 and MUL8. Further analysis by complementation using BACs and cosmid clones should reveal the exact position of *cvs1*.

Table 5.5 Fine mapping analysis of *cvs1* between AtSO191 and DFR

Marker	AthPHYC	AthSO191	MBB18	K15E6	MUL8	MKM21	MYH19	MSN9R	MNF13R	DFR	LFY3
Chromosome	V	V	V	V	V	V	V	V	V	V	V
Position	12.8	13.7	14.25	14.3	14.6	14.7	14.8	14.9	15.15	15.9	24.1
Lines	Polymorphism										
33	H	C	C	C	C	C	C			C	H
45	C	C	C	C	C	C	C	C	H	H	O
46	H	H	C	C	C	C	C	C	C	C	C
48	H	C	C	C	O	C	C			C	C
55	C	C	C	C	H	H	H	H	H	H	L
67	O	H	O	C	H	H	H	H	H	H	C
68	O	C	C	C	O	C	C			C	C
87	C	C	C	C	C	C	C	C	C	H	H
89	C	C	C	O	C	C	C	C	C	O	H
109	C	C	C	C	C	C	O	C	C	O	H
120	H	C	C	C	C	C	O			C	C
124	H	C	C	C	C	C	C			C	C
130	H	H	H	C	C	C	C	C	C	C	C
132	H	H	C	C	C	C	C	C	C	C	H

Markers developed for this region were used in mapping analyses and the results are tabulated. Lines: numbers represent individual segregant DNA sample. H = Heterozygous polymorphism (Red). C = Columbia polymorphism (Blue). L = Landsberg *erecta* polymorphism (Yellow). 0 = No amplification. Blank = DNA was not sampled. Marker = SSLP or CAPS marker used in developing the polymorphism. Position = Mbp along chromosome V.

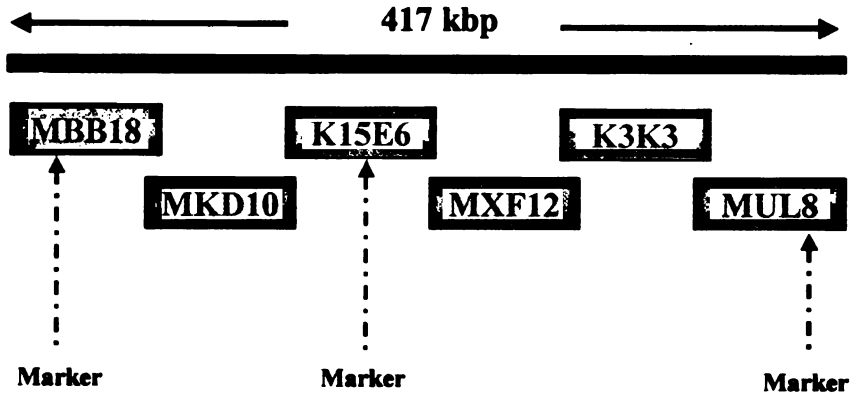

Figure 5.6 BAC contig spanning *cvs1* locus
Light blue boxes represent BACs and broken arrows indicate the approximate position of the mapping markers used.

MATERIALS AND METHODS

Crosses

Line ACWS was crossed into *Arabidopsis thaliana* ecotype Landsberg *erecta* (L*er*) as follows. Seedlings were grown under controlled conditions of 22 °C, 110 μEinsteins of light at an 18-hour light duration for 4 weeks until 3-4 siliques had fruited. Stamens from ACWS flowers were dissected and used to pollinate emasculated L*er* flowers. The pollinated flowers were covered with plastic film for 4 days until the silique started developing and then the film was removed. Seeds (F_1) were harvested and stored at 4 °C for further use. The self-fertilized progeny from the F_1 line was harvested (F_2).

F_2 seeds were sterilized in 50 % commercial bleach with a drop of Tween 20, and rinsed several times. These seeds were plated on *Arabidopsis* germination media [2 % (w/v) sucrose, Gamborg's B5, 300 mg/L carbenicillin] for seven days and stained with 5-bromo-4-chloro-3-indolyl β-D-glucuronide [0.1 % (w/v) in 0.05 % (w/v) agarose] for 5

days. Blue seedlings were transferred to soil and leaves from individual seedlings harvested.

Genomic DNA isolation

Leaves from individual blue seedlings were homogenized in liquid nitrogen using a Kontes motorized pestle and the leaf resuspended in extraction buffer [2% cetyltrimethylammonium bromide (CTAB), 100 mM Tris-HCL, pH 8.0, 1.4 M NaCl, 20 mM EDTA, 0.2 % β-mercaptoethanol] and incubated at 65 ^{0}C for 1 hour. The suspension was extracted with chloroform. The aqueous phase was then treated with ethanol to precipitate genomic DNA. The DNA pellet was air dried at room temperature overnight and dissolved in TE (10 mM Tris-HCl pH 7.0, 1 mM EDTA) containing 10 µg/ml bovine pancreatic RNAse (Roche, Palo Alto, CA, USA).

PCR Amplifications: SSLP markers

SSLP analysis was performed according to Bell and Ecker, 1994. Briefly, genomic DNA samples were amplified in a 20 µl volume of 20 mM Tris-HCl pH 8.4, 50 mM KCl, 0.2 mM dNTPs, 2 mM $MgCl_2$ containing 50 ng genomic DNA, 1 unit Taq DNA polymerase (Gibco-BRL, Rockville, MD, USA), 5 pmoles forward primer and 5 pmoles reverse primer. The following cycling profile was used to amplify the reactions: denaturation at 94 °C for 2 minutes, amplification cycle (40 cycles) at 94 °C for 15 seconds, 55 °C for 15 seconds, 72 °C for 30 seconds and extension at 72 °C for 7 minutes. SSLP primers were obtained from Research Genetics, Huntsville, AL, USA. Amplified

products were resolved by TBE-agarose gel electrophoresis and visualized by ethidium bromide staining.

PCR Amplifications: CAPS markers

CAPS analysis was performed according to Konieczny and Ausubel (1993). Briefly, genomic DNA samples were amplified in a 10 µl volume of 20 mM Tris-HCl pH 8.4, 50 mM KCl, 0.125 mM dNTPs containing 50 ng genomic DNA, 1 unit Taq DNA polymerase (Gibco-BRL, Rockville, MD, USA), 3.2 pmoles forward primer and 3.2 pmoles reverse primer. The following cycling profile was used to amplify the reactions: denaturation at 94 °C for 2 minutes, amplificaton cycle (50 cycles) at 94 °C for 1 minute, 56 °C for 1 minute, 72 °C for 3 minutes and extension at 72 °C for 10 minutes. CAPS primers were obtained from Research Genetics, Huntsville, AL, USA. Amplified products were digested with the appropriate restriction enzymes (as recommended by the manufacturer) and resolved by TBE-agarose gel electrophoresis and visualized by ethidium bromide staining.

Development of additional markers

BAC sequences between ATSO191 and DFR on chromosome V (available at http://www.kazusa.or.jp/arabi/) were compared with sequences of a Lansdberg *erecta* random genomic shot gun library available at the Institute for Genomic Research, (TIGR, Rockville, MD, USA) by multiple FASTA (Pearson and Lipman, 1988) analyses using do_fasta.pl (kindly provided by Dr. J White, Michigan State University). Sequence comparisons were used to identify polymorphic regions between Columbia loci and

Landsberg *erecta* loci and PCR primers were designed to amplify these regions. In situations were a fragment length polymorphism was not available, a restriction site was identified which would provide the required polymorphism. Using this approach, polymorphic markers for BACs MUL8, MKM21, K15E6 and MBB18 were developed.

Chapter 6

Identification of ORFs and complementation of *cvs1*

INTRODUCTION

Classically mutants were mapped and then cloned by chromosome walking. This process required the identification of several overlapping YACs (yeast artificial chromosomes) spanning the two closest RFLP markers (Restriction Fragment Length Polymorphism) on either side of the mutation. This was achieved by using the RFLP marker to probe a YAC library and identify YACs that contain the RFLP marker. Typically, the average distance between two RFLP markers was 400 kbp (Gibson and Somerville, 1992). This could be covered by about 5 overlapping YACs depending on the sizes of the inserts. Once the YACs bearing the RFLP markers were identified, the YAC ends were used once again to identify overlapping YACs until the mutation was spanned. The YAC ends would also be used as RFLP markers to identify segregants with chromosomal break points closer to the mutation. This process was repeated until a YAC was found which covered the mutant locus. The YAC was then subcloned into a library of cosmids for complementation analysis. Eventually the cosmid that complemented the mutation was sequenced and the gene identified.

With the sequencing of the *Arabidopsis* genome, the process of "walking YACs" has been replaced with *in silico* operations that identify overlapping BACs spanning the region that includes the mutation. Previously, mutation *cvs1* was mapped to a region covered by BACs: MBB18, MKD10, K15E6, MXF12, K3K3 and MUL8 (see Chapter 5). The analysis of the sequence of this region is presented.

RESULTS AND DISCUSSION

Mutant analysis

The mapping analysis revealed that the mutation *cvs1* lay between markers ATSO191 and DFR and more specifically on a small contig of 6 BACs: MBB18, MKD10, K15E6, MXF12, K3K3 and MUL8. Since the polymorphism markers for MBB18 and MUL8 revealed heterozygous bands in certain segregants, mutation *cvs1* should lie "south" of position 25021 on BAC MBB18 and "north" of position 48901 on BAC MUL8. The length of this region is about 417 kbp (Table 6.1).

Table 6.1 List of BACs that map close to locus *cvs1*

BACs	Size (bp)
MBB18	79537
MKD10	47460
K15E6	71736
MXF12	66237
K3K3	68889
MUL8	83450
Total Span	417309

BACs spanning the region between the SSLP marker for BAC MBB18 and BAC MUL8 have been listed along with their sizes (bp).

Complementation analysis

To verify that a mutation has been mapped, it is essential to complement the mutation by transforming the mutant with a transgene that encodes the wild type gene. The loss of the mutant phenotype in the primary transformant (in case of a recessive mutation) will indicate complementation. Such an event has to be verified by following the segregation pattern in the selfed progeny (T_2) to the primary transformant and ensure

that the loss of the mutant phenotype segregates with the complementing transgene in the T_2 lines (Dormann et al, 1999).

As BAC K15E6 (Liu et al, 1999) was available in a plant transformation vector, mutant *cvs1* has been transformed with clone K15E6 using *Agrobacterium*-mediated transformation (Bent et al, 1994). Primary transformants have been selected on hygromycin media. Results from this experiment are awaited. Complementation with BAC K15E6 should indicate if the mutation lies on that BAC. However, this step is not easy to accomplish. The transformation efficiency of plants is low. Once, this step has been accomplished, the BAC will be subcloned into cosmid vectors and these will be used for transforming mutant *cvs1*. Upon identifying a complementing cosmid clone, the corresponding region will be amplified out of *cvs1* genomic DNA by PCR amplification and sequenced. A comparison of the *cvs1* sequence to that of the wild type Columbia sequence should indicate the site of the mutation. The gene can then be cloned from a root specific cDNA library and this clone used to complement the mutant once again.

ORF Analysis

A parallel strategy is based on the sequence of the *Arabidopsis* genome. The sequencing of the genome has revealed a lot of information about genes present in the genome and their putative functions. Using this information, putative proteins encoded by open reading frames (ORF) as annotated by the Kazusa DNA Research Institute and The Institute for Genomic Research (TIGR) have been listed in Appendix C: Table C.1 (BAC MBB18), Table C.2 (BAC MKD10), Table C.3 (BAC K15E6), Table C.4 (BAC

MXF12), Table C.5 (BAC K3K3) and Table C.6 (BAC MUL8). A number of candidate genes appear to have potential roles in the secretory pathway (Table 6.2). This locus

Table 6.2 List of genes neighboring *cvs1*

Clone(s)	Similarity
K15E6.14,K15E6.16,K15E6.18, K15E6.20,K3K3.2,K3K3.4	Oxalate oxidase (germin protein)-like protein
MXF12.13,MXF12.14,MXF12.15, MXF12.16,MXF12.18,K3K3.6	Germin-like protein
K15E6.19	Similar to oxalate oxidase
K15E6.6	Disease resistance protein-like
MBB18.12,MBB18.14	Similar to heat shock transcription factor HSF 30
MBB18.7,MBB18.8	Myrosinase binding protein-like, similar to jasmonate induced protein
MKD10.5	Similar to Salt-inducible protein
MUL8.2	AHP3
MUL8.21	v-SNARE AtVTI1a

List of genes present on the BAC contig spanning MBB18 to MUL8 encoding proteins with a potential role in the secretory pathway. Annotation performed by the Kazusa DNA Research Institute (Chiba, Japan) and The Institute for Genomic Research, (TIGR, Rockville, MD, USA). See Appendix C for a detailed list of ORFs present on the BACs.

appears to encode many stress related proteins including germin-like proteins, HSF 30s, myrosinase-binding protein-like, salt-inducible proteins and signal transduction components. The only known secretory pathway protein in this locus is the soluble N-ethyl maleimide sensitive factor adaptor protein receptor (SNARE) AtVTI1a (Zheng et al, 1999).

Germins

Germins are a family proteins typically found in monocots such as wheat and barley. They are similar to desiccation tolerant vacuolar proteins such as vicilin and

legumin. They have been found to have oxalate oxidase and super oxide dismutase activity (Gane et al, 1998; Woo et al, 2000). They are induced during wheat seed maturation and germination and have been implicated in pathogen defense and resistance to abiotic stresses such as heat, salt, submergence and aluminium toxicity.

Myrosinase-binding protein

Myrosinase binding protein (MBP) is a protein that binds to myrosinase and activates the enzyme to hydrolyze glucosinolates. Glucosinolates are a group of thioglucosides found in crucifer species that are broken down into glucose, sulfate, and nitriles, isothiocyanates, thiocyanates, or epithionitriles depending on the species. These catabolites are highly toxic to insect and fungal pathogens (Geshi and Brandt, 1998). MBP is a vacuolar protein that has a putative ctVSS and is induced by wounding and by jasmonic acid.

Heat Shock induced factor 30 (HSF30)

HSF 30 belongs to a group of heat shock induced transcription factors first identified in tomato (Scharf et al, 1990). These have an N-terminal DNA-binding domain and form dimers or trimers by the interaction of their leucine zipper motifs. The C-terminal domains are not conserved and are implicated in specificity. Scharf et al (1990) demonstrated that the DNA-binding domain interacts with the heat stress consensus element.

AHP3

The sequence of ORF MUL8.2 is similar to AHP3. AHP3 is one of three proteins identified with a histidine phosphotransfer domain (Suzuki et al, 1998). This protein is implicated in the signal transduction pathway involving ETR1 (Chang et al, 1993), a plant sensor His kinase, and a response regulator such as ARR (Imamura et al, 1998). AHP3 is postulated to be a signal transducer in this pathway.

AtVTI1a

While the above described proteins have a role in stress responses, disease resistance and development, they do not have a clearly defined role in the secretory pathway. The only gene in this BAC contig with a known role in the plant secretory pathway is AtVTI1a (Sanderfoot and Raikhel, 1998; Zheng et al, 1999; Bassham et al, 2000). This protein is a homologue of the yeast SNARE Vti1p (Fischer von Mollard and Stevens, 1999; Tishgarten et al, 1999; Antonin et al, 2000; Sato et al, 2000). In plants, AtVTI1a is implicated in vesicle traffic. By subcellular fractionation studies and immunoelectron microscopy AtVTI1a has been found to be co-localized with the ssVSS receptor, AtELP, at the TGN and with the prevacuolar compartment syntaxin AtSYP21 (Zheng et al, 1999). Its role in protein transport has not been clearly demonstrated but its association with other proteins known to be involved in the secretory pathway such as AtELP and AtSYP21 makes it likely to have a role in the secretory pathway.

Since *cvs1* is a root specific mutation, it was interesting to speculate that this protein must be preferentially expressed in roots. Expression analyses by Zheng et al

(1999) revealed that AtVTI1a message was most abundant in roots. Also, the TIGR database shows that AtVTI1a is found in root cDNA libraries. Complementation analysis and RT-PCR analysis of *cvs1* root mRNA should reveal if the *cvs1* mutation corresponds to AtVTI1a.

ORF analysis summary

Proteins possessing ctVSS are predominantly ancillary vacuolar storage proteins (Marty, 1999). They have been found to be important against salt-stress, pathogen attack and insect attack. It is interesting to note that a locus potentially encoding a protein involved in the sorting of ctVSS-bearing proteins should also encode proteins involved in defense against salt and heat stress and pathogen attack. Considering that *Arabidopsis* chitinase is induced by ethylene (Samac et al, 1990), it is interesting that an ethylene signal transduction factor, AHP3, is encoded by this locus. The most promising candidate in this locus is AtVTI1a, a plant SNARE, which has been co-localized with other secretory pathway proteins.

CONCLUSION

The isolation of mutants impaired in ctVSS mediated protein transport to the vacuole has generated a collection of mutants, of which *cvs1* was characterized and mapped. The mutant was narrowed down to a contig of BACs that include heat and salt stress proteins, pathogen and insect defense proteins, an ethylene signal transduction factor and AtVTI1a, a SNARE implicated in plant vesicle formation. Further studies including complementation with BACs, followed by complementation with cosmids

derived from a complementing BAC should help in cloning *cvs1*. In parallel, sequencing the locus in the region of AtVTI1a should reveal if *cvs1* is really AtVTI1a. If *cvs1* is found to be AtVTI1a this will be the first major protein identified in the ctVSS-mediated transport pathway.

MATERIALS AND METHODS

Complementation Analysis

Clone K15E6 (Liu et al, 1999) was electroporated into *Agrobacterium tumefaciens* GV3101 pMP90 and selected for transformation. This clone was used in subsequent steps for plant transformation. Line ACWS (*cvs1* back-crossed and self fertilized, see Chapter 5) was germinated on vertical media plates and stained with X-GlcU. ACWS plants were transformed by vacuum infiltration with K15E6 (Agro) according to Bent et al., (1994). Seeds (K15E6-T1) were harvested and selected by germination on media containing 12.5 mg/l hygromycin and 250 mg/L carbenicillin.

Chapter 7

Conclusions and Future

CONCLUSIONS

Vacuolar proteins have been the focus of plant biology since the 1970s, as they perform a myriad of functions essential to the plant cell. A major function of vacuolar proteins is plant defense: defense-related macromolecules are stored in the vacuole to be released upon wounding or pathogen attack. A number of these defense-related proteins are ctVSS-bearing proteins. These proteins appear to be induced in the aerial tissues of plants during wounding or pathogen attack. However, in root tissues, these are constitutively expressed. This phenomenon might be related to the fact that roots are continuously exposed to abiotic stresses in the course of developing in the substratum.

Vacuolar proteins are synthesized on the rough endoplasmic reticulum and are co-translationally inserted into the ER lumen. By vesicle-mediated transport, these proteins are thought to exit the ER and reach the *cis*-Golgi. Eventually, via vesicle-mediated transport, these proteins are delivered to the *trans*-Golgi network. In this organelle, vacuolar proteins are sorted away from other secretory proteins and packaged into vesicles. The cargos of these vesicles reach the vacuole via a prevacuolar compartment (Jiang and Rogers, 1998; Miller et al, 1999). Proteins such as barley lectin and tobacco chitinase possess a C-terminally located vacuolar targeting signal (ctVSS). Dombrowski et al (1993) and Neuhaus et al (1994) demonstrated that the hydrophobic residues in the targeting signal are important and that two Gly residues added to the C-terminus of the ctVSS could abrogate vacuolar sorting. Although glycosylation per se is not necessary for vacuolar targeting, the addition of a glycosylation site at the C-terminus of the ctVSS abrogates vacuolar targeting. This property of the glycan is probably due to a steric hindrance preventing the ctVSS sorting machinery from correctly recognizing the ctVSS.

Roos-Runeberg et al (1994) have demonstrated that the ctVSS of barley lectin can be processed by the aspartic proteinase, phytepsin, *in vitro*. This activity has been postulated to occur in the vacuole. Pharmacologically, ctVSS-mediated sorting can be inhibited by the application of 33 μM of Wortmannin (an inhibitor of the mammalian phosphatidyl inositol-3 kinase). In plants, however, the *in vivo* activity of tobacco PI-3 kinase is not inhibited by 33 μM of wortmannin. The missorting of BL or tobacco chitinase in the presence of wortmannin has been suggested to be due to the inhibition of phospholipid biosynthesis by the drug.

To summarize the above, a ctVSS is a C-terminally located hydrophobic targeting motif. This motif losses activity if two Gly residues are added at the C-terminus or a glycosylation site is introduced at the C-terminus. This pathway is sensitive to wortmannin. Gal and Raikhel (1994) demonstrated that the ctVSS pathway is unique to plants in that ctVSS bearing proteins are not targeted to the yeast vacuole.

These studies have revealed the properties of the ctVSS, but very little has been elucidated regarding proteins interacting with ctVSS-bearing proteins and the machinery involved in transporting these proteins to the vacuole. Using classical genetics, an attempt has been made to understand the processes involved in the transport of ctVSS-bearing proteins by developing a screen for mutants impaired in the transport of ctVSS-bearing proteins. This screen involved the development of a single locus line homozygous for a reporter construct encoding barley lectin and rat preputial β-glucuronidase with a ctVSS (Rat-GUS-ctVSS) from tobacco chitinase. This line was mutagenized and screened for mutants with a sorting defect. Using an assay designed to visually score for secreted Rat-GUS activity, several putative mutants were identified. These putative mutants were

allowed to self-fertilize and the progeny was scored for secretion. Most of the progeny were found to have a heritable phenotype. To differentiate between *cis*-mutations and real sorting defects, the localization of barley lectin was examined by immunoelectron microscopy using anti-wheat germ agglutinin antibodies. All the mutants analyzed accumulated barley lectin in the vacuole. In addition, the majority of the mutants analyzed accumulated a secreted pool of barley lectin in spaces between adjacent cells and along the wall in root sections. However in leaf sections, barley lectin was localized exclusively in the vacuole. These results indicated the identification of root specific mutants, which displayed a partial mislocalization of BL in the apoplast.

By studying the expression profile of Rat-GUS and BL, mutants were classified into "over-expressers" and "low-expressers". The over-expressers were not pursued further as the mislocalization observed could have arisen due to a saturation of the ctVSS transport pathway resulting in secretion. One of the low expressers, designated *cvs1*, was analyzed further and found to have a single gene recessive defect. *cvs1* was then mapped to a contig of BACs between AtSO191 and DFR on chromosome V. It has been narrowed down to a stretch of 6 overlapping BACs (MBB18, MKD10, K15E6, MXF12, K3K3 and MUL8) covering a span of 417 kbp. Preliminary complementation analysis has been endeavored with BAC K15E6 and the results are awaited. This contig of BACs encodes, salt and heat stress proteins, defense related proteins, ethylene signal transduction proteins and the SNARE AtVTI1a.

Taken together, a model has been proposed (Figure 7.1). Proteins bearing a ctVSS

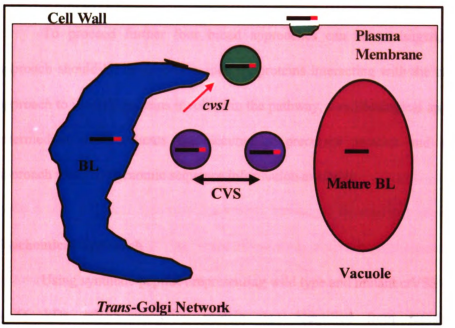

Figure 7.1 Model of the ctVSS sorting pathway
Proteins bearing a ctVSS (red) such as barley lectin (black) are sorted and packaged into vesicles (magenta). These vesicles eventually deliver their cargo to the vacuole (red) where the ctVSS is proteolytically cleaved by an enzyme such as phytepsin. The factors involved in this pathway may include several CVS gene products. This pathway can be blocked in mutants such as *cvs1* resulting in the secretion (green) of the protein to the extracellular space.

such as barley lectin, may be sorted at the *trans*-Golgi network and transported to the vacuole via a process involving several CVS gene products. In the vacuole, the ctVSS may be proteolytically cleaved by an enzyme such as phytepsin. This pathway can be partially blocked in mutants such as *cvs1* and *cvs2* resulting in the secretion of the protein to the extracellular space, or by drugs such as wortmannin. Although this model incorporates the implied functions of the mutants characterized, it cannot be confirmed until the genes encoding these factors are identified and analyzed.

THE FUTURE

To proceed further four broad approaches can be envisaged, a biochemical approach should be developed to identify proteins interacting with the ctVSS, a genetic approach to identify proteins involved in the pathway, a cell biological approach to purify intermediate compartments that accumulate precursor proteins and a bioinformatic approach using the genomic sequence information available.

Biochemical Approach

Using synthetic peptides representing wild type and mutant ctVSS of barley lectin, a 66 kDa ctVSS-interacting protein was identified from radish root tissues (Venkataraman and Raikhel, unpublished). To the best of our knowledge, this was the first demonstration of a ctVSS-specific interacting protein. Although the identity of the protein is not known, this approach has demonstrated that *in vitro* methods can be pursued to identify interacting factors. Boller and Vogeli (1984) have reported that tobacco chitinase can accumulate in vacuoles to the extent of 1% of the total leaf protein if the leaf is treated with ethylene. Mauch and Staehelin (1989) have demonstrated that in ethylene-treated leaves, chitinase is exclusively localized in the vacuole. Samac et al (1990) have demonstrated that the *Arabidopsis* basic chitinase is highly expressed in ethylene-treated leaves. While the identity of the interacting factor is not known, it is logical to assume that a tissue that over-expresses a cargo protein and does not saturate the pathway, should over-express the receptor as well. These evidences suggest that ethylene-treated leaves may be a suitable source of ctVSS-interacting proteins. The specificity of the protein can be verified by applying the protein preparation to an ssVSS

synthetic peptide matrix or synthetic peptide matrices representing known ctVSSs such as those of tobacco chitinase, kiwi actinidin and tobacco osmotin. Once a ctVSS-sorting receptor is characterized, further research should attempt to identify other factors involved in sorting, by examining factors that interact with this sorting receptor.

Based on the structure of AtELP (Ahmed et al, 1997), it can be postulated that the ctVSS-interacting factor may have a lumenal ctVSS-interacting domain, a transmembrane domain and a cytosolic domain for vesicle budding. One can then proceed to identify factors that interact with the cytosolic domain and use a proteomic approach to characterize these proteins.

Genetic Approach

The identification of proteins involved in the traffic of ctVSS-bearing proteins through biochemical means will be limited to factors involved early in post-Golgi traffic. To identify factors involved in later steps, a mutant approach may be useful. The approach described in chapters 2 through 6 was helpful in identifying two mutants *cvs1* and *cvs2*. The characterization of *cvs1* led to the identification of a locus potentially involved in ctVSS-mediated protein sorting. Considering the large number of proteins likely to be involved in protein traffic, this mutagenesis screen is clearly not saturated. By performing a larger screen on the same collection of M_2 seeds it will be possible to identify more loci involved in the pathway. Also by crossing mutants identified, it will be possible to determine if the mutations are allelic or novel. Considering some of the technical drawbacks in the screen performed with respect to the use of the double 35S promoter, it will be better to repeat a mutant screen with a population of mutagenized

seeds derived from a homozygous line of BL/Rat-GUS plants where each reporter is driven by a different promoter such as a single 35S promoter or a NOS promoter. Such an approach will limit silencing events due to promoter homology and result in a more robust screen.

To complement a genetic screen, a battery of endogenous reporters will have to be identified. A number of candidate ctVSS-bearing proteins have been identified in *Arabidopsis* such as AtOsmotin (Capelli et al, 1997), myrosinase-binding protein (Takechi et al, 1999), vacuolar invertases such as AtFruct3 and AtFruct4 (Haouazine-Takvorian et al, 1997), and a vacuolar chitinase, AtChib (Samac et al, 1990). However, the role of the putative ctVSS in these proteins has not been characterized. The characterization of these proteins will provide endogenous markers to follow the ctVSS pathway in *Arabidopsis thaliana*. These markers will prove invaluable in characterizing mutants identified by a genetic screen.

Cell biology

Currently, no intermediate compartment has been identified containing Golgi modified ctVSS-bearing precursor proteins. By electron microscopy, barley lectin has been localized in the Golgi stacks and in the vacuole. However, the intermediate compartments through which the protein must progress have not been identified. Using a combination of rate-zonal and isopycnic centrifugation steps, techniques will have to be optimized to purify membrane-bound organelles which contain precursor proteins. Certain drugs such as the ionophore, monensin, and nucleotide analogs may be useful in

trapping precursor in intermediate compartments or in slowing down the transport process.

Bioinformatics

With a large body of sequence information and the recent completion of the *Arabidopsis* genome (Walbot, 2000), an *in silico* design may be used to identify potential ctVSS pathway proteins. Currently about 30 % of the 25498 predicted gene products have not been assigned a function (The *Arabidopsis* Genome Initiative, 2000). These are either unique to plants or are homologous to unidentified gene products in other organisms. Considering that the ctVSS pathway is unique to plants, one might start with a subset of these unknown gene products that are unique to plants and identify proteins which appear to have a large lumenal domain, a transmembrane domain and a cytosolic domain. Such an analysis might lead to the identification of novel sorting receptors. Recently, Sanderfoot et al (2000) have identified a large number of SNARE homologues in *Arabidopsis* using sequence information available. Once candidate proteins are identified, they will have to be verified by examining at their *in vivo* localization and activity. Such an approach may be useful in the near future when prediction programs become more accurate.

The exploration of some of these avenues may one day complete the map that was initiated with the identification of the ctVSS in barley lectin and tobacco chitinase.

REFERENCES

Ahmed SU, Bar-Peled M, Raikhel NV (1997) Cloning and subcellular location of an *Arabidopsis* receptor-like protein that shares common features with protein-sorting receptors of eukaryotic cells. Plant Physiol **114:** 325-336

Ahmed SU, Rojo E, Kovaleva V, Venkataraman S, Dombrowski JE, Matsuoka K, Raikhel NV (2000) The Plant vacuolar sorting receptor AtELP is involved in transport of NH2-terminal propeptide-containing vacuolar proteins in *Arabidopsis thaliana.* J Cell Biol **149:** 1335-1344

Amidon WJ, Pfeil JE, Gal S (1999) Modification of luciferase to be a substrate for plant aspartic proteinase. Biochem J **343:** 425-33.

Andreeva AV, Kutuzov MA, Evans DE, Hawes CR (1998) The structure and function of the Golgi apparatus: a hundred years of questions. J Expt Bot **49:** 1281-1291

Andreeva AV, Zheng H, Saint-Jore CM, Kutuzov MA, Evans DE, Hawes CR (2000) Organization of transport from endoplasmic reticulum to Golgi in higher plants. Biochem Soc Trans **28:** 505-512

Antonin W, Riedel D, von Mollard GF (2000) The SNARE Vti1a-beta is localized to small synaptic vesicles and participates in a novel SNARE complex. J Neurosci **20:** 5724-5732

Balandin T, Castresana C (1997) Silencing of a beta-1,3-glucanase transgene is overcome during seed formation. Plant Mol Biol **34:** 125-137

Barlowe C (1998) COPII and selective export from the endoplasmic reticulum. Biochim Biophys Acta **1404:** 67-76

Bar-Peled M, Raikhel NV (1997) Characterization of AtSEC12 and AtSAR1: Proteins likely involved in endoplasmic reticulum and Golgi transport. Plant Physiol **114:** 315-24

Bassham DC, Gal S, Conceição AS, Raikhel NV (1995) An *Arabidopsis* syntaxin homologue isolated by functional complementation of a yeast pep12 mutant. Proc Natl Acad Sci USA **92:** 7262-6

Bassham DC, Sanderfoot AA, Kovaleva V, Zheng H, Raikhel NV (2000) AtVPS45 complex formation at the *trans*-Golgi network. Mol Biol Cell **11:** 2251–2265

Battey NH, James NC, Greenland AJ, Brownlee C (1999) Exocytosis and endocytosis. Plant Cell **11:** 643–659

Bell CJ, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis.* Genomics **19:** 137-144

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ (1994) RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science **265:** 1856-1860

Boller T, Vogeli U (1984) Vacuolar localization of ethylene-induced chitinase in bean leaves. Plant Physiol **74:** 442-444

Bonfanti L, Mironov AA Jr, Martinez-Menarguez JA, Martella O, Fusella A, Baldassarre M, Buccione R, Geuze HJ, Mironov AA, Luini A (1998) Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. Cell **95:** 993-1003

Brown RM (1969) Observations on the relationship of the Golgi apparatus to wall formation in the marine chrysophycean alga, *Pleurochrysis scherffelii.* J Cell Biol **41:** 109-123

Cabanes-Macheteau M, Fitchette-Laine AC, Loutelier-Bourhis C, Lange C, Vine ND, Ma JK, Lerouge P, Fayl L (1999) N-Glycosylation of a mouse IgG expressed in transgenic tobacco plants. Glycobiology **9:** 365-372

Capelli N, Diogon T, Greppin H, Simon P (1997) Isolation and characterization of a cDNA clone encoding an osmotin-like protein from *Arabidopsis thaliana.* Gene **191:** 51-56

Carroll AD, Moyen C, Kesteren PV, Tooke F, Battey NH and Brownlee C (1998) Annexins and GTP modulate exocytosis from maize root cap protoplasts. Plant Cell **10:** 1267–1276

Chang C, Kwok SF, Bleecker AB, Meyerowitz EM (1993) *Arabidopsis* ethylene-response gene ETR1: similarity of product to two-component regulators. Science **262:** 539-544

Choi S-B, Wang C, Muench DG, Ozawa K, Franceschi VR, Wu Y, Okita TW (2000) Messenger RNA targeting of rice seed storage proteins to specific ER subdomains. Nature **407:** 765-767

Clark GB, Roux SJ (1999) Isolation and characterization of two different *Arabidopsis* annexin cDNAs (Accession Nos. AF083913 and AF083914). Plant Physiol **120**: 339

Coupland G (1992) Transposon tagging In Koncz C, Chua N-H and Schell J (eds.) Methods in *Arabidopsis* Research. World Scientific, Singapore. pp 290-309.

Di Sansebastiano GP, Paris N, Marc-Martin S, Neuhaus J-M (1998) Specific accumulation of GFP in a non-acidic vacuolar compartment via a C-terminal propeptide-mediated sorting pathway. Plant J **15:** 449-457

Dombrowski JE (1995) Identification of components involved in plant vacuolar protein targeting. PhD Thesis, Michigan State University, East Lansing MI, USA.

Dombrowski JE, Raikhel NV (1996) Protein targeting to the plant vacuole - a historical perspective. Braz J Med Biol Res **29:** 413-430

Dombrowski JE, Schroeder MR, Bednarek SY, Raikhel NV (1993) Determination of the functional elements within the vacuolar targeting signal of barley lectin. Plant Cell **5:** 587-596

Dormann P, Balbo I, Benning C (1999) *Arabidopsis* galactolipid biosynthesis and lipid trafficking mediated by DGD1. Science **284:** 2181-2184

Dupree P, Sherrier DJ (1998) The plant Golgi apparatus. Biochim Biophys Acta. **1404:** 259-270

Fischer von Mollard G, Stevens TH (1999) The *Saccharomyces cerevisiae* v-SNARE Vti1p is required for multiple membrane transport pathways to the vacuole. Mol Biol Cell **10:** 1719-1732

Frigerio L, Virgilio de M, Prada A, Faoro F, Vitale A (1998) Sorting of phaseolin to the vacuole is saturable and requires a short C-terminal peptide. Plant Cell **10:** 1031-1042

Gal S, Raikhel NV (1994) A carboxy-terminal plant vacuolar targeting signal is not recognized by yeast. Plant J **6:** 235-240.

Galili G, Sengupta-Gopalan C, Ceriotti A (1998) The endoplasmic reticulum of plant cells and its role in protein maturation and biogenesis of oil bodies. Plant Mol Biol **38:** 1-29

Gane PJ, Dunwell JM, Warwicker J (1998) Modeling based on the structure of vicilins predicts a histidine cluster in the active site of oxalate oxidase. J Mol Evol **46:** 488-493

Geshi N, Brandt A (1998) Two jasmonate-inducible myrosinase-binding proteins from *Brassica napus* L. seedlings with homology to jacalin. Planta **204:** 295-304

Gibson SI, Somerville C (1992) Chromosome walking in *Arabidopsis thaliana* using yeast artificial chromosomes In Koncz C, Chua N-H and Schell J (eds.) Methods in *Arabidopsis* Research. World Scientific, Singapore. pp 119-143.

Guerineau F, Mullineaux P (1993) Plant transformation and expression vectors. In Croy RRD (ed.) Plant Molecular Biology Labfax, Academic Press, Oxford, UK. pp 121-147.

Haouazine-Takvorian N, Tymowska-Lalanne Z, Takvorian A, Tregear J, Lejeune B, Lecharny A, Kreis M (1997) Characterization of two members of the *Arabidopsis thaliana* gene family, At beta fruct3 and At beta fruct4, coding for vacuolar invertases. Gene **197:** 239-251

Hara-Nishimura I, Shimada T, Hatano K, Takeuchi Y, Nishimura M (1998) Transport of storage proteins to protein storage vacuoles is mediated by large precursor-accumulating vesicles. Plant Cell **10:** 825–836

Harlow E, Lane D (1988) Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. pp 471-510.

Haughn GW, Somerville C (1986) Sulfonylurea-resistant mutants of *Arabidopsis thaliana*. Mol Gen Genet **204:** 430-434

Herman EM, Larkins BA (1999) Protein storage bodies and vacuoles. Plant Cell **11:** 601-614

Houdebine L-M, Puissant C (1990) An improvement of the single-step method of RNA isolation by acid guanidinium Thiocyanate-phenol-chloroform extraction. BioTechniques **8:** 148-149

Imamura A, Hanaki N, Umeda H, Nakamura A, Suzuki T, Ueguchi C, Mizuno T (1998) Response regulators implicated in His-to-Asp phosphotransfer signaling in *Arabidopsis*. Proc Natl Acad Sci USA **95:** 2691-2696

Jauh GY, Fischer AM, Grimes HD, Ryan CA Jr, Rogers JC (1998) d-Tonoplast intrinsic protein defines unique plant vacuole functions. Proc Natl Acad Sci USA **95:** 12995–12999

Jefferson RA (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol Biol Reporter **5:** 387-405

Jiang L, Rogers JC (1998) Integral membrane protein sorting to vacuoles in plant cells: evidence for two pathways. J Cell Biol **143:** 1183-1199

Jiang L, Rogers JC (1999) Functional analysis of a Golgi-localized Kex2p-like protease in tobacco suspension culture cells. Plant J **18:** 23-32

Johnson KF, Chan W, Kornfeld S (1990) Cation-dependent mannose 6-phosphate receptor contains two internalization signals in its cytoplasmic domain. Proc Natl Acad Sci USA **87:** 10010-10014

Kirsch T, Paris N, Butler JM, Beevers L, Rogers JC (1994) Purification and initial characterization of a potential plant vacuolar targeting receptor. Proc Natl Acad Sci USA **91:** 3403-3407

Kirsch T, Saalbach G, Raikhel NV, Beevers L (1996) Interaction of a potential vacuolar targeting receptor with amino- and carboxyl-terminal targeting determinants. Plant Physiol **111:** 469-474

Kjemtrup S, Borkhsenious O, Raikhel NV, Chrispeels MJ (1995) Targeting and release of phytohemagglutinin from the roots of bean seedlings. Plant Physiol **109:** 603-610

Klionsky DJ, Herman PK, Emr SD (1990) The fungal vacuole: composition, function, and biogenesis. Microbiol Rev **54:** 266-292

Koide Y, Hirano H, Matsouka K, Nakamura K (1997) The N-terminal propeptide of the precursor to sporamin acts as a vacuole-targeting signal even at the C-terminus of the mature part in tobacco cells. Plant Physiol **114:** 863-870

Koncz C, Schell J, Redei GP (1992) T-DNA transformation and insertion mutagenesis In Koncz C, Chua N-H and Schell J (eds.) Methods in *Arabidopsis* Research. World Scientific, Singapore. pp 224-273.

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers Plant J **4:** 403-410

Koornneef M, Stam P (1992) Genetic Analysis. In Koncz C, Chua N-H and Schell J (eds.) Methods in *Arabidopsis* research, World Scientific Publishing, Singapore. pp 83-99

Lawrence MC, Izard T, Beuchat M, Blagrove RJ, Colman PM (1994) Structure of phaseolin at 2.2 A resolution: Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. J Mol Biol **238:** 748-776

Lerouge P, Cabanes-Macheteau M, Rayon C, Fischette-Laine AC, Gomord V, Faye L (1998) N-glycoprotein biosynthesis in plants: recent developments and future trends. Plant Mol Biol **38:** 31-48

Liu YG, Shirano Y, Fukaki H, Yanai Y, Tasaka M, Tabata S, Shibata D (1999) Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. Proc Natl Acad Sci USA **96:** 6535-6540

Marsh JL, Erfle M, Wykes EJ (1984) The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation Gene **32:** 481-485

Marty F (1999) Plant Vacuoles. Plant Cell. **11:** 587-599

Mason HS, Haq TA, Clements JD, Arntzen CJ (1998) Edible vaccine protects mice against *Escheria coli* heat-labile enterotoxin (LT): potatoes expressing a synthetic LT-B gene. Vaccine **16:** 1336-1343

Matsuoka K (2000) C-terminal propeptides and vacuolar sorting by BP-80-type proteins: not all C-terminal propeptides are equal. Plant Cell. **12:** 181-182

Matsuoka K, Bassham DC, Nakamura K, Raikhel NV (1995) Different sensitivity to wortmannin of two vacuolar sorting signals indicates the presence of distinct sorting machineries in tobacco cells. J Cell Biol **130:** 1307-1318

Matsuoka K, Neuhaus J-M (1999) Cis-element of protein transport to the plant vacuoles. J Expt Bot **50:** 165-174

Mauch F, Staehelin LA (1989) Functional implications of the subcellular localization of ethylene-induced chitinase and β-1-3 glucanase in bean leaves. Plant Cell **1:** 447-457

Miller EA, Lee MCS, Anderson MA (1999) Identification and characterization of a prevacuolar compartment in stigmas of *Nicotiana alata*. Plant Cell **11:** 1499-1508

Movafeghi A, Happel N, Pimpl P, Tai GH, Robinson DG (1999) *Arabidopsis* Sec21p and Sec23p Homologs. Probable coat proteins of plant COP-coated vesicles. Plant Physiol **119:** 1437–1445

Neuhaus J-M, Pietrzak M, Boller T (1994) Mutational analysis of the C-terminal vacuolar targeting peptide of tobacco chitinase: low specificity of the sorting system and gradual transition between intracellular retention and secretion into the extracellular space. Plant J **5:** 45-54

Neuhaus J-M, Pietrzak M, Freydl E, Boller T (1995) Expression and correct localization of animal secretory proteins in plant cells. J Cell Biochem **146:** 19A

Neuhaus J-M, Rogers JC (1998) Sorting of proteins to vacuoles in plant cells. Plant Mol Biol **38:** 127-144

Neuhaus J-M, Sticher L, Meins F Jr, Boller T (1991) A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. Proc Natl Acad Sci USA **88:** 10362-10366

Nishimura Y, Rosenfeld MG, Kreibich G, Gubler U, Sabatini DD, Adesnik M, Andy R (1986) Nucleotide sequence of rat preputial gland beta-glucuronidase cDNA and in vitro insertion of its encoded polypeptide into microsomal membranes. Proc Natl Acad Sci USA **83:** 7292-7296

Okita TW, Rogers JC (1996) Compartmentation of proteins in the endomembrane system of plant cells. Annu. Rev. Plant Physiol. Plant Mol Biol **47:** 327–350

Palauqui JC, Elmayan T, de Borne FD, Crete P, Charles C, Vaucheret H (1996) Frequencies, timing, and spatial patterns of co-suppression of nitrate reductase and nitrite reductase in transgenic tobacco plants. Plant Physiol **112:** 1447-1450

Paris N, Stanley CM, Jones RL, Rogers JC (1996) Plant cells contain two functionally distinct vacuolar compartments. Cell 85:563-572

Park YD, Papp I, Moscone EA, Iglesias VA, Vaucheret H, Matzke AJ, Matzke MA (1996) Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity. Plant J **9:** 183-194

Paul W, Amiss J, Try R, Praekelt U, Scott R, Smith H (1995) Correct processing of the kiwifruit protease actinidin in transgenic tobacco requires the presence of the C-terminal propeptide. Plant Physiol **108:** 261-268

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA **85:** 2444-2448

Pimpl P, Movafeghi A, Coughlan S, Denecke J, Hillmer S, Robinson DG (2000) In situ localization and in vitro induction of plant COPI-coated vesicles. Plant Cell **12:** 2219–2235

Raventos D, Meier C, Mattsson O, Jensen AB, Mundy J (2000) Fusion genetic analysis of gibberellin signaling mutants. Plant J **22:** 427-438

Raymond CK, Howald-Stevenson I, Vater CA, Stevens TH (1992) Morphological classification of the yeast vacuolar protein sorting mutants: evidence for a prevacuolar compartment in class E *vps* mutants. Mol Biol Cell **3:** 1389-1402

Redei GP, Koncz C (1992) Classical mutagenesis. In Koncz C, Chua N-H and Schell J (eds.) Methods in *Arabidopsis* Research. World Scientific, Singapore. pp 16-82

Richter LJ, Thanavala Y, Arntzen CJ, Mason HS (2000) Production of hepatitis B surface antigen in transgenic plants for oral immunization. Nat Biotech **18:** 1167-1171.

Runeberg-Roos P, Kervinen J, Kovaleva V, Raikhel NV, Gal S. (1994) The aspartic proteinase of barley is a vacuolar enzyme that processes probarley lectin in vitro. Plant Physiol **105:** 321-329.

Samac DA, Hironaka CM, Yallaly PE, Shah DM (1990) Isolation and characterization of the genes encoding basic and acidic chitinase in *Arabidopsis thaliana* Plant Physiol **93:** 907-914

Sanderfoot AA, Ahmed SU, Marty-Mazars D, Rapoport I, Kirchhausen T, Marty F, Raikhel NV (1998) A putative vacuolar cargo receptor partially colocalizes with AtPEP12p on a prevacuolar compartment in *Arabidopsis* roots. Proc Natl Acad Sci USA **95:** 9920-9925

Sanderfoot AA, Assaad FF, Raikhel NV (2000) The *Arabidopsis* genome. An abundance of soluble N-ethylmaleimide-sensitive factor adaptor protein receptors. Plant Physiol **124:** 1558-1569

Sanderfoot AA, Raikhel NV (1998) The specificity of vesicle trafficking: coat proteins and SNAREs. Plant Cell **11:** 629-642

Sato F, Koiwa H, Sakai Y, Kato N, Yamada Y (1995) Synthesis and secretion of tobacco neutral PR-5 protein by transgenic tobacco and yeast. Biochem-Biophys-Res-Comm **211:** 909-913

Sato TK, Rehling P, Peterson MR, Emr SD (2000) Class C Vps protein complex regulates vacuolar SNARE pairing and is required for vesicle docking/fusion. Mol Cell **6:** 661-671

Scharf KD, Rose S, Zott W, Schoffl F, Nover L, Schoff F (1990) Three tomato genes code for heat stress transcription factors with a region of remarkable homology to the DNA-binding domain of the yeast HSF. EMBO J **9:** 4495-4501

Shimada, T., Kuroyanagi M, Nishimura M, Hara-Nishimura I (1997) A pumpkin 72-kDa membrane protein of precursor-accumulating vesicles has characteristics of a vacuolar sorting receptor. Plant Cell Physiol **38:** 1414-1420

Sutter JU, Homann U, Thiel G (2000) Ca^{2+}-stimulated exocytosis in maize coleoptile cells. Plant Cell **12:** 1127–1136

Suzuki T, Imamura A, Ueguchi C, Mizuno T (1998) Histidine-containing phosphotransfer (HPt) signal transducers implicated in His-to-Asp phosphorelay in *Arabidopsis*. Plant Cell Physiol **39:**1258-1268

Swanson SJ, Bethke PC, Jones RL (1998) Barley aleurone cells contain two types of vacuoles: characterization of lytic organelles by use of fluorescent probes. Plant Cell **10:** 685-698

Takechi K, Sakamoto W, Utsugi S, Murata M, Motoyoshi F (1999) Characterization of flower-specific gene encoding a putative myrosinase binding protein in *Arabidopsis thaliana*. Plant Cell Physiol **40:** 1287-1296

Takeuchi M, Ueda T, Sato K, Abe H, Nagata T, Nakano A (2000) A dominant negative mutant of sar1 GTPase inhibits protein transport from the endoplasmic reticulum to the Golgi apparatus in tobacco and *Arabidopsis* cultured cells. Plant J **23:** 517-525

The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796-815

Tishgarten T, Yin FF, Faucher KM, Dluhy RA, Grant TR, von Mollard GF, Stevens TH, Lipscomb LA (1999) Structures of yeast vesicle trafficking proteins. Protein Sci **8:** 2465-2473

Trull MC, Deikman J (1998) An *Arabidopsis* mutant missing one acid phosphatase isoform. Planta **206:** 544-550

van Houdt H, Ingelbrecht I, van Montagu M, Depicker A (1997) Post-transcriptional silencing of a neomycin phosphotransferase II transgene correlates with the accumulation of unproductive RNAs and with increased cytosine methylation of 3' flanking regions. Plant J **12:** 379-392

Venkataraman S, Raikhel NV (1998) Mechanisms for the transport of soluble proteins to the vacuole in plants. In Lo Schiavo F, Last RL, Morelli G and Raikhel NV (eds,) Cellular Integration of Signalling Pathways in Plant Development, NATO ASI Series, Vol. H 104. Springer-Verlag, Berlin, pp 1-9.

Vitale A, Denecke J (1999) The endoplasmic reticulum - gateway of the secretory pathway. Plant Cell **11:** 615–628

Vitale A, Raikhel NV (1999) What do proteins need to reach different vacuoles? Trends Plant Sci **4:** 149-155

von Heijne G (1985) Signal sequences: the limits of variation. J Mol Biol **184:** 99-105

Walbot V (2000) A green chapter in the book of life. Nature **408:** 794-795

Woo EJ, Dunwell JM, Goodenough PW, Marvier AC, Pickersgill RW (2000) Germin is a manganese containing homohexamer with oxalate oxidase and superoxide dismutase activities. Nat Struct Biol **7:** 1036-1040

Wozniak CA, Owens LD (1994) Native beta-glucuronidase activity in sugar-beet (*Beta vulgaris*). Physiol Plantarum **90:** 763-771

Zarembinski TI, Theologis A (1993) Anaerobiosis and plant growth hormones induce two genes encoding 1-aminocyclopropane-1-carboxylate synthase in rice (Oryza sativa L.). Mol Biol Cell. **4:** 363-373

Zheng H, von Mollard GF, Kovaleva V, Stevens TH, Raikhel NV (1999) The plant vesicle-associated SNARE AtVTI1a likely mediates vesicle transport from the trans-Golgi network to the prevacuolar compartment. Mol Biol Cell **10:** 2251-2264

Ziegelhoffer EC, Medrano LJ, Meyerowitz EM (2000) Cloning of the *Arabidopsis* WIGGUM gene identifies a role for farnesylation in meristem development. Proc Natl Acad Sci USA **97:** 7633-7638

APPENDIX A

Mapping and Linkage Analysis made simple: A web-based approach
http://www.msu.edu/~venkata1/sslpfull.htm

ABSTRACT

With the establishment of molecular genetics resources to study basic physiological and developmental functions in plants, mutant analysis has become extremely important in understanding the role of a particular gene product. Such studies involve the design of comprehensive screens to identify mutants defective in a given function. Once identified, the gene aberrant in function is to be cloned, typically by a map-based approach. Such an approach requires the evaluation and interpretation of recombination frequencies between the mutant locus and linked or unlinked markers. A JavaScript program has been developed to provide scientists all the tools required to calculate recombination frequencies and mapping distances along with appropriate statistics. This program is optimized to operate on any operating system and in a networked manner.

INTRODUCTION

In order to determine the genes that code for properties of organisms, mutants aberrant in a given function are convenient approaches to identify the key players in these processes. Once a mutant is identified, approaches have to be taken to clone the gene aberrant in the mutant. The model dicotyledonous plant, *Arabidopsis thaliana* has been a model genetic system providing a convenient system to study plant biology.

Historically, mutants have been developed in *Arabidopsis thaliana* ecotype Columbia (Col) or Landsberg *erecta* (L*er*). The mutants are then crossed into L*er* or Col respectively for mapping purposes. In the case of non-lethal recessive phenotypes, the second filial generation of the cross (F2) is screened for the mutant phenotype. Individual plants expressing the mutant phenotype are used for mapping the mutation using molecular markers such as CAPS or SSLPs among others.

RESULTS

In order to map mutations using the principle classical genetics with molecular markers, one often has to calculate various parameters to evaluate the performance of a particular molecular marker in determining the location of the mutation. Calculations of recombination frequency (RF) are extremely important in determining the distance between a marker and the locus of interest. The RF needs to be tested, its deviation from an unlinked RF evaluated and finally the mapping distance calculated. In the description that follows we have considered the case of a recessive mutation identified in a *Arabidopsis thaliana* ecotype Columbia background and crossed into *Arabidopsis thaliana* ecotype Landsberg *erecta* to develop mapping populations.

Recombination Frequency (RF)

The recombination frequency (RF) is the proportion of recombinants chromosomes observed among the total number chromosomes present. This can be equated to:

$RF = (Het + 2\ Ler)/(2 * (Col + Het + Ler))$

where Het = number of mapping individuals heterozygous for a particular marker, Col = number of mapping individuals homozygous for the Columbia allele for a particular marker and Ler = number of mapping individuals homozygous for the Lansdberg *erecta* allele for a particular marker

Standard Deviation of RF

The standard deviation of the RF value can be of interest to give an indication of the spread of the results. This has been evaluated using $((RF * (1 - RF))/T)^{1/2}$

where T = Col + Het + Ler

Chi-Square test of goodness of fit of RF

This test allows one to determine whether the observed frequency of marker alleles deviates from an unlinked situation of 1:2:1 (Col: Het: Ler). Since the Chi-Square test is performed on the Observed results, in this case the frequency of individuals with Col, Het, Ler, genotypes at a given locus and the

χ^2 statistic = $\Sigma (O - E)^2/e$ with n-1 degrees of freedom (n = 3 in this case).

The χ^2 statistic can be equated to $(4\ Col^2 + 2\ Het^2 + 4\ Ler^2 - T^2) / T$ where T = Col + Ler + Het.with 2 degrees of freedom.

For convenience, the observed χ^2 is compared to theoritical χ^2 at $\alpha = 0.1, 0.05$ and 0.01 (χ^2 values of 4.605, 5.991 and 9.210 respectively). A null hypothesis of no significant deviation from an unlinked condition is tested under that above conditions. One can conclude that the larger the χ^2 value the greater the deviation of the observed frequency from an unlinked condition and hence greater indication of linkage between the test locus and the mutation.

Mapping Distance

During meiosis, a large number of cross-overs occur which can be classified into different classes such as single, 2, 3, ... , n cross-overs per meiosis within a particular region. Since the number of possible cross-overs is very large over the entire chromosome but small within the region of interest i.e. between loci considered, the probability of a particular number of cross-overs occurring within a particular distance on a chromosome follows a Poisson distribution. .....(1)

Since non-zero number of cross-overs per meiosis, results in a 50 % rate of recombinants (Stahl, 1969 for review), and a zero cross-over meiosis produces a 0 percent rate of recombinants, it follows that the probability of obtaining at least one cross-over within a particular distance, is the summation of the probabilities of obtaining non-zero number of cross-overs. (2)

This is equivalent to 1- p(0) where p(0) is the probability of obtaining 0 cross-overs within a particular distance. Since only half of those cross-overs will be recombinant, the frequency of recombination (RF) = ½(1-P(0))

An event following the Poisson distribution can be described by $p(i) = (e^{-x} x^i)/(i!)$ where p(i) is the probability of occurrence of i events and x the average number of events occurring in the given experiment.

The probability of 0 cross-overs occurring within the 2 loci follows $p(i) = (e^{-x} x^i)/(i!)$ where in this case i = 0 and x is the average number of Cross-overs between the 2 loci.

With i = 0, $p(0) = e^{-x}$

Subsequently p(at least 1 non-zero Cross-over) = $1 - e^{-x}$ and p(recombinants) = ½x p(at least 1 non-zero Cross-over) = ½$(1 - e^{-x})$.

$e^{-x} = 1 - x + x^2 - x^3/3! + x^4/4! + ...(-x)^n/n!$ Ad infinitum

when $x << 1$, all the exponents of x tend to 0 and can be neglected.

Thus when $x << 1$, $e^{-x} = 1 - x$

Therefore p(recombinants) = ½$(1 - e^{-x})$ = ½$(1 - (1 - x))$ = x/2.

If the "distance" between the two loci is defined as *d*, such that *d* is proportional to x and equal to p(recombinants) (in the case of closely linked loci). Then

d = x/2 or x = 2d

Therefore p(recombinants) = ½$(1 - e^{-x})$ = ½$(1 - e^{-2d})$.

p(recombinants) in reality is the recombination frequency or RF.

Thus RF = ½$(1 - e^{-2d})$

This can be rearranged to calculate the mapping distance d

d = - ½Ln (1 - 2RF) with d expressed in Morgans.

A more convenient unit is the cM = 1/100 Morgans.

Again d = -100/2 Ln (1 - 2RF) = -50 Ln (1 - 2RF), where d is expressed in centiMorgans (cM).

To summarize, the recombination frequency RF = ½(1- $e^{-d/50}$) where d is the map distance (in cM) between the two loci with a recombination frequency of RF and e is the base of the natural logarithm and d = -50 Ln (1-2RF) in centiMorgans.

With a large number of samples and with large number of measurements, these calculations can be cumbersome. We have developed a java script code to perform these calculations and test the frequencies obtained for all the above parameters and present them in a fashion useful to the user. Using this code it is possible analyze one's mapping data on any operating system (Macintosh or IBM) and using Netscape, or Internet Explorer the two standard World Wide Web browsers.

This program allows one to go directly from raw mapping data such as counts of occurrences of Col, Het and Ler allele profiles on electrophoresis gel as in the case of CAPS and SSLP markers to obtain the final results such as distance of the mutant locus from marker loci.

To complement these analytical tools and exploit the networked resources, a number of world-wide web-based tools have also been "linked" to the program. This program is available at http://www.msu.edu/~venkata1/sslpfull.htm

Recently Raventos et al (2000) reported the use of this program for analyzing linkage data.

REFERENCES

Haldane JBS (1919) J. Genet 8: 299-309.

Koopmans LH (1987) Introduction to contemporary statistical methods (2nd ed).

Koornneef M, Stam P (1992) Genetic Analysis. In Koncz C, Chua N-H and Schell J (1992) Methods in *Arabidopsis* research, World Scientific Publishing.

Stahl FW (1969). The Mechanics of Inheritance. 2nd edition, Englewood Cliffs, NJ: Prentice-Hall. A short introduction to genetics, including some advanced material presented with a novel approach.

Suzuki DT, Griffiths AJF, Miller JH, Lewontin RC (1989) An introduction to genetic analysis. (4th ed)

```
 HTML code for sslpfull.htm
<.HTML><.HEAD><.TITLE>PCReactor <./TITLE>
<.SCRIPT>function PCR (Intro, NR,Prim,B10,prog)
{
if (B10 == "Sep0")
{
B10X = NR * 2
B10Xreal = Math.ceil (B10X)
K2O = "10 X PCR Buffer = " + B10Xreal + " ul      "
MG = NR * 0.6
MGreal = Math.ceil (MG)
K3O = "50 mM MgCl2 = " + MGreal + " ul    "
taq = NR * 0.4
Primer = NR * 0.25
dntp = NR * 0.25
dd = (NR * 19) - ( B10X + MG + taq + Primer + Primer + dntp)
ddreal = Math.ceil(dd)
}
if (B10 == "Sep")
{
B10X = NR * 2
B10Xreal = Math.ceil (B10X)
K2O = "10 X PCR Buffer = " + B10Xreal + " ul      "
MG = NR * 1.2
MGreal = Math.ceil (MG)
K3O = "50 mM MgCl2 = " + MGreal + " ul    "
taq = NR * 0.4
Primer = NR * 0.25
dntp = NR * 0.25
dd = (NR * 19) - ( B10X + MG + taq + Primer + Primer + dntp)
ddreal = Math.ceil(dd)
}
if (B10 == "Combo")
{
B10X = NR * 3.2
B10Xreal = Math.ceil (B10X)
K2O = "6.25 X PCR-Mg Buffer = " + B10Xreal + " ul  "
K3O = "                    "
taq = NR * 0.4
Primer = NR * 0.25
dntp = NR * 0.25
dd = (NR * 19) - ( B10X + taq + Primer + Primer + dntp)
ddreal = Math.ceil(dd)
}
if (prog == "SVSSP4")
```

```
{
prog = "              | Cycling Program = " + prog
S1= "|  Step 1: 94 C for 7 min"
S2= "     | Step 2: 94 C for 30 sec"
S3= "  |  Step 3: 55 C for 15 sec"
S4= "     | Step 4: 72 C for 30 sec"
S5= "     | Step 5: Back to Step 2 40 times"
S6= "       | Step 6: 72 C for 7 min"
S7= "      | Step 7: 4 C forever"
}
if (prog == "SVSSP3")
{
prog = "              | Cycling Program = " + prog
S1= "|  Step 1: 94 C for 2 min"
S2= "     | Step 2: 94 C for 30 sec"
S3= "  |  Step 3: 55 C for 15 sec"
S4= "     | Step 4: 72 C for 30 sec"
S5= "     | Step 5: Back to Step 2 40 times"
S6= "       | Step 6: 72 C for 7 min"
S7= "      | Step 7: 4 C forever"
}
if (prog == "SVSSP5")
{
prog = "              | Cycling Program = " + prog
S1= "|  Step 1: 94 C for 5 min"
S2= "     | Step 2: 94 C for 1 min"
S3= "  |  Step 3: 55 C for 1 min"
S4= "     | Step 4: 72 C for 1 min"
S5= "     | Step 5: Back to Step 2 40 times"
S6= "       | Step 6: 72 C for 7 min"
S7= "      | Step 7: 4 C forever"
}
TexSep = Intro + "\n\nName of Primer = " + Prim +
"       | Number of reactions = " + NR + "\nWater = " + ddreal
+ " ul" + prog + "\n" + K2O  + S1 + "\n" + K3O + S2 +
"\ndNTPs (10 mM each) = " + dntp + " ul" + S3 + "\n" +
Prim + " Forward = " + Primer + " ul" + S4 + "\n" + Prim
+ " Reverse = " + Primer + " ul" + S5 + "\nTaq Enzyme = "
 + taq + " ul" + S6 + "\n-----> 19 ul per RXN" + S7 +
"\n+ 1ul Template DNA"

document.PCRForm.TextBox.value = TexSep
}
function FindMarker (marker) {
fullQuery = "http://genome-www3.stanford.edu/cgi-bin/Webdriver
```

```
?MIval=atdb_locus_max&name=" + escape (marker)
location.href=fullQuery
}
function MasMix () {
location.href = "http://www.msu.edu/~venkata1/sslpmas.htm"
}
function RFcM () {
location.href = "http://www.msu.edu/~venkata1/cminvort.htm"
}
function labl () {
location.href = "http://www.msu.edu/~venkata1/labsslst.htm"
}
function Eckers () {
location.href = "http://genome.bio.upenn.edu/SSLP_info/SSLP.html"
}
function EckerMaps () {
location.href = "http://genome.bio.upenn.edu/SSLP_info/SSLP_map.html"
}
function AraSer () {
location.href = "http://genome-www.stanford.edu/Arabidopsis/AT-arabgen.html"
}
function LocatMarker (Smarker)
{
fulQuery = "http://genome-www3.stanford.edu/cgi-bin/AtDB/RImap?locus=" + escape
(Smarker)
location.href=fulQuery
}
function Chisqhet(Intro,C,H,L,Al)
{
C = C * 1
H = H * 1
L = L * 1
T = C + H + L
CHI = ((4 * C * C) + (2 * H * H) + (4 * L * L) - (T * T))/ T
if (Al == 0.1)
        {CHITAB = 4.605}
if (Al == 0.05)
        {CHITAB = 5.991}
if (Al == 0.01)
        {CHITAB = 9.210}
if (CHI >= CHITAB)
        {R = "Failed"}
if (CHI <. CHITAB)
```

```
    {R = "Passed"}
CHISHO = (Math.floor (1000 * CHI))/1000
document.PCRForm.Hourf.value = Intro + "\nObserved Segregation
 Ratio of Col to Het to Ler = " + C + ":" + H + ":" + L +
"\nExpected ratio of Col : Het : Ler if unlinked is = " + T/4 + ":"
+ T/2 + ":" + T/4 + "\n The Chisquare value is " + CHISHO + "\nand has "
 + R + " the null hypothesis at alpha = " + Al
}
function CalculateLink (Intro, HC, Het, HL) {
N = (HC * 1) + (HL * 1) + (Het * 1)
RF = ((Het * 1) + (HL * 2))/(2 * N)
M1 = (Math.log(1- (2*RF))) * (-50)
Mrounded = (Math.round (100 * M1))/100
if (Mrounded <. 1)
{Mrounded = "0" + Mrounded}
if (RF >= 0.5)
    {Mrounded = "Undefined & Unlinked"}
if (RF >= 0.3)
        {Linko = "Not Linked"}
if (RF <. 0.3)
        {Linko = "Linked"}
SD = Math.sqrt((RF * (1 - RF))/N)
R2F = RF - SD
M2 = (Math.log(1- (2*R2F))) * (-50)
M2R = (Math.round (100 * M2))/100
if (M2R <. 1)
{M2R = "0" + M2R}
SDcM = SD
SDcMR = (Math.round (100 * SDcM))/100
if (SDcMR <. 1 )
        {SDcMR = "0" + SDcMR}
RF = (Math.round (100 * RF))/100
if (RF <. 1 )
        {RF = "0" + RF}
document.PCRForm.result0.value = RF
document.PCRForm.result1.value = Mrounded
document.PCRForm.result2.value = Linko
document.PCRForm.result3.value = SDcMR
C = HC * 1
H = Het * 1
L = HL * 1
T = C + H + L
CHI = ((4 * C * C) + (2 * H * H) + (4 * L * L) - (T * T))/ T
if (CHI >= 9.21)
        {R = "failed the null hypothesis at alpha = 0.01"}
```

```
if ((CHI <. 9.21) && (CHI >= 5.991))
      {R = "failed the null hypothesis at alpha = 0.05"}
if ((CHI <. 5.991) && (CHI >= 4.605))
      {R = "failed the null hypothesis at alpha = 0.1"}
if (CHI <. 4.605)
      {R = "passed the null hypothesis at alpha = 0.1"}
CHISHOT = (Math.floor (1000 * CHI))/1000
CHISHO = CHISHOT
if (CHISHOT <. 1)
CHISHO = "0" + CHISHOT
document.PCRForm.Hourf.value = Intro + "\n\nObserved Segregation Ratio of
 Col to Het to Ler = " + C + ":" + H + ":" + L +
"\nExpected ratio of Col : Het : Ler if unlinked is = " + T/4 + ":"
+ T/2 + ":" + T/4 + "\nThe Chisquare value is " + CHISHO + "\nand has "
 + R + "." +
"\nThe Recombination frequency is = " + RF + " +/- " + SDcMR + "." +
"\nThe mapping distance is = " + Mrounded + " cM." +
"\nThe lower limit (RF - SD) mapping distance is = " + M2R + " cM."
}
today = new Date ()
document.write ("<.H1>SSLP Package <./H1><.DT>This program estimates materials
 to be used for a PCR SSLP reaction and the analysis of the results<.DT>
<.FONT SIZE = 1> A freeware product of Iyer-Venk: Developed by Meera Iyer
 and Sridhar Venkataraman<./DT><./FONT><.HR>Date: " + today)
<./SCRIPT>
<.BODY>
<.FORM NAME = "PCRForm"><.FONT FACE = "tahoma" SIZE =2><.DT>Details of
the experiment <.INPUT TYPE = "text" NAME = "ibox" SIZE= 50> <.DT>Number of
standard SSLP - 20 ul Reactions <.INPUT TYPE = "text" NAME = "NumR" SIZE= 3>
 Name of Primer <.INPUT TYPE = "text" NAME = "Prim" SIZE= 8>
<.dt>Buffer and Mg
<.SELECT NAME= "Buf10Mg">
<.OPTION VALUE = "Sep0">Separate 10 X PCR buffer and MG (1.5 mM
final)<./OPTION>
<.OPTION VALUE = "Sep">Separate 10 X PCR buffer and MG (3 mM
final)<./OPTION>
<.OPTION VALUE = "Combo">Combo PCR buffer and MG (6.25 X and 18.75
mM)<./OPTION>
<./SELECT>
Cyling Program
<.SELECT NAME= "Prog">
<.OPTION VALUE = "SVSSP3">SVSSP3<./OPTION>
<.OPTION VALUE = "SVSSP4">SVSSP4<./OPTION>
<.OPTION VALUE = "SVSSP5">SVSSP5<./OPTION>
```

<./SELECT>
<.dt><.INPUT TYPE = "button" VALUE = "Polymerize" onClick = " PCR
(document.PCRForm.ibox.value, document.PCRForm.NumR.value,
document.PCRForm.Prim.value,document.PCRForm.Buf10Mg.options
[document.PCRForm.Buf10Mg.selectedIndex].value,document.PCRForm.Prog.options
[document.PCRForm.Prog.selectedIndex].value)"><./FONT>
<.INPUT TYPE = "button" VALUE = "Master Mixes" onClick = " MasMix ()"
target = "_new_window_"><./FONT>
<.INPUT TYPE = "button" VALUE = "Explain" onClick = " FindMarker
(document.PCRForm
.Prim.value)" target = "_new_window_"><./FONT><.INPUT TYPE = "button" VALUE
=
"Locate on RI Map" onClick = " LocatMarker (document.PCRForm.Prim.value)"
target = "_new_window_"><./FONT>
<.INPUT TYPE = "button" VALUE = "cM <.=> RF conversions" onClick = " RFcM ()"
target = "_new_window_"><./FONT>
<.dt><.INPUT TYPE = "button" VALUE = "GoTo Ecker's SSLP page" onClick = "
Eckers()
" target = "_new_window_"><./FONT>
<.INPUT TYPE = "button" VALUE = "Ecker's SSLP Map page" onClick = " EckerMaps
()
" target = "_new_window_"><./FONT>
<.INPUT TYPE = "button" VALUE = "Raikhel's SSLPs" onClick = " labl ()" target =
"_new_window_"><./FONT>
<.FONT SIZE =3><.dt><.BR><.textarea name="TextBox" rows=15
cols=70><./textarea><./BR><.DT>
<.p><.b><.font size = 3> Analysis of Data <./font><./b>
<.dt><.FONT SIZE =3>
Details of the experiment <.INPUT TYPE = "text" NAME = "box0" SIZE= 50><./dt>
<.DT>Number of homozygous lines in Columbia Ecotype <.INPUT TYPE = "text"
NAME = "HomCol" SIZE= 5> <./DT>
<.DT>Number of heterozygous lines <.INPUT TYPE = "text" NAME = "Het" SIZE=
15> <./DT>
<.DT>Number of homozygous lines in Lansberg Ecotype <.INPUT TYPE = "text"
NAME =
"HomLer" SIZE= 5> <./DT>
<.dt><.INPUT TYPE = "button" VALUE = "Evaluate" onClick = "CalculateLink
(document.PCRForm.box0.value,document.PCRForm.HomCol.value,
document.PCRForm.Het.value,
document.PCRForm.HomLer.value)">
<.P>Recombination Frequency (RF) <.INPUT TYPE = "text" NAME = "result0"
VALUE="" SIZE=15>
<.DT>Standard Deviation in "~RF" <.INPUT TYPE = "text" NAME = "result3"
VALUE="" SIZE=15> <./DT>
<.DT>Linked (<. 0.3 RF)/ Unlinked (> 0.3 RF) <.INPUT TYPE = "text" NAME =

```
 "result2" VALUE="" SIZE=15> <./DT>
<.DT>Mapping Distance (cM)  <.INPUT TYPE = "text" NAME = "result1" VALUE=""
SIZE=30> <./DT>
<.BR>
<.textarea name="Hourf" rows=12 cols=70><./textarea><.BR>
<.FONT FACE="Tahoma"><.P>You are visitor no.
<.IMG SRC="http://counter.digits.com/wc/-d/4/svsslp" ALIGN=middle
    WIDTH=60 HEIGHT=20 BORDER=0 HSPACE=4 VSPACE=2>
<.A HREF="http://www.digits.com/"><.FONT
SIZE=2>http://www.digits.com<./FONT><./A>
<.P> Should you find this program useful in your analysis,
please cite "Sridhar Venkataraman and Natasha V. Raikhel,
Arabidopsis Mapping Package using SSLP and CAPS data, 1999"
<.P><.A HREF="mailto:venkata1@pilot.msu.edu"><.FONT FACE="Times,Times New
Roman">Please send Comments to Sridhar Venkataraman<./A><./DT>
<.DT><.A HREF="http://www.msu.edu/~venkata1/index.htm"><.FONT FACE="Times,
Times New Roman">Back to the main page<./A><./DT><./FORM>
<.strong>A search of all messages contained in the BioSci Arabidopsis Genome
mailing or newsgroup.<./strong>
<.FORM ACTION="http://genome-www.stanford.edu/cgi-bin/AtDB/AT-arabgenbiosci
search.cgi" METHOD=POST>
<.TABLE><.TR><.TD><.INPUT NAME="search" VALUE="" size=50><./TD>
<.TD><.INPUT TYPE="image" SRC="http://genome-www.stanford.edu/Excite/pictures/
AT-search_button.gif" NAME="searchButton" HEIGHT=20 WIDTH=75 ALT="Search"
BORDER=0><./TD><./TR><./TABLE>
<.INPUT TYPE="hidden" NAME="sp" VALUE="sp">
<./FORM>
<.FORM NAME= "main" METHOD="POST" ACTION="http://www.williamstone.com/
primers/calculator/calculator.cgi">
Primer Evaluation at www.williamstone.com
 <.dt>   <.td><.INPUT TYPE="text" NAME="fprimer" SIZE=40><./td>
    <.td><.SELECT NAME="forient"><.OPTION SELECTED VALUE="0">5' to 3'
<.OPTION VALUE="1">3' to 5'<./SELECT><./td>
   <./tr><./table><./font><./td>
 <.INPUT TYPE="submit" NAME="nextbutton" VALUE="Evaluate Primer">
<.INPUT TYPE="hidden" NAME="next" VALUE="evaluate">
<./FORM>
<./BODY>
<./HTML>
```

APPENDIX B

Bradford Analysis made simple: A web-based approach

http://www.msu.edu/~venkata1/bradfordhsa.htm

INTRODUCTION

A key step in protein studies is to quantify the concentration of protein present in a given solution. A variety of assays are available for determination of protein concentrations such as the Lowry procedure, the Biuret method, the bicinchonic acid procedure and the Bradford assay (see Dunn, 1989 for review). The Bradford assay is a convenient assay to measure concentrations of protein in the range of 200 μg/ml and 1,400 μg/ml. This assay is based on the binding of Coomassie Brilliant Blue G-250 to primarily basic (especially arginine) and aromatic amino acid residues (Bradford, 1976). The reaction can be detected using colorimetrically at 595 nm.

RESULTS AND CONCLUSION

To measure the concentration of a protein sample, the absorption of light at 595 nm is compared to the absorption of known concentrations of human serum albumin (HSA). To facilitate the calculation of the concentrations an HTML code has been written. This code carries out a regression analysis to determine the slope and y-intercept of the standards used in the assay. These results are then used to calculate the concentration of the test sample.

The use of this program has facilitated the evaluation of protein concentrations in a convenient manner.

REFERENCES

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72:** 248-54.

Dunn MJ (1989) Determination of total protein concentration. In Harris ELV and Angel S (eds.) Protein purification methods: a practical approach. Oxford University Press, Oxford, UK. pp 10-17

HTML code for Bradford Analysis available at www.msu.edu/~venkata1/bradfordhsa.htm

```
<.HTML>
<.HEAD>
<.TITLE>BradFord<./TITLE>
<.SCRIPT>function Brad
(Intro,NS,X1,X2,X3,X4,X5,X6,X7,X8,X9,X10,Y1,Y2,Y3,Y4,Y5,Y6,Y7,Y8,Y9,Y10,
NU,VU,AU,N1,N2,N3,N4,N5,N6,N7,N8,N9,N10,U1,U2,U3,U4,U5,U6,U7,U8,U9,U10)
{
ARRAYNBIG = new Array(N1,N2,N3,N4,N5,N6,N7,N8,N9,N10)
ARRAYXBIG = new Array(X1,X2,X3,X4,X5,X6,X7,X8,X9,X10)
ARRAYYBIG = new Array(Y1,Y2,Y3,Y4,Y5,Y6,Y7,Y8,Y9,Y10)
ARRAYUBIG = new Array(U1,U2,U3,U4,U5,U6,U7,U8,U9,U10)
ARRAYUX = new Array(NU)
SIGX = 0
for (a = 0; a <. NS; a++)
        {
  SIGX = SIGX + (1 * ARRAYXBIG[a])
        }
SIGX2 = 0
for (i=0; i <. NS; i++)
        {
  SIGX2 = SIGX2 + (ARRAYXBIG[i])*(ARRAYXBIG[i])
     }
SIGY = 0
for (b = 0; b <. NS; b++)
        {
  SIGY = SIGY + (1 * ARRAYYBIG[b])
        }
SIGXY = 0
for (j=0; j <. NS; j++)
        {
  SIGXY = SIGXY + (ARRAYXBIG[j])*(ARRAYYBIG[j])
     }
B = (SIGXY - ((SIGX) * (SIGY))/NS)/ (SIGX2 - (SIGX * SIGX)/NS)
A = (SIGY/NS) - (B * SIGX/NS)
for (k=0; k <. NU; k++)
        {
ARRAYUX[k] = (Math.round ((ARRAYUBIG[k] - A) * 10000/(B * VU)))/10000
}
STDTAB = "\n\nStandard's Table" + "\n---------------------------------------"
+ "\n|HSA (ug/ul)| OD 595 |" + "\n---------------------------------------"
for (p=0; p <. NS; p++)
{
```

```
TX = ARRAYXBIG[p]
if (ARRAYXBIG[p] <. 1)
 { TX = "0" + ARRAYXBIG[p]}
TY = ARRAYYBIG[p]
if (ARRAYYBIG[p] <. 1)
 { TY = "0" + ARRAYYBIG[p]}
STDTAB = STDTAB + "\n|   " + TX + "    |  " + TY + "  |"
}
UNKTAB = "\n-----------------------------------" + "\n\nUnknown's Table" + "\n----------
--------------------------"
+ "\n Sample Vol = " + VU + " ul."
+ "\n|Sample| OD 595 | Conc (ug/ul)|  Vol (ul) for " + AU + " ug protein" + "\n----------
-----------------------"
for (q=0; q <. NU; q++)
{
TU = ARRAYUBIG[q]
if (ARRAYUBIG[q] <. 1)
 { TU = "0" + ARRAYUBIG[q]}
TUX = ARRAYUX[q]
VTUX = (Math.round (10 * AU / TUX))/10
if (ARRAYUX[q] <. 1)
 { TUX = "0" + ARRAYUX[q]}
UNKTAB = UNKTAB + "\n|  " + ARRAYNBIG[q] + "  |  " + TU + "   |   " + TUX + "
|  " + VTUX + "   |"
}
document.ecoliForm.Brader.value = Intro + STDTAB +  UNKTAB + "\n----------------
------------------"
}
today = new Date ()
document.write ("<.H1>Bradford test by Regression Analysis<./H1><.DT>This program
evaluates protein concentrations based on OD595<.DT><.FONT SIZE = 2>Protocol
<.dt> 1)  Add 5 ml Bradford reagent (Bio-Rad 500-0006) to 20 ml H2O. <.dt> 2)  Start
the Spec and set at OD 595 nm. <.dt> 3)  Aliquot 1 ml of the diluted reagent to sample
tubes, add 0,2,4,6,8 and 10 ul of <.dt>  1 mg/ml HSA (BSA can also be used but HSA is
more linear and reproducible) for standards and say 3 ul of protein extract for analysis.
<.dt>4)  Quantitate and calculate below <.dt> <.FONT SIZE = 1> A freeware product of
Iyer-Venk: Developed by Meera Iyer and Sridhar
Venkataraman<./DT><./FONT><.HR>Date: " + today)
<./SCRIPT>
<.BODY>
<.FORM NAME = "ecoliForm">
<.P><.FONT FACE="Tahoma" SIZE=3>Details of the experiment <.INPUT TYPE =
"text" NAME = "Intro" SIZE= 50>
<.DT>Number of Standards (including blank) <.INPUT TYPE = "text" NAME = "NoS"
SIZE= 3>
```

<.DT>Standard Blank (0 ug/ul HSA) <.INPUT TYPE = "text" NAME = "S1P" SIZE= 4>
OD <.INPUT TYPE = "text" NAME = "SOD1" SIZE= 6>
<.DT>Standard 1 (ug/ul HSA)<.INPUT TYPE = "text" NAME = "S2P" SIZE= 4> OD
<.INPUT TYPE = "text" NAME = "SOD2" SIZE= 6>
<.DT>Standard 2 (ug/ul HSA)<.INPUT TYPE = "text" NAME = "S3P" SIZE= 4> OD
<.INPUT TYPE = "text" NAME = "SOD3" SIZE= 6>
<.DT>Standard 3 (ug/ul HSA)<.INPUT TYPE = "text" NAME = "S4P" SIZE= 4> OD
<.INPUT TYPE = "text" NAME = "SOD4" SIZE= 6>
<.DT>Standard 4 (ug/ul HSA)<.INPUT TYPE = "text" NAME = "S5P" SIZE= 4> OD
<.INPUT TYPE = "text" NAME = "SOD5" SIZE= 6>
<.DT>Standard 5 (ug/ul HSA)<.INPUT TYPE = "text" NAME = "S6P" SIZE= 4> OD
<.INPUT TYPE = "text" NAME = "SOD6" SIZE= 6>
<.DT>Standard 6 (ug/ul HSA)<.INPUT TYPE = "text" NAME = "S7P" SIZE= 4> OD
<.INPUT TYPE = "text" NAME = "SOD7" SIZE= 6>
<.DT>Standard 7 (ug/ul HSA)<.INPUT TYPE = "text" NAME = "S8P" SIZE= 4> OD
<.INPUT TYPE = "text" NAME = "SOD8" SIZE= 6>
<.DT>Standard 8 (ug/ul HSA)<.INPUT TYPE = "text" NAME = "S9P" SIZE= 4> OD
<.INPUT TYPE = "text" NAME = "SOD9" SIZE= 6>
<.DT>Standard 9 (ug/ul HSA)<.INPUT TYPE = "text" NAME = "S10P" SIZE= 4> OD
<.INPUT TYPE = "text" NAME = "SOD10" SIZE= 6>
<.P>
<.DT>Number of Unknowns <.INPUT TYPE = "text" NAME = "NoU" SIZE= 3>
Volume of Unknowns (ul) <.INPUT TYPE = "text" NAME = "VoU" SIZE= 4>
Required amount of Unknowns (ug)<.INPUT TYPE = "text" NAME = "AU" SIZE= 4>
<.DT>Unknown 1 <.INPUT TYPE = "text" NAME = "U1P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD1" SIZE= 6>
<.DT>Unknown 2 <.INPUT TYPE = "text" NAME = "U2P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD2" SIZE= 6>
<.DT>Unknown 3 <.INPUT TYPE = "text" NAME = "U3P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD3" SIZE= 6>
<.DT>Unknown 4 <.INPUT TYPE = "text" NAME = "U4P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD4" SIZE= 6>
<.DT>Unknown 5 <.INPUT TYPE = "text" NAME = "U5P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD5" SIZE= 6>
<.DT>Unknown 6 <.INPUT TYPE = "text" NAME = "U6P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD6" SIZE= 6>
<.DT>Unknown 7 <.INPUT TYPE = "text" NAME = "U7P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD7" SIZE= 6>
<.DT>Unknown 8 <.INPUT TYPE = "text" NAME = "U8P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD8" SIZE= 6>
<.DT>Unknown 9 <.INPUT TYPE = "text" NAME = "U9P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD9" SIZE= 6>
<.DT>Unknown 10 <.INPUT TYPE = "text" NAME = "U10P" SIZE= 5> OD 595
<.INPUT TYPE = "text" NAME = "UOD10" SIZE= 6>

<.P><.INPUT TYPE = "button" VALUE = "Submit" onClick = "Brad
(document.ecoliForm.Intro.value,
document.ecoliForm.NoS.value,document.ecoliForm.S1P.value,
document.ecoliForm.S2P.value, document.ecoliForm.S3P.value,
document.ecoliForm.S4P.value, document.ecoliForm.S5P.value,
document.ecoliForm.S6P.value, document.ecoliForm.S7P.value,
document.ecoliForm.S8P.value, document.ecoliForm.S9P.value,
document.ecoliForm.S10P.value, document.ecoliForm.SOD1.value,
document.ecoliForm.SOD2.value,document.ecoliForm.SOD3.value,
document.ecoliForm.SOD4.value, document.ecoliForm.SOD5.value,
document.ecoliForm.SOD6.value, document.ecoliForm.SOD7.value,
document.ecoliForm.SOD8.value, document.ecoliForm.SOD9.value,
document.ecoliForm.SOD10.value,document.ecoliForm.NoU.value,document.ecoliForm.
VoU.value,document.ecoliForm.AU.value,
document.ecoliForm.U1P.value, document.ecoliForm.U2P.value,
document.ecoliForm.U3P.value, document.ecoliForm.U4P.value,
document.ecoliForm.U5P.value, document.ecoliForm.U6P.value,
document.ecoliForm.U7P.value, document.ecoliForm.U8P.value,
document.ecoliForm.U9P.value, document.ecoliForm.U10P.value,
document.ecoliForm.UOD1.value, document.ecoliForm.UOD2.value,
document.ecoliForm.UOD3.value, document.ecoliForm.UOD4.value,
document.ecoliForm.UOD5.value, document.ecoliForm.UOD6.value,
document.ecoliForm.UOD7.value, document.ecoliForm.UOD8.value,
document.ecoliForm.UOD9.value, document.ecoliForm.UOD10.value)">
<./FONT>
<.BR>
<.textarea name="Brader" rows=50 cols=80 ><./textarea>
<.BR>
<.DT>
<.P><.A HREF="mailto:venkata1@pilot.msu.edu"><.FONT FACE="Times,Times New
Roman">Please send Comments to Sridhar Venkataraman<./A><./DT>
<.DT><.A HREF="http://www.msu.edu/~venkata1/index.htm"><.FONT
FACE="Times,Times New Roman">Back to the main page<./A><./DT>
<./FORM>
<./BODY>
<./HTML>

APPENDIX C

Table C.1 Genes present on the BAC MBB18 and the proteins encoded

Clone	Similarity
MBB18.2	Unknown Protein
MBB18.3	Unknown Protein
MBB18.4	Unknown Protein
MBB18.5	Unknown Protein
MBB18.6	Tryptophan synthase beta chain
MBB18.7	Myrosinase binding protein-like, similar to jasmonate induced protein
MBB18.8	Myrosinase binding protein-like, similar to jasmonate induced protein
MBB18.9	Unknown Protein
MBB18.10	similar to protein kinase
MBB18.11	Unknown Protein
MBB18.12	Similar to heat shock transcription factor HSF 30
MBB18.13	Unknown Protein
MBB18.14	Similar to heat shock transcription factor HSF 30
MBB18.15	Unknown Protein
MBB18.16	Unknown Protein
MBB18.17	MADS-box protein-like
MBB18.18	Cytochrome B-561
MBB18.19	Similar to guanine nucleotide exchange factor
MBB18.20	Unknown Protein
MBB18.21	Unknown Protein
MBB18.22	Unknown Protein
MBB18.23	Unknown Protein
MBB18.24	Unknown Protein
MBB18.25	Similar to non-LTR retroelement reverse transcriptase

Annotation by the Kazusa DNA Research Institute (Chiba, Japan) and The Institute for Genomic Research, (TIGR, Rockville, MD, USA).

Table C.2 Genes present on the BAC MKD10 and the proteins encoded

Clone	Similarity
MKD10.2	Unknown Protein
MKD10.3	Proline oxidase precursor
MKD10.4	Unknown Protein
MKD10.5	Similar to Salt-inducible protein
MKD10.6	Unknown Protein
MKD10.8	Unknown Protein
MKD10.9	Similar to pollen coat protein
MKD10.10	S-adenosyl methionine:salycylic acid carboxlmethyltranferase-like protein
MKD10.11	Unknown Protein
MKD10.12	Unknown Protein

Annotation by the Kazusa DNA Research Institute (Chiba, Japan) and The Institute for Genomic Research, (TIGR, Rockville, MD, USA).

Table C.3 Genes present on the BAC K15E6 and the proteins encoded

Clone	Similarity
K15E6.1	Similar to bZIP transcription factor
K15E6.2	Unknown Protein
K15E6.3	Amino acid transporter-like
K15E6.4	Cysteine-tRNA ligase
K15E6.5	Unknown Protein
K15E6.6	Disease resistance protein-like
K15E6.7	DNA-binding protein-like
K15E6.8	Unknown Protein
K15E6.9	Unknown Protein
K15E6.10	Unknown Protein
K15E6.11	Unknown Protein
K15E6.12	FrnE protein-like
K15E6.13	TRNA-Ile(AAT)
K15E6.14	Oxalate oxidase (germin protein)-like protein
K15E6.15	Similar to non-LTR retroelement reverse transcriptase
K15E6.16	Oxalate oxidase (germin protein)-like protein
K15E6.17	Similar to retroelement pol
K15E6.18	Oxalate oxidase (germin protein)-like protein
K15E6.19	Similar to oxalate oxidase
K15E6.20	Oxalate oxidase (germin protein)-like protein
K15E6.21	Cytochrome P-450
K15E6.22	Similar to retroelement pol
K15E6.23	Unknown Protein
K15E6.24	Receptor protein kinase-like protein

Annotation by the Kazusa DNA Research Institute (Chiba, Japan) and The Institute for Genomic Research, (TIGR, Rockville, MD, USA).

Table C.4 Genes present on the BAC MXF12 and the proteins encoded

Clone	Similarity
MXF12.2	Receptor-like protein kinase
MXF12.3	Unknown Protein
MXF12.4	Receptor protein kinase-like protein
MXF12.5	Receptor protein kinase-like protein
MXF12.6	ABC transporter-like protein
MXF12.7	Anthocyanin acyltransferease-like protein
MXF12.8	Transposase-like protein
MXF12.9	Similar to Ac-like transposase
MXF12.10	Anthocyanin acyltransferease-like protein
MXF12.11	Anthocyanin acyltransferease-like protein
MXF12.12	Similar to retroelement pol
MXF12.13	Germin-like protein
MXF12.14	Germin-like protein
MXF12.15	Germin-like protein
MXF12.16	Germin-like protein
MXF12.17	Unknown Protein
MXF12.18	Germin-like protein
MXF12.19	Receptor protein kinase-like protein

Annotation by the Kazusa DNA Research Institute (Chiba, Japan) and The Institute for Genomic Research, (TIGR, Rockville, MD, USA).

Table C.5 Genes present on the BAC K3K3 and the proteins encoded

Clone	Similarity
K3K3.2	Oxalate oxidase (germin protein)-like protein
K3K3.3	Unknown Protein
K3K3.4	Oxalate oxidase (germin protein)-like protein
K3K3.5	Similar to retroelement pol
K3K3.6	Germin-like protein
K3K3.7	Unknown Protein
K3K3.8	Unknown Protein
K3K3.9	Unknown Protein
K3K3.10	Unknown Protein
K3K3.11	Unknown Protein
K3K3.12	Unknown Protein
K3K3.13	Unknown Protein
K3K3.14	Unknown Protein
K3K3.15	non-LTR retroelement reverse transcriptase-like protein
K3K3.16	Unknown Protein
K3K3.17	expansin-like protein
K3K3.18	expansin-like protein
K3K3.19	expansin-like protein
K3K3.20	expansin-like protein
K3K3.21	expansin-like protein
K3K3.22	expansin-like protein
K3K3.23	UDP-glucose dehydrogenase

Annotation by the Kazusa DNA Research Institute (Chiba, Japan) and The Institute for Genomic Research, (TIGR, Rockville, MD, USA).

Table C.6 Genes present on the BAC MUL8 and the proteins encoded

Clone	Similarity
MUL8.1	Unknown Protein
MUL8.2	AHP3
MUL8.3	Selenium-binding protein-like
MUL8.4	Unknown Protein
MUL8.5	Similar to S60 S-locus receptor kinase
MUL8.6	Unknown Protein
MUL8.7	Receptor kinase-like protein
MUL8.8	Similar to Protein-tyrosine phophatase
MUL8.9	Unknown Protein
MUL8.10	Cyclin-dependent protein kinase-like protein
MUL8.11	Unknown Protein
MUL8.12	AKin11
MUL8.13	Unknown Protein
MUL8.14	Unknown Protein
MUL8.15	Similar to CHP-rich Zinc Finger protein
MUL8.16	Similar to CHP-rich Zinc Finger protein
MUL8.17	Unknown Protein
MUL8.18	Unknown Protein
MUL8.19	Unknown Protein
MUL8.20	Pattern formation protein
MUL8.21	v-SNARE At-VTI1a
MUL8.22	Unknown protein
MUL8.23	Unknown protein
MUL8.24	UDP-glucose dehydrogenase (K3K3.23)

Annotation by the Kazusa DNA Research Institute (Chiba, Japan) and The Institute for Genomic Research, (TIGR, Rockville, MD, USA).