5-2002

Promoter trap analysis of the light signaling circuitry in Arabidopsis thaliana

Olga Ruiz Kopp

To the Graduate Council:

I am submitting herewith a dissertation written by Olga Ruiz Kopp entitled "Promoter trap analysis of the light signaling circuitry in Arabidopsis thaliana." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Botany.

Albrecht Von Arnim, Major Professor

We have read this dissertation and recommend its acceptance:

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Olga Ruiz Kopp entitled " Promoter trap analysis of the light signaling circuitry in Arabidopsis thaliana." I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Botany.

Albrecht von Arnim, Major Professor

We have read this dissertation
and recommend its acceptance:

Acceptance for the Council:

Vice Provost and Dean of Graduate Studies

PROMOTER TRAP ANALYSIS OF THE LIGHT SIGNALING CIRCUITRY IN ARABIDOPSIS THALIANA

A Dissertation
presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Olga Ruiz Kopp
May, 2002

Dedication

This dissertation is dedicated to the loving memory of

my father Augusto Ruiz and my brother Jairo Ruiz

thoughts of whom are still alive within me.

To Paul for all his love and support.

Acknowledgments

This thesis has benefited from the contributions of a number of people. I am indebted to my research advisor Dr. Albrecht von Arnim for his guidance, patience, and encouragement during my graduate career. I could not have completed this work without Dr. von Arnim's help. I am very grateful to the members of my doctoral committee, Dr. Beth Mullin, Dr. Mary Ann Handel and Dr. John Dunlap for their suggestions and encouragement throughout the course of this study.

I would like to thank Dr. Robert Martienssen and the members of his laboratory at Cold Spring Harbor for allowing me to screen the promoter trap lines they had generated.

Sincere thanks to Hui Sun for creating the Database.

Special thanks to the University of Tennessee for my education, particularly the Department of Botany and the Division of Biology for their financial support. I also would like to thank the Arthur E. Yates Foundation for a Graduate Fellowship in dissertation writing. I would like to acknowledge the Statistical Consulting Center for their valuable input in the analysis of the data.

I would like to express my gratitude to my fellow graduate students at the Botany Department. Special thanks to the secretaries, Eileen Coltharp and Eunice Turner for their friendship and support.

My sincere gratitude is extended to my mother, Magdalena de Ruiz and my sister Fabiola for their love, and for believing in me. My special thanks to Dee Ann Ostby for being a wonderful friend. I will always be grateful to Dr. Otto Kopp and Helen Kopp and my extended family for opening their hearts to me and for their love and support.

Finally, I would like to thank my dear husband, Paul, for his unflagging emotional support and for brightening everyday of my life with his love. I will always be grateful to him and 'los animales' that cheered me up when I needed it the most.

Abstract

Light is one of the most important environmental factors that control plant growth and development. Research concerning the effect of light quality and light intensity has focused on the photoreceptors that perceive the light signals, the genetics of photoreceptor-specific signaling pathways, and the developmental responses to light. However, little is known about the integration of signals from the various light signaling pathways.

Light signaling pathways and their interactions were investigated in *Arabidopsis thaliana* by measuring light responsive transcription in a collection of promoter trap lines. Each promoter trap line carries a single, transposon-mediated fusion between a randomly selected endogenous *Arabidopsis* sequence and the *E. coli* β-glucuronidase (GUS) reporter gene. Light response profiles, defined as the variation in GUS expression of a gene over a range of environmental conditions, were characterized for a large number of individual promoter traps by histochemical GUS staining and by whole-plant enzyme assays. By applying either constant light or darkness or shifts between light and darkness, I identified 286 light responsive promoter traps. Interestingly, a large fraction of the lines screened displayed tissue-specific light responsiveness. Few examples of this phenomenon have been published based on transgenic promoter:reporter fusions in plants.

The light responsive lines were screened for their response under constant red, constant far-red light, and four types of light pulse regimes in order to define the relative sensitivity of the lines to phytochrome A (phyA) and phytochrome B (phyB) photoreceptors. All tissues surveyed, including the root, were capable of supporting a variety of light response profiles, suggesting that the interaction between light signaling pathways are similarly complex in most cell types. It was very rare to find stronger expression in constant far red than constant red light. This finding suggests that the phyA pathway, active under constant far red, plays a major role in the regulation of genes that are repressed by light and a minor role in genes that are active by light. Because certain light response profiles occurred repeatedly, it was possible to group the profiles using hierarchical cluster analysis. This revealed groups of genes with antagonistic or synergistic interactions between phyA and phyB.

The chromosomal flanking sequences that drive the light responsive expression profiles were analyzed in detail for fifteen promoter traps. The typical light responsive promoter trap was light repressible, yet contained basal promoter sequence motifs, as well as

motifs previously implicated in light inducible gene expression. Many of the light responsive flanking sequences were bona fide cryptic promoters, rather than regulatory regions for native *Arabidopsis* genes, indicating that cryptic promoters may play an important role in the light-signaling network.

Microarray analyses of individual *Arabidopsis* cDNAs in response to shifts between light and darkness were performed as a first step to correlate expression profiles identified by promoter trapping with profiles defined at the level of mRNAs. For both light shift experiments, the most commonly activated genes coded for proteins involved in metabolism, and among these, primary metabolism outweighed secondary metabolism. After the shift from constant dark to light, inducible genes preferentially coded for proteins involved in photosynthetic light reactions. In contrast, a shift from light to darkness elicited the expression of genes involved in cell expansion and transport, as expected. More unexpectedly, the shift to darkness also triggered expression of genes involved in translation and stress responses.

Table of contents

Page

List of Figures

Page

List of Tables

List of abbreviations

AS1/*as1*	Asparagine synthetase 1 gene
ASF1	Activation sequence factor 1
ATHB-2	*Arabidopsis thaliana* homeobox-2 gene
BAS2	Phytochrome B activation-tagged suppressor
CAB	Light harvesting chlorophyll A/B binding protein gene
CCT	Carboxyl terminal domain of cryptochromes
cGMP	Cyclic GMP
CHS	Chalcone synthase
CIP4	COP1-interacting partner
COP1/COP9	Constitutively photomorphogenic
CRY/*cry*	Cryptochrome
CUE/*cue*	CAB protein underexpressed
DD	Differential display
DE1/*de1*	Dark inducible element
DET/*det*	De-etiolated
DOC1/*doc1*	Dark overexpression of CAB1
EGF	Epidermal growth factor
ELF/*elf*	Early flowering
eid	Empfindlicher im dunkelroten Licht (mutant sensitive to far-red light)
EOD	End of day responses
ET	Enhancer trap
FAR1/*far1*	Far-red impaired response
fed1	Ferrodoxin 1
fhy	Far-red elongated hypocotyl
FIN/*fin*	Far-red induction
FNR	Ferredoxin NADP+ oxidoreductase
GI	GIGANTEA
GPA1	G-protein alpha subunit 1
GT	Gene trap
GUS	β-glucuronidase
HIR	High irradiance response
HY/*hy*	Long hypocotyl
HFR1/*hfr1*	Long hypocotyl in far-red
IAAH	Indole acetic acid hydrolase
KAN	Kanamycin

KNAT3	Knotted1-like
LED	Light emitting diodes
LFR	Low Fluence Response
Lhcb	Light harvesting chlorophyll a/b binding protein gene
LRE	Light responsive elements
MAPK	Mitogen-associated protein kinases
MU	Methyl umbelliferone
NAM	Naphthalene acetamide
NEDD8	Neural precursor cell-expressed developmentally down-regulated
NDPK2	Nucleotide diphosphate kinase
ORF	Open reading frame
PAB1	Poly(A) binding protein gene 1
PAT1/*pat1*	Phytochrome A signal transduction
PCR	Polymerase chain reaction
PEF/*pef*	Phytochrome early flowering
phy	Phytochrome
phyA	Phytochrome A
phyB	Phytochrome B
phyC	Phytochrome C
phyD	Phytochrome D
PIF3/*pif3*	Phytochrome interacting factor
PKS1	Phytochrome kinase substrate 1
Pr	Red absorbing form of phytochrome
Pfr	Far-red absorbing form of phytochrome
PKS1	Phytochrome kinase substrate 1
PRA1/*pra1*	Pea small GTPase gene phytochrome regulated
PRP1/prp1	Pathogenesis related protein 1
PSI2/*psi2*	Phytochrome signaling
QTL	Quantitative trait loci
rbcS	Ribulose biphosphate carboxylase/oxygenase
RED1/*red1*	Red light specific extragenic revertant
RSF1/*rsf1*	Reduced sensitivity to far-red inhibition of hypocotyl elongation
SAGE	Serial analysis of gene expression
SCF	Skp1, Cdc53/Cullin, Fbox receptor family of ubiquitin ligase complex
SCN	Cop9 signalosome
shy2	Suppressor of *hy2* mutation
SPA1/*spa1*	Suppressor of phyA-105
srl1	Short hypocotyl in red light

TAIL PCR	Thermal asymmetric interlaced PCR
Tic22	Inner membrane translocon
TUB1/*tub1*	Beta tubulin
VLFR/*vlfr*	Very Low Fluence Response
X-GLUC	5-bromo-4-chloro-3-indole-β-D-glucuronide

Light treatments

cL	Constant light
cD	Constant dark
cR	Constant red
cFR	Constant far-red
Rp	Red pulses
R-FRp	Red/ far-red pulses
FRp	Far-red pulses
FR-Rp	Far-red/ red pulses
L=D	Equal staining in light and dark
L>D	Stronger staining in light than dark
D>L	Stronger staining in dark than light
LD	Shift from constant light to dark
DL	Shift from constant dark to light

List of Plates

Chapter 1. Introduction

Organisms acquire information about their surroundings by processing a variety of environmental signals and using that information to modulate various aspects of growth and development. Plants are sessile organisms and cannot move to more favorable settings when the environmental conditions are adverse. Instead, plants have evolved extremely sophisticated mechanisms to monitor their environment, transducing the information to the gene expression machinery to respond to environmental changes.

Light is one of the most important environmental factors that control plant development. The ability of a plant to maximize its photosynthetic productivity depends on its capacity to sense and respond not only to the quantity but also to quality and direction of light. The perception of light quantity involves photon counting (i.e. response to total fluence). The perception of the light quality (spectral distribution of radiation) involves the estimation of ratios of photons in two or more wavelength bands. The perception of the direction of light depends upon the detection of photon gradients in space, including photon counting and comparison at spatially separated points in the organism. Plants are able to perceive the duration of exposure to light by timing the light-dark transitions; this involves not only photon counting but also detection of changes in spectral distribution of radiation (Smith, 1994). The response to the fluence rate obeys the Bunsen-Roscoe Reciprocity Law, in which a response is proportional to the quantity of photoproduct, irrespective of whether that quantity is produced by brief pulses of high photon irradiance, or longer periods of low photon irradiance (Smith, 1994).

Following absorption of light, photoreceptors interact with other signal transduction elements, which eventually leads to many molecular and morphological responses. While a complete signal transduction cascade is not known yet, molecular genetic studies using the model plant Arabidopsis have led to substantial progress in dissecting the signal transduction network. Important gains have been made in determining the function of the photoreceptors, the terminal response pathways, and the intervening signal transduction components (Fankhauser and Chory, 1997).

1.1.Types of light responses in plants

1.1.1. Photoreceptors

Light modulates the development and physiology of plants. For example, plants grown in the dark exhibit a typical etiolated appearance (skotomorphogenesis) characterized by yellow, unopened cotyledons and an elongated stem. Plants grown under light conditions show a short hypocotyl and open, green cotyledons. Light influences other developmental

processes including germination and floral induction, which are collectively referred to as photomorphogenesis. In order to check the whole light spectrum, plants use several sensor pigments such as phytochromes, blue light/UV-A and UV-B photoreceptors. Phytochrome (phy) was the first photoreceptor identified. It is a red, far-red photoreversible chromoprotein that operates predominantly within the red (R) and far-red (FR) range. Cryptochromes (cry's) and phototropin are blue light photoreceptors. Cryptochromes sense the blue and UV-A part of the spectrum (Ahmad and Cashmore, 1993, Salomon et al., 2000) and phototropin, a flavoprotein, is involved in phototropism (Liscum and Briggs, 1995). UV-B photoreceptors are sensitive to UV-B irradiation and are involved in the regulation of the formation of UV-shielding pigments (Christie and Jenkins, 1996) but they are not well characterized at the molecular level.

Cryptochromes (cry's), represented by *cry1* and *cry2* in Arabidopsis, are flavin and pterin-containing blue light receptors. They are evolutionarily related to the blue-light activated DNA photolyases, which cleave thymidine dimers, yet cry's have no known enzymatic activity. The carboxyl terminal domain of the cryptochromes (CCT) mediates a constitutive light response (Yang et al., 2000). The CCT may function via an interaction with repressors of light signaling, such as COP1 (Wang et al., 2001a), and/or one of the COP9 signalosome subunits, or by interacting with signaling molecules upstream or downstream of the COP1 and/or the COP9 signalosome. Blue light receptors have long been implicated as important modulators of phytochrome signaling pathways (reviewed in Casal, 2000), although phy-independent activities are also evident (Poppe et al., 1998). More recently, physical interactions between cry's and phy's have been identified. Specifically, cry1 interacts with at least phyA (Ahmad et al., 1998), and cry2 interacts with phyB (Mockler et al., 1999, Mas et al., 2000). Detailed analysis of the phy-cry interactions may eventually lead to a better understanding of the crosstalk between blue and red/far-red specific signaling chains.

Phytochrome exists in two spectrally distinct, photointerconvertible forms: Pr, a red absorbing form and Pfr, a far red absorbing form (Quail et al., 1995). Pr absorbs maximally in the red wavelength, with a peak around 660 nm; Pfr absorbs maximally in the far-red region, with a peak at about 730 nm (Vierstra and Quail, 1983). By measuring the amount of Pfr and the ratio of Pr to Pfr, plants assess the intensity, duration, and spectral quality of the light environment. The photoconversion of Pr to Pfr induces diverse morphogenetic responses whereas the reverse conversion from Pfr to Pr cancels the induction of responses. Therefore, Pfr is considered the active form and Pr the inactive form of the phytochrome photoreceptor. The molecular mechanisms by which Pfr induces the downstream developmental responses have not been elucidated. One hypothesis states that conformational changes associated with Pr and Pfr photoconversion

result in differential interactions with downstream components of signal transduction chains linking phytochrome to physiological responses.

Phytochrome possesses a carboxyl terminal domain with similarity to the histidine kinase module of bacterial two-component signaling systems. The His-kinase module is not enzymatically active, however, and it is partially dispensable for phy function (Krall and Reed, 2000). However, phy has serine/threonine kinase activity (Yeh and Lagarias, 1998). Potential substrates of the phy kinase activity have been identified in the form of phytochrome interacting proteins such as PKS1 (Fankhauser et al., 1999), PIF3 (Ni et al., 1998) and NDPK2 (Choi et al., 1999). This area of research made a significant leap forward when Ni and coworkers could demonstrate that the interaction between phytochrome and one of its interactors, PIF3, is entirely dependent on the active Pfr form of phy (Ni et al., 1999). Moreover, PIF3, which is a basic helix-loop-helix protein, binds to the G-box, a specific DNA sequence element found in many light regulated promoters. Given that phytochrome appears to be associated with the PIF3-promoter complex in a light-dependent fashion, a simple pathway is emerging: Upon light absorption, phy may function as a transcriptional coactivator of light-inducible target genes (Martínez-García et al., 2000).

Arabidopsis has five phytochrome genes, called *phyA-phyE* that diverge from each other as much as 50%. Genetic and physiological studies have shown that these phytochrome genes regulate distinct light responses (Fankhauser and Chory, 1997). The functional phytochrome photoreceptors are homodimers with each subunit containing the linear tetrapyrrole chromophore phytochromobilin attached to an approximately 120-kD polypeptide. Phytochrome-mediated responses can be divided into Low Fluence Responses (LFR), Very Low Fluence Responses (VLFR), High irradiance Responses (HIR), photoperiodic responses and End-of day responses (EOD). The LFRs are the classical phytochrome mediated responses induced by short exposures to R and these show R-FR reversibility. The photon fluences required for saturation of the response by R vary from 1 to 1000 μmole m^{-2}. The VLFR can be induced by very low photon fluences (10^{-4} to 10^{-1} μmole m^{-2}). The HIR requires exposures to light over a long period of time of relatively high photon flux (>1000 μmole m^{-2}) for maximum expression. The HIR does not show R-FR reversibility and does not obey the reciprocity law (Mancinelli, 1994). Photoperiodic responses are responses to the duration of the light and dark periods in a 24-hour cycle. For example, many plants flower only during certain time of the year, in response to the daylength. The EOD responses include those responses to the state of phytochrome established at the end of the daily light period (Mancinelli, 1994). The EOD responses show R-FR reversibility and obey the reciprocity law.

Even though phyA and phyB both absorb R and FR light to equal extents, they monitor distinct aspects of the light environment. PhyA is necessary for continuous FR perception (FR-HIR, Whitelam et al., 1993) but it is not necessary for the response to continuous red. PhyA mediates germination and partial de-etiolation under very low fluences (VLFR, Neff et al., 2000) and mediates inhibition of hypocotyl elongation under FR-HIR. Thus, the VLFR and HIR appear to correspond to two branches of phyA signal transduction (Hennig et al., 2001, Casal and Sánchez, 1998). In contrast, phyB is light stable and it may constitute the predominant molecular species of phytochrome in light grown tissue (Somers et al., 1991, Clack et al., 1994). PhyB operates in the low fluence response mode but it is not necessary for continuous Far-red (cFR) perception (Quail et al., 1995) and it mediates germination and inhibition of hypocotyl elongation under pulses of low fluences of red (LFR) which are photoreversible by FR pulses (Casal et al., 1998). Rather little is known about the elements of the light signaling machinery that distinguish between these different response modes. The VLFR appears to be mediated by phyA and the LFR can be mediated by phyB.

Even though the photophysical properties of phyA and phyB are barely distinguishable, differences in their concentrations or signaling partners ensure that phyA and phyB play very distinct roles in the plant. Analysis of responses such as inhibition of hypocotyl elongation and cotyledon unfolding has shown that phyA and phyB operate synergistically or antagonistically depending on the light conditions (Cerdán et al., 1999). Thus, there is a synergistic interaction between phyA and phyB in seedlings exposed to FR pulses (FRp) followed by R pulses (Rp) (Casal and Boccalandro, 1995, Smith et al., 1997, Casal, 2000, Lin, 2000). However, there is an antagonistic interaction between phyA and phyB when seedlings are exposed to continuous red or repeated red pulses (Mazzella et al., 1997). The negative interaction between phyA and phyB HIR occurs in the cytoplasm. It is a property of the light signaling network and it probably occurs early in the signaling cascade of phyA. However, activation of the phyB signaling cascade is not required to inhibit phyA action as shown by a *phyB* mutant unable to inhibit the phyA HIR (Hennig et al., 2001).

1.1.2. Ecological and physiological significance of light responses in plants

Light plays a critical role throughout the life cycle of plants by regulating shade avoidance, seed germination and transition to flowering. Light environments under vegetation canopies vary in quantity and quality. Using the suite of phy's and other photoreceptors introduced above, plants can tell whether they are in the light or the shade. Low ratios of R:FR are characteristic of radiation transmitted through or reflected from vegetation. These low R:FR ratios elicit a set a morphogenic changes such as shoot elongation, stem bending, suppression of axillary bud initiation and branching and

redistribution of assimilates (Smith, 1982) that is known as a shade avoidance response (Schmitt et al, 1995). Evidence for the involvement of phytochromes in the control of responses to canopy density and the elicitation of tropisms and other morphological responses that influence the ability of plants to capture light in patchy canopies has been gathered from experiments altering either the canopy light environments or the light sensing mechanisms in plants (Ballaré and Scopel, 1997). Phytochromes control branching (Casal et al., 1986) and elongation responses to canopy density (Ballaré, et al., 1990). Phytochrome B (Ballaré et al, 1992, Yanovsky et al., 1995) and PhyC, D and E (Devlin et al., 1999) play a role in the control of these responses and therefore in controlling key aspects of plant acclimation to the plant community environment. Mutants that lack functional phyB had reduced response to plant density confirming that light dependent mechanisms play a critical role in the elicitation of morphological responses to crowding and that phyB cannot be substituted by other phytochromes or other photoreceptors in the perception of proximity photosignals (Ballaré and Scopel, 1997).

Light promotion of seed germination is mediated by a phytochrome pool that is highly stable in the Pfr form (Casal et al., 1991). This suggests that even very weak light received by seeds left under a few millimeters of soil might induce germination because the light signal could be integrated as Pfr over several photoperiods (Ballaré et al., 1992). This phenomenon is known as a very low fluence (VLFR) response mechanism (Scopel et al., 1991). The VLFR is common in several processes mediated by phytochrome and it is triggered by light exposures that would form very small amounts of the far-red absorbing form of phytochrome (Pfr) (Cone et al., 1985). The VLFR may allow seeds to detect submillisecond exposures to sunlight when the soil is being disturbed (Scopel et al, 1991). PhyA and PhyB modulate the timing of dormancy break in seeds in a different way. PhyA photoirreversibly triggers the photoinduction of seed germination upon irradiation at a very low fluence with light of the UV-A, visible and far-red range. PhyB mediates the photoreversible reaction, responding to red and far-red light at 10^4 fold higher fluences than those to which phyA responds. This novel action spectrum for phyA-specific induction of seed germination demonstrates that phyA is the photoreceptor for the VLFR (Shinomura et al., 1996).

1.2. Reductionist approaches to the dissection of light signaling

The light signal transduction pathways provide the means by which information from specific wavelengths of light may be amplified and coordinated, resulting in complex responses. The perception of environmental light signals by phytochromes controls many processes including seed germination, seedling establishment, the proper development of photosynthetic machinery, the architecture of the vegetative plant, the timing of

flowering, tuberization and bud dormancy, the responses to neighbor competition and the allocation of resources to root, stem, leaf, reproductive or storage structures (Smith, 2000).

A lot of effort has been dedicated to the analysis of photoreceptors and the light regulated responses in plants. Given the presence of at least five phytochrome species, two cry's, as well as ancillary photoreceptors, and given the diversity of responses at every level of organization, from the gene to the whole organism, perhaps a very complicated network could be expected consisting of a great number of signaling molecules. The precise structure of this network remains very much in the dark and this is where the major challenge for future research resides. Moreover, this network is not only affected by light but also by additional factors such as other environmental stimuli or endogenous signals like phytohormones (Chory et al., 1996, Chory and Wu, 2001). Genetic approaches and biochemical/ cell biological approaches (McNellis and Deng, 1995; Chory et al., 1996, Barnes et al., 1997, Chory and Wu, 2001, Fankhauser, 2001) have been used to characterize the light signal transduction network, each giving rise to a different kind of model to explain the complexity of the light signaling process.

1.2.1. Genetic approaches

Genetic studies indicate that light responses are not simply endpoints of linear signal transduction pathways but result from the integration of information from a network of interacting signaling components. The signaling components include the photoreceptors themselves, as well as positive and negative regulatory elements that act downstream of these photoreceptors (Fankhauser and Chory, 1997). The signaling network involves direct interactions of the photoreceptors as well as cross talk and integration of pathways both early and late in the signaling (Chory and Wu, 2001).

More than 50 genes acting downstream of photoreceptors have been identified by genetic and molecular screens (Deng and Quail, 1999, Neff et al., 2000). Mutations in Arabidopsis *phyA* and *phyB* genes have been identified (Reed et al., 1993, Whitelam, et al., 1993). Plants mutant for *phyA* are not responsive to far-red light-mediated inhibition of hypocotyl elongation. Mutants for *phyB* show reduced red-light-mediated inhibition of hypocotyl elongation. Mutants lacking phyB have altered regulation of hypocotyl elongation, flowering time and morphology in response to low fluence red light or the ratio of R/FR (Reed et al., 1993) whereas the *phyD* mutation does not affect significantly the rosette morphology or development. This example demonstrates that different phytochromes have different functions: phyA mediates responses under cFR, i.e. the FR-HIR as well as the VLFR and phyB is responsible for the low fluence-response and R-HIR conditions (Batschauer, 1998).

Mutant screens for morphogenetic responses such as the supression of hypocotyl elongation in the light have been used to identify plants that show diminished responsiveness to red, far-red or blue light. For example, mutations in one out of five phytochromes may have rather subtle effects (e.g. phyD Devlin et al., 1999) whereas double, triple and quadruple mutant combinations are increasingly severe. This indicates the importance of multiple signal inputs from various types and members of receptor families, suggesting that the pathways are partially redundant (Mulligan et al., 1997). Therefore, distinguishing the roles of individual phytochromes is not trivial. Once mutations in all five phys are known one new approach will be to study the phenotypes of plants that possess only one single phytochrome and are mutant for the other four (Casal, 2000).

The isolation of Arabidopsis mutants defective in light-regulated morphogenesis has led to the identification of three classes of key players (insensitive, hypersensitive and constitutive mutants (Chamovitz and Deng, 1996). Light insensitive mutants display etiolated phenotypes (elongated hypocotyls) in a range of light conditions. These mutants are recessive and have mutations in either the photoreceptor (*phyA, phyB*, *cry*), the phy chromophore biosynthesis, i.e. *hy1, hy2, hy6* (reviewed in Møller and Chua, 1999), or some positively acting downstream component of the signal transduction pathway, i.e. *hy5, red1, pef, cue1* (reviewed in Khurana et al., 1998). Mutants that are hypersensitive to light have been found that suggest the presence of negative light signaling components for example, *spa1, eid1, pat1, fhy1* (reviewed in Fankhauser, 2001). Constitutive mutants display partially photomorphogenic phenotypes in the dark, i.e. the *cop/det/fus*-group, *det2* and *shy2* (reviewed in Fankhauser, 2001; Hardtke and Deng, 2000)

Mutations such as *hy1* and *hy2* (chromophore biosynthetic mutants) (Parks et al., 1989), *hy3* (*phyB*) (Somers et al., 1991), *hy8* (*phyA*), *hy4* (*cry1*) (Parks and Quail, 1991; Somers et al., 1991) are considered insensitive mutants because they display an elongated phenotype in the light and they are believed to have defects in phytochrome chromophore biosynthesis or attachment (Nagatani et al., 1993). Mutations in *HY5* seem to be deficient in red-, far-red and blue light-mediated hypocotyl inhibition (Chory, 1993) and they are certainly not photoreceptor mutants because they have normal levels of phytochrome (Barnes et al., 1997). HY5 is a b-zip transcription factor that promotes photomorphogenesis (Hardtke et al., 2000). The mutant *hy5* is impaired in its responsiveness to different wavelengths and therefore the mutation appears to be in a downstream element of the phytochrome and blue-light receptor transduction pathways (Batschauer, 1998). Since the mutant is not resistant to the block in de-etiolation caused by exposure of seedling to constant FR, this indicates an additional phyA signaling

branching point between pathways regulating hypocotyl elongation and de-etiolation (Barnes et al., 1997).

The hypersensitive mutants exhibit negative regulation of either the phyA-signaling pathway such as *spa1* (Hoecker et al., 1999) or both, the phyA and phyB signaling pathways such as *psi2* (Genoud et al., 1998). SPA1 is a novel WD-40 repeat-containing protein localized to the nucleus (Hoecker et al., 1999). PSI2 might act as a kinase that could desensitize phyA and phyB by phosphorylation (Genoud et al., 1998).

Mutants that exhibit aspects of a photomorphogenic phenotype when grown in the dark identify negative regulators of light signal transduction (Hardtke and Deng, 2000). One subgroup of these genes codes for enzymes responsible for the synthesis of brassinosteroid plant hormones (e.g. *DET2*), thus implicating this class of growth regulator in the etiolation response. The second subgroup has been termed the *COP/DET/FUS* group, due to the constitutively photomorphogenic, or deetiolated appearance and the dark purple (*fusca*) coloration of the mutant embryos. *Cop/det/fus* mutants have characteristics of light grown plants even when grown in darkness including the expression of CAB genes (Millar et al., 1994). It has been shown that *cop1* mutations are epistatic to the long hypocotyl mutations *hy1*, *hy2*, *hy3* and *hy4*, which suggests that COP1 acts downstream of the phytochromes and a blue light receptor (Ang and Deng, 1994). DET and COP proteins may be negative regulators of a wide range of developmental pathways, integrating signals from various environmental stimuli including light and plant growth regulators (Bowler and Chua, 1994). The *COP/DET/FUS* genes appear to play a role in the proteolytic turnover of light regulatory signaling components, as evident from two key findings. First, the COP1 protein is required for the dark-mediated destabilization of the HY5 transcription factor (Osterlund et al., 2000, Wang et al., 2001a). Consistent with this notion, COP1 bears a Ring-finger domain, a hallmark of E3 ubiquitin ligases, proteins that effect the ubiquitination of future proteolytic substrates. A second group of COP/DET/FUS proteins assemble to form the COP9 signalosome (CSN), a nuclear protein complex whose evident biochemical activity is the deconjugation of the small ubiquitin-like peptide, NEDD8, from a multisubunit ubiquitin ligase complex termed SCF (Skp1, Cul-1, R0C1 and F-box protein, Schwechheimer et al., 2001; Lyapina et al., 2001). NEDD8 enhances the ubiquitylating activity of the SCF complex by accelerating the formation of the E2-E3 complex, which stimulates protein polyubiquitylation (Takayuki et al., 2001).

Light exerts its regulation of gene expression at the levels of transcription, mRNA stabilization, translation and post-translational events. In plants, many genes are known to be regulated by light and only a few of those genes such as *CAB*, *rbcS*, and *phyA* have been well characterized. Light regulated genes fall into two categories: down- and up-

regulated genes. Examples of phytochrome down-regulated genes include *DE1* (Inaba et al., 2000), *PRA1* (Inaba et al., 1999), *TUB1* (Leu et al., 1995), *AS1* (Ngai et al., 1997), *ATHB2* (Carabelli et al., 1996), *NPR* (Okubara et al., 1993) and *PHYA* itself (Dehesh et al., 1990). There are many phytochrome up-regulated genes of which *rbcS* and Lhcb (light harvesting chlorophyll a/b binding protein gene) are the best-characterized (Batschauer et al., 1994). Northern blot analysis and nuclear run-off experiments have been used to study transcriptional regulation by light. For example, *CAB* genes are regulated mainly at the level of transcription (Tobin and Silverthorne, 1985) and this regulation is mediated by different photoreceptors. CAB mRNA levels are affected by the VLF and LF range (Kaufman et al., 1984) and the HIR fluences (Wehmeyer et al., 1990). However, there are some differences in response to light within the family of CAB genes. For example, Arabidopsis *CAB1*, *CAB2* and *CAB3* show a strong mRNA accumulation in response to red, but specific transcripts accumulate to different levels (Sun and Tobin, 1990). An interesting example of the complexity of light regulation is the Ferredoxin 1 (*Fed1*) gene, a nuclear gene from pea whose mRNA levels are regulated at the level of transcription initiation in etiolated seedlings (Gallo-Meagher et al., 1992) but in green leaves it is regulated by changes in mRNA stability (Petracek et al., 1998) and/or by regulation at the level of translation (Hansen et al., 2001, Dickey et al., 1998).

Simplified models of the light signal transduction network have been described (Figure 1, Chory and Wu, 2001, Fankhauser, 2001). Signaling downstream of phyA and phyB splits in at least three branches (Neff et al., 2000, Fankhauser, 2001). The phyA-specific branch is defined by mutants such as *fhy3* (Whitelam et al., 1993), *fin2* (Soh et al., 1998), *eid1* (Buche et al., 2000), *spa1* (Hoecker et al., 1998), *pat1* (Bolle et al., 2000), *far1* (Hudson et al., 1999) and the alleles *rsf1/hfr1* (Spiegelman et al., 2000, Fairchild et al., 2000). In addition, the *vlf1* and *vlf2* loci (Yanovsky et al., 1997) identified by Quantitative Trait Loci (QTL) define the phyA-specific brand. The phyB-specific signaling branch is defined by mutants such as *red1* (Wagner et al., 1997), *pef2*, *pef3* (Ahmad and Cashmore, 1996), *poc1* (Halliday et al., 1999), *elf3* (Liu et al., 2001), and *srl1* (Huq et al., 2000). Moreover, a branch implicated in both phyA and phyB signaling includes the mutants/genes *pef1* (Ahmad and Cashmore, 1996), *psi2* (Genoud et al., 1998), *pks1* (Fankhauser et al., 1999), *ndpk2* (Choi et al., 1999) and the PIF3 (Ni et al., 1998). All the pathways are postulated to converge at or upstream of the COP/DET/FUS regulators that integrate the various light signals and modulate the activity of downstream effectors such as HY5 and other unknown factors. These factors direct the changes in metabolism and gene expression that will eventually result in photomorphogenic responses (Hardtke and Deng, 2000).

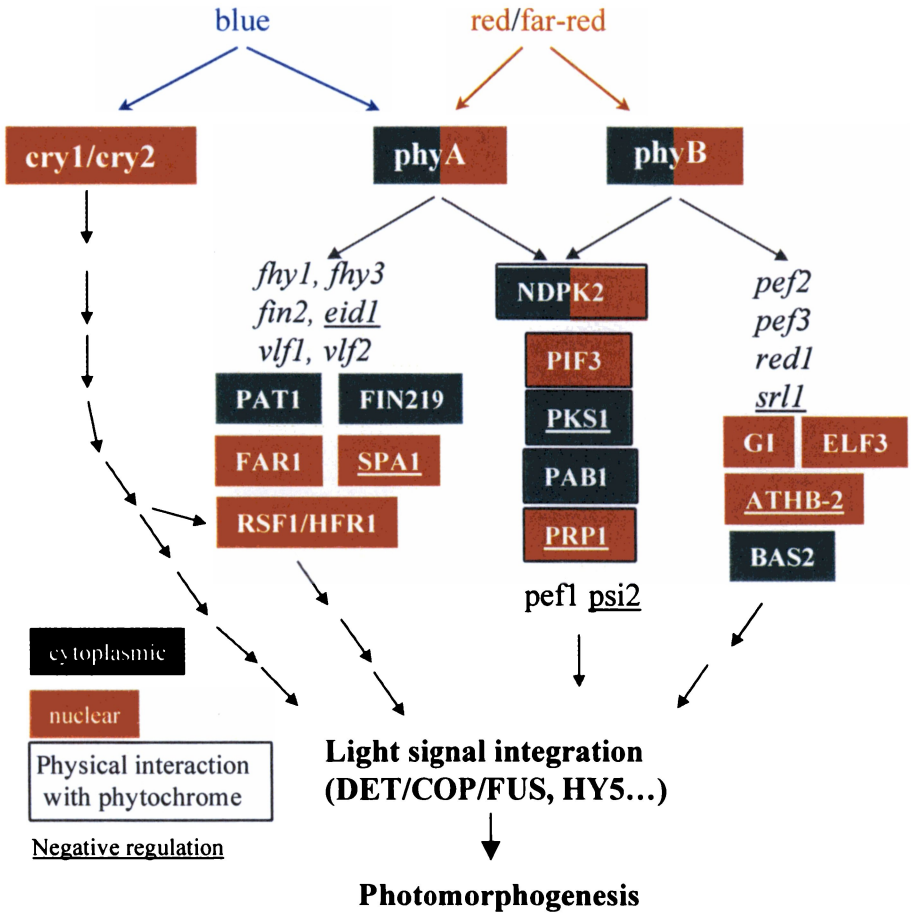


Figure 1. A simple model for phytochrome-mediated light signaling (according to Fankhauser, 2001, Chory and Wu, 2001). Genetic loci affecting specific branches of phytochrome signaling are indicated in italics. Cloned genes are in capitals; proteins that interact with phytochrome are boxed and negative regulators are underlined. Some signaling components act downstream of a single photoreceptor and other ones act downstream of multiple photoreceptors reflecting the fact that light signals perceived by different photoreceptors must be integrated.

1.2.2. Studies of light regulated promoters

Regulated gene expression is driven by regulatory DNA sequences that interact with specific protein factors which are in turn governed by information perceived by signaling pathways triggered by environmental light conditions (Khurana et al., 1998). In an attempt to dissect light signaling pathways at the molecular level, much emphasis has been placed on the structure and function of light regulated promoters, with the argument that transcriptional activation from a single promoter marks an early, fairly simple and quantitative output branch of the light signaling network. One might expect that different light signaling pathways may merge into separable promoter sequence elements, and the signals are then integrated into a transcriptional activation profile by elements of chromatin structure and associated DNA binding proteins. This expectation has been only partially fulfilled. One of the rare examples is the promoter region of the *CAB* gene where VLFRs and HIRs are mediated by phytochrome A and LFRs are mediated by spatially separable phyB specific elements (Cerdán et al., 2000). More typically, a single promoter element mediates a fairly complex response profile that may include responsiveness to light, the circadian clock and tissue specific factors (Anderson and Kay, 1995, Millar and Kay, 1996). Within these promoters, one can discern light regulatory sequence motifs that serve as binding sites for nuclear transcription factors. Among these are the G-box, bound by HY5 and PIF3 (Chattopadhyay et al., 1998a); the GT element, bound by GT-1 (reviewed by Zhou, 1999); and the GATA motif bound by the GATA-binding protein, CGF1, which is closely related to GT-1 (Teakle and Kay, 1995), and other proteins. Individually, these elements are often necessary but not sufficient for light responsiveness.

Promoters of different genes have been found to exhibit different sensitivities to different types of light (Lubberstedt et al., 1994). No single element has been found to be a universal feature for all light regulated promoters (Khurana et al., 1998). Therefore, light-responsive elements and/or the factors interacting with them may play a role as signal integration points in the network mediating both light and developmental control of gene expression (Puente et al., 1996). Analysis of well conserved light-responsive elements (LREs) indicated that pairwise interaction of multiple promoter elements, but not the individual elements alone, constitute the minimal promoters that regulate light responsiveness and cell type specificity (Puente et al., 1996, Chattopadhyay et al., 1998b). Recently, a small cis-regulatory element of 12 base pairs was found to be capable of conferring light responsiveness to a minimal promoter (Inaba et al., 2000). Curiously, this element regulates induction in darkness rather than repression by light and it is up to date the smallest sequence unit reported, sufficient to induce light responsiveness to a minimal promoter (the region around the TATA-box in the CaMV 35S promoter). Longer elements such as the 52-bp element of chalcone synthase (Weisshaar et al., 1991)

have been reported to be sufficient to confer light inducibility. Deletion analyses of promoter regions have defined the regions necessary for conferring light-inducible and tissue-specific expression of several genes (Ha and An, 1988, Block et al., 1990, Orozco and Ogre, 1993).

Puente et al. (1996) reported that combinatorial interactions of two distinct light regulatory promoter elements constitute an autonomous light responsive determinant in the promoter. This phenomenon has been termed combinatorial interplay of the promoter elements. At present, the molecular basis for this phenomenon is not known. At least two general systems can be envisioned: cooperative binding of two DNA binding proteins that recognize the individual promoter elements and the involvement of coactivators that simultaneously recognize both promoter element binding proteins. The emergence of putative coactivators such as CIP4 and CIP7 involved in the light control of development might support the latter system. It is possible that CIP4 and CIP7 could activate transcription of the target genes by interacting cooperatively with more than one protein that is bound to distinct promoters. CIP4 and CIP7 were first identified as COP1-interacting proteins. HY5 is another COP1 interacting protein (Ang et al., 1998). The COP1 interacting partners promote overlapping but distinct combinations of light regulated processes. For example: HY5 affects primarily hypocotyl elongation and anthocyanin accumulation whereas CIP7 promotes chloroplast development and anthocyanin accumulation in the light (Yamamoto et al., 1998). Another interacting partner, CIP4, promotes inhibition of hypocotyl elongation and chloroplast development (Yamamoto et al., 2001). Recently a model (Figure 2) for the COP1-mediated light control in Arabidopsis was described (Yamamoto et al., 2001). This model suggests that light signals are perceived by different photoreceptors and transduced to inactivate COP1 and negatively regulate its nuclear abundance. COP1 may negatively regulate diverse light induced processes by instigating the degradation of HY5 (Osterlund et al., 2000) and possibly other CIPs.

1.2.3. Cellular pharmacology of light signaling

Pharmacology approaches have been used to gain some understanding about the nature of the light signal transduction pathways. Neuhaus et al. (1993) designed a system to demonstrate that single cells of tomato *aurea* mutants (deficient in phyA but wild type for phyB) could respond to microinjected phyA. When dark grown *aurea* seedlings were transferred to light for 48 hours there was no significant chloroplast development and neither anthocyanin biosynthesis in the hypocotyl cells. When phyA was microinjected into *aurea* hypocotyl cells, chloroplast development and anthocyanin accumulation was observed. Therefore, this system allowed the selective analysis of the effects of a single phytochrome species at the single cell level. A range of compounds was microinjected

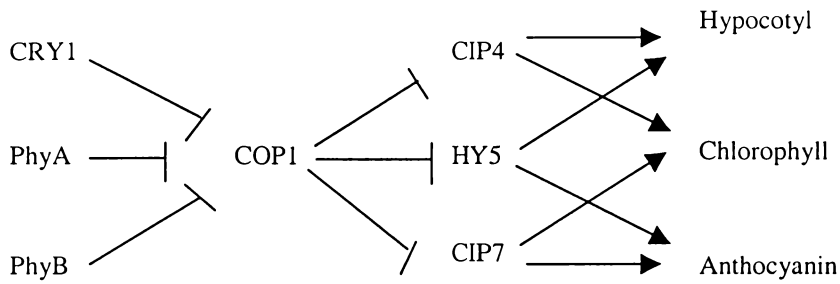


Figure 2: Model described by Yamamoto et al. (2001) for the COP1-mediated light control of Arabidopsis development. CIP4, HY5 and CIP7 are negatively regulated by COP1. Each of these targets is a positive regulator of distinct sets of light-controlled processes.

Into *aurea* hypocotyl cells in order to test putative signal-transduction components. Three distinct phytochrome signal-transduction pathways were described (Figure 3). The first is a calcium-dependent pathway that regulates the expression of the *CAB* gene and other components of photosystem II. The second is a cGMP-dependent pathway that regulates the expression of the gene encoding chalcone synthase (CHS) and the production of anthocyanin pigments. The third is a calcium and cGMP dependent pathway that regulates the expression of FNR (Ferredoxin NADP+ oxidoreductase) genes encoding the components of photosystem I, and is necessary for the production of mature chloroplasts (Neuhaus et al., 1993, Bowler et al., 1994). Analysis of transgenic plants overexpressing the heterotrimeric Gα protein provided physiological evidence of the involvement of heterotrimeric G-proteins in light-regulated seedling development. Gα may be involved in only a branch of the phyA signaling pathway (Okamoto et al., 2001). G-protein signaling is also important for plant responses to phytohormones (Ulla et al., 2001, Wang et al., 2001b) as shown by an increase in the frequency of cell division, a phenotype that mimics the effect of auxin on wild type cells, when the Gα-protein (GPA1) gene was overexpressed in tobacco (Ulla et al., 2001).

These microinjection results are important for two reasons: First, they pinpoint the contribution of biochemical signaling compounds in pathways leading to the activation of individual genes. Second, and perhaps more importantly, they imply a testable model for the architecture ('the wiring') of the light-signaling network. Based on these results, one might expect that genes residing on different output branches (*CHS*, *CAB*) may show

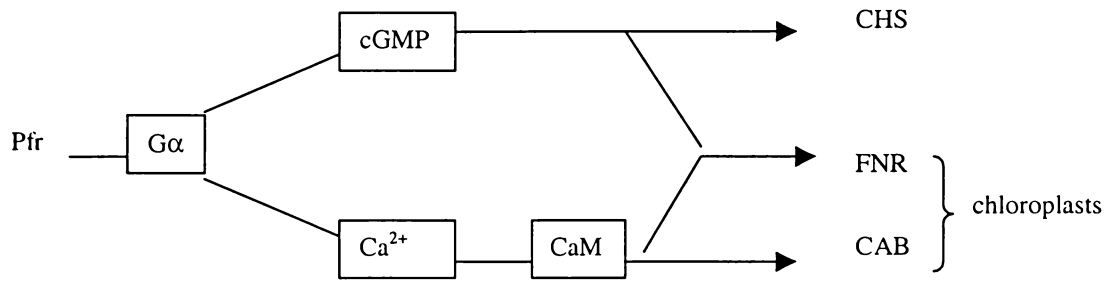


Figure 3. Model of Phytochrome signal transduction pathways described by Neuhaus et al (1993) depicting the mediation by heterotrimeric G proteins followed by calcium, cGMP and calcium/cGMP-dependent pathways.

different response profiles to light whereas genes residing on the same branch (*CAB*, *rbcS*) may show related response profiles. Unfortunately, this prediction cannot be tested easily because the model does not account for the contribution of phyB and other photoreceptors.

Some Arabidopsis mutants have been related to the biochemically defined phytochrome-signaling pathways in tomato. For example, the increased chalcone synthase expression (*icx1*) mutant shows elevated CHS expression but normal expression of the genes encoding the CAB protein (Jackson et al., 1995) and may designate a negative regulator of the cGMP-dependent pathway.

1.2.4. Mutants in specific branches of light signaling pathways

Transgenic plants have been developed to isolate mutants that are defective in light-regulated gene expression. Arabidopsis plants are transformed with reporter genes under the control of light-responsive promoters, followed by mutagenesis of homozygous transgenic offspring. This method allows the isolation of mutants with deficiencies in

signaling, based upon the gene expression. For example dark over-expression of CAB (*doc*) mutants have increased CAB mRNA levels, but normal levels of rbcS (Li et al, 1994). This is an indication of a branch point in the signaling pathways that regulate the expression of the two genes. The *doc* mutation may affect only the repression of the CAB genes in the dark but not the primary signaling pathway that stimulates their expression in the light since there is an induction of normal levels of CAB genes by light in the *doc* mutants (Barnes et al., 1997). The *doc1* mutation is due to a defect in the *BIG* gene that encodes a 560 KD protein containing several putative zinc finger domains. Multiple light-regulated genes in *doc1* mutants can be suppressed by elevated levels of auxin caused by overexpression of an auxin biosynthetic gene. Auxin transport is required for phytochrome-regulated hypocotyl elongation in light grown plants (Jensen et al., 1998). BIG appears to be essential for proper position of the auxin efflux carrier at the plasma membrane. This suggests that *BIG* is required for normal auxin efflux which in turn is required for the low expression of the *CAB* genes in the dark (Gil et al., 2001).

As described above, several approaches have been taken in order to study phytochrome signaling including promoter analysis, mutagenesis, cell biology and biochemistry. These approaches are reductionists in the sense that they focus on a specific branch of the light signaling network, either a specific input (e.g. phyA), output (e.g. CAB expression), or signaling intermediate (e.g. HY5). As a result we are now familiar with numerous detailed steps that are part of light signaling, but we know little of the overall level of complexity of the network. For example, even though we are aware that expression of *CHS*, *rbcS* and *CAB* are regulated differently, we don't know how many distinct ways there are for genes to respond to light.

1.3. Complex systems are amenable to modeling

A genetic network is defined as all the genes and their products participating in one regulatory event (Vohradsky, 2001). These networks can be viewed as maps of the cell, delineating potential signaling pathways (Tucker et al., 2001). They form intricate webs when information such as transcriptional and translational regulation and spatial and temporal expression patterns are included.

In order to define a comprehensive model of the light signal transduction network more information about the downstream components is necessary. Despite the advances in knowledge about the mechanism by which light sensors such as phytochromes and cryptochromes perceive light, we know very little about the complexity of the signaling network. Most of the described models of light signal transduction pathways are based only on one type of information: biochemical/cell biological data, mutants, or transcriptional regulation.

One approach to understand this complexity is to start with a simplistic and reductionist view of the signaling and add details that introduce new levels of complexity (Weng et al., 1999). Once enough variables have been defined experimentally, complex systems, including the light signaling machinery, should be amenable to modeling. As of today, this is illustrated by the following examples: the metabolic network of galactose utilization in yeast (Ideker et al., 2001), the transcriptional activation of a developmental gene in sea urchin (Yuh et al., 1998) and the epidermal growth factor stimulation of mitogen-associated protein kinases (MAPK) in mammals (Bhalla and Iyengar, 1999).

1.3.1. The genetic network of galactose metabolism in yeast

An integrated approach including microarrays, quantitative proteomics and databases of protein interactions was used to build a model of galactose (GAL) metabolism in yeast (Ideker et al., 2001). The authors defined all the genes, subsets of genes, proteins and small molecules that were involved in the galactose pathway. Then, they perturbed each pathway component through a series of genetic (gene deletions or overexpressions) or environmental (changes in growth conditions or temperature) manipulations. Then, the observed mRNA and protein responses were integrated with the current pathway-specific model and with the global network of protein-protein, protein-DNA and other known physical interactions. The authors used microarrays to measure the mRNA expression profiles and tested whether the changes in mRNA expression also reflected the changes in protein abundance. A catalog of previously observed physical interactions in yeast including protein-protein interactions and protein-DNA interactions was assembled. The genes associated with interactions in this catalog were affected in mRNA or protein expression by at least one perturbation or they were involved in two or more interactions with affected genes.

A graphical representation of these genes along with their associated interaction was displayed as a physical interaction network. In this network, genes linked by physical interactions had more strongly correlated expression profiles than genes chosen at random. The authors postulate that these correlations identify network interactions that are likely to have transmitted a change in expression from one gene (or protein) to another over the 20 perturbations applied. This means that a protein-DNA interaction may be responsible for directly transmitting a change on expression from a transcription factor to a highly correlated target gene. Conversely, correlated expression may be evidence for an interaction between two gene products. Alternatively two genes (A and B) may be under a control of a common transcription factor C and therefore coexpression of A and B provides evidence that C transmits these changes even if no detectable change in gene expression of C is observed.

To identify putative interactions the authors analyzed the Gal4p-binding site upstream of genes in three expression clusters found in the microarray, which contained all seven genes with established Gal4p-binding sites. Additional genes with Gal4p-binding sites were identified in this array. Based on previous genetic and biochemical research the authors defined an initial model of the molecular interactions governing pathway function and then integrated the observed mRNA and protein responses with that model. Observations that were not predicted by the model suggested new regulatory phenomena that could be tested (Ideker et al., 2001) by designing additional perturbation experiments. The model was refined through further iterations of perturbation and global measurements. These types of approaches are very powerful for suggesting new hypotheses about the regulation of a pathway and for understanding the interaction with other networks.

1.3.2. Quantitative modeling of the wiring diagram of a developmental promoter: The sea urchin Endo16 gene

A detailed structure-function analysis of a sea urchin promoter led to a model for the regulatory network that controls a gene encoding a secreted protein of the midgut (Yuh et al., 1998). This analysis revealed that the promoter that mediates complex developmental patterns of expression is modular in organization. Each cis-regulatory module contains target sites for DNA binding factors (approximately four to eight different factors per module). A computational model of the promoter was constructed which reflected logical functions hard-wired into the DNA (Yuh et al., 1998). One aspect of this model is of specific interest since it could be generally applicable: Immediately upstream of the TATA box module, but separate from it, the promoter contains exactly one 'master switchboard' module, whose function it is to accept signal input from four or five second-tier modules. The second-tier modules, which lie further upstream, have either positive or negative effects and are largely responsible for specific aspects of the tissue specificity and developmental time course of Endo16 gene expression.

1.3.3. Quantitative modeling of mammalian signal transduction pathways

Models for simple networks consisting of up to four signaling pathways were developed in order to determine whether the network has properties not present in the individual pathways and whether networking results in persistent activation of protein kinases after a transient stimulus (Bhalla and Iyengar, 1999). As an example of a signaling pathway, the authors modeled the epidermal growth factor (EGF) stimulation of Mitogen-Associated Protein Kinases, MAPK). Experimental data of (i) protein-protein interactions and enzymatic reactions such as protein phosphorylation and dephosphorylation and (ii)

protein degradation or production of intracellular messengers was included. A network model was developed in stages: First, individual pathways were modeled and then experimentally defined combinations of two or three pathways were tested and compared against published data. The pathways were linked by two kinds of interactions: (1) second messengers such as arachidonic acid and diacylglycerol produced by one pathway were used as inputs to another pathway and (2) enzymes whose activation was regulated by one pathway were used with substrates belonging to other pathways and the enzymatic reactions measured. The authors conclude that information for learned behavior of biological systems may be stored within intracellular biochemical reactions that comprise signaling pathways. These finding may also be relevant for plant light signal transduction because signal integration over time and adaptation are expected to be important there.

The complexity of signal transduction networks is exemplified by the intricate models described above. Analysis of linear pathways provides valuable information about the system properties such as the stimuli required to trigger a response and time courses for signal output (Weng et al., 1999). Tools such as expression of reporter genes in combination with high-resolution visualization techniques should allow the semi-quantitative estimation of molecular interactions (Weng et al., 1999). Integration of diverse sets of data such as microarray analysis, proteomics, and large-scale genetic analysis are invaluable for integrating and modeling signal transduction networks.

1.4. Holistic approaches to light signaling

Light signaling has been studied not only by analyzing the gene expression of individual genes, but also using holistic approaches. 'Holistic' in this context stands for a comprehensive, system-wide, description, with the premise that the resulting 'big picture' will provide constraints to our models of light signaling that would be difficult or impossible to gather from the reductionist, i.e. single-pathway investigation. Probably the first attempts in this direction were analyses of eleven genes for their detailed light-dependent changes in mRNA expression (Thompson et al., 1983, Kaufman et al., 1985). More recently, several new methods for high-throughput gene expression analysis have been developed, including serial analysis of gene expression (SAGE, Velculescu et al., 1995 and 1997), differential display and microarray technologies. The latter two techniques have been used to analyze light regulated genes (Ito et al., 1994, Kuno et al., 2000) and circadian-regulated genes in Arabidopsis (Harmer et al, 2000, Schaffer et al., 2001).

1.4.1. Analysis of mRNA expression

Earlier studies of light regulated genes were focused mainly on one or two genes. More ambitious studies included multiple transcripts. Phytochrome regulated changes in transcript abundance were analyzed for 11 different light regulated mRNAs in pea buds (Kaufman et al., 1985). Fluence response curves using red pulses showed that low fluence (LFR) and very low fluence (VLFR) responses could be observed for different transcripts (Kaufman et al., 1985). The fluence-dependent accumulation of mRNA in response to single pulses of R, the ability of FR to reverse the R effect and the ability of FR to induce mRNA accumulation in the absence of prior R treatments was evaluated. Based on these treatments it was possible to identify groups of genes that had similar response profiles and it was possible to set them apart from genes with distinct response profiles. For example, no response to red was observed in two transcripts, eight transcripts showed only a LF response and one transcript showed a small increase in response to FR alone.

In a subsequent study, the same transcripts were analyzed in response to two additional criteria, a time course of transcript accumulation in response to a red pulse and the 'escape kinetics' (Kaufman et al., 1986). The escape kinetics describes the extent to which a red light response can be reversed by a subsequent FR pulse as a function of the intervening period of darkness between R and FR. Escape kinetics studies used to be interpreted as an indication of how rapidly the 'active phytochrome' signal was being passed on to the subsequent light signaling step. Four general classes were observed. First, with respect to the time course, six transcripts accumulated at a linear rate during 24 hours in darkness following a red pulse. Two transcripts increased rapidly at first but then stabilized after 3 hours remaining at that level for the next 21 hours. Two transcripts exhibited a long lag period before they started to accumulate, reaching significant accumulation only after 12 to 16 hours after the red light pulse. One transcript displayed little or no change in the mRNA level in response to the red pulses. Second, with respect to the escape kinetics four different profiles were observed among 9 transcripts tested. Two transcripts began to escape from Pfr control immediately after the R pulse. Four transcripts had a lag between the inductive R pulse and the time at which signal transduction began to become irreversible. Finally, three transcripts did not display escape (i.e. they remained fully reversible by FR) even after 7 hours of induction. The authors concluded that the different transcripts are coupled to the phytochrome system in different ways and that the diversity in the light responses and the accumulation patterns of different transcripts reflects the underlying diversity of the signal transduction pathways involved. Even transcripts that accumulated with similar time courses differed in their escape kinetics which suggests that in most cases there is considerable independence between different regulatory programs (Kaufman et al., 1986).

1.4.2. Differential display

Differential display (DD) of mRNA is a powerful PCR based tool for identifying differentially expressed genes. In this technique, multiple samples of RNA can be screened for up-regulated and down-regulated genes (Liang and Pardee, 1992). DD is not quantitative and is primarily a gene discovery technique. Hence, the gene expression profile of individual candidate genes is subsequently determined by northern blotting.

Fluorescent differential display has been applied to identify genes under phyA control (Kuno et al., 2000). On the basis of 13 upregulated and two downregulated genes, the authors were able to demonstrate overlapping effects of phyA and phyB signaling pathways at the level of mRNA accumulation. The quantitative contribution of phyA and phyB appeared to be quite similar for all the genes tested. This result echoes results of partially redundant yet additive relationships between phyA and phyB pathways at the morphological and physiological level.

1.4.3. Microarray analysis

Microarrays involve printing of large numbers of different DNAs on a glass microscope slide in a gridded pattern. The expression level of individual mRNAs that are complementary to the arrayed DNA species is determined by hybridizing a fluorescently labeled copy of a biological mRNA sample to the glass array, followed by reading of the fluorescent signal using a scanning device (the same principle as a reverse northern blot). Microarray experiments typically compare the relative mRNA levels between two different biological samples by labeling them with different colors and hybridizing them to the same array. DNA microarray technology provides the means to measure the expression levels of thousands of genes in a single experiment. This makes this technology a powerful tool for gene discovery and the study of gene expression. Moreover, because clusters of co-regulated genes may be functionally related and are likely to be regulated by common pathways, microarrays also have the potential to reveal gene interactions (Basset et al., 1999; Brown and Botstein, 1999).

The application of microarrays to the field of light signaling in plants is still in its infancy. However, as expected, the first published accounts confirm that microarrays will become a powerful resource for light signaling (Tepperman et al., 2001, Wang et al., 2001a). Microarrays have the potential to reveal a truly global image of gene expression profiles under a large number of light conditions. However, in their current form,

microarrays do not distinguish reliably between closely related gene family members. In addition, in the process of mRNA preparation from small Arabidopsis plants, information about the tissue specific pattern of the mRNA is usually lost.

Light is one of the regulatory components of the circadian clock and clusters of circadian regulated genes have been found in pathways involved in plant responses to light and other metabolic pathways (Harmer et al., 2000). Temporal patterns of gene expression in Arabidopsis plants under constant light conditions were analyzed using microarrays in order to determine the genes that exhibited a circadian pattern of expression. Identification of genes with transcripts regulated in a diurnal cycle allowed the identification of clusters of genes regulated only by the circadian clock. A highly conserved promoter motif was identified to be required for circadian control of gene expression (Schaffer et al., 2001).

1.5. Promoter trap methods

Promoter trapping is an efficient method to screen for genes or regulatory regions based on their expression pattern. Promoter trapping involves the random integration of a promoterless reporter gene coding region at multiple independent sites in the genome. Each promoter trap line harbors one individual integration event and each line thus reports the properties of the locus-specific gene regulatory sequences within which the reporter gene has come to reside. Because the reporter gene also serves as a molecular tag, promoter traps that are of particular interest can then be analyzed easily at the molecular level (O'Kane and Gehring, 1987). Among the promoter trap constructs, gene traps carry a reporter gene without a transcriptional start site. In contrast, for enhancer trapping, the reporter gene comes with a minimal transcriptional start site, which only requires an enhancer in order to activate transcription. In plants, the most popular reporter gene is β-glucuronidase (GUS). In a promoter trapping project a large number of individual transformants that have the reporter gene integrated into different sites throughout the genome are generated and their progeny are analyzed for the expression of the reporter gene. Usually the assumption is made that the reporter gene expression mimics the expression of a neighboring chromosomal gene.

Distribution of the promoter trap element throughout the genome is commonly accomplished by direct transformation or by transposon-mediated mobilization. In Arabidopsis, direct transformation is now possible at high throughput, but one disadvantage is that *Agrobacterium* T-DNA transformation often results in multiple insertions per genome and in complex arrangements of multiple T-DNAs per locus. Transposon-mediated mobilization usually results in single insertions. Moreover, transposons can be re-mobilized, while T-DNAs are fixed. One of the earliest workable

systems developed was a two-component system in which a Ds-GUS transposon derived from the non-autonomous maize Ds element is under the control of an unlinked Ac transposase gene (Sundaresan et al., 1995). This system, which was used in the present work, offered the additional advantage of selection against transpositions to linked sites, a common problem with transposons. However, because the transposon is not highly active, generating large numbers of transposon containing lines (transposants) is labor intensive.

In this research project a collection of enhancer and gene trap lines was screened for their gene expression patterns in response to up to ten different light conditions in order to analyze the complexity of the light regulatory gene expression patterns. The analysis of promoter trap lines in response to environmental stimuli is a novel approach. Promoter traps give a direct measure of transcriptional activity and give access to a large panel of responses. Moreover, promoter traps are not subject to the problem of cross-hybridization between gene family members. One goal of this work was to identify different clusters of light response profiles as a way to define output branches of the light-signaling network. This is important in order to understand the "wiring" of the network, i.e. the connection between photoreceptors and responses.

Chapter 2. Generation and screening of gene and enhancer trap lines.

2.1. Introduction

Promoter trap analysis allows the identification and cloning of novel genes based on their expression patterns. Gene trap vectors consist of a reporter gene (in plants usually GUS) without any promoter and with a 3' splice acceptor site upstream of the reporter gene. The reporter gene is activated when inserted inside a transcription unit (Skarnes, 1990, Nussaume et al. 1995, Campisi et al., 1999). Enhancer trap vectors consist of a reporter gene with a minimal promoter, that is, a TATA box and transcription start site. The minimal promoter is not sufficient to drive expression of the reporter gene (Skarnes, 1990). The enhancer trap system reveals the expression pattern of genes in the vicinity of the insertion site.

For Arabidopsis, Sundaresan et al. (1995) developed a transposon-mediated promoter trapping system. Enhancer trap and gene trap elements were constructed and subcloned into T-DNA vectors for Agrobacterium-mediated transformation of Arabidopsis. These transposons are referred to as 'donor sites'. Transposons that have moved to a new position are referred to as 'gene trap lines' or 'transposant lines' or simply 'transposants'.

For this work, enhancer trap and gene trap lines were screened for their response to different light conditions. Some of the lines were newly generated for this project and other had been generated previously at Cold Spring Harbor Laboratory. A total of 2098 enhancer trap (1048) and gene trap (1050) lines were screened under constant light or constant dark conditions in order to determine which lines showed any expression under constant light and/or constant dark. A set of lines was identified in which the GUS reporter gene was responsive to constant light and constant dark. In the second step, a group of these lines was then analyzed for their response to a shift from light to dark and from dark to light. The purpose of this experiment was to distinguish between the genes that respond rapidly to either light or dark and those genes that respond slowly to those conditions. A set of rapidly responsive genes was observed. A more detailed characterization of these lines for their light regulation is the subject of chapter 3.

2.2. Rationale

The identification of a large set of light regulated genes should eventually lead to a comprehensive characterization of the types of light responsiveness that are encoded within the light signaling network. This in turn will give us a clue about the "wiring" between photoreceptors and responses, i.e. the internal structure of the light signal transduction network by providing information about the genes that are regulated

together in response to a stimulus. Unlike typical RNA expression profiling methods, gene and enhancer trap systems preserve the tissue-specificity of expression. In addition, one can observe how genes respond to different environmental conditions. In the following, the term ***'pattern'*** shall denote a tissue-specific, ***spatial*** variation in the response of an individual gene. In contrast, the term ***'profile'*** shall denote a variation in expression of a gene over a *range of environmental conditions*. The characteristics of the gene trap and enhancer trap elements make them useful for studying the light signaling pathways by analyzing the tissue specific patterns of expression as well as the light specific response profiles of the genes.

2.3. Materials and methods

2.3.1. Generation of gene and enhancer trap lines

Two types of transgenic starter lines were obtained from the Arabidopsis Stock Center at Ohio State University. One set of lines carries the promoter trap, linked to a kanamycin resistance gene (NPTII) as a selectable marker, between the 5' and 3' ends of the non-autonomous transposable element Ds ('donor lines'). Another line expresses the cognate transposase of the maize transposable element Ac under the control of the CaMV 35S promoter. The construction of the Ac and Ds elements is fully described in Sundaresan et al. (1995). Sundaresan also inserted a negative selectable marker gene, the IAAH gene (indole acetic acid hydrolase), next to the Ds element to select for transposition of the Ds element to unlinked sites. The IAAH gene confers sensitivity to naphthalene acetamide (NAM) by conversion of NAM to the synthetic auxin naphtalene acetic acid which inhibits seedling growth (Karlin-Neumann et al., 1991). An IAAH gene was also placed next to the 35S: Ac gene to counter select against Ac if necessary. Plants homozygous for the transposon ([Ds::GUS; Kan-R]; IAAH) were crossed to plants homozygous for the transposase enzyme gene (35S-Ac; IAAH). The seeds of this cross (F1) were then sown on soil and allowed to self (Figure 4). The F2 seed batches were selected for the small fraction of progeny that carries the transposed Ds::GUS element but does not carry the original Ds::GUS site (donor site), nor the 35S-Ac transposase gene. These seedlings are the ones able to grow on germination medium (4.3 g/l MS salts [GIBCO/BRL], 10 g/l sucrose [tissue culture grade, Sigma], pH to 5.7 with 1 M KOH and 7 g/l agar) containing 0.65 µg/ml NAM and 50 µg/ml kanamycin. The IAAH gene also serves to select against F2 seedlings carrying an Ac element, thus stabilizing the transposed Ds element against secondary transposition. Plants resistant to kanamycin and NAM (transposants) were selected and a second screening of the seedlings was performed in fresh medium with NAM and KAN, in order to assure the selection of transposants. Approximately 10 days after the transfer, the resistant seedlings (F2 generation) were transferred to soil and

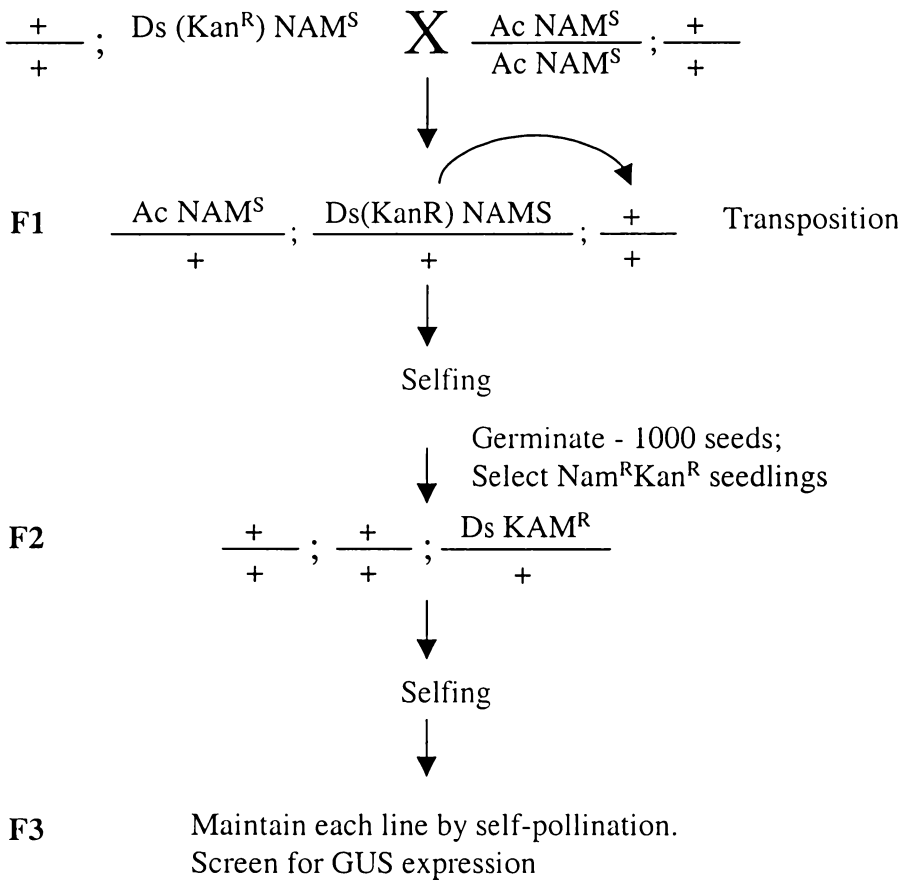


Figure 4. Generation of lines carrying transposed Ds element by selection of unlinked transposition events (Sundaresan et al., 1995).

allowed to self (F3 generation, Figure 4). Each family of F3 seeds and their progeny are referred to as a 'transposant line'.

2.3.2. Response of gene/enhancer trap lines to constant light or constant dark

In order to find a group of genes with light responsiveness, F3 Arabidopsis seeds of the transposant lines were plated on germination medium and incubated for 24 hours at 4°C to improve germination (stratification). Plates were then incubated for 5 five days under either constant light or constant dark after which at least 4 seedlings were stained for GUS in order to determine the degree of gene expression under those conditions. The constant light treatment was provided in a Percival Scientific ® incubator Model CU-32L with General Electric ® bulbs (Trimline T8, F17T8.SP41 17 watt) at a distance of 16 cm. For the staining, seedlings were incubated in 300 µl GUS solution (100 mM sodium phosphate pH 7, 10 mM EDTA, 0.1% Triton X-100, and 1 mg/ml 5-bromo-4-chloro-3-indole-β-D-glucuronide, X-GLUC) in a 24-well microtitre plate. The plates were placed in a vacuum dessicator for 10 minutes, sealed to prevent evaporation and then incubated in the dark at 37°C for 48 hours. After this period of time the staining solution was removed and replaced with 70% ethanol (clearing). The clearing process was repeated two or three times when necessary to remove residual chlorophyll. Stained seedlings were examined under a dissecting microscope.

2.3.3. Light-Dark shift experiment

Seeds were plated on germination medium and after 24 h of cold treatment the plates were incubated either for 5 days in cL or cD, 4 days of light and a shift to darkness for 24 hour or 4 days, 12 hours of darkness followed by transfer to light for 12 hours. After the light treatments, at least 4 seedlings of each line and each treatment were stained for their GUS expression as described before. In order to evaluate if lines that were GUS negative under both cL and cD might respond to the shift, 66 lines with this pattern were evaluated for their response to the shift. No staining was observed in these lines after the shift. Therefore, they were not used for further experiments. The data were organized by using the 'Access®' database (Appendix 4, Plate 1).

2.4. Results

2.4.1. Generation of transposant lines

Transposon-mediated gene/enhancer trapping gives a direct measurement of the transcriptional activity at a large number of chromosomal sites and it allows the study of a large panel of genes. A transposon system developed by Sundaresan et al. (1995) was

used for the generation of transposant lines in our lab. Table 1 shows the number of F2 families screened for transposition events and the number of families that contained transposants. Approximately equal numbers of DsG and DsE lines were screened. The percentage of transposants recovered was similar for the DsE (20%) and DsG (15%) elements.

The kanamycin resistance (Kan^R) segregation of 40 of the F3 lines was evaluated in order to determine if the transposants selected contained single insertions. Of these, 24 lines had a segregation of ¾ Kan^R and ¼ Kanamycin sensitive (Kan^S) seedlings, confirming that the lines selected contained single insertions. Sixteen lines were completely Kan^R as expected for homozygous transposants. The transposant lines generated at the University of Tennessee were named AJ144, AJ146 (5 lines), AJ170 (3 lines), LH201 (7 lines) LH202 (2 lines), LH203 (1 line), LH204 (4 Lines), LH205 (7 lines), LH206 (6 lines), LH207 (3 lines), LH208 (10 lines), LH209 (5 lines), LH210 (41 lines), LH211 (10 lines), LH212 (31 lines), OK001 (15 lines), OK003 (11 lines), OK004 (2 lines), OK005 (11 lines), OK006 (1 line), OK007 (14 lines), OK008 (4 lines), OK009 (10 lines), OK010 (2 lines), OK011 (6 lines). In addition to these 214 transposant families, 1884 transposant families were screened that had previously been generated under the direction of Drs. R. Martienssen and V. Sundaresan at Cold Spring Harbor Laboratory. These enhancer trap and gene trap lines are referred to as ET5125 to ET5838; ET6361 to ET6659 and GT5870 to GT6740.

2.4.2. Response of gene/enhancer trap lines to constant light or constant dark

Studying the transcriptional activity of transposon-tagged genes under different light conditions can give us a clue about the internal complexity of the light signaling network and the connection between photoreceptors and responses. A total of 1048 enhancer trap and 1050 gene trap lines were screened under constant light or constant dark conditions. At least 4 seedlings from each line and each treatment were stained for GUS and the staining patterns analyzed under a stereomicroscope. Levels from 0 (no staining) to 3 (strongest staining) were assigned to the degree of staining of a particular organ (Figure 5).

The GUS staining was recorded for the following organs and tissues: Root: root tip (RT), vasculature (V), root non-vascular tissue such as cortex or epidermis (CE), root hairs (RH) and colet (C, i.e. the root base). Hypocotyl: vasculature (V) and hypocotyl non-vascular tissue such as cortex or epidermis (CE). Cotyledon: vasculature (V), cotyledon non-vascular tissue such as mesophyll or epidermis (ME), hydathodes (H) and petiole (P). Primary leaves: vasculature (V), leaves non-vascular tissue such as mesophyll or epidermis (ME), hydathodes (H), trichomes (T) and petiole (P). Shoot apex: General

Table 1: Number of F2 families screened and F3 transposants observed in Arabidopsis seedlings screened for transposition events. DsG and DsE lines were crossed to Ac lines and the F2 generation families screened for NAM/KAN resistance in order to isolate those seedlings carrying transpositions.

Element	No. of F2 families screened	No. of transposants (F3 generation)	Percentage
DsG	699	102	15%
DsE	563	112	20%
Total	**1262**	**214**	**17%**

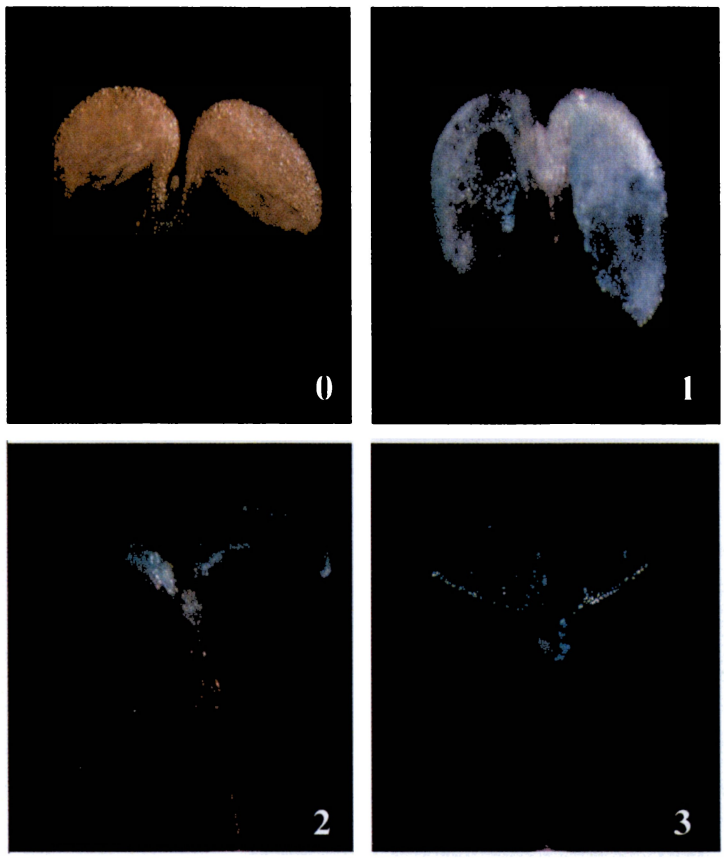


Figure 5. Classification of the degree of GUS staining (0 to 3) of five-day old seedlings. Seedlings were stained at 37°C for 48 hours, clarified with 70% ethanol and analyzed under the stereoscope.

staining (SA) and shoot apex vasculature (V). Lines were grouped according to their response to constant light and dark. The seven groups included: "no stain" (NS), if there was no GUS staining under constant light and constant dark. "cL" or "cD" if staining was observed only under light or dark, respectively, but not under both conditions. "L=D" if both treatments showed the same level of staining; "L>D" for stronger staining in light than dark; "D>L" for stronger staining in dark than light and "C" for complex. In the complex group, one organ may respond stronger under one condition (for example constant light) whereas another organ responds stronger under the other condition (for example constant dark, Figure 6).

One example of a complex pattern is given in Table 2 and Figure 7 (line ET5627). In this line, the cotyledons were stained under cL but not under cD. The tissues that stained were mesophyll/epidermis (ME), vascular tissue and petiole. In contrast, in the hypocotyl, there was equal staining in the cortex/epidermis (CE) in cL and cD, but the vascular tissue stained only under cD conditions and not under cL (Table 2). The root showed staining in the cortex/epidermis (CE) under both cL and cD conditions. These data show that the light response of ET5627 is tissue specific.

For both the enhancer trap and gene trap lines, the most common pattern of expression was no staining in light or dark (NS) comprising 73% of the lines analyzed, followed by the complex pattern, 12%, followed by cL, cD, L=D, D>L and L>D (Table 3, Figure 8). The large number of NS lines is not unexpected. Given the gene density in the Arabidopsis genome (the Arabidopsis initiative, 2000), enhancer elements are expected to be limited in the distance over which they can activate transcription. Therefore, most insertions remain silent. However, it is also possible that a fraction of the 75% silent lines will show GUS expression at stages of development or under environmental conditions other than those tested in this work.

In this screening, 28% (297/1048) of the ET lines and 25.6% (269/1050) of the GT lines screened showed some gene expression under the conditions evaluated (Table 3). In a previous experiment, Sundaresan et al. (1995) reported that 48% of the ET lines and 26% of the GT lines displayed some GUS expression at the same seedling stage. The reasons for the lower percentage of GUS positive enhancer traps in this study are not clear.

The representation of the various modes of light responsiveness among the subset of GUS positive lines is shown in a histogram in Figure 8. The large fraction of lines showing tissue-specific light responsiveness (complex pattern) was surprising because few examples for this phenomenon have been published on the basis of transgenic promoter: reporter fusions in plants. Are the ET lines, which only require a nearby enhancer element, more likely to show a tissue-specific light response than the GT lines,

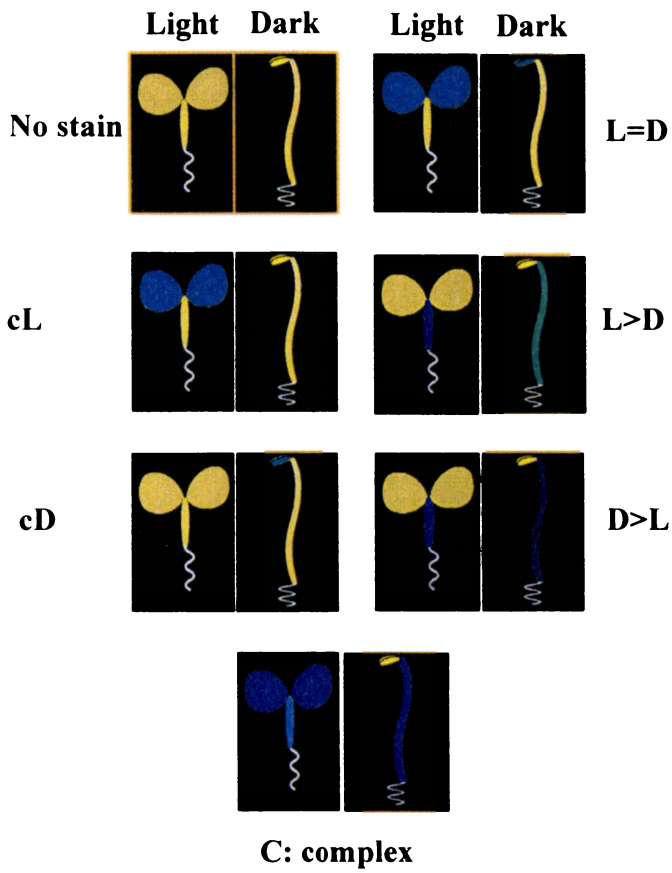


Figure 6. Classification of the GUS staining patterns of five day old seedlings. This classification is used throughout the experiments described in this document. Seedlings were stained at 37°C for 48 hours, clarified with 70% ethanol and analyzed under the stereoscope.

Table 2. Example of a complex staining pattern (line ET5627). There is staining in the cotyledon under cL but not under cD and equal staining in the hypocotyl and root cortex-epidermis under cL and cD. The vascular tissue of the hypocotyl stains under cD but not under cL.

Treatment	Cotyledons	hypocotyl	Root
Constant Light	Mesophyll-Epidermis 3 Vascular tissue 3 Petiole 2	Cortex-Epidermis 2	Cortex-Epidermis 1
Constant Dark	no stain	Cortex-Epidermis 2 Vascular tissue 3	Cortex-Epidermis 1

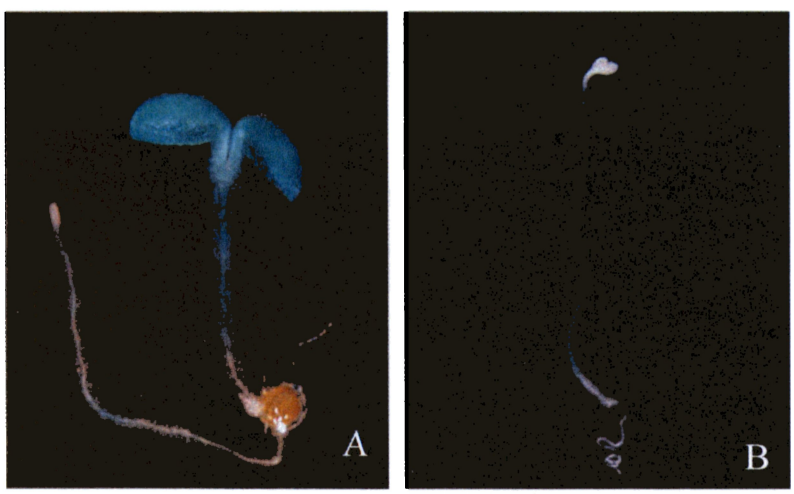


Figure 7. Example of a complex staining pattern (line ET5627). As shown in Table 2, the cotyledon stains under cL but not under cD. There is equal staining in the hypocotyl cortex-epidermis in both cL and cD. The vascular tissue of the epidermis stains only under cD but not under cL (as seen under the microscope). There is also faint equal staining in light and dark in the root.

Table 3: Number of transposant lines responding to the constant light or constant dark. See text for details.

Trap	NS	CL	cD	L=D	L>D	D>L	complex	Total
Enhancer	751	84	46	40	8	7	112	1048
Gene	781	37	63	16	2	8	143	1050
Total	1532	121	109	56	10	15	255	2098

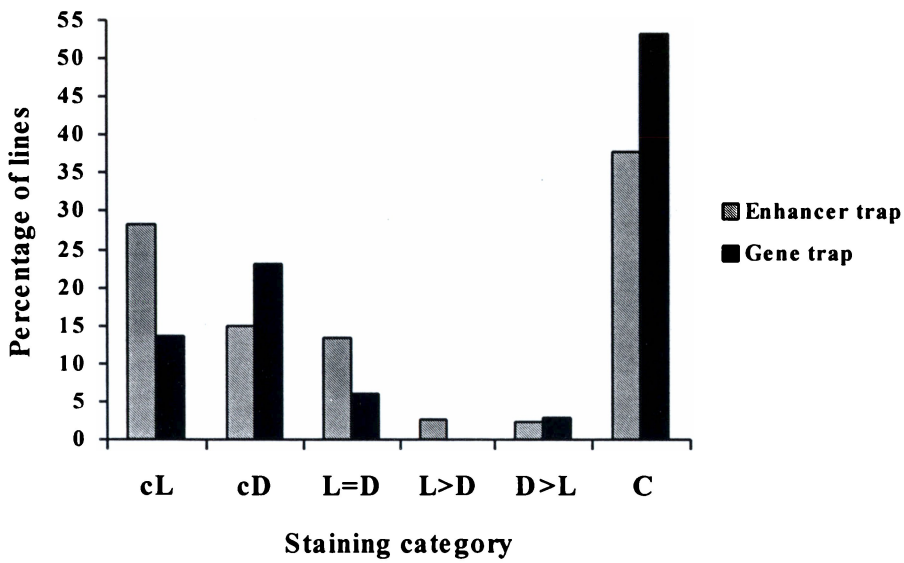


Figure 8. Characterization of the light responsiveness of GUS positive GT and ET lines under constant light and constant dark. For details see text.

which must be inserted into a transcription unit and are therefore more likely to reflect the expression pattern of the flanking endogenous gene? Figure 8 suggests that ET and GT inserts are equally likely to be subject to tissue-specific light responsiveness. This in turn suggests that numerous Arabidopsis genes show tissue-specific light responsiveness.

2.4.3. Light-Dark shift experiment

In order to distinguish between genes that respond rapidly to either light or dark and the genes that respond very slowly to these conditions, a total of 365 lines that responded differentially to the cL and cD conditions were analyzed for their response to a shift from light to dark (LD) or from dark to light (DL). A strong response was defined as at least a two-step increase or decrease in GUS staining (for example from level 1 to level 3; Figure 4). Classification in response to the shift (Figure 9) was as follows: (1) *No response* to the shift, (2) *Inducible*, if the shift from light to dark or dark to light induced a response that is in line with predictions from the cD and cL treatments (two examples are given); and (3) *'Other'* if the response did not fit the 'no response' or 'inducible' categories. For the purpose of illustration, in the example shown, the response to the LD shift in the cotyledon is not in line with the results from the cD and cL treatments.

No response to the shift was observed in 21.6% of the lines (Table 4, Figure 10), indicating that those genes show either slow responses to the environmental conditions evaluated. Alternatively, the gene may be light responsive early in development (for example during the first 2 days), leading to differential accumulation of the stable GUS protein. However, the gene may have lost its responsiveness to light by the time the light shifts were performed on the fifth day. Since the purpose of this experiment was to identify a set of genes that showed a rapid response to changes in the light conditions, the genes that did not respond to the shift were not considered for further experiments. The most common pattern of expression observed was the "other" pattern (Table 4, Figure 10).

2.5 Discussion

Transposition frequency is measured as the percentage of F2 progeny with a transposed element (Schmidt and Willmitzer, 1989, Altmann et al., 1992). Out of 1262 lines screened, 17% contained transposants. Different frequencies of transposition have been reported in different systems. Previously observed frequencies of transposition ranged between 5-45% in Arabidopsis (Long et al., 1993). Another element, the maize transposable element En-1 proliferates rapidly in *A. thaliana* as shown by rapid changes of the number of En-1 insertions through selfed generations (Wisman et al., 1998). This system is proposed to be especially suitable for high frequency mutagenesis, but the

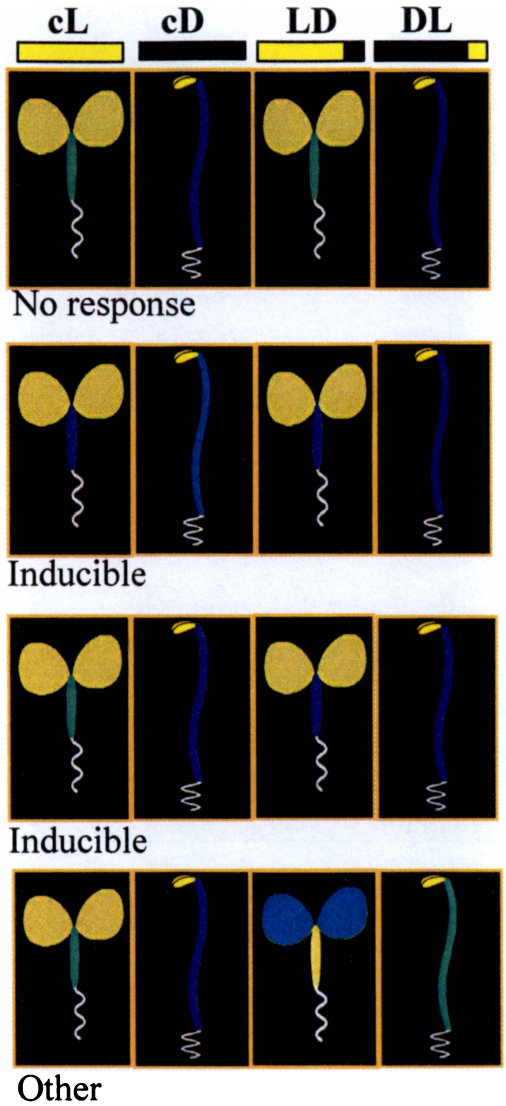


Figure 9. Classification of the patterns of expression in response to a shift from constant light to dark (LD shift) or from constant dark to light (DL shift) of five day-old seedlings. Lines were classified in 3 groups: (1) *No response* to the shift, (2) *Inducible*, and (3) *Other*. The yellow and black bars represent light regimes (see details in the text).

Table 4. Number of transposant lines responding to the shift from cL to dark and cD to light. Seedlings were grown either 5 days in cL or cD, or 4 days in cL and a shift to darkness for 24 hours or 4 days, 12 hours in cD and transfer to light for 12 hours. Lines from the cL, L>D and the cD, D>L categories in Table 3 are reported as belonging to the L>D and D>L categories respectively.

		Response to constant light or constant dark			
		L>D	**D>L**	**Complex**	**Total**
Response to the shift	No response	8	23	48	79
	Inducible	13	39	78	130
	Other	18	26	112	156
	Total	**39**	**88**	**238**	**365**

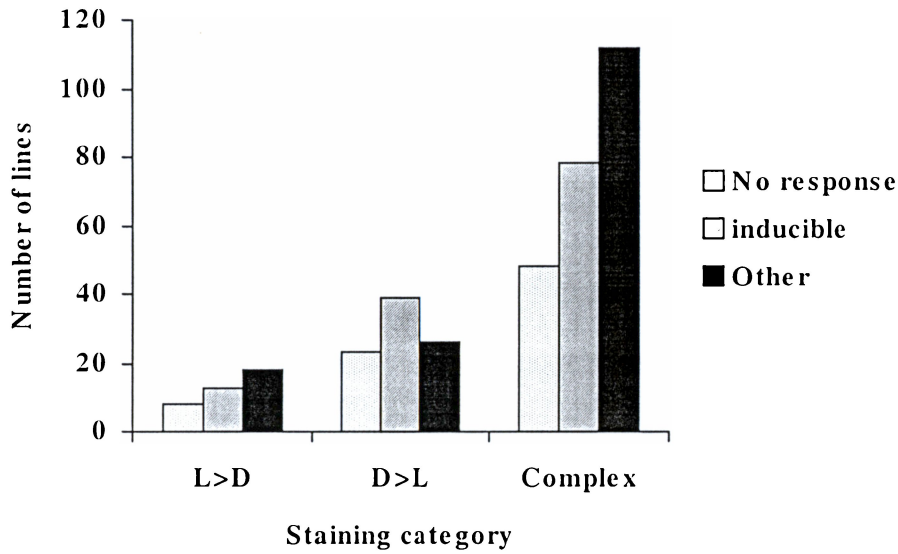


Figure 10. Characterization of the light responses to the shift. The experimental conditions were described in Table 4 and in the text. The X-axis displays the response to the constant light and dark and the colored bars indicate the response to the shift. L>D, stronger in light that dark or staining only in constant light. D>L, stronger in dark than light or staining only in cD. Complex, if one organ stains stronger in one treatment and another organ stains stronger in another treatment. Other: If the response does not fit the 'no response' or 'inducible' category.

increased number of transposons per genome is problematic. In rice, transposition events between 3-20% have been reported (Nakagawa et al., 2000, Enoki et al., 1999). The transposition frequency observed in my work falls within the range of transposition observed in other systems.

I was surprised that many promoter trap lines showed tissue-specific light responses because few examples for this phenomenon have been published based on transgenic promoter:reporter fusions in plants. Perhaps a more in depth analysis of existing conventional promoter:reporter lines under diverse light conditions would confirm the regularity of such tissue-specific responses. Our models of the light signal transduction network must eventually incorporate these tissue-specific transcriptional inputs that lead to the complex expression patterns observed in this work.

Differential light responsiveness in the expression of a gene in different tissues has been reported occasionally (Sheng et al., 1993; Walter et al., 1996; Serikawa et al., 1997), for example, for the knotted1-like KNAT3 promoter:GUS fusion (Serikawa et al., 1997). Cotyledons, upper hypocotyl and roots stained stronger for GUS under continuous white light than under continuous red light. Under continuous far red, there was stronger staining in hypocotyl and cotyledons but not staining in the roots.

In my work, some errors in the appreciation of the staining in a particular cell type against a strong background staining in neighboring tissues are likely to occur. This was a qualitative analysis and therefore the data should be taken as just an approximation of the tissues being stained. For example, in most of the cases it was impossible to determine whether there was staining in the vascular tissue if the surrounding non-vascular tissue was stained very strongly. In these cases, the vascular tissue was recorded as being positively stained as well. However, in some instances it was clearly observed that even though the non vascular tissue stained, the vascular tissue did not. In those cases, the degree of staining for vascular tissue was recorded as '0' (as illustrated in Figure 5).

Enhancer and gene trap lines were generated and screened for their response to constant light or dark and to shifts from light to dark and dark to light. A set of 286 rapidly responsive genes was identified (Table 4: lines in the category 'Inducible' and 'Other' in response to the shift). Therefore, the Ac/Ds elements are indeed suitable to identify light regulated promoter traps. Compared to other large-scale screens for light responsive genes, such as microarrays or differential display, promoter traps have the advantage of preserving the tissue-specificity of a transcriptional response.

Chapter 3. Analysis of the tissue specificity of light responses

3.1. Introduction

Tissue-specific factors and light must work in concert to regulate the expression of light regulated genes (Chory et al., 1995) as shown by genes that are activated in cell-type specific patterns such as the light-harvesting chlorophyll a/b binding proteins (CAB), the small subunit of the ribulose biphosphate carboxylase/oxygenase (rbcS) and the chalcone synthase (Li et al., 1993). The rbcS and CAB mRNA's are present in low amounts in etiolated tissues and increase in abundance after red light activation of phytochrome (Fluhr et al., 1986). The *rbcS* genes are expressed abundantly in leaves, moderately in stems and at low levels in roots (Coruzzi et al., 1984). The tissue-specificity of rbcS can be traced to specific promoter fragments (Meier et al., 1995). Transcripts of the *CAB* gene family are abundant in leaves but lower or undetectable in roots and hypocotyl (Nagy et al., 1986, Gilmartin et al., 1990). The light and tissue specific regulation of the CAB and rbcS genes may be mediated through overlapping cis-regulatory elements (Chory and Peto, 1990). Little emphasis has been placed on how tissue specific expression patterns respond to environmental conditions. Therefore we know little about the integration of the signaling pathways that mediate tissue-specific and environmental responses.

The photomorphogenic process can be divided into light reception, signal transduction and selective induction of gene expression. Phytochrome photoreceptors play an important role in light signaling in plants. Individual phytochromes do not act independently but are part of a complex network of interactions with other members of the phytochrome family and with blue light photoreceptors (Casal, 2000). Therefore individual genes often respond to combinations of phyA, phyB and blue-light receptors. It is thought that most, if not all plant cell types are light responsive. Accordingly, promoter:GUS constructs (Somers and Quail, 1995) have revealed that both *phyA* and *phyB* are expressed in roots, shoots and flowers during the entire life cycle of *Arabidopsis*. However, there are striking differences in the expression levels and photoregulation of both genes, in that phyA expression is more prominent in etiolated seedlings and is repressed by light. As outlined in the main introduction, constant far-red light (cFR) operates through phyA, whereas constant red light (cR) is processed via the phyB pathway.

3.2. Rationale

Light responses occur in a diverse range of tissues in Arabidopsis and other plants. While most light inducible gene expression events have been identified in photosynthetic cell

types, other cell types also respond to light. For example, gravitropism in the root is dependent on light conditions, and hypocotyl cells respond to phytochrome activation with inhibition of cell elongation. It is also evident that particular organs, such as the hypocotyl, can display many of the photomorphogenic response modes, including VLFR, LFR, and R and FR-HIR. What is not clear however, is whether all organs, tissues, or cell types basically possess the same capacity to process light signals, or whether certain response modes are restricted to a subset of cell types. One prerequisite for a full set of response modes in all cell types would be that the underlying photoreceptors are ubiquitously expressed. phyA and phyB are indeed expressed in essentially all cell types investigated (Somers and Quail, 1995), but for cryptochromes and other phytochromes this question has yet to be answered. In any event, restrictions on the usage of different response modes might exist even if all phy's and cry's were ubiquitously expressed.

The gene/enhancer trap lines were subjected to cR and cFR in order to characterize the relative sensitivity of the lines to the pathways gated by phyA and phyB. cR and cFR initiate the so-called high irradiance responses (see Introduction). In addition, the phytochrome signaling network mediates two other response modes, the inductive low fluence response (LFR) and the very low fluence (VLFR) response. Experimentally, the LFR and VLFR are conveniently activated by light pulse treatments of dark grown seedlings. Therefore, selected promoter trap lines were also subjected to four different light pulse treatments (Chapter 3). Table 5 summarizes the rationale for the design of the ten light treatments used to define light response profiles. The tissue specificity was analyzed only for the HIR (cL, cD, cR and cFR). The results from the pulse treatments in terms of tissue-specificity were included in Appendix 1. However, the analysis was focused on cL, cD, cR, and cFR.

3.3. Materials and methods

3.3.1. Red and Far-red experiment

Seeds were plated on germination medium and after the cold treatment (24 h) the plates were incubated under different light conditions: cL (60 μmoles/m^2.s), cD, LD: shift from light to dark, DL: shift from dark to light (refer to chapter 1 for details), constant red (cR, 0.5 μmoles/m^2.s), constant far-red (cFR, 15 μmoles/m^2.s). Red pulses (Rp, 4 μmoles/m^2.s) or far-red pulse (FRp, 130 μmoles/m^2.s) treatments were given as ten hourly sets of pulses of 5 minutes duration followed by incubation in the dark for 16 hours. The R-FRp and FR-Rp treatments consisted of 10 hourly sets of pulses of either 3 min of red or far red followed by 5 min of FR or R respectively. The white light was provided in a Percival Scientific incubator as described in chapter 1. The red and far-red treatments were provided with Snap-Light ® LEDs (high efficiency light emitting

Table 5. Rationale for the design of the ten light treatments used to define the light responsive expression profiles of promoter traps.

Treatment	Response mode	Rationale
cL	HIR	Multiple photoreceptors including phyA, phyB and cryptochromes.
cD	-	absence of light
LD	HIR	Test for 'rapid' dark responsiveness during a 24h time window
DL	HIR	Test for 'rapid' light responsiveness during a 12 hour time window
cR	HIR	Mediated by phyB and phyB-like phy but not phyA
cFR	HIR	Mediated by phyA
Rp	LFR	Inductive response mediated by phy
R-FRp	LFR	Test for reversibility of the inductive LFR response to Rp
FR-Rp	LFR	Control for R-FRp: same photon fluence but different time course
FRp	VLFR	Mediated by phyA

diodes) from Quantum Devices at a distance of 14 cm from the petri plates. Light intensities were determined with a LI-COR-1800/22 spectroradiometer. After the treatments the seedlings were stained for GUS and the data recorded as described in Chapter 1. The complete data was included in Appendix I.

3.4. Results

In order to define the types and numbers of light response profiles among Arabidopsis promoter trap lines, I focused on those lines that showed a strong response to a light-to-dark shift or a dark-to-light shift response (at least a two step increase or decrease in GUS staining). The rationale was that these lines were less likely than others to reflect gene expression responses at early developmental time points. In contrast, these lines are likely to represent immediate outputs of the light signaling network, which might respond to brief pulses of light during a ten hour time period, as planned for the following experiment. In a preliminary characterization, 77 promoter traps that had been identified as shift-responsive (Table 4 'Inducible and Other' in Chapter 1) were analyzed for their histochemical responses to cR and cFR.

From among 286 lines (Table 4 Chapter 1) with recognizable light responsiveness, 33 ET and 43 GT lines were chosen that had moderate or strong staining, rather than weak staining, and that stained in a substantial fraction of the seedling. Of the lines selected, 6 lines were 'light inducible', 22 'dark inducible' and 48 had the 'Other' pattern as described. Histochemical staining patterns were recorded as previously described after the light treatments specified in the 'rationale' and 'methods'. Data from previously described experiments (Chapter 1) are presented together with those from new experiments. The raw data for all individual experiments are detailed in Appendix 1. Most light treatments were repeated, typically 2 or 3 times and sometimes up to 6 times.

3.4.1. Experimental variability in staining patterns

An example of the in situ GUS staining pattern is illustrated by line ET5627 (Table 6). For the purpose of discussing the 'between-experiment' differences in staining, the differences are color coded in this table. Data that are consistent between all four experiments are in black. Data points with only one repeat are in green (DL and LD shift). Slight discrepancies in the degree of staining are in blue. Finally, results that represent a substantial departure from the consensus are highlighted in red. It is clear that the majority of the observations were reproducible. However, differences in the degree of staining were recorded quite frequently (blue). It is possible that most of these differences arise from the difficulty of maintaining a consistent standard for staining intensities over many weeks of data collection and hundreds of lines. Major discrepancies in the staining

Table 6. Line ET5627 serves as an example of the in situ GUS staining pattern in response to ten different light treatments. This line is also representative for the 'between-experiment' variation of staining. Data are presented as a letter depicting a tissue within an organ followed by the level of staining. The staining tissues include: cortex-epidermis including non vascular tissue (CE), petiole (P), Vascular tissue (V) and root hairs (RH). For color-codes see text.

Experiment 1					
Treatment	**Shoot apex**	**Cotyledon**	**Hypocotyl**	**Root**	**Immature leaf**
cL	-	CE3 V3 P3	CE2	-	-
cD	-	-	CE3 V3	V3	-
LD	-	**CE3 V3 P3**	**CE3 V3**	**CE1**	-
DL	-	**CE1**	**CE3 V3**	**V3**	-
		Experiment 2			
cL	-	CE3 V3 P3	CE2	V1	-
cD	-	-	CE2 V3	CE1 V3	-
cR	-	CE2 V3	CE2 V3	CE1	-
cFR	-	CE2 V3	CE2 V3	CE1 V3	-
		Experiment 3			
cL	-	CE3 V3 P3	CE2	CE1 RH1	-
cD	-	-	CE2 V3	CE1	-
cR	-	CE1 V2 P1	CE2 V3	CE1	-
cFr	-	CE1 V2 P1	CE1 V2	CE1	-
Rp	-	CE1 V1	CE2 V3	CE1 RH2	-
R-FRp	-	-	CE2 V3	CE1	-
FRp	-	-	CE1 V3	CE1 V1	-
FR-Rp	-	-	CE2 V3	CE1 V1	-
		Experiment 4			
cL	-	CE3 V3 P2	CE2	CE1	-
cD	-	-	CE2 V3	CE1	-
cR	-	CE1 V2 P1	CE2 V3	CE1	-
cFr	-	CE1 V1 P1	none	CE1	-
Rp	-	-	CE2 V3	CE1 RH2	-
R-FRp	-	-	CE2 V3	CE1	-
FRp	-	-	CE1 V3	CE1 V2	-
FR-Rp	-	-	CE2 V3	CE1 V2	-

pattern also occurred, but less frequently (red, which stands for discrepancies in the level of staining by more than one level, including absence of staining). ET5627 is a representative example but similar cases can be found in Appendix 1. For example, the strong staining of the hypocotyl in cFR in Exp. 2 and 3 was not recorded for Exp. 4. Moreover, strong staining of the root vasculature in cD and cFR was seen in Exp. 1 and 2 but not in Exp. 3 and 4. In principle, both minor differences (blue) and major discrepancies (red) may be due to human data collection errors or variation in the plant material per se. However, it seems likely that human error accounts for a larger fraction of the minor differences than the major discrepancies. (note for example, that the spurious root vascular staining in ET5627 is confined to all non-photosynthetic conditions in experiments 1 and 2, suggesting an underlying regularity that is difficult to reconcile with human error).

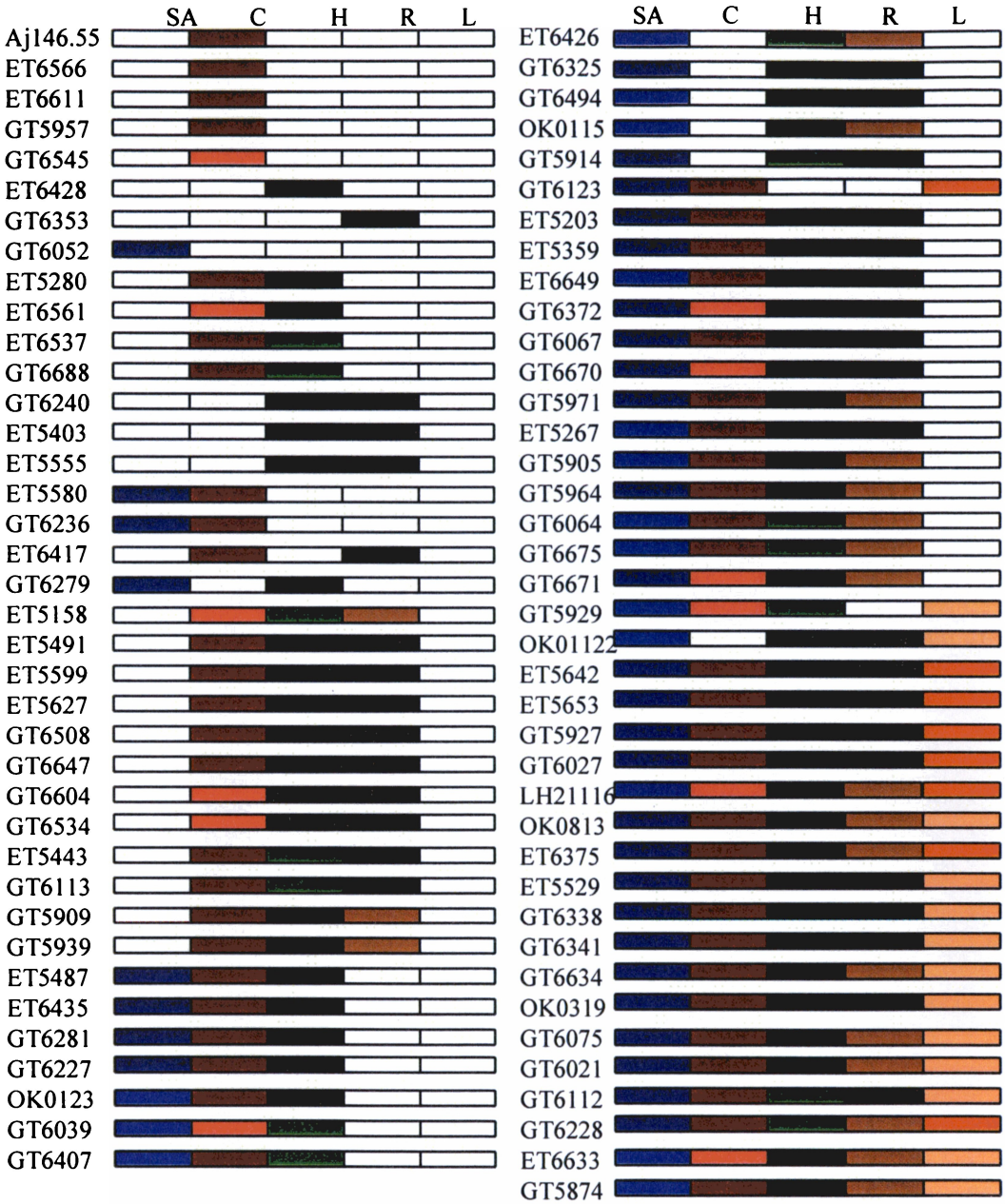
Variations in the plant material, which may account for at least some of the differences, may be due to genetic factors. The subsequent analysis tried to account for the observed variation in two ways. First of all, when establishing a light response profile across different light treatments, only changes in expression levels of 2 or more were considered significant (as already mentioned in Chapter 1). Second, spurious major discrepancies in staining, such as those highlighted in red in Table 6, were not used as the basis to establish response profiles. Rather, the consensus emerging from the data as a whole was used.

3.4.2. Organ-specificity of GUS expression

A graphical representation of the organ specific staining is shown in Figure 11. Each of the organs was given a color: blue for the shoot apex, red for cotyledon, green for hypocotyl, brown for root and orange for leaves. Note that for the purpose of this figure, the tissue specificity of staining was disregarded. Lines were organized according to the number of organs stained, beginning with those lines that stain only in one organ and finishing with the lines that stain in all five organs. Within each organ, the lighter shade of the color indicates major discrepancies in the staining pattern (those depicted in red in Table 6).

The percentage of lines with either one, two, three, four or five organs stained was calculated. The most common patterns include those ones in which more than one organ is stained (Figure 12). Thus, 23.3% of the lines stain in five organs (Figure 12B), 19.5% stain in 4 organs, 32.5% stain in 3 organs, 14.3% stain in 2 organs and 10.4% stain in 1 organ. Of the 8 lines that stain in only one organ, 5 of them stain in the cotyledon and 1 line each stain in either, the hypocotyl (H), root (R) or shoot apex (SA).

Figure 11. Graphical representation of organs stained in 77 lines analyzed for their GUS specific patterns of expression. SA = shoot apical meristem, C = cotyledon, H = hypocotyl, R = root, L = leaf. Each organ is represented by a color. Lighter colors within an organ refer to major variability in the staining patterns as described in the text.


SA C H R L
Aj146.55
ET6566
ET6611
GT5957
GT6545
ET6428
GT6353
GT6052
ET5280
ET6561
ET6537
GT6688
GT6240
ET5403
ET5555
ET5580
GT6236
ET6417
GT6279
ET5158
ET5491
ET5599
ET5627
GT6508
GT6647
GT6604
GT6534
ET5443
GT6113
GT5909
GT5939
ET5487
ET6435
GT6281
GT6227
OK0123
GT6039
GT6407
SA C H R L
ET6426
GT6325
GT6494
OK0115
GT5914
GT6123
ET5203
ET5359
ET6649
GT6372
GT6067
GT6670
GT5971
ET5267
GT5905
GT5964
GT6064
GT6675
GT6671
GT5929
OK01122
ET5642
ET5653
GT5927
GT6027
LH21116
OK0813
ET6375
ET5529
GT6338
GT6341
GT6634
OK0319
GT6075
GT6021
GT6112
GT6228
ET6633
GT5874

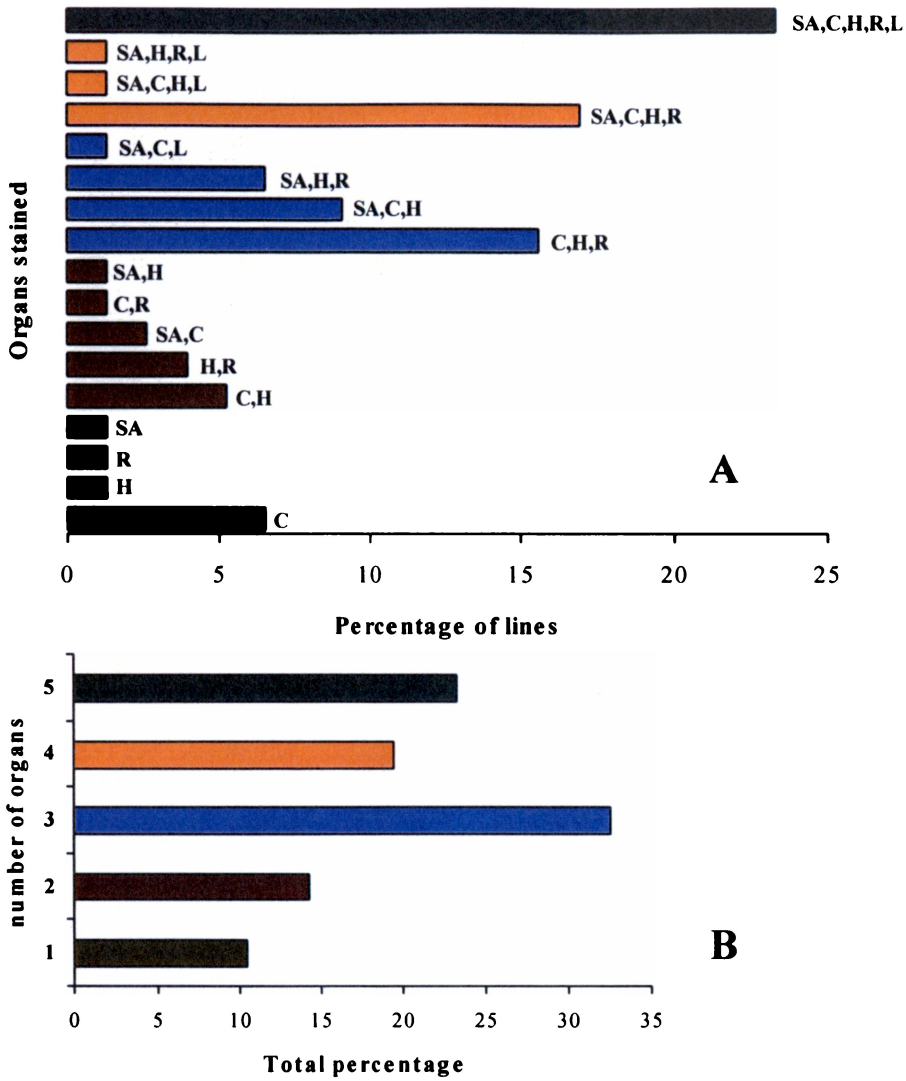


Figure 12. Organ specific staining of five day-old *Arabidopsis thaliana* seedlings. A. Percentage of lines with staining in a particular organ: SA: shoot apex, C: cotyledon, H: hypocotyl, R: root, L: leaves. B. Total percentage of lines with staining in 1 to 5 organs.

Sundaresan et al. (1995) also found a higher fraction of lines with expression in multiple organs and a lower proportion of the lines with staining restricted to one organ.

3.4.3. Response to constant red and far-red - histochemical assay

Seventy seven light shift responsive promoter trap lines were analyzed for their response to constant red and far-red. At a superficial level, Table 7 shows that the light shift responsive lines differed in their responsiveness to cR and cFR. When looking at the 'inducible' lines, it becomes apparent that typically both cR and cFR result in some level of expression. Interestingly, it is very rare to see stronger expression in cFR than cR. Instead, the expression level is often higher in cR than cFR. This finding is of interest in light of the fact that the 'inducible' category contains both 'light-inducible' (6 lines) and 'dark-inducible' (22) lines. Therefore, for the light inducible lines one must infer that cR (i.e. the phyB pathway) is more powerful than cFR (the phyA pathway) in mimicking the inducing effect of a dark-to-light shift. In contrast, in dark-inducible lines, which can be regarded as light repressed, it is cFR (phyA) rather than cR (phyB) that effectively represses GUS gene expression.

In general, the tendency for cR-inducible but cFR-repressible expression, which was observed among the 'light inducible' and 'dark-inducible' lines was replicated among the lines with a more complex response to light shifts (Table 7). Because the 'complex' classification implies some tissue-specificity of the light response, it is not surprising that the most common pattern of expression among the 48 complex lines was the 'other' pattern (34/48). As explained before, this pattern includes lines in which the response to cR and cFR is tissue-specific.

3.4.4. What is the capacity of Arabidopsis tissues to distinguish between different light conditions?

The tissue specific GUS staining patterns of 77 lines have been described under ten different light treatments (Appendix 1). In a given organ (for instance the root), the level of expression of a line as a function of the light treatment represents a response *profile*. The types of profiles observed in a given tissue thus open a window into the capabilities of the light signaling network as it operates in that tissue. Taking, for example, line ET5627, the data in Table 6 confirm, not surprisingly, that the cotyledon is able to distinguish cL from cD. On the other hand, ET5627 does not answer the question, whether the cotyledon can distinguish between cR and cFR, because the expression level of this line is similar under these two light conditions. Rather, ET5627 tells us that the light signaling network can process cR and cFR in such away that an output response is

Table 7: Number of lines with a particular GUS expression pattern in response to constant Red and Far-red light. The top of the table describes the response to the light shifts and the left of the table includes the responses to cR and cFR. R=FR, equal staining in R and FR. R>FR = stronger staining in cR than cFR. FR>R = stronger staining in cFR than cR.

		Light Inducible	Dark inducible	Complex	Total
Response to constant red and constant far-red	No stain	0	5	0	5
	R=FR	2	4	9	15
	Red only	0	1	0	1
	Far red only	0	0	0	0
	R>FR	3	7	5	15
	FR>R	0	0	0	0
	Other	1	5	35	40
	Total	**6**	**22**	**49**	**77**

essentially indistinguishable. Other lines might inform us that the cotyledon could distinguish between cR and cFR as evident by different responses.

Applying a similar logic on a comprehensive scale, this work focuses on profiles incorporating the four light treatments cL, cD, cR and cFR. For simplicity, only two expression states were distinguished: active (GUS staining level 1 to 3, denoted as '+') and inactive (level 0, denoted as '-'). Thus simplified, a total of 16 different profiles of light responsiveness are possible (Table 8). The analysis included only cotyledon, hypocotyl and root to avoid the problem of differential development and variable GUS staining of the shoot apex and young leaves. The classification was done as follows: Each organ was analyzed to determine how many of the 77 lines displayed each possible profile of light responsiveness. For example, for the staining of line ET5627 described in Table 5, the cotyledon has a #14 (+ - + +) profile because there is staining of the cotyledon under cL, no staining under cD and staining under cR and cFR. ET5627 is one of five lines that showed profile #14 in the cotyledon (6.5% of all 77 lines). Summed over all 77 lines, it becomes clear which response profiles were found at which frequency in which organ. Each of the organs (cotyledon, hypocotyl and root) had similar number of profiles (10, 10 and 11 respectively, Table 8).

Table 8. Light response profiles observed in each of three different organs. The response of each organ to these light conditions was determined in five-day-old Arabidopsis seedlings grown under constant light (cL), dark (cD), red (cR) and far-red (cFR). (+) refers to any level of expression above zero in at least one tissue type within this organ. The percentage of lines with a specific light response profile was determined separately for each organ.

Profile number	cL	cD	cR	cFR	cotyledon	hypocotyl	root
1	-	-	-	-	25.9	20.8	33.8
2	+	-	-	-	1.3	0	1.3
3	-	+	-	-	6.5	13	5.2
4	-	-	+	-	0	1.3	0
5	-	-	-	+	0	0	0
6	+	+	-	-	0	2.6	3.9
7	+	-	+	-	3.9	0	1.3
8	+	-	-	+	0	0	1.3
9	-	+	+	-	5.2	15.6	3.9
10	-	+	-	+	2.6	1.3	0
11	-	-	+	+	0	2.6	0
12	+	+	+	-	1.3	0	5.2
13	+	+	-	+	0	0	0
14	+	-	+	+	6.5	1.3	1.3
15	-	+	+	+	10.4	9.1	6.5
16	+	+	+	+	40.3	32.5	36.4
Number of profiles per organ					**10**	**10**	**11**

In Table 8, the overall staining of an organ was taken into account, including all the tissues. Therefore, in an organ (for example cotyledons), when a line displayed a profile such as (+ + - +) in one tissue (for example epidermis) and the profile (+ + + +) in another tissue (such as vasculature), the overall profile for that organ was indicated as (+ + + +). Therefore, Table 8 is a conservative estimate of light response profiles. In contrast, Tables 9, 10 and 11 present the same analysis at the level of tissue-specific resolution. Caution should be used when interpreting Table 8 because a missing profile (for example #13) could be due to the pooling of the data from different tissues. In fact, even though profile #13 is absent from the pooled data (Table 8), it is present in the epidermis of the cotyledon of line GT5927 and in the vascular tissue of GT6604 (Table 9). These novel response profiles were masked by constitutive expression in other cotyledon tissues and hence counted under profile #16 in Table 8. It is clear that the majority of 16 theoretically possible profiles are indeed realized in all three organs and most tissues examined. These data suggest that the cell types examined all have fairly complex light signaling networks. The profiles that were not found in an organ in any of the tissues are highlighted by blue.

The most common profiles in all the organs are profiles #1 (no staining) and #16 (light independent staining). Profile #5 is the only profile that was not found in any of the lines in any organ. The fact that certain profiles are represented by just one example suggests that the analysis has not yet reached saturation. Therefore, analysis of additional lines might turn up evidence for the existence of profile #5.

3.5. Discussion

Many diverse profiles were present in different organs and tissues. One critical question is: how do we interpret the different frequencies of occurrence of the response profiles among the promoter trap lines? What do we make out of profiles that surfaced in only one of the three organs? Shall we hypothesize different structures of the light signaling machinery in different organs? The best candidate for an organ specific profile is #11 (cR and cFR specific expression), which appeared four times in the hypocotyl but not once in the cotyledon or root. Obviously, scaling the analysis from 2,000 insertion sites (and 77 lines enriched for light-responsiveness) to the entire genome may well lay this hypothesis to rest, but at this point it remains a possibility.

Another question for future analysis is whether individual promoter traps usually adopt the same or different response profiles in the different tissues in which they are expressed. At a superficial level, the large fraction of genes that showed tissue-specific light responses after light shifts or after cR/cFR ('complex' in Chapter 2, Table 2) and 'Other' in this Chapter, Table 7) would suggest the latter.

Table 9. Light response profiles observed in different tissues of the cotyledon to cL, cD, cR and cFR in five day-old Arabidopsis seedlings. (+) refers to any level of expression above zero. The percentage of lines with a specific profile was determined for the ET and GT lines. CE= cortex epidermis, V= vascular tissue, P= petiole, H= hydathodes.

					Cotyledon							
					CE		V		P		H	
#	cL	cD	cR	cFR	ET	GT	ET	GT	ET	GT	ET	GT
1	-	-	-	-	16.9	13	20.8	20.8	32.5	53.2	41.6	53.2
2	+	-	-	-	0	0	0	0	1.3	0	0	0
3	-	+	-	-	0	5.2	3.9	2.6	1.3	0	0	0
4	-	-	+	-	0	0	1.3	1.3	0	0	0	1.3
5	**-**	**-**	**-**	**+**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
6	+	+	-	-	0	0	0	0	1.3	0	0	0
7	+	-	+	-	1.3	1.3	0	0	1.3	0	1.3	0
8	+	-	-	+	0	0	0	0	0	0	0	1.3
9	-	+	+	-	0	3.9	1.3	5.2	0	0	0	0
10	-	+	-	+	0	2.6	1.3	1.3	0	0	0	0
11	**-**	**-**	**+**	**+**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
12	+	+	+	-	1.3	1.3	0	2.6	0	0	0	0
13	+	+	-	+	0	1.3	0	1.3	0	0	0	0
14	+	-	+	+	5.2	0	2.6	0	3.9	0	0	0
15	-	+	+	+	5.2	5.2	1.3	3.9	0	1.3	0	0
16	+	+	+	+	13	23.4	10.4	18.2	1.3	2.6	0	1.3

Table 10. Light response profiles observed in different tissues of the hypocotyl to cL, cD, cR and cFR in five day-old Arabidopsis seedlings. (+) refers to any level of expression above zero. The percentage of lines with a specific profile was determined for the ET and GT lines.

					Hypocotyl			
					Cortex-epidermis		Vascular tissue	
#	cL	cD	cR	cFR	ET	GT	ET	GT
1	-	-	-	-	11.7	11.7	15.6	20.8
2	**+**	**-**	**-**	**-**	**0**	**0**	**0**	**0**
3	-	+	-	-	6.5	7.8	5.2	6.5
4	-	-	+	-	0	1.3	0	1.3
5	**-**	**-**	**-**	**+**	**0**	**0**	**0**	**0**
6	+	+	-	-	0	0	0	1.3
7	**+**	**-**	**+**	**-**	**0**	**0**	**0**	**0**
8	**+**	**-**	**-**	**+**	**0**	**0**	**0**	**0**
9	-	+	+	-	5.2	9	5.2	5.2
10	-	+	-	+	0	1.3	0	0
11	-	-	+	+	2.6	0	2.6	0
12	+	+	+	-	0	1.3	0	1.3
13	+	+	-	+	0	2.6	0	0
14	+	-	+	+	0	1.3	0	0
15	-	+	+	+	5.2	7.8	5.2	6.5
16	+	+	+	+	11.7	13	9	14.3

Table 11. Light response profiles observed in different tissues of the root to cL, cD, cR and cFR in five day-old Arabidopsis seedlings. (+) refers to any level of expression above zero. The percentage of lines with a specific profile was determined for the ET and GT lines. CE= cortex-epidermis, V= vascular tissue, LP = lateral root primordia, RT= root tip, C= colet.

					Root									
					CE		V		LP		RT		C	
#	CL	cD	cR	cFR	ET	GT	ET	GT	ET	GT	ET	GT	ET	GT
1	-	-	-	-	20.8	35.1	32.5	48	36.4	41.6	40.3	48	31.2	36.4
2	+	-	-	-	2.6	2.6	0	0	1.3	1.3	0	0	1.3	0
3	-	+	-	-	1.3	1.3	1.3	0	1.3	1.3	0	2.6	2.6	0
4	**-**	**-**	**+**	**-**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
5	**-**	**-**	**-**	**+**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
6	+	+	-	-	2.6	0	1.3	0	0	1.3	0	0	0	0
7	+	-	+	-	0	0	0	0	0	0	0	0	0	1.3
8	+	-	-	+	0	0	0	0	0	0	1.3	0	0	1.3
9	-	+	+	-	1.3	1.3	0	0	1.3	2.6	0	1.3	0	1.3
10	-	+	-	+	0	0	0	0	1.3	1.3	0	0	0	0
11	**-**	**-**	**+**	**+**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
12	+	+	+	-	1.3	0	0	2.6	0	1.3	0	1.3	1.3	1.3
13	**+**	**+**	**-**	**+**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
14	+	-	+	+	0	0	0	0	0	0	0	0	1.3	0
15	-	+	+	+	0	2.6	2.6	2.6	1.3	1.3	0	0	0	5.2
16	+	+	+	+	13	14.3	5.2	3.9	0	5.2	1.3	3.9	5.2	10.4

Further analysis of 'rare' profiles indicated that each organ had profiles that were represented by only one line, for instance profile #2 in the cotyledon (line ET5203). These profiles seem to be reliable since most of them were represented by more than one experiment (see Appendix 1).

Many diverse profiles were present in different organs and tissues. One profile was not observed: #5 (- - - +) and several profiles have only one representation, suggesting, as stated before, that the analysis has not yet reached saturation. It seems to be difficult or impossible for the plant to generate a positive promoter trap response to cFR without at the same time generating a positive response to another constant condition (cL, cD or cR). Given our molecular understanding of light signaling, a positive response to cFR will probably require nuclear translocation of phyA (Kircher et al., 1999 and 1999a). Any of the promoter trap responses that are sensitive to nuclear phyA (i.e. cFR) also appeared to be sensitive to at least constant light (profile #8, observed in root tips and colet and in hydathodes) or cR (#11, observed in hypocotyl, CE and vascular tissue). This might suggest that the phyA signaling pathway is wired such as to 'condition' other signaling pathways, such as cry and phyB pathways active under cL or phyB pathways active under cR. Alternatively, the missing profiles might exist in certain cell types that are hidden among the background of a more constitutive response in neighboring cell types. For example, a line with a (+ + + +) profile in the root CE might have a #5 (- - - +) profile in the vasculature, but it is difficult to distinguish the vasculature staining when the epidermis stains stronger.

The absence of profile #5 is surprising because a conceptually simple signaling pathway, phyA specific gene induction, should result in this profile. The data do in fact give evidence of a phyA specific pathway, but it is wired to repression, not induction of GUS expression, i.e. profile #12. This indicates that Arabidopsis is clearly able to interpret cFR signals differently from all three other light conditions tested. With this argument in mind, our challenge of the light signaling machinery was mastered very well by Arabidopsis. Reports of genes expressed only under cFR were not found. The Arabidopsis Athb2 that encodes for a HD- Zip protein is strongly expressed under cFR but it is also expressed under cD (Carabelli et al., 1993) A high steady-state level of Athb-2 mRNA was found in dark-adapted plants and the far-red rich light treatment resulted in a strong induction of the gene expression.

The analysis was limited by condensing the expression level results into binary categories using an arbitrary threshold, i.e. ('+' equals any staining) versus ('-' equals no staining). In order to find out if this classification was responsible for the absence of profile #5, a

different threshold of + versus – was used to analyze the data. This analysis focused on lines that showed significantly lower degree of staining under cL, cD and cR (i.e. cortex-epidermis CE1) than under cFR (CE3). No examples were found that fit profile #5 under these parameters either.

The GUS staining patterns showed some variability for most of the lines evaluated. More frequent variability was found in the degree of staining, while major discrepancies in the degree or pattern of staining were less frequent. As discussed previously, these discrepancies could be due to human error when recording the data or to variations due to genetic or environmental factors. Genetic factors might include second-site mutations that may affect the GUS staining pattern or secondary Ds:GUS promoter trap elements that may segregate in the families subjected to staining, even though all 24 lines examined in detail had only one insertion. However, a more likely source of variation is environmental factors. It is evident from this analysis that a majority of lines are highly sensitive to the environmental factor light. Even though the seedlings were grown under highly controlled light conditions, it is possible that other environmental factors might affect the GUS staining patterns. For example, early stages of growth of microorganisms would have remained undetected but the promoter traps might be sensitive to signals from such disturbances. Even factors such as relative proximity of the individual seedlings might affect gene expression in certain instances. This analysis took into account the variability at every stage. For example, light response profiles in Tables 8-11 were always based on reproducible observations. In addition, the general conclusions reached in this chapter would not be sensitive to occasional fluctuations in the GUS response patterns between experiments. The negative results are difficult to interpret. However, the positive results show that many diverse profiles exist and are present in different tissues, not just the photosynthetic ones.

Chapter 4. Analysis of enhancer and gene trap lines – global definition of light responses

4.1. Introduction

The process of phototransduction involves two basic steps: photoperception and signal transduction (Casal et al., 1998). Efforts to dissect the signal transduction cascade from photoreceptor to light-dependent responses have provided important clues about the function of the photoreceptors, terminal response agents, and intervening signal transduction components (Chory and Wu, 2001). Following absorption of light, the photoreceptors must interact with other components in the signal transduction chain, and that will eventually lead to the observed responses. Such responses are likely to be mediated by changes in gene expression. The responses usually entail changes in the physiological states of the cells in which the signaling pathway operates. Complex, highly networked, signaling systems are more versatile than simple systems in terms of signal input-signal output possibilities and more robustly buffered against environmental changes that have no apparent informational value (Meagher et al., 1999).

Even though much research has been done on plant photoreceptors, little is known about light signal transduction networks. Given the complexity of the input signals and the diverse array of developmental events regulated by light, it seems likely that light responses result from integration of a variety of signals through an intricate network of interacting signaling components (Chory et al., 1995). A complete understanding of light signaling networks will require information at many different levels: transcriptional, translational and post-translational regulation, gene interactions and tissue specificity. Enhancer and gene traps allow the analysis of primarily transcriptional regulation along with the tissue-specificity of gene expression, as determined by the GUS staining patterns.

Despite much progress in the characterization of individual phytochromes, we have only an incomplete picture about the contribution of crosstalk between the pathways. Genetic analyses have shown that crosstalk exists not only between phytochromes but also between phytochromes and cryptochromes. A complex web of interactions between the photoreceptors includes redundancy, antagonism and effector/modulator relationships (Chory and Wu, 2001). The current picture has emerged from a combination of genetic, biochemical and cell biological approaches (McNellis and Deng, 1995; Chory et al., 1996, Barnes et al., 1997). As a result several models, such as those illustrated in figures 1, 2 and 3 (of the main introduction) have been proposed to explain the complexity of the light signaling process (Ni et al., 1998, Yamamoto et al., 2001, Neuhaus et al., 1993,

Bowler et al., 1994). However, these are simple models that are based on a small set of genes and are unlikely to reflect the entire complexity of the light signaling network.

Even though phyA and phyB pathways are activated by different light conditions, the end-point responses (e.g. hypocotyl growth, cotyledon unfolding, flowering, etc.) controlled by these photoreceptors are largely the same. This indicates that the phototransduction pathways of phyA and phyB converge at some point (Casal et al., 1998). Analysis of the kinetics of phototransduction (Casal et al., 1998) and of the loci affecting phyA or phyB mediated responses (Wagner et al., 1997) have shown that phyA and phyB probably have different transduction chains. Under certain conditions, a single photoreceptor is implicated in a response (for example phyA controls cotyledon expansion in far-red light), indicating a simple signal transduction pathway. In other cases, multiple photoreceptors can contribute to a growth response in an additive manner (for example, cry1, cry2, phyA and phyB contribute to hypocotyl growth inhibition in blue light). This suggests the convergence of receptor-specific pathways downstream of the signal. There is also the possibility of parallel completely independent pathways affecting the growth response through different mechanisms (Neff and Chory, 1998).

PhyA can exhibit both antagonistic (negative) and synergistic (positive) effects on phyB signaling. For example analyses of Lhcb1*2-gusA expression, hypocotyl growth and cotyledon unfolding have shown that under red light, phyA and phyB are antagonistic, but under far-red light, followed by red light pulses, phyA and phyB are synergistic. These interactions are likely to result from cross-talk of the receptor signaling pathways (Casal, 2000). As introduced earlier, there are three light response modes: the LFR (R/FR reversible responses mediated by phyB), the VLFR (responses to pulses of FR or R but not reversible by FR, mediated by phyA) and the HIR (the response to continuous light, including FR and hourly FR pulses that are mediated by phyA). The negative or positive effects of phyA on phyB mediated responses could be accounted for by a scenario where phyA initiates two divergent signaling pathways. the VLFR pathway interacting antagonistically with phyB signaling and the HIR pathway interacting synergistically with phyB signaling (Cerdán et al., 1999). An example of negative regulation of phyB activity by phyA was illustrated by an exaggerated LFR (GUS expression, hypocotyl elongation and cotyledon unfolding) of the phyA mutant in seedlings exposed only to red pulses. A positive (synergistic) response was illustrated by an enhanced LFR response when given a far red pulse pre-treatment to a seedling treated with red pulses (Cerdán et al., 1999).

4.2 Experimental outline and rationale

In light of the interactions between photoreceptor pathways as determined at the morphological level and at the level of individual genes such as Lhcb*2, I set out to determine types of photoresponse interactions on a larger, genome-wide scale. The subset of light responsive promoter trap lines identified in the previous chapters might lend itself to investigate quantitatively how phyA and phyB signaling pathways may intersect. There are no reports of a quantitative analysis of the light signaling network at the scale described in this work.

The GUS protein activity was determined using a fluorimetric in vitro assay for a group of gene trap and enhancer trap lines under eight different light conditions: cL, cD, cR, cFR, red pulses (Rp), far-red pulses (FRp), red followed by far-red pulses (R-FRp) and far-red followed by red pulses (FR-Rp). Emphasis was placed on promoter traps that had shown activation in darkness and repression of activity by light (Table 4). The resulting data represent light response profiles, as previously described in Chapter 3. Compared to the histochemical in situ assay adopted for experiments in Chapters 2 and 3, which is semi-quantitative, I anticipated that the fluorimetric assay would lead to a more quantitative set of data that would be more amenable to statistical evaluation. Several statistical methods were explored for the clustering of light response profiles. Clusters of co-regulated genes were indeed observed. Genes showing phyA or phyB dependent expression or repression were identified, along with genes for which there was antagonism or redundancy of phyA and phyB.

4.3. Materials and methods

4.3.1. Measurement of GUS activity profiles by fluorimetric assay

Seedlings were grown as described on Chapter 1. After the treatments (cL, cD, cR, cFR, Rp, R-FRp, FRp, FR-Rp) the GUS protein expression levels were measured by a fluorimetric assay. In addition, as a control, a subset of seedlings was examined for their histochemical GUS staining patterns, as described in Chapter 2. For the fluorimetric assay approximately 30 seedlings were harvested and the tissue was ground in 100 µl extraction buffer (50 mM sodium phosphate buffer pH 7, 1 mM Na_2EDTA, 0.1% Sodium-lauryl sarcosine and 0.1% Triton X-100). The homogenate was spun for 10 min at 4°C and the supernatant was kept on ice. 20 µl of the extract was mixed with 80 µl of GUS assay mix (70 µl extraction buffer and 10 µl substrate solution (1 mM 14-methyl-umbelliferyl-beta-D-glucuronide)). The mixture was incubated at 37°C for 1 hour after which 0.9 ml of 0.2M sodium carbonate were added and the fluorimetric measurements carried out using a fluorimeter (Turner, model 450). The protein content was measured

with the BCA protein assay reagent (Pierce) according to the manufacturer's instructions. Standardization of the GUS assay was done by relating the relative fluorescence units to the amount of methyl-umbelliferone (MU) using a 1 mM standard of MU. In addition to the selected transposant lines, non-transgenic seedlings (Landsberg ecotype) were frequently included as controls in order to define the minimal detectable GUS activity levels. For the purpose of comparison, the GUS activity profiles were analyzed for a line with a minimal Arabidopsis CAB1 promoter (Hou et al., 1993). This line was grown under the same conditions described for the promoter trap lines analyzed in this work.

The fluorimetric data were analyzed as follows: The raw GUS activities (in nmole MU per minute per milligram protein) were normalized against the activity under constant white light, which was taken to represent 100%. This was done in part to compare between profiles irrespective of the absolute expression levels and in part to account for variation among experiments. One source of such variation is the dosage of the GUS reporter, which may be either heterozygous or homozygous. For example, two experimental repeats might have been conducted first with heterozygous F3 seedlings and later with homozygous F4 seedlings. Next, the normalized values were log transformed, which was necessary for statistical purposes, and means and standard deviations were calculated. Analysis of variance and a Tukey test was performed in order to define pairs of treatments that showed significant differences. Recommendations given by the UT Statistics Consulting Center were taken into account for the statistical analysis. SPSS© software was used for the final analysis.

4.4. Results

4.4.1. Whole seedling response profiles

GUS activity levels were determined using a fluorimetric assay for 85 transposant lines, as well as non-transgenic wild type and the CAB:GUS fusion, under up to eight different light conditions. Statistical analysis of the GUS protein levels was performed in 51 lines for which the experiments were repeated between two and six times. The data from the other 34 lines analyzed (out of a total of 85 described above) did not have replicates so these data were not suitable for statistical analysis and are not presented here.

As explained in 'methods', the raw GUS activity levels were normalized to eliminate variability between experiments (Table 12) that might be due to differences in gene dosage. The normalized (protein percentage) GUS activities (Table 12) were displayed as shown in Figure 13.

Table 12. Example of the transformation of the GUS activity levels in line ET5203. The second column indicates the GUS activity in nmole MU per minute per milligram protein, the third column the normalization of those values, taking cL as 100% and the fourth column the logarithm of the normalized data.

	GUS activity	Normalization	Log
Experiment 1			
cL	15.45	100	2.00
cD	28.78	186.28	2.27
cR	77.96	504.60	2.70
cFR	9.05	58.58	1.77
Rp	14.51	93.92	1.97
R-FRp	27.94	180.84	2.26
FRp	27.90	180.58	2.26
FR-Rp	16.40	106.15	2.03
Experiment 2			
cL	14.42	100	2.00
cD	28.48	197.50	2.30
cR	60.69	420.87	2.62
cFR	8.72	60.47	1.78
Rp	21.77	150.97	2.18
R-FRp	15.92	110.40	2.04
FRp	26.28	182.25	2.26
FR-Rp	20.52	142.30	2.15
Experiment 3			
cL	22.15	100	2.00
cD	38.13	172.14	2.24
cR	84.48	381.40	2.58
cFR	10.42	47.04	1.67
Rp	31.35	141.53	2.15
R-FRp	90.43	408.26	2.61
FRp	89.05	402.03	2.60
FR-Rp	57.60	260.05	2.42
Experiment 4			
cL	13.96	100	2.00
cD	18.92	135.53	2.13
cR	62.08	444.70	2.65
cFR	3.64	26.07	1.42
Rp	24.69	176.86	2.25
R-FRp	39.56	283.38	2.45
FRp	37.57	269.13	2.43
FR-Rp	29.79	213.40	2.33

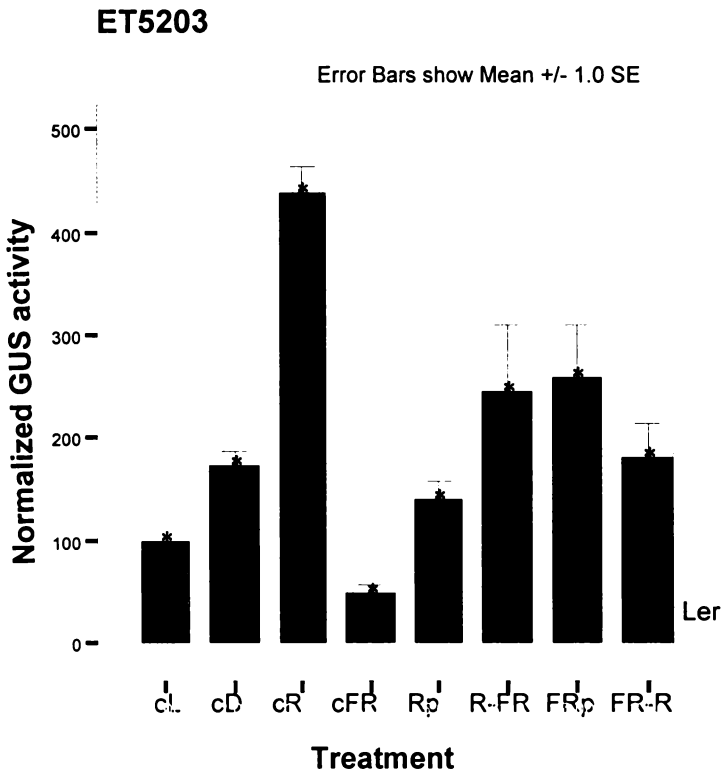


Figure 13. Example of the normalized GUS activity (Line ET5203). The background GUS expression was measured in Landsberg wild type seedlings and indicated in the figure as a black line.

Subsequently, analyses of variance (ANOVA, Table 13) and Tukey analysis of the data were performed (Table 14). This is a two-step process. In the first step, ANOVA is conducted in order to determine whether the data contain at least a single pair of data points that are significantly different. In the second step, a 'multiple comparison procedure' is used to determine which pair(s) of data points are significantly different and which pairs are not. Different algorithms have been devised for this purpose. The Tukey method was suggested by statistics experts from the 'Statistics Consulting Center' at the University of Tennessee. In the ANOVA, a significant 'F' value tells only that the population means are probably not all equal but it does not tell which pairs of data-points have different means. Within the ensuing Tukey test, the maximum number of significantly different pairs is (7+6+5+4+3+2+1) = 28. This would mean that all eight treatments resulted in significantly different GUS activity levels, and the GUS activities would fall into eight distinct groups. Typically, however, a group comprises more than one treatment, because no significant differences can be detected between them; these groups are referred to as 'homogeneous subsets'. Note that a single treatment can belong to more than one homogeneous subset. For example, in line ET5203 four overlapping homogeneous subsets are evident (Table14). Subset 'a' comprises only the low activity level displayed under cFR. Subset 'b' includes cL, Rp, FR-Rp and cD whereas subset 'c' includes Rp, FR-Rp, cD, Rp, R-FRp and FRp. Thus, there is an overlap between 'b' and 'c' subsets. It is evident that GUS levels in the cFR experiment are significantly lower than GUS levels in all the other treatments (Figure 13). The Tukey analysis is a very stringent analysis tool that protects against the danger of falsely assigning 'significance' to differences that arise when multiple (in this case 28) comparisons are made.

Out of up to 28 possible 'significantly different pairs' the analysis for ET5203 renders eleven: 7 (cFR is different from the other 7 treatments) + 3 (cL vs R-FRp, cL vs FRp, cL vs cR) + 1 (Rp vs cR). Because the analysis was done with a confidence level of 95%, the statistical probability that one or more pairs differ due to chance, i.e. to be irreproducible, is 5%. This level of stringency appears to be very high. Therefore, the pairs of significantly different means were calculated at both the 95% and 75% confidence level. At the 95% level, only 1 in 20 analyses will contain one or more significant pairs that are due to chance. In other words, 19 out of 20 lines would be expected to return the exact same light response profile if the entire experiment was repeated. At the 75% level, the same is true for only 16 out of 20 lines. For the other four, one would expect those single pairs of differences, or -less likely- double or triple pairs of differences would fail to be reproduced. In the case of ET5203, three homogeneous subsets were observed at the 95% level whereas the 75% level resulted in 4 homogeneous subsets. In line LH211.16 (Table 15), 3 homogeneous subsets were distinguished at the 95% confidence level whereas at the 75% level there were 4 homogeneous subsets. In addition, while the 95% level yielded six significant pairs, the 75% level yielded eleven.

Table 13. Example of descriptive statistics and ANOVA (Line ET5203). In this example there are significant differences between the treatments.

ET5203				
	N	**Mean**	**Standard deviation**	**Standard Error**
cL	4	2.0000	0.0000	0.0000
cD	4	2.2334	0.0072	0.0035
cR	4	2.6391	0.0050	0.0025
cFR	4	1.6595	0.1693	0.0084
Rp	4	2.1375	0.1171	0.0058
R-FRp	4	2.3409	0.2457	0.1228
FRp	4	2.3879	0.1653	0.0083
FR-Rp	4	2.2308	0.1748	0.0087

ANOVA					
	Sum of squares	**Degrees of freedom**	**Mean square**	**F**	**Significance 0.05**
Between groups	2.344	7	0.335	15.912	.000
Within groups	0.505	24	0.0021		
Total	2.849	31			

Table 14. Tukey analysis of normalized GUS activities for line ET5203 at the 95% significance level. For visualization purposes, the normalized GUS activity was included but the statistical analysis was done with the Log of the normalized GUS.

Trt	N	Homogeneous subsets for alpha = 0.05		
		a	b	c
cFR	4	48.04		
cL	4	100.00	100.00	
Rp	4	140.82	140.82	
FR-Rp	4	180.47	180.47	
cD	4		172.86	
R-FRp	4		245.72	
FRp	4		258.49	
cR	4			437.89

Means for groups in homogeneous subsets are displayed

Table 15: Tukey analysis of normalized GUS activities of line LH211.16 at two different confidence levels.

Tukey HSD

Treatment	N	Subset for alpha = 0.05		
		a	b	c
cR	2	10.44		
cD	2	17.13		
FRp	2	18.11	18.11	
cFR	2	25.77	25.77	
Rp	2	27.87	27.87	27.87
FR-Rp	2	40.60	40.60	40.60
R-FRp	2		69.62	69.62
cL	2			100

		Subset for alpha = 0.25			
		a	b	c	d
cR	2	10.44			
cD	2	17.13	17.13		
FRp	2	18.11	18.11		
cFR	2	25.77	25.77		
Rp	2		27.87	27.87	
FR-Rp	2		40.60	40.60	40.60
R-FRp	2			69.62	69.62
cL	2				100

Means for groups in homogeneous subsets are displayed

Obviously, the 75% level provides higher resolution in the definition of light response profiles, but with a calculated risk of returning false positives.

4.4.2. Comparison of in situ staining and fluorimetric GUS expression profiles

Most of the tissue specific patterns of GUS staining coincided with the fluorimetric profiles but in some cases small differences were found. Line LH211.16 is presented as an example to discuss the source of discrepancies between the in situ data and the fluorimetric data. While LH211.16 is typical in its complex in situ expression profile, it is also of particular interest because of its apparent responsiveness to light pulses. As evident from Table 16, there is light-dependent staining in the root; the absence of root staining in cR and cFR suggests that multiple photoreceptors (either phyA and phyB together, or cryptochromes) must be activated to induce root expression. In the hypocotyl there is nearly constitutive expression –in the hypocotyl cortex-epidermis (CE) and a striking induction of expression in the vasculature under cR. Interestingly, red pulses appear to induce this response as well, and subsequent FR pulses revert it. FR-Rp pulses may also induce this response, clearly implicating a phytochrome LFR in the hypocotyl vasculature. In the cotyledon, expression is dark-dependent. It may be partially suppressed by some of the pulse treatments.

When comparing the whole-seedling fluorimetric data and the in situ staining, two questions arise: (A) which pairs of fluorimetric data display significant differences in GUS activity, and are the in situ data consistent with these differences? And (B): Are there striking differences in the in situ staining that do not show up as significant in the fluorimetric analysis? In line LH211.16 there are six significantly different pairs of data points: cL > cR, cD, cFR, FRp; and R/FRp > cR, cD. The four pairs involving cL (Table 15) make sense, considering the strong root expression under cL. However, there is a discrepancy between the R-FRp pulses and the cD data given that the fluorimetric data suggest stronger staining in the R-FRp treatment when the in situ data suggest the opposite. This is probably due to the fact that the vasculature in this line stains only in the region close to the cotyledon, a piece of information that was recorded in the raw data but not for the tables in the Appendix.

In summary, any apparent discrepancies between the in situ staining patterns and the fluorimetric profiles can usually be explained based on the volumes of stained tissue, patches of staining or variability in the pattern of staining (see Chapter 2). It is likely that the quantitative expression levels are reflected more accurately by the fluorimetric data than by the in situ data.

Table 16. Tissue-specific pattern of expression of line LH211.16. Five-day-old seedlings were grown under constant L, D, R, FR or red or far-red pulses or R-FRp and FR-Rp pulse combinations as described in materials and methods.

Treatment	Cotyledon	Hypocotyl	Root
cL	-	CE1	CE2 V2 RH2
cD	ME2	CE1	-
cR	-	V2	-
cFR	-	CE1	-
Rp	ME2	CE1 V2	-
R-FRp	ME1	CE1	-
FRp	-	CE1	-
FR-Rp	ME2	CE1 V1	-

4.4.3. Cluster analyses of the whole-seedling GUS gene expression profiles

Out of the 51 lines for which multiple sets of fluorimetric data were obtained, 16 lines showed no significant differences between the treatments. Among the remaining 35 lines, 9 lines had low activity, close to the background. Even though these lines had significantly different pairs, they were not included in the following analysis because they stained only in small organs, such as stipules or lateral root primordia. The remaining 26 lines were used for further analysis.

The statistically validated GUS expression data were represented graphically in a color-coded scheme (Figure 14). This more intuitive display of the data, together with the histograms of the profiles exemplified in Figure 13, formed the basis for various attempts at cluster analysis that will be discussed below.

A red color was assigned to the treatment resulting in the lowest GUS activity, green to the highest activity, and other colors to intermediate levels as needed (Figure 14). More specifically, each homogeneous subset (a,b,c,d) emerging from the ANOVA was given a color. Each promoter trap line is represented by two entries. For each entry, the top layer refers to the ANOVA conducted at the 95% confidence level and the lower layer to the 75% level (e.g. Table 15). For example, for LH211.16 at the 75% confidence level, the treatment cR is in the 'low activity' homogeneous subset and is therefore represented by

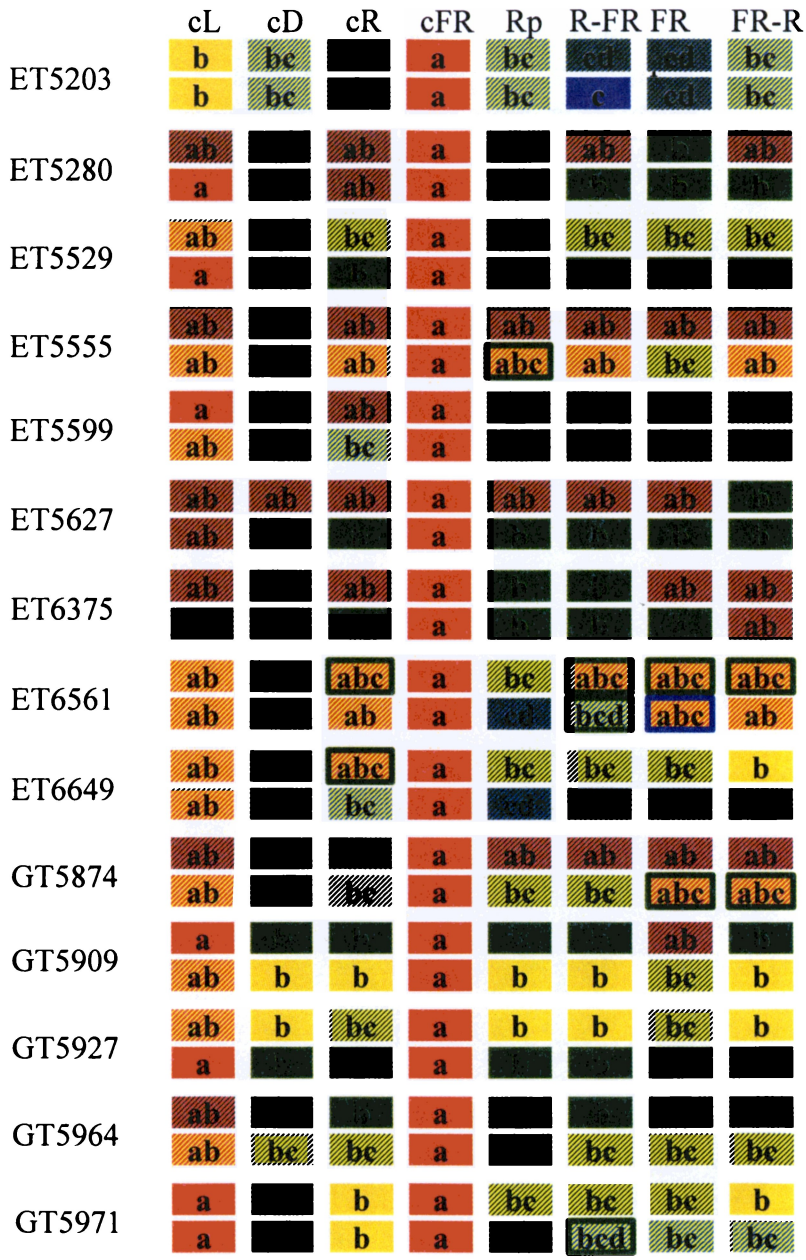


Figure 14. Colorimetric representation of the ANOVA/Tukey analyses of the fluorimetric GUS expression data from five-day-old seedlings of the indicated promoter trap lines. The results of the Tukey analyses were represented in color, with the lowest value being red and the highest value green. Treatments that share a color or a letter are not significantly different. For each line, the top refers to 95% and the bottom to 75% confidence level.

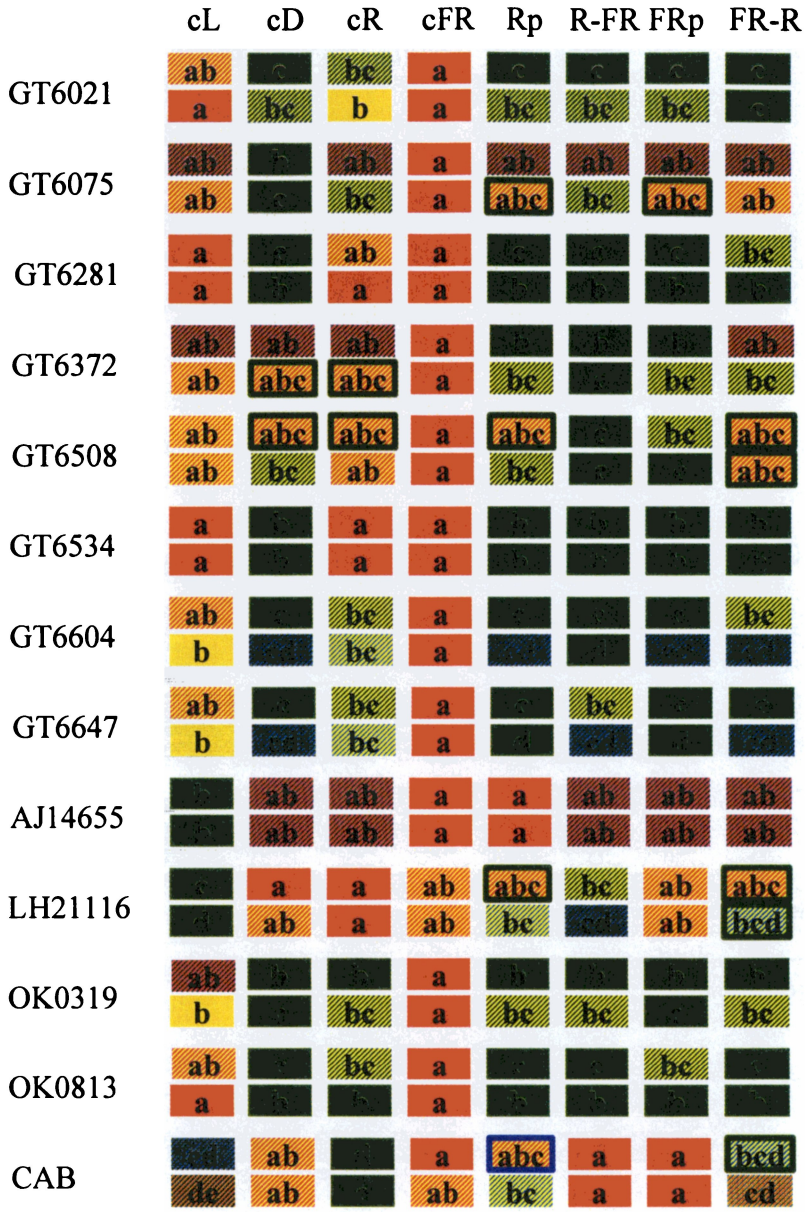


Figure 14 continued. Colorimetric representation of the ANOVA/Tukey analyses of the fluorimetric GUS expression data from five-day-old seedlings of the indicated promoter trap lines. The results of the Tukey analyses were represented in color, with the lowest value being red and the highest value green. Treatments that share a color or a letter are not significantly different. For each line, the top refers to 95% and the bottom to 75% confidence level.

a red color and the letter 'a'. The data from treatments cD, FRp and cFR are part of homogeneous subsets 'a' and 'b', and are therefore represented by 'ab' and a red/yellow striped pattern, and so forth.

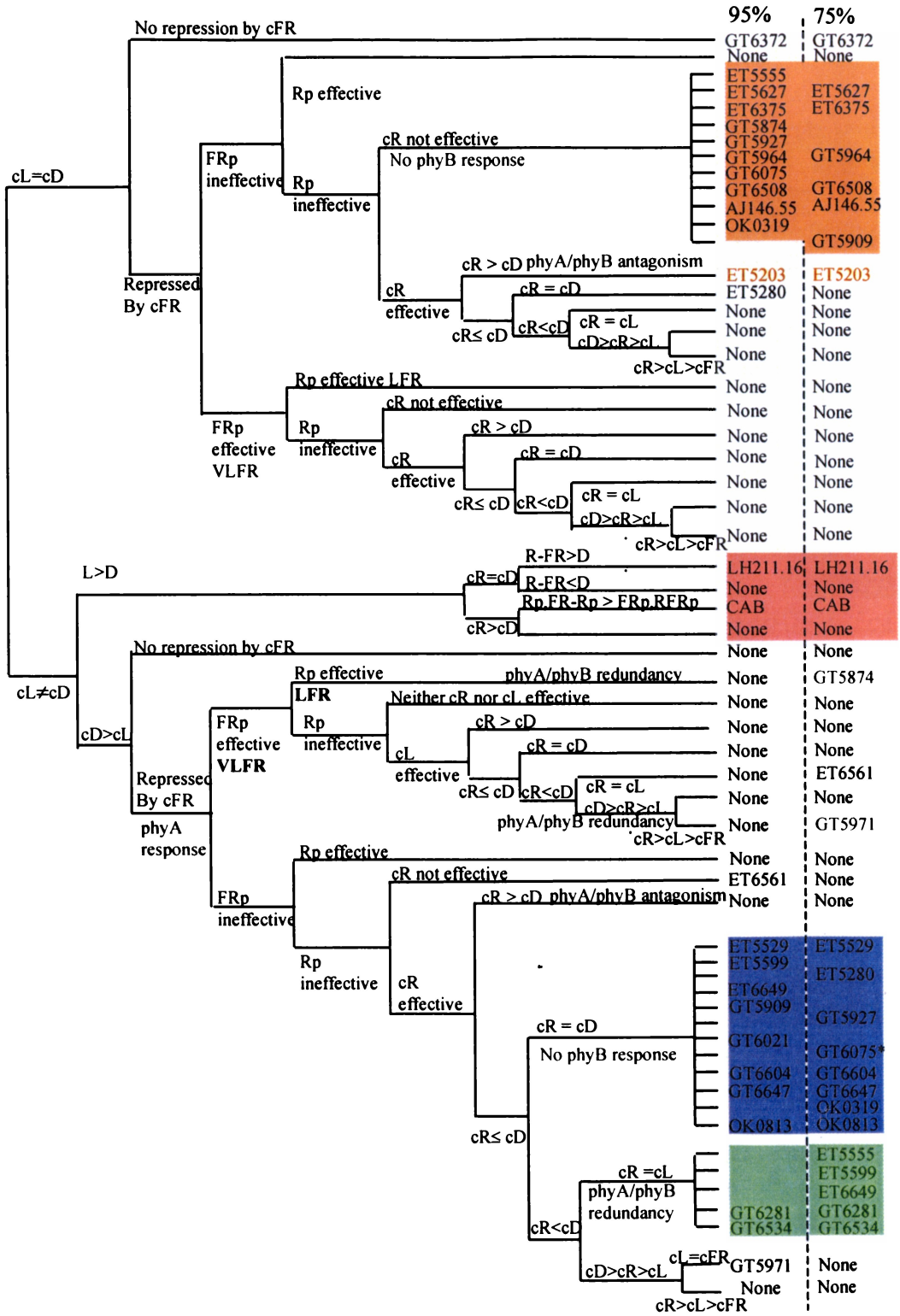
Even a cursory look at Figure 14 reveals similarities between subsets of light response profiles of the promoter trap lines. In order to relate the response profiles to each other, various cluster analysis algorithms available through SPSS® were explored. However, these algorithms do not take into account the known elements of the light signaling network and are therefore limited in their predictive value. Moreover, the algorithms work with the mean GUS activities and do not place penalties on data points with high error margins. This leaves it up to the investigator to assign a level of confidence to the clustered data. For these reasons, the promoter trap lines were clustered in a more empirical fashion. Specifically, the lines were placed onto a rooted tree in which each branch point corresponds to a binary distinction between two expression characteristics (Figure 15).

In detail, using Figure 14, I first visually grouped lines with indistinguishable profiles into clusters. Subsequently, I identified binary questions that would distinguish between these clusters. The questions were then ordered into a hierarchical structure ('tree'). A number of different tree designs were evaluated in an attempt to minimize the number of nodes. As evident from Figure 15, the binary questions query (i) a difference between cD and cL, (ii) the effect of cFR, (iii) of FR pulses, (iv) of R pulses, (v) and finally the relative effects of cR and cL, in that order. The responses to R/FRp and FR/Rp were disregarded; (in fact, lines within a single cluster may differ with respect to their response to these composite pulses).

Importantly, this clustering method incorporates the results from the statistical analysis. When addressing the effect of a given treatment, for example cD versus cL, the first question to be asked is whether there is a significant difference between the two treatments. The second question addresses which treatment gave the higher or lower gene expression. None of the available computational clustering algorithms incorporated the statistical significance.

For the purpose of comparison, hierarchical clustering analysis was carried out using Ward's algorithm, which is included in the SPSS® program. This algorithm has been recommended over others in several studies (El Hamdouchi and Willett, 1986, Figure 16). The hierarchical clustering begins by finding the closest pair of objects according to a distance measure and combines them to form a cluster. Ward's distance measure uses

Figure 15. Hierarchical clustering of light response profiles according to an empirical procedure that is sensitive to statistical significance either at the 95% or 75% level (see text). In order to compare this tree with Ward's cluster from the SPSS package (Figure 15), lines belonging to the same cluster were highlighted with the same color. Lines with red font belong to the phyA/phyB antagonism group and those with green font to the phyA/phyB redundancy group. *: FR-Rp ineffective means that this line can be distinguished from the other in the same cluster based on the response to the pulses)


95%
75%
No repression by cFR
GT6372
None
Rp effective
ET5555
ET5627
ET6375
GT5874
GT5927
GT5964
GT6075
GT6508
AJ146.55
OK0319
GT5909
cR not effective
No phyB response
FRp
ineffective
Rp
ineffective
cL=cD
cR > cD phyA/phyB antagonism
ET5203
ET5280
cR = cD
cR
effective
cR≤ cD
cR<cD
cR = cL
cD>cR>cL
cR>cL>cFR
Repressed
By cFR
Rp effective LFR
cR not effective
FRp
effective
VLFR
Rp
ineffective
cR > cD
cR = cD
cR
effective
cR≤ cD
cR<cD
cR = cL
cD>cR>cL
cR>cL>cFR
L>D
R-FR>D
cR=cD
R-FR<D
Rp.FR-Rp > FRp.RFRp
cR>cD
LH211.16
CAB
No repression by cFR
cL≠cD
Rp effective
phyA/phyB redundancy
GT5874
LFR
Neither cR nor cL effective
cD>cL
FRp
effective
VLFR
Rp
ineffective
cR > cD
cR = cD
cL
effective
cR≤ cD
cR<cD
cR = cL
cD>cR>cL
ET6561
Repressed
By cFR
phyA
response
phyA/phyB redundancy
cR>cL>cFR
GT5971
Rp effective
cR not effective
ET6561
FRp
ineffective
cR > cD phyA/phyB antagonism
Rp
ineffective
cR
effective
ET5529
ET5599
ET5280
ET6649
GT5909
GT5927
cR = cD
No phyB response
GT6021
GT6075*
GT6604
GT6647
OK0319
OK0813
cR≤ cD
ET5555
ET5599
ET6649
cR =cL
phyA/phyB
redundancy
GT6281
GT6534
cR<cD
cL=cFR
cD>cR>cL
GT5971
cR>cL>cFR

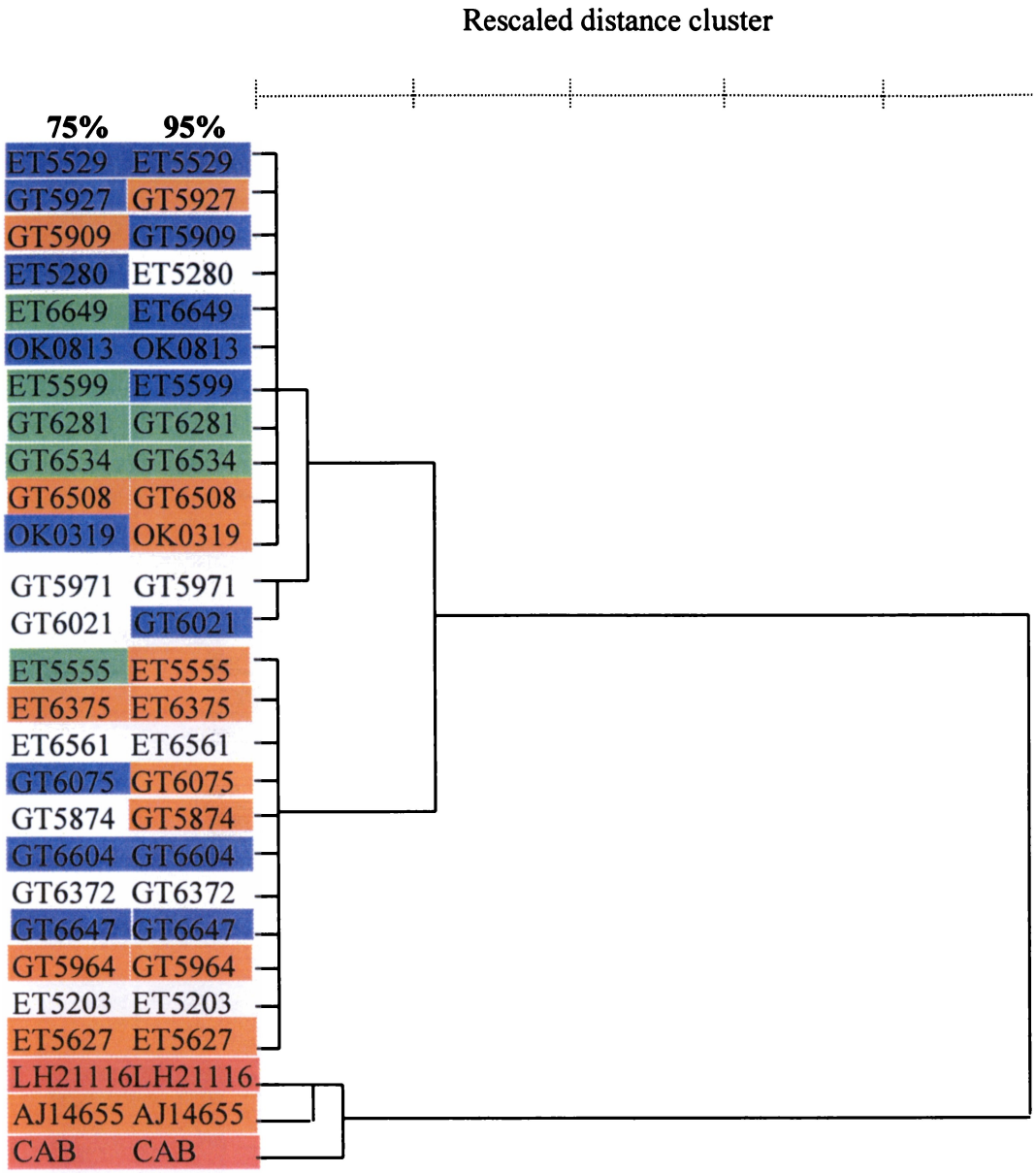


Figure 16. Hierarchical cluster analysis of GUS expression profiles according to Ward's algorithm (SPSS). Lines indicated by the same highlighted color, belonged to the same cluster in Figure 14. For details see text.

the Square Euclidean distances. That is, the distance between two objects (e.g. promoter trap lines x and y) is the sum of the squared differences between the values of the clustering variables (e.g. eight light treatments): $Sxy = \Sigma_i(X_i - Y_i)^2$. The algorithm continues one step at a time, joining pairs of objects, pairs of clusters, or an object with a cluster until all the data are in one cluster. Smaller distances indicate that fairly homogeneous clusters are joined, while larger distances are an indication that the members of the clusters are more dissimilar (SPSS®, 1999). The criterion for joining is that it should produce the smallest possible increase in the sum of squares (Sxy). Note that Ward's algorithm does not consider significance levels: the columns labeled 75% and 95% in figure 16 refer to the placement of the lines in Ward's clusters in the empirical cluster analysis (Figure 15).

For the purpose of comparison, lines belonging to the same cluster in Figure 15 were highlighted with the same color in Figure 16. The results of both clustering analyses are quite different although there were general similarities. For example, at the 95% level (i) the 'pink' lines CAB and LH21116 also cluster close to each other with Ward's method; (ii) five out of eight lines from the blue cluster (ET5529, GT5909, ET6649, OK0813 and ET5599) cluster together, (iii) six out of the ten lines from the 'yellow' cluster (ET5555, ET6375, GT6075, GT5874, GT5964 and ET5627) and (iv) the two lines (GT6281 and GT6534) from the 'green cluster' also cluster together. However, the yellow line AJ14655 appears along with LH21116 and CAB. This illustrates how the algorithm based-tree is not able to distinguish the lack of statistically significant differences between the cL and cD data in line AJ14655.

The lines belonging to the two biggest clusters from Figure 15 are scattered throughout the first three clusters from Figure 16, both at the 75% and 95% significance level. The interpretation of the algorithm-based cluster is complicated given that it is not taking into account the biological basis of light signaling, nor does it make use of the statistical confidence intervals defined by the ANOVA and Tukey tests.

4.5. Discussion

The expression profiles of gene/enhancer trap lines in response to cL, cD, cR, cFR, Rp, R-FRp, FRp and FR-Rp were analyzed quantitatively at the whole plant level in order to explore contributions of various phyA and phyB mediated light signaling modes at the scale of the entire signaling network. Some discrepancies between the in situ staining patterns and the fluorimetric whole-plant profiles were observed. However, these discrepancies can usually be explained based on the sizes of stained tissue, the staining only in small patches in a particular tissue or the occasional variability in the patterns of

staining (See Chapter 2 for discussion and Appendix 1 for individual patterns of staining).

The fluorimetric GUS activity data for the CAB:GUS fusion were compared with the well established profile in the published literature. The CAB gene is known to be repressed in darkness. It shows a weak FR-HIR under cFR. It also responds in a LFR mode to pulses of red light. This response is reversible by FR pulses. CAB is expressed strongly under cR and cL conditions (Kaufman et al., 1984, Nagy et al., 1986, Horwitz et al., 1988,). The analysis of the transgenic CAB promoter:GUS fusion was consistent with these findings: Low expression in cFR and after R-FRp and FRp was observed. Low but more variable expression was also observed in cD and after Rp. As expected, the highest expression was seen in cR and cL. However, it should also be pointed out that it was not possible to distinguish the established VLFR and FR-HIR of CAB because cD, FRp and cFR were not significantly different. More surprisingly, the LFR was not directly detected, i.e. when comparing Rp versus cD treated plants. Yet, an inhibitory effect of FR pulses following R pulses was significant at the 75% level (R-FRp), as expected for the phytochrome mediated LFR. In addition, R pulses did result in a significant (75%) induction compared to cD when preceded by FR pulses (FR-Rp), as expected for a LFR. Taken together, the CAB:GUS results suggest that the majority of conclusions reached at the 75% confidence level may be justified. The CAB:GUS results also underscore the wisdom of focussing the promoter trap analysis on dark-induced, light repressed lines. In these lines the GUS expression levels are in their majority well above background, which makes them better targets for statistical analysis.

From the discussion of the comparison between the in-situ and fluorimetric data of line LH211.16 (section 4.4.2) it is evident that the fluorimetric data fail to capture many of the intricate details of light regulation that are apparent from the in situ staining profiles. However the in situ data do not lend themselves to any meaningful statistical analysis. Moreover, the statistical differences apparent from the fluorimetric data are meaningful regardless of the underlying complexities of tissue-specificity. One possibility for preserving the depth of light regulatory information from the in situ assay for statistical analysis would be to extract GUS activity separately from different organs or tissues. Unfortunately, it is impractical to do this in five day-old Arabidopsis seedlings due to their small size.

PhyA and phyB share in the morphogenic control of different processes such as hypocotyl growth inhibition and cotyledon unfolding (Quail et al., 1995). Synergistic and antagonistic interactions between phyA and phyB have been described. PhyA and phyB act synergistically in hypocotyl growth and cotyledon unfolding in etiolated Arabidopsis seedlings exposed to FR followed by a terminal R pulse (Casal and Boccalandro, 1995),

whereas under cR or repeated Rp, phyA and phyB act antagonistically (Mazzella et al., 1997). It is important to understand whether phyA and phyB operate independently or show some interaction, not only with respect to the relatively small number of responses that can be scored morphologically, but also with respect to the much larger number of independent responses that are apparent at the level of gene expression. Analysis of the response of gene/enhancer trap lines to constant red and far-red identified genes affected by phyB or phyA, respectively, and also defined diverse types of interactions between them, as discussed below, by positioning 26 lines in 10 clusters in a classification tree of light response profiles (Figure 15).

Clearly, the clustering is in some way limited by the statistical resolution power of the GUS activity data. If a line is placed on a branch labeled 'equal response' to two different conditions, more detailed analysis of this line in the future might still discern a difference in the response, leading to a reappraisal of the clustering.

The positioning of the lines in the cluster (Figure 15) revealed some interesting facts:

(1) Without exception, all dark-inducible, i.e. light repressed lines were repressed by cFR (i.e. the phyA mediated FR-HIR), suggesting that light repression of dark-inducible genes is primarily a function of the phyA pathway. This interpretation makes sense in light of the high expression level of phyA in dark-grown seedlings (Casal et al., 2000).

(2) Concerning the relative contribution of phyA and phyB to the repression of gene expression, three different situations were distinguished given that phyA was always effective: (a) an opposite effect of cR in comparison to cFR and cL suggests an antagonistic interaction between phyA and phyB (ET5203). (b) A concerted effect of cR compared to cFR and cL suggests redundant or at least additive activities of phyA and phyB (green cluster). (c) One line shows a graded response (cD>cR>cL=cFR, GT5971). Compared to the green cluster (2b), this line suggests that phyB can operate at a reduced-power mode, but in a concerted fashion with phyA. (d) No effect of cR despite a repression by cFR (and cL) means that the phyB pathway can remain uncoupled from other light signaling pathways (blue cluster). This suggests that the repression of gene expression by phyA (see point 1) can operate in the absence of active phyB. If this repressory function of phyA was physiologically important, one might predict that phyA mutants, which show few abnormalities when grown in the light, have a deetiolation defect upon shift from cD to light.

(3) An effect by far-red pulses on gene expression suggests that a VLFR might be discernable here (GT5874, ET6561, GT5971), although the statistical support is weak.

(4) In a few cases the R-FRp or FR-Rp, which were not used for Figure 15, could distinguish between sub-clusters of lines.

In summary analysis of promoter trap lines allows the identification of clusters of genes with similar profiles. Lines with phyA response only, lines in which phyA/phyB acted redundantly and lines displaying phyA/phyB antagonism were found.

Chapter 5. Analysis of genomic flanking sequences of promoter trap insertions

5.1. Introduction

Light is an essential environmental factor that controls plant development. One of the primary steps in which light exerts its control is regulation of gene expression at the transcriptional level. Deletion and mutagenesis analysis of light regulated promoters has led to the identification of several light responsive elements (LRE's) (Anderson et al., 1994, Kehoe et al., 1994, Bruce and Quail, 1990, Ha and An, 1988). Some LREs such as the G (CACGTG), GT1 (GGTTAA) and GATA boxes are commonly found in the promoters of light-regulated genes and these LREs have been shown to be necessary for light-controlled transcriptional activity (Terzaghi and Cashmore 1995). No single LRE has been found in all light-regulated promoters (Arguello-Astorga and Herrera-Estrella, 1996), and some of the LREs can be present in promoters that are not light regulated (Chattopadhyay et al., 1998b). Moreover, certain LRE's function to integrate the signals from different pathways. For example, the binding site for the factor CCA1 may be required for the response to phytochrome (Wang et al., 1997) and the circadian clock (Wang and Tobin, 1998). This indicates the variety of LRE's known to have a role in transcriptional activity.

Diverse patterns of gene expression can be generated from a small set of differentially regulated factors as shown by the dependence of tissue specific expression on factor concentration and the copy number of the binding sites in tobacco (Lam and Chua, 1990). For example, the presence of a single binding site for ASF-1 (Activation sequence factor 1) confers expression preferentially in the root whereas a tetramer of the same site enhances leaf expression considerably.

Generally, a regulated promoter consists of multiple ubiquitous elements in a defined arrangement and the organization of these elements plays an important role in the signal responsiveness of the promoter (Khurana et al., 1998). The minimal sequence requirement necessary for the integration of light and developmental signals controlling promoter activity has been defined by analyzing individual LREs and selected combinations of LREs for their ability to confer light responses to non-light regulated basal promoters (Puente et al., 1996). It was found that combinatorial interactions of the cis-elements within a promoter define the tissue specificity as well as the light responsivenes of the synthetic promoters analyzed. Only pair-wise combinations of distinct LREs but not an individual LRE were able to direct light inducible expression of the GUS reporter gene attached to a minimal promoter. The combinatorial LREs, or the

trans-acting factors interacting with those LREs may be regarded as signal integration points in the network mediating both light and developmental control of gene expression.

Light can have a positive or negative effect in transcription and promoter elements involved in both types of effects have been described (Castresana et al., 1988). Compared to elements responsible for light-inducibility, few elements have been characterized that mediate downregulation in response to light or upregulation in darkness. Among the latter are a TGGG sequence, active in the down-regulation of phyA and asparagine synthase 1 (AS1) by phytochrome (Bruce and Quail, 1990, Neuhaus, et al., 1997). Another example is a 12 bp element (GGATTTTACAGT) of the pra2 gene. This element is involved in the down-regulation by phytochrome (Inaba et al., 1999) and is capable of receiving signals from various photoreceptors and conferring dark induction to a minimal promoter.

Promoter trapping is a powerful method for the analysis of gene regulatory regions in relation to their expression patterns. Sundaresan et al. (1995) developed an insertional mutagenesis system for the identification of genes by their patterns of expression during development. A modified maize Dissociation (Ds) transposable element carrying a GUS reporter gene was used as either a gene trap or an enhancer trap for detection of genes by their gene expression patterns. The flanking sequences of the tagged genes can be readily identified by different methods (Hui et al., 1998, Liu and Whittier, 1995). In this experiment TAIL-PCR (Liu and Whittier, 1995) was used for the identification of the flanking sequences of promoter traps generated.

Thermal asymmetric interlaced (TAIL)-PCR involves three rounds of PCR with nested sequence-specific primers together with a single degenerate primer of arbitrary sequence, so that the relative amplification efficiencies of specific and nonspecific products can be thermally controlled. This control is possible due to the lower melting temperature (Tm) of the degenerate primer. In the first round of PCR, a low-stringency PCR cycle is conducted to amplify one or more annealing sites for the degenerate primer in the targeted sequence. The specific product is then preferentially amplified over nonspecific ones by alternating two high-stringency PCR cycles with one low-stringency PCR cycle. The PCR products from the first round then serve as a template for a second round of PCR, in which more specific amplification of the flanking sequence is achieved with a nested PCR primer and the arbitrary primer (Liu and Whittier, 1995). Finally, a third round of PCR is done with a third nested primer. The final PCR product, which is expected to contain a portion of the GUS coding region, as well as the genomic flanking sequence, is sequenced. By matching the flanking sequence against the entire Arabidopsis genome sequence, one can then build a picture of the chromosomal context that gave rise to the specific pattern of GUS expression in the line analyzed.

Subsequently, further analysis of the putative promoter sequences is possible. For example, one can search for common elements in the promoters of trapped genes that share common expression characteristics, or one can ask whether the expression profile of a promoter trap line can be understood on the basis of known promoter sequence motifs. Eukaryotic promoters are usually composed of multiple discrete functional modules that contain one or more recognition sites for proteins triggering or repressing the transcription of a gene. Identification of those recognition sites can provide information about the regulation of the gene in response to different environmental conditions.

The flanking sequences of 42 promoter traps were analyzed. Although most of the genes analyzed were light repressed, few motifs involved in light repression were found, suggesting the action of unknown motifs that confer induction by darkness or repression by light.

5.2. Rationale

Chapters 2 and 3 have described that groups of promoter traps shared similar light response profiles. Moreover, many of the lines had expression profiles that differed sharply from the standard, well-studied, expression profile of the CAB promoter. Therefore, it was of particular interest to examine the genomic flanking sequences of a subset of promoter trap lines in order to gain insight into the promoter sequences driving these diverse profiles of light responsiveness. Some of the questions to be addressed include:

a) Are the promoter trap lines typically reflecting the expression pattern of endogenous genes?
b) Do dark-inducible promoters differ from light-inducible promoters with respect to recognizable sequence elements?
c) Are the light responsive promoter traps flanked by well-known LRE sequences?
d) Are there conserved sequence elements in promoters with similar expression profiles?

The initial goal was to identify the flanking sequences of all those promoter traps for which a detailed light response profile had been determined (Chapter 3). The flanking sequences were determined for a total of 42 promoter traps. The availability of flowering stages and the success in the isolation and sequencing of the PCR products determined which lines were sequenced. Even though, in some cases apparently good DNA was obtained from a particular line, technical difficulties in the PCR reaction or the sequencing reaction did not allow the recovery of a reliable sequence.

5.3. Materials and methods

5.3.1. DNA preparation and TAIL-PCR

Genomic DNA from the transposon lines was prepared as follows: 3 or 4 inflorescences or 10 to 15 seedlings were ground in the presence of 100 µl of homogenization buffer (100 mM Tris/Cl, pH 7, 100 mM NaCl, 10 mM EDTA pH 8, 1% SDS). The homogenate was spun for 30 seconds and the supernatant was transferred to a fresh tube and extracted twice with phenol. Nucleic acid was ethanol-precipitated and the pellet re-suspended in 15 µl TE buffer. The DNA was stored at -20°C for further analysis.

TAIL-PCR was performed as described by Liu and Whittier (1995) using a Peltier PTC-200 (MJ Research) thermocycler. Three successive rounds of amplification were performed using three semi-nested primers specific for either the 5' or 3' end of the Ds element and one or two arbitrary degenerate (AD) primers:
Ds5-1: 5'-ACGGTCGGGAAACTAGCTCTAC-3'
Ds5-2: 5'-CCGTTTTGTATATCCCGTTTCCGT-3'
Ds5-3: 5'-TACCTCGGGTTCGAAATCGAT-3'
Ds3-1: 5'-ACCCGACCGGATCGTATCGGT-3'
Ds3-2: 5'-CGATTACCGTATTTATCCCGTTC-3'
Ds3-3: 5'-GTATTTATCCCGTTCGTTTTCGT-3'
AD-2: 5'-NGTCGA(G/C)(A/T)GANA(A/T)GAA-3'
AD-5: 5'-(A/T)CAGNTG(A/T)TNGTNCTG-3'

A master mix was prepared in an eppendorf tube by adding 11 µl of sterile water, 2 µl Ds5-1 (or Ds3-1) primer (2 µM stock), 3µl AD-2 or AD-5 primer (20 µM stock), 2 µl 10x buffer, 2 µl 10x dNTPs and 0.2 µl Taq polymerase for each DNA sample. 19 µl of the master mix were added to 1 µl DNA from each sample. The solution was mixed by pipetting up and down and the PCR was performed as described by Liu and Whittier (1995). The primary PCR products were diluted 1:50 in sterile water and 1 µl of this dilution was used for the secondary PCR reaction with 11 µl of water, 2 µl of Ds5-2 (or Ds3-2) primer stock, 2 µl AD2 primer stock, 2 µl 10X buffer, 2 µl 10X dNTPs, and 0.2 µl Taq polymerase for each sample tube. The secondary PCR products were diluted 1:50 in sterile water and 1 µl of this dilution was used for the tertiary PCR reaction that 11 µl of water, 2 µl of Ds5-2 primer stock, 2 µl AD2 primer stock, 2 µl 10X buffer, 2 µl 10X dNTPs, and 0.2 µl Taq polymerase for each sample tube. The product of this PCR reaction was purified with a Qiagen column and sequenced with the PRISM Ready reaction DyeDeoxy Terminator Cycle sequencing kit (ABI) using the Ds5-3 (or Ds3-3) primer (Figures 17 and 18).

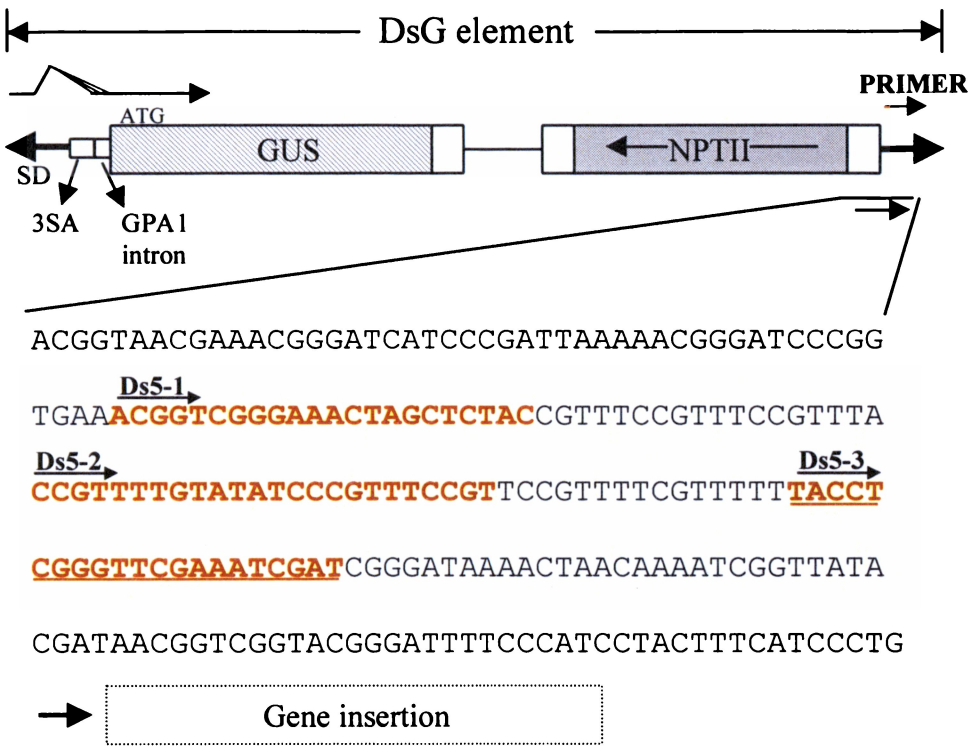


Figure 17. Location of specific Ds5 primers used for TAIL-PCR (red) and sequencing (red underlines). At the top is a schematic diagram of the construct indicating the position of the primers (red arrow). SD= splice donors (D1, D2, D3); 3SA splice acceptors (A1, A2, A3); NPTII= Neomycin phosphotransferase; GPA1= intron for G-protein α subunit from Arabidopsis.

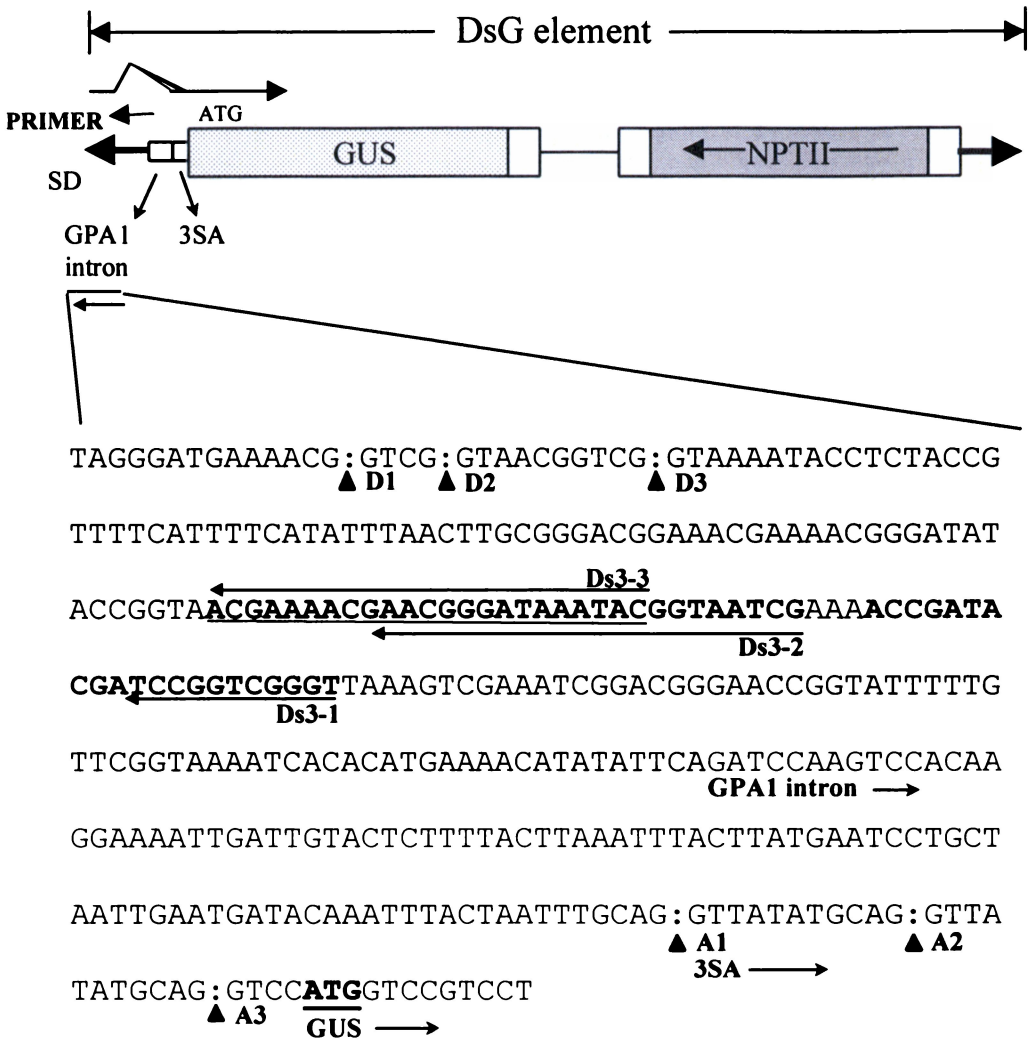


Figure 18. Location of specific Ds3 primers used for TAIL-PCR (red) and sequencing (red underlines). At the top is a schematic diagram of the construct indicating the position of the primers (red arrow). SD= splice donors (D1, D2, D3); 3SA splice acceptors (A1, A2, A3); NPTII= Neomycin phospho-transferase; GPA1= intron for G-protein α subunit from Arabidopsis.

The sequence was compared against the completed Arabidopsis genome sequence using the BLAST program from the National Center for Biotechnology Information, NCBI (http://www.ncbi.nlm.nih.gov/BLAST/), the Munich information center for protein sequences, MIPS database (http://mips.gsf.de/proj/thal/db/index.html) and the Institute of Genomic Research, TIGR database (http://www.tigr.org/). Graphic representations of the insertion points (Appendix 2) were generated based on sequence annotations found at NCBI or MIPS. In few cases when there were possible problems with the annotation, the sequence of the genes identified with the BLAST search was analyzed by using gene prediction programs, including:
GeneFinder (http://argon.cshl.org/genefinder/ARAB/arab.htm),
Genescan (http://bioweb.pasteur.fr/seqanal/interfaces/genscan.html) and the
Grail exon prediction program (http://grail.lsd.ornl.gov/grailexp).

Analysis of the promoter elements was done using the database PLACE (Higo et al., 1999) (http://www.dna.affrc.go.jp/htdocs/PLACE/) to search for regulatory motifs 500 bp upstream of the insertion point. This database contains motifs found in cis-acting regulatory DNA elements of vascular plants.

5.3.2. Statistical analysis of promoter motifs

The frequency of selected motifs among 15 GT and 14 ET flanking sequences was tested for significant discrepancies with the frequencies expected by chance alone. First, the likelihood ($p^{(I)}$) of the motif occurring at a given position in the genome was calculated. The fractions of A(29%), T(29%), C(21%) and G(21%) in the Arabidopsis genome (Leutwiler et al., 1984) were used as a basis to calculate $p^{(I)}$. For example for the ASF1 site, TGACG, $p^{(I)} = 0.29x0.21x0.29x0.21x0.21 = 0.78 \times 10^{-3}$. Then the average frequency of the motif in a stretch of 500 bp was calculated as m = 500 times p(I). Next, the Poisson distribution was applied to derive the likelihood (k) that a given line has at least one copy of the motif (see below). Next, a binomial expansion was used to estimate the likelihood of finding a given number of lines (x) among all the GT (n=15) or ET (n=14) with at least one copy of the motif. This probability was calculated as follows:

$$\begin{bmatrix} n \\ x \end{bmatrix} \cdot k^x \cdot (1-k)^{n-x} = \frac{n!}{x!\,(n-x)!} \cdot k^x \cdot (1-k)^{n-x}$$

Where n = total number of lines
x = given number of lines where the motif was found
k = probability that a given flanking sequence has at least one copy of the motif:
$k = 1 - p^{(o)}$
$p^{(o)} = e^{-m}$

5.4. Results

5.4.1. Chromosomal environment

To learn more about the genomic context contributing to the light response profiles of selected promoter trap elements, the flanking sequence of a total of 42 insertion lines was determined (see Methods) by sequencing the TAIL-PCR products. Out of 42 analyzed, 55% were enhancer trap (ET) and 45% gene trap (GT) lines. Data from 6 lines were obtained from CSHL (Lines ET5280, ET5359, ET5487, ET5627, ET5653, ET6633). Two lines were sequenced at both the University of Tennessee and CSHL (ET5403 and ET6428), resulting in the same sequence. Three sequences were taken from the NCBI database (GT5909, GT5957 and GT6227). One additional line (GT6228) had been sequenced here resulting in a different sequence than one retrieved from the NCBI database. The sequence obtained at the University of Tennessee was used because it was of better quality and it had the Ds element included whereas the sequence from NCBI had a poorer quality and lacked the Ds element sequence suggesting a PCR artifact. In summary, a total of 33 lines were sequenced at the University of Tennessee, 6 at CSHL and 3 sequences were taken from the NCBI database. Of these 42 lines, the GUS fluorimetric expression profiles had been determined for 30 lines. The other 12 sequenced lines did not have enough replicates of the fluorimetric data to be used in the statistical analysis described in chapter 4.

For the purpose of quality control, the sequences of the TAIL-PCR fragments were examined to confirm the presence of a string of expected basepairs from the 3' or 5' end of the Ds:GUS element. An average of 9 base pairs of the Ds-GUS sequence were readable with high confidence. The Blast search was able to match the remaining TAIL-PCR sequence with a unique sequence from the Arabidopsis genome in all cases. Nevertheless, one line (OK011.22 was omitted from subsequent analysis because of discrepancies in the annotation of the sequence in the database. No evidence for insertion into repetitive DNA was found. Aside from technical sequence ambiguities, small discrepancies between the experimental and database sequences were explained by ecotypic differences (the promoter traps are in Landsberg and the database genome sequence is from the Columbia ecotype).

Insertions occurred in all the chromosomes (Figure 19). The chromosomal distribution was as follows: 38% in chromosome I, 19% each in chromosomes II and III and 12% each in chromosomes IV and V. Some clustering of the insertions at the end of the chromosomes was observed (Figure 20). For this figure, the total length of each chromosome was taken as 100%, then this percentage was divided in 10% increments

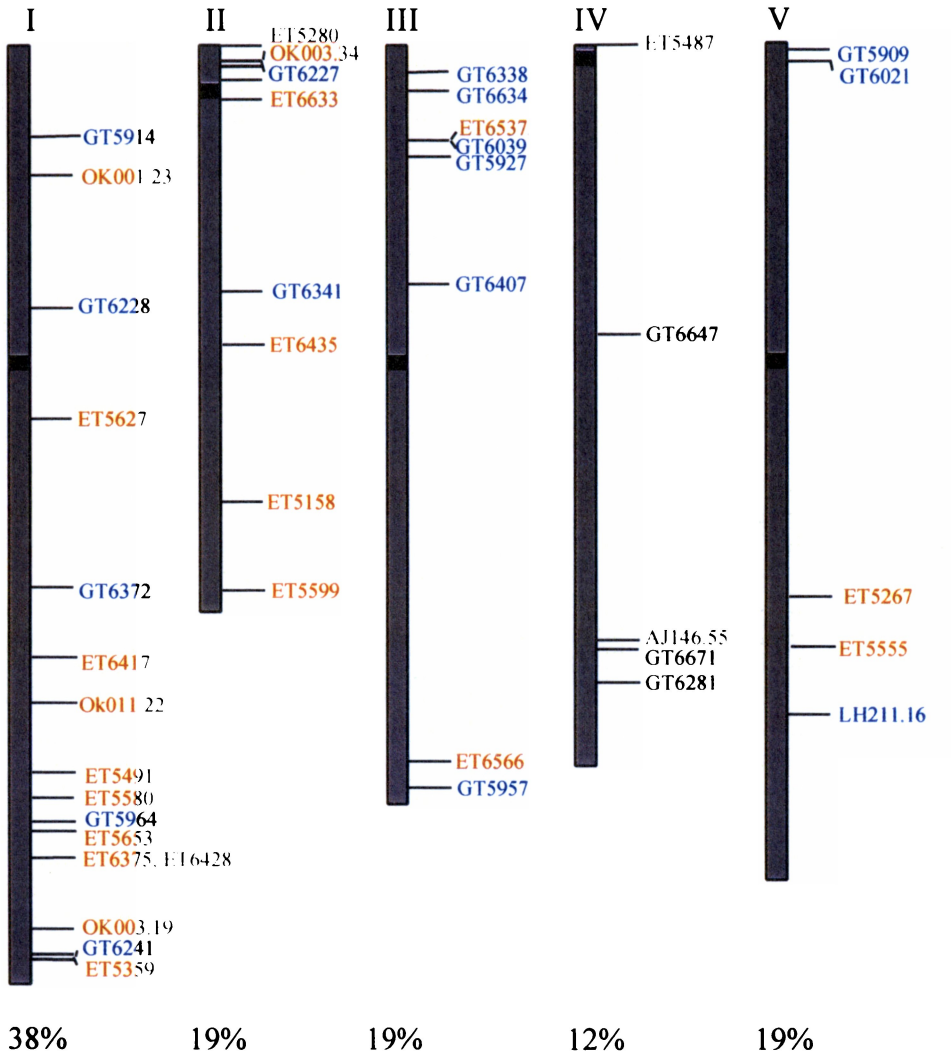


Figure 19. Localization of the insertion points of gene trap (blue) and enhancer trap (red) elements drawn to scale in the Arabidopsis genome.

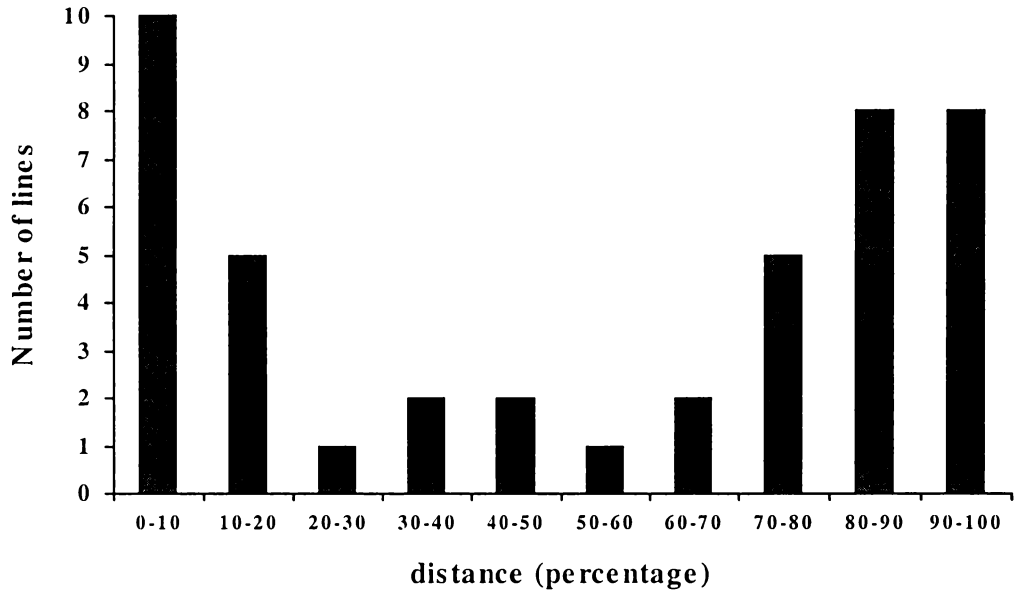


Figure 20. Number of inserted genes at different positions in the chromosome illustrating clustering of inserted genes at the end of the chromosomes. The total length of each chromosome was taken as a 100% and the chromosome was divided in 10% increments. The number of lines with insertions at each region was recorded.

and the number of insertions in each region was determined. A bias towards insertion at the ends of the chromosomes was evident. Clustering of the insertion points and a higher frequency of insertions at the end of the chromosomes has been observed previously (R. Martienssen, personal communication). This bias was only partially explainable by the tendency of Ds elements to hop to linked sites.

Graphical displays of the insertion sites of the Ds:GUS elements including the coordinates of predicted genes, are shown in Appendix 2. Two examples of these graphical representations are given in Figure 21, including the insertion points of GT6228, with an insertion within a gene, and GT6281 with an insertion between two genes. The clone where the insertion occurred was represented by a line with arrows representing the direction of the annotated ORFs drawn to scale. A closer detail of the insertion within a gene in line GT6228 is given in Figure 22.

Analysis of the insertion sites revealed that 55% of the lines had inserts outside of a predicted ORF ('between genes') and 45% of the lines displayed insertions between the start and stop codons of a predicted ORF ('within genes' Table 17). Lines with insertions within an ORF were classified according to the direction of the insertion (sense or antisense). Examples of GT and ET lines in any orientation (sense and antisense) were found.

There was no apparent preference for insertion of the Ds element into the amino terminal half or carboxyl-terminal half of a predicted coding region (Figure 23). In a previous analysis, Parinov et al., 1999 found a preference for insertions into the 5' ends of genes of Arabidopsis transposant lines. In other systems such as the Drosophila P and yeast Ty1 elements, the insertions occurred preferentially into upstream regions of the genes (Liebman and Newman, 1993, Spradling et al., 1995).

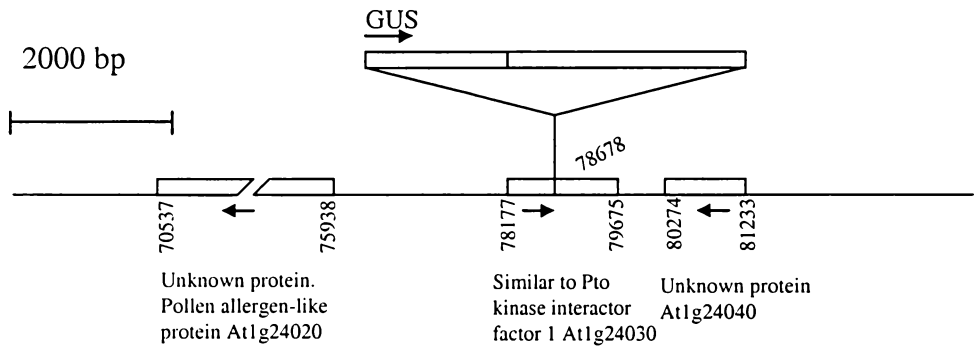
5.4.2 Description of the lines with insertions within a gene

Tables 18 and 19 show the current annotations of genes that have promoter trap insertions within the coding region.

5.4.2.1. Gene expression

All of the insertion points corresponded to predicted unknown or hypothetical proteins and therefore the lack of expression data made it impossible to compare the observed GUS expression of the promoter traps with expectations from published research. Insertions of the GUS reporter in the sense direction with respect to the tagged gene may reflect the expression pattern of the tagged gene. Only two of the insertions were in frame

A. GT6228: A. thaliana chromosome 1 BAC T23E23 (Ds5-3)



B. GT6281: A. thaliana chromosome 4 BAC M4E13 (Ds5-3)

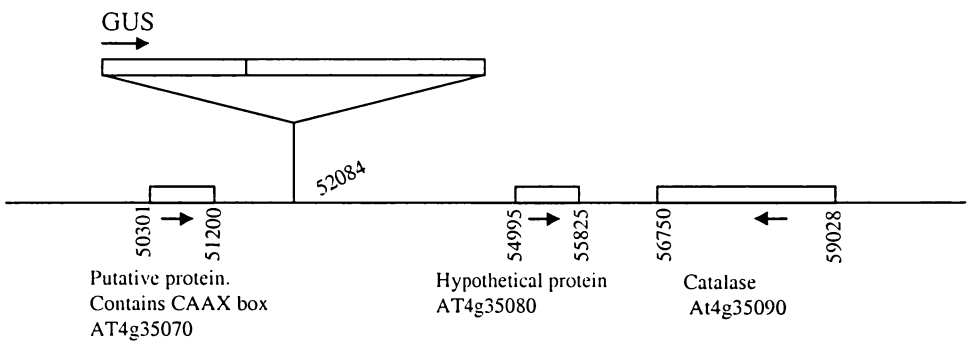
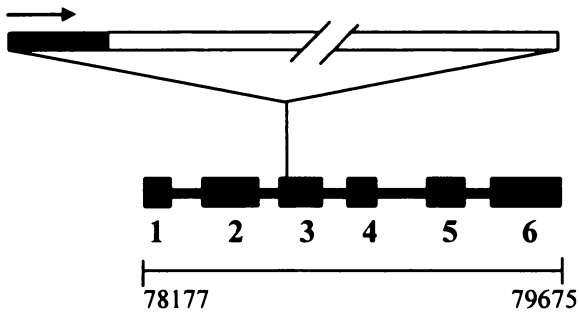


Figure 21. Graphic representation of the insertion points of two gene trap lines: A. GT6228 with an insertion within a gene and B. GT6281 with an insertion between two genes. The single long line represents the region of the clone where the insertion occurred. Each box represents the annotated gene drawn to scale with the coordinates being the putative start and stop codons. The arrows indicate the direction of the ORF. A scale of 2000 bp is shown for reference.



1. Exon: 78177-78269
2. Exon: 78388-78586
3. Exon: 78655-78818
4. Exon: 78908-79012
5. Exon: 79211-79333
6. Exon: 79436-79675

Figure 22. Detail of the exon-intron organization ('gene model') of gene At1g24030) (Similar to Pto kinase interactor factor 1), the insertion point of line GT6228. The exons are indicated in red. The ORF is between base pairs 78177 and 79675 of BAC T23E23.

Table 17. Classification of gene and enhancer trap lines with insertions within a gene or between two genes. The lines with insertions within a gene were also classified as having an insertion in the sense or antisense direction.

Summary of insertions			
Within a gene		Between two genes	
Sense	Antisense		
ET lines			
ET5599	ET5280	ET5158	ET6417
ET5653	ET5580	ET5267	ET6435
ET6537	ET6428	ET5359	ET6566
OK001.23		ET5487	ET6633
OK011.22		ET5491	OK003.19
		ET5555	OK003.34
		ET5627	AJ146.65
		ET6375	
GT lines			
GT6228	GT5909	GT5927	
GT6241	GT5914	GT5964	
GT6671	GT5957	GT6021	
	GT6039	GT6281	
	GT6338	GT6227	
	GT6372	GT6341	
	GT6647	GT6407	
	LH211.16	GT6634	

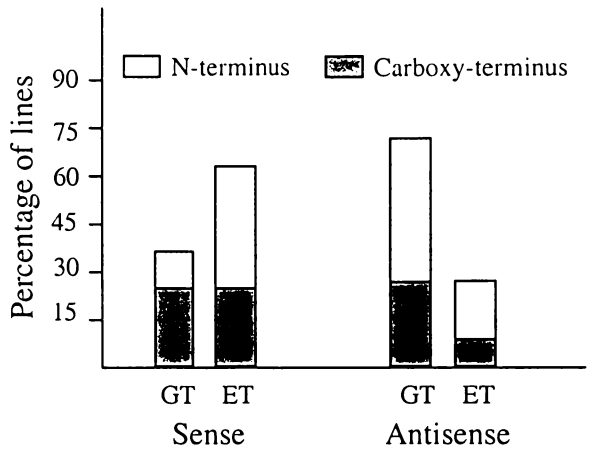


Figure 23. Percentage of ET and GT lines with insertions at either the amino-terminal half or the carboxy-terminal half of a protein-coding region.

Table 18. Annotation (from MIPS and NCBI) of the insertion points of Ds:GUS elements in the sense direction. Note that none of the annotations have been verified experimentally.

Sense insertions					
	Gene	**Protein**	**Exon/ Intron**	**In frame?**	**Prediction**
Enhancer trap					
ET5599	At2g48010	Putative protein kinase	exon	no	Two trans-membrane (tm) domains
ET5653	At1g69870	similar to peptide transporter	exon	yes	8 tm domains Peptide transporter
ET6537	At3g09600	MYB related	exon	no	Chloroplast targeted. Transcriptional control
OK001.23	At1g11960	Putative protein	exon	no	6 tm domains
OK011.22	At1g60100	Plutative protein	exon	no	
Gene trap					
GT6228	At1g24030	Similar to Pto kinase interactor factor 1	exon	yes	1 tm domain. Lipid- fatty acid metabolism
GT6241	At1g80760	Nodulin-like protein	exon	no	5 tm domains (transport facilitator)
GT6671	At4g33040	Putative protein similar to Glutaredoxin protein	exon		Possible function: Electron transport, detoxification

Table 19. Annotation (form MIPS and NCBI) of the insertion points of Ds:GUS elements in the antisense direction. Note that none of the predictions have been verified experimentally.

Antisense insertions				
	Gene	**Protein**	**Exon/ Intron**	**Prediction**
Enhancer trap				
ET5580	At1g67330	unknown	exon	one tm domain.
ET5280	At2g01340	unknown	exon	one tm domain
ET6428	At1g72170	unknown	exon	One tm domain
Gene trap				
GT5909	At5g01540	Similar to receptor like protein kinase	exon	Signal transduction protein
GT5914	(not numbered)	Similar to phloem specific lectin	exon	
GT5957	At3g62700	ABC transporter-like protein	exon	13 predicted tm domains (transport facilitation)
GT6039	At3g09530	Putative protein	exon	extracellular transport (secretion)
GT6338	(not numbered)	predicted exon	exon	
GT6372	At1g48670	Similar to GH3 auxin-responsive protein	exon	Cell division and DNA synthesis
GT6647	At4g14180	Putative protein	exon	three tm domains (electron transport)
LH211.16	At5g54730	Similar to unknown protein (chloroplast targeting)		Assembly of protein complexes

fusions between the tagged ORF and GUS (ET5653 and GT6228). However, even out of frame fusions may result in expression of a fusion protein because the splice acceptor sites that reside upstream of the GUS ORF in the Ds element may lead to an alternatively spliced fusion transcript containing an ORF with endogenous sequences and GUS.

5.4.2.2. Function

Predicted functional categories of the genes with insertions in the sense direction include (Table 18) signal transduction (line ET5599), peptide transport (ET5653), electron transport and detoxification (GT6671), lipid metabolism (GT6228) and transcriptional control (ET6537). Functional categories for the antisense insertions (Table 19) include signaling (ET5580), secretion (GT6039), secondary metabolism (GT6228), cell division and DNA synthesis (GT6372), electron transport (GT6647), and signal transduction (GT5909). Among the eight sense insertions, transmembrane domains were predicted for five of them (ET5599, ET5653, OK001.23, GT6228, GT6241). This variety of functions illustrates the variety of genes that can be tagged using promoter trapping.

5.5 Gene expression profiles and chromosomal environment

Gene trap lines that are inserted either between genes or within a gene in the antisense orientation are not expected to result in GUS expression. However, based on the flanking sequences identified and their annotations, 8 out of 19 GT lines fell into this class (Table 17). One possible explanation for the GUS expression of these GT lines may be the proximity of cryptic promoters (Irniger et al., 1992; Foster et al., 1999). A cryptic promoter is defined as a DNA sequence that has fortuitous promoter activity when queried experimentally even though it does not drive the expression of a recognizable gene in its original chromosomal context. Alternatively, the annotation of the genome may be flawed. Given that few cryptic promoters have been described in the literature, it was important to know whether the GT lines in question shared a particular characteristic in their expression profiles. As shown in Figure 24, there was no recognizable commonality among the GT lines speculated to be driven by cryptic promoters, and no correlation between the specific gene expression profiles and the type of insertion. Therefore, if cryptic promoters are responsible, their expression characteristics can be highly variable.

5.6 Promoter analysis

For 15 GT and 14 ET lines, the sequence upstream (500 bp) of the insertion point was analyzed for the presence of promoter motifs using the Web Signal Scan program with

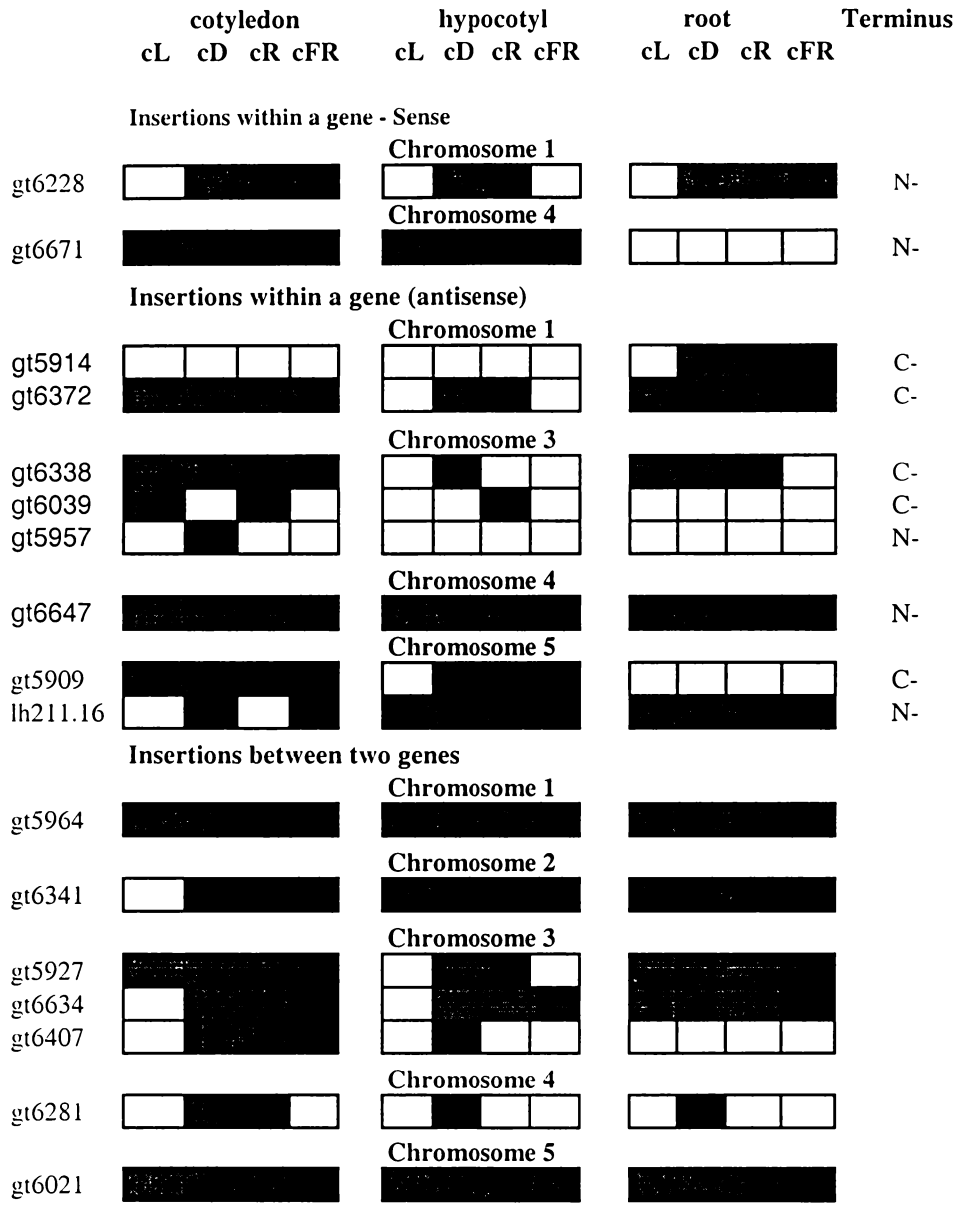


Figure 24. Comparison of organ-specific light response profiles for GT lines surmised to be regulated by cryptic promoters. A dark box stands for GUS positive in the given organ and light condition. Insertions at the N- and C-terminal half were indicated for the lines with insertions within a gene.

The PLACE database (http://www.dna.affrc.go.jp/htdocs/PLACE/, Figures 25 and 26). This database contains plant promoter motifs and is updated regularly. For the 8 GT lines with insertions within genes, three lines with insertions in the sense direction (GT6228, GT6241 and GT6671) and five lines with antisense insertions (GT5914, GT6338, GT6039, GT6647, GT6372) were inspected. Additionally, seven lines with insertions between genes (GT6634, GT5927, GT6341, GT6281, GT5964, GT6021, GT6407) were included in the analysis (Figure 25). No fluorimetric data were available for line GT6241. For the ET lines, the upstream sequence of 8 lines with insertions between two genes (ET5158, ET5267, ET5359, ET5491, ET5555, ET6417, ET6566, ET6633) and six lines with insertions within a gene (ET6537, ET5653, ET5599, ET5280, ET6428, ET5580) was analyzed (Figure 26).

In order to understand how the flanking sequences can drive any kind of gene expression of the promoter traps, the presence of two core promoter motifs, the TATA and CAAT boxes, was analyzed. All of the GT and ET lines have at least one CAAT box and most of the lines have 2 or more CAAT boxes. Two different probabilities p(a) and p(b) were calculated. p(a) is the probability of having 'x' or more number of lines showing a particular motif by chance alone. p(b) is the probability of having 'x' or less number of lines with a particular motif (where 'x' indicates the observed number of lines displaying the motif). Therefore, a p(a) value below 0.05 indicates that the motif is found particularly frequently among the promoter trap insertion sites analyzed, whereas a p(b) value below 0.05 indicates that a motif is less frequent than expected by chance alone.

Among the general transcriptional activation motifs, the frequency of the CCAAT motif was clearly higher than expected by chance alone p=0.013) in the GT lines but this that was not the case in the ET lines. For the CAAT motif p(a) was not significant for the ET or GT lines. The frequency of the TATA box (combined) was somewhat higher than expected by chance alone among the GT lines (p=0.067) to find 10 or more GT lines with a TATA box). Interestingly, the frequency of assorted TATA boxes was also clearly elevated among the ET lines (p(a)=0.002), even though the endogenous TATA motif within the ET element should make this element independent of a chromosomal TATA box.

Among the motifs known to be involved in light regulation, the GATA box, the GT1 consensus and the IBOX core motifs were frequently found in the GT and ET lines analyzed. The GATA box is required for high level, light regulated and tissue specific expression. It is conserved in the promoters of all CAB type I genes (Gilmartin et al., 1990; Lam and Chua, 1989). The GT1 consensus is a protein binding site in many

Figure 25. Number of promoter motifs found in the 500-bp region upstream of the insertion point of 15 GT lines containing insertions between and within genes. P(II): Average number of occurrences of the motif within any 500 bp of Arabidopsis sequence. p(a): Probability that the observed number *or more* of 15 lines contain the motif. p(b): Probability that the observed number *or fewer* of the 15 lines contain the motif. M(A/ C), N(A/C/G/T), R (A/G), W(A/T).

Description	Sequence	Average in 500 bp p(II)	Between GT6634	GT5927	GT6341	GT6281	GT5964	GT6021	GT6407	Within sense GT6228	GT6241	GT6671	Within antisense GT5914	GT6338	GT6039	GT6647	GT6372	p(a)	p(b)
CAAT boxes																			
CAAT box	CAAT	2.56																0.299	1
CCAAT box	CCAAT	0.538																0.013	0.997
TATA boxes																			
TATABOX2	TATAAAT	0.086																0.030	0.994
TATABOX3	TATTAAT	0.086																0.355	0.879
TATABOX4	TATATAA	0.086																0.355	0.879
TATABOX5	TTATTT	0.297																0.165	0.935
TATA box Phaseolus	TTTATATA	0.025																0.313	0.948
combined		0.581																0.067	0.979
Light regulation																			
AT1 box	AATATTTTTATT	1.8E-04																1	0.997
light regulationASF1	TGACG	0.389																0.976	0.093
phytochrome regulation	AAMAATCT	0.031																0.076	0.990
Circadian expression	CAANNNNATC	0.156																0.371	0.840
Circadian expression	AAAATATCT	0.005																1	0.929
GATABOX	GATA	2.561																0.976	0.104
G-box	MCACGTGGC	1.8E-03																1	0.973
G-box	ACCACGTGGC	2.2E-04																1	0.997
G-box	CACGTG	0.082																1	0.318
Coupling G-box	CGACG	0.282																1	0.019
GT1CONSENSUS	GRWAAW	1.485																0.115	0.979
GT1CORE	GGTTAA	0.156																0.660	0.629
GT1-rbcS	KWGTGRWAAWRW	7.6E-03																1	0.899
H-box consensus	CCTACCNNNNNNNC	5.0E-03																1	0.928
IBOX	GATAAG	0.156																0.660	0.629
IBOXCORE	GATAA	0.743																0.759	0.425
rbcS light regulation	AAATTAACCAA	3.2E-04																1	0.995
CAB1 promoter	ATACGTGT	9.5E-03																1	0.875
T BOX	ACTTTG	0.156																1	0.113
BOXII-light regulation	ACGTGGC	0.017																1	0.786
LRE	ACGTGGCA	5.0E-03																1	0.928
LRE	AACCTAACCT	5.8E-04																1	0.991
LRE	TCCACGTGGC	2.2E-04																1	0.997
LRE	ACTATTTTCACTAT	1.2E-06																1	0.999
rbcS consensus	AATCCAA	0.045																0.140	0.974
Phytochome regulation	AACCAA	0.156																0.660	0.629
Phytochome regulation	CGGATA	0.113																0.211	0.932
Light repression																			
ATHB2-binding site	TAATMATTA	0.013																1	0.829
Box B- repression	AAACGACACCGTTT	2.1E-06																1	0.999
Box C- repression	CTCCCAC	0.017																1	0.773
Box C- repression	TCCCGGTACACAC TTCTT	5.8E-09																1	0.999
DE1- light repression	GGATTTTACAGT	4.9E-05																1	0.999
GT2-Light repression	GCGGTAATT	2.0E-03																1	0.971
Light repression	GAAATAGCAAATG TTAAAAATA	2.0E-10																1	0.999
Light repression	CAGCTCCCATGGC TCTCCCATCCGCG CCGGT	8.9E-18																1	
RE1-light repression	CATGGGCGCGG	3.3E-05																1	0.999
Miscelaneous																			
Core DNA binding	AAAG	2.561																0.299	1
Pollen specific	AGAAA	0.743																0.085	0.973
Root motif	ATATT	1.026																0.536	0.674
C-box-transcription	GACGTC	0.082																0.330	0.893

Figure 26. Number of promoter motifs found in the 500-bp region upstream of the insertion point of 14 ET lines containing insertions between and within genes. For explanations see legend to Figure 25. The number of the motifs is indicated. K(G or T), M(A or C), N(A,C,G,or T), R(A/G), W(A or T).

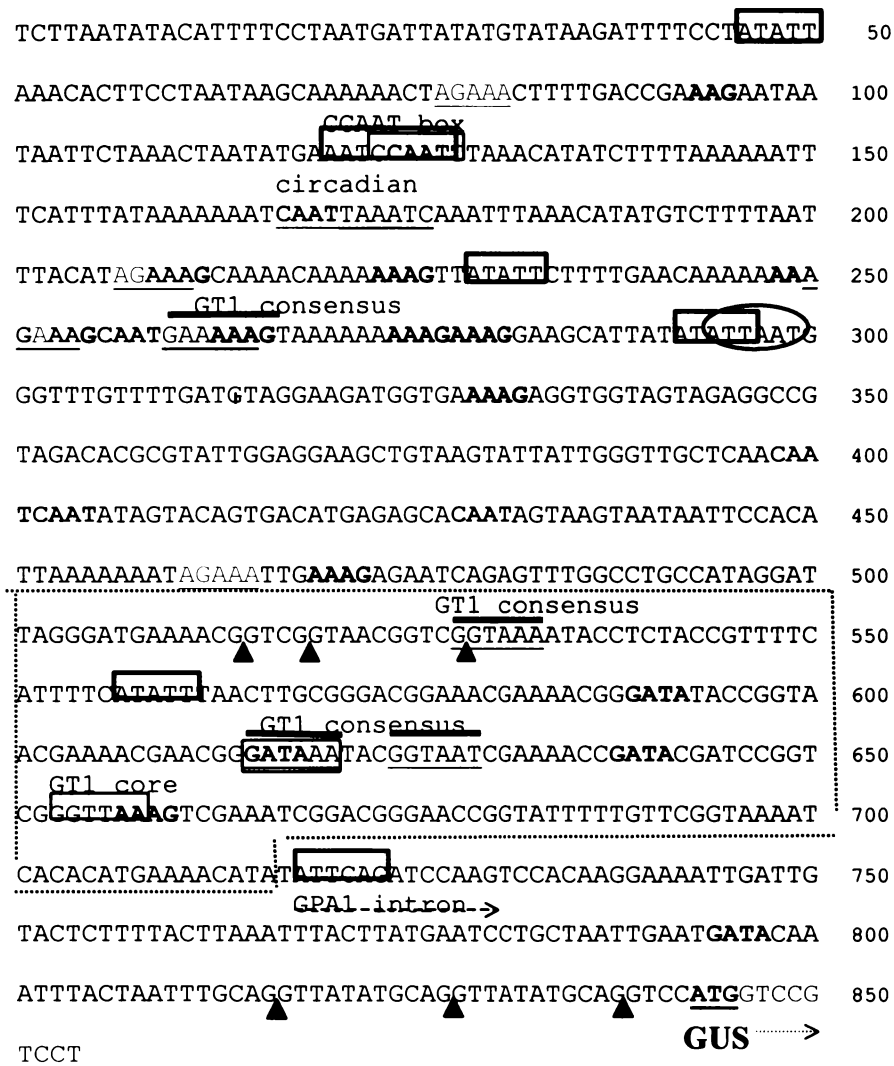
Description	Sequence	Average in 500 bp p(II)	between ET5158	ET5267	ET5359	ET5491	ET5555	ET6417	ET6566	ET6633	within ET6537	ET5653	ET5599	ET5280	ET6428	ET5580	p(a)	p(b)
CAAT boxes																		
CAAT box	CAAT	2.560															0.912	0.296
CCAAT box	CCAAT	0.538															0.899	0.240
TATA boxes																		
TATABOX2	TATAAAT	0.086															0.701	0.676
TATABOX3	TATTAAT	0.086															0.103	0.976
TATABOX4	TATATAA	0.086															0.324	0.897
TATABOX5	TTATTT	0.297															0.280	0.875
TATA box Phaseolus	TTTATATA	0.025															0.295	0.954
combined		0.581															0.002	0.999
Light regulation																		
AT1 box	AATATTTTTATT	1.8E-04															1	0.998
light regulationASF1	TGACG	0.389															0.280	0.870
phytochrome regulation	AAMAATCT	0.031															0.067	0.992
Circadian expression	CAANNNNATC	0.156															0.887	0.379
Circadian expression	AAAATATCT	0.005															0.071	0.998
GATABOX	GATA	2.561															0.981	0.088
G-box	MCACGTGGC	1.8E-03															1	0.975
G-box	ACCACGTGGC	2.2E-04															1	0.997
G-box	CACGTG	0.082															0.302	0.908
Coupling G-box	CGACG	0.282															0.981	0.107
GT1CONSENSUS	GRWAAW	1.485															0.140	0.973
GT1CORE	GGTTAA	0.156															0.887	0.379
GT1-rbcS	KWGTGRWAAWRW	7.6E-03															0.101	0.995
H-box consensus	CCTACCNNNNNNNC	5.0E-03															1	0.933
IBOX	GATAAG	0.156															0.040	0.990
IBOXCORE	GATAA	0.743															0.837	0.326
rbcS light regulation	AAATTAACCAA	3.2E-04															1	0.996
CAB regulation	ATACGTGT	9.5E-03															0.125	0.992
T-BOX	ACTTTG	0.156															0.132	0.960
BOXII-light regulation	ACGTGGC	0.017															0.214	0.977
LRE	ACGTGGCA	5.0E-03															1	0.933
LRE	AACCTAACCT	5.8E-04															1	0.992
LRE	TCCACGTGGC	2.2E-04															1	0.997
LRE	ACTATTTTCACTAT	1.2E-06															1	0.999
rbcS consensus	AATCCAA	0.0452															0.125	0.978
Phytochome regulation	AACCAA	0.1560															0.621	0.671
Phytochome regulation	CGGATA	0.1129															0.450	0.818
Light repression																		
ATHB2-binding site	TAATMATTA	0.013															0.161	0.987
Box B- repression	AAACGACACCGTTT	2.1E-06															1	0.999
Box C- repression	CTCCCAC	0.0172															1	0.786
Box C- repression	TCCCGGTACACAC TTCTT	5.8E-09															1	0.999
DE1- light repression	GGATTTTACAGT	4.9E-05															1	0.999
GT2-Light repression	GCGGTAATT	2.0E-03															1	0.972
Light repression	GAAATAGCAAATG TTAAAAATA	2.0E-10															1	0.999
Light repression	CAGCTCCCATGGC TCTCCCATCCGCG CCGGT	8.9E-18															1	
RE1-light repression	CATGGGCGCGG	3.3E-05															1	0.999
Miscelaneous																		
Core DNA binding	AAAG	2.561															0.325	1
Pollen specific	AGAAA	0.743															0.123	0.957
Root motif	ATATT	1.026															0.396	0.799

light regulated genes such as rbcS and bean CHS. However, the likelihood of having the GT1 consensus present in 14 or more out of 15 GT lines is not statistically significant. The GATA and IBOX core were also not enriched among the lines evaluated. The full IBOX was present in five of 14 ET lines (p(a)=0.04). However, it should be noted that this statistical analysis did not take into account the fact that multiple comparisons were made. Therefore, the elevated frequency of the full IBOX is strictly speaking not significantly different from what is expected by chance alone.

Motifs associated with light repression include Box B, Box C, DE1, RE1 and other motifs involved in light repression (Figures 25 and 26). The probabilities of finding these motifs are very low since these are usually long motifs with a relatively high GC content. None of those motifs were found in any of the GT lines analyzed. However, one motif involved in light repression, the binding site for the Arabidopsis homeobox gene ATHB2 (Ohgishi et al., 2001) was found in line ET5267. This line has an insertion between two genes in chromosome 5.

Among the motifs not suspected to be involved in light regulation (Miscellaneous) none were enriched. Even the AAAG motif, present in all flanking sequences is in fact expected to be found in most if not all lines by chance alone. However, for this motif the copy number per upstream sequence seems to be elevated (70 copies in 15 lines compared to an expected 38 copies). This motif is a core site required for binding of Dof proteins (DNA binding proteins unique to plants). This motif enhances transcription from the promoters of both cytosolic orthophosphate kinase and a non-photosynthetic phosphoenol pyruvate gene (Yanagisawa, 2000).

A graphical representation of the promoter flanking sequence of line GT6228 is indicated in Figure 27. The sequence consists of two parts (i) 500 bp of flanking sequence and (ii) 350 bp of sequence originating from the Ds element and containing intron, splice donor and splice acceptor sites. It is evident that the Ds sequence itself contains multiple copies of potentially regulatory motifs, such as four GATA and three GT1 consensus motifs. However, no TATA boxes or CAAT boxes are present in the Ds portion. Figure 27 also highlights the position of sequence motifs in the chromosomal flanking sequence, including one TATA box motif (circled) as a plausible controller of a transcription start site (arrow). There are also numerous (C)CAAT motifs as well as motifs not known to be associated with light regulation. However, with the exception of one GT1 consensus and a putative 'circadian element', none of the motifs in the flanking sequence portion belong to the group of light regulatory motifs indicated in Figures 25 and 26. Although the Ds element alone is thought to be transcriptionally silent (Solis et al., 1999), it may contribute to the expression, including the light regulatory characteristics of the GUS reporter in the various GT and ET lines.



▲ Splice donors (D1, D2, D3)

▲ Splice acceptors (A1, A2, A3)

Figure 27. Upstream sequence of the insertion point of line GT6281. The sequence boxed with a discontinuous line indicates the beginning of the Ds. The GUS sequence is indicated in purple font.

5.6.1. Analysis of common regulatory motifs

Under the assumption that co-regulated genes share cis-acting regulatory elements, it is important to investigate the upstream sequences controlling the transcription of these genes. In order to evaluate for the presence of common regulatory motifs upstream of the genes that clustered together in the tree (Figure 15), the upstream sequences (500 bp) of the genes corresponding to the blue and green cluster (Figure 15) were analyzed using an extension of the Gibbs sampling (Lawrence et al., 1993) and the expectation maximization algorithms (Bailey and Elkan, 1995) for motif finding with a higher-order background model. These analyses use a selected data set of intergenic sequences from *A. thaliana* to construct a reliable higher-order background model of gene upstream regions of this plant (Thijs et al., 2001). This background model includes a high-order Markov model of non co-regulated genes that improves the discrimination between motifs that are over-represented by chance and motifs that are biologically functional. The Markov models are a class of probabilistic models well suited for representing profiles of multiple sequence alignments (Krogh et al., 1994).

The algorithm works as follows: first, it creates an initial alignment vector that describes the putative start position of the motif over the different sequences. Then, it initializes the probabilities of observing the motif in each sequence. Then, it samples and selects one sequence and calculates the motif model based on the current set of positions. A weight is assigned to each segment by calculating the probability of the motif being at a given position, the motif model and the background model. A new alignment vector is generated according to the calculated weights and the process is repeated several times until a stable motif is found. A motif is called stable if the probability matrix does not change significantly between two iterations.

The sequences (in FASTA format) of the blue and green clusters were imported into Motif Sampler (http://www.esat.kuleuven.ac.be/~thijs/Work/MotifSampler.html). Note that sequences were only available from a subset of the clustered lines. The shared motifs between these sequences were tabulated (Tables 20, 21, 22, 23, column 1). Only motifs present in all the sequences evaluated for each cluster were included. Completely conserved motifs and motifs that have one base difference were reported. When there was only one base difference between the motifs in the sequences evaluated, only those motifs that changed from one base into a single other base were included. Motifs in which the base-change corresponded to more than one different base were not included. For example, a motif nCGAG (n=A, T, C or G) was not included, but a motif such as GCkCC (k= G/T) was included.

Table 20. Common regulatory motifs found in the blue cluster (Figure 15, 95% significance, lines ET5599, GT5909, GT6021 and GT6647). k = G/T, m= C/A, r = G/A, y= C/T, w= A/T, s= C/G

Shared motif	Similarity to other motifs in the database		
	Motif	**Function**	**Sequence**
CCACG	G-box	Light responsiveness *Arabidopsis thaliana*	CTT**CCACG**TGGCA
AGAGAGT GAGmG GAGGw mGAGAG AGAGAGw GmGAG	GAG-motif	Light responsive element (LRE) *A. thaliana*	**AGAGAGT**
GCkCC	RE-1	Repressing element *Avena sativa*	GGGCGCGGAACAAGGATCG GC**GCGCC**ACGCC
GyAAC	Chs-CMA2b	Part of light responsive element *Daucus carota*	ATT**GCAAC**TCAA
kAGGT	GT1-motif	LRE *A. sativa*	ACAGAAGTAGGTTAATCAA TTT**CAGGT**TAATCAAAGTG GAGATCGGAAAG

Table 21. Common regulatory motifs found in the blue cluster (Figure 15, 75% significance, lines ET5280, GT5927, GT6647 and OK0319). k = G/T, m= C/A, r = G/A, y= C/T, w= A/T, s= C/G

Shared Motif	Similarity to other motifs in the database		
	Motif	**Function**	**Sequence**
ACGAG	C-box	Light responsiveness *Hordeum vulgare*	**ACGAG**CACCGCC
ATAAC AATAm	I-box	Part of a LRE *Pisum sativum*	TAG**ATAAC**C
	Pc-CMA2b	Part of a LRE *P. sativum*	AG**ATAAC**CCACTTTA
GGCTT	No information		
GGkGG	TGG-motif	Part of a LRE *Helianthus annuus*	T**GGTGG**CTA
	GT-1 motif	Part of a LRE *Solanum tuberosum*	AT**GGTGG**TTGG
	rbcS-CMA7c	Part of a LRE *Zea mays*	ACGCGACATGT**GGTGG**CGGA
GTrCG	G-box	Light responsiveness *Oryza sativa*	**GTACG**TG
	Chs-CMA2c	Part of a LRE *Petroselinum crispum*	AT**GTACG**TGGAGG
AsCGA	DRE motif	Involved in stress responses *A. thaliana*	T**ACCGA**CAT
CGAGr	No information		
TAwATA	TATA box	*A. thaliana*	**TATATA**A **TATATA** **TATATA**TA TA**TAAATA**TAAA
ATATTs	No information		

Table 22. Common regulatory motifs found in the green cluster (Figure 15, 75% significance, lines ET5555, ET5599 and GT6281). k = G/T, m= C/A, r = G/A, y= C/T, w= A/T, s= C/G

Shared	**Similarity to other motifs in the database**		
Motif	**Motif**	**Function**	**Sequence**
CTCAA	CAAT box	*Daucus carota*	AG**CTCAA**TTTCA
	Chs-CMA2b	Part of a LRE - *D.carota*	ATTGCAA**CTCAA**
GTTAT	No information		
ATTTA	TATA box	*A.thaliana*	T**ATTTA**AA
ATTTwA			T**ATTTAA**A
wTTTA	DREP	Part of a LRE	**ATTTA**TCCTCCAAAAATCA
	module	*Spinacia oleracea*	
AAATCm	DREP	Part of a LRE,	ATTTATCCTCCAA**AAATCA**
	module	*S. oleracea*	
AAGTA	GT1-motif	LRE – *A. sativa*	ACAG**AAGTA**GGTTAATCAA
			TTTCAGGTTAATCAAAGTG
			GAGATCGGAAAG
TAGTA	AT rich	Chalcone synthase gene *P.*	TAAAATACTATCCATTCGTT
	sequence	*sativum*	AA**TAGTA**AAATACT
TGTAA	Prolamin box	Seed storage protein gene	tgcca**TGTAA**AGatgac
TGTAw		promoter *O. sativa*	
AAGTAA	Prolamin box	Seed storage	tgtagTGTA**AAGtaa**aa
wAGTAA		*H. vulgare, Zea mays*	tgacgTGTA**AAGtaa**at
ATAATA	Box II	Part of a LRE *A. thaliana*	AC**ATAATA**GCCACATATT
	Pc-CMA2b	Part of a LRE *S. oleracea*	AG**ATAATA**CCCTTTA
ACACr	Box II	Part of a LRE	**ACACG**TAGA
	G-box	Light responsiveness	**ACACG**TGT
ATGTCTT	No information		

Table 22 continued. Common regulatory motifs found in the green cluster (Figure 15, 75% significance, lines ET5555, ET5599 and GT6281). k = G/T, m= C/A, r = G/A, y= C/T, w= A/T, s= C/G

Shared Motif	Similarity to other motifs in the database		
	Motif	**Function**	**Sequence**
ACwTT	OBP-1 site	Cis-acting regulatory Element, *A. thaliana.*	TAC**ACTTT**TGG
	PcCMA2b	Part of a LRE *P. Sativum*	AGATAACCC**ACTTTA**
ACAsT		*P.sativum* element involved in phytochrome down-regulation expression	GGATTTT**ACAGT**
TTwGAT	No information		
wAGTAT	Endosperm box	Endosperm regulation *Z. mays*	GGAAGGT**TAGTAT**GATGAC ATG
TAATwC	Pc-CMA2b	Part of a LRE *S. oleracea*	AGA**TAATAC**CCTTTA
ATwAGAT	I-box	Part of a LRE *A. thaliana*	G**ATAAGAT**T
TAwATAT	TATA box	*A. thaliana*	**TATATAT**A TA**TAAATAT**AAA tc**TATATAt**t
AATArTA	AT-rich sequence	Element for maximal elicitor-mediated activation *P. sativum*	TAAAATACTATCCATTCGTT **AATAGTA**AAATACT

Table 23. Shared regulatory motifs found in the pink cluster (Figure 15, LH21116, and CAB). k = G/T, m= C/A, r = G/A, y= C/T, w= A/T, s= C/G

Shared Motif	Similarity to other motifs in the database		
	Motif	**Function**	**Sequence**
CTACT	H-Box	Part of a LRE. *Glycine max*	ACCTACC**CTACT**TCCTA
ACTTA AsTTA ACCAk	H-box	Part of a LRE. *Daucus carota*	AGCTACCAT**ACTTA**TTA AGCT**ACCAT**ACTTATTA
TGTTA	Endosperm box	*Zea Mays*	TTGGA**TGTTA**GTGGGATGA CATG
CTTAG	JERE	Jasmonate elicitor responsive element	CT**CTTAG**ACCGCCTTCTTTG AAAG
GTTTT	HSE	Heat shock element *Brassica oleracea*	GAAGAAAAT**GTTTT**AAAAA CT
TTTAC		Involved in phytochrome down-regulation	GGAT**TTTAC**AGT
CTTTAC	DREP module	Part of a LRE	A**CTTTAC**CTCCAAAATTCA
CTTwT	I-box	Part of a LRE. *A.thaliana*	C**CTTAT**CCT
GCTACT	No information		
TAACTT	No information		
TTTGGGT	No information		
AAGTATTA	No information		
TCGyT	AT-rich sequence	*Pisum sativum*	TAAAATACTATCCAT**TCGTT** AATAGTAAAATAC
TwCTC	Chs-CMA2b	Part of a LRE. *Petroselinum crispum*	GTATC**TACTC**AC
TAyTC	Z-motif	Element conferring high activity in dark-grown *A. thaliana*	ATC**TATTC**GTATACGTGTCA C
	Chs-CMA2b	Part of LRE. *Petroselinum crispum*	GTATC**TACTC**AC

Table 23 continued. Shared regulatory motifs found in the pink cluster (Figure 15, LH21116, and CAB). k = G/T, m= C/A, r = G/A, y= C/T, w= A/T, s= C/G

Shared Motif	Similarity to other motifs in the database		
	Motif	**Function**	**Sequence**
AGCTr	ATC motif	Part of a conserved DNA module involved in light responsiveness	**AGCTA**TCCA
	H-Box	Part of a LRE	**AGCTA**CCATACTTATTA
TTAGy	GATT	Part of a LRE. *Zea mays*	CTCCTGA**TTAGC**
AyTAC	ACA	Part of gapA in (gapA-CMA1) involve in light responsiveness	A**ATTAC**AGCCATT
ACCAk			AATCACA**ACCAT**A
	H-box	Part of a LRE *D. carota*	AGCT**ACCAT**ACTTATTA
CTyTA	Sbp-CMS1	Part of a LRE *A. thaliana*	CTTTATCTCTTCCA
	LAMP element	Part of a LRE *Pisum sativum*	CTTTATCA
CAsTG	rbcS-CMA7c	Part of a LRE. *Lemna gibba*	ACG**CAGTG**TGTGGAGGAGC A
yTACTT	Chs-CMA2a	Part of a LRE *Daucus carota*	**TTACTT**AA
	H-box	Part of a LRE *Glycine max*	ACCTACC**CTACTT**CCTA
CwGCTA	No information		
TTTkTC	5UTR Py-rich	Cis-acting element conferring high transcription levels	TTTCTCTCCTC**TTTTTC**CTC

The following were the lines included in the analysis:
- Blue cluster 95%, stronger response to dark, no phyB response (Table 20): ET5599, GT5909, GT6021 and GT6647 (lines ET5529, ET6649, GT6604 and OK0813 were not included because no sequence was available). Blue cluster 75% (Table 21): ET5280, GT5927, GT6647 and OK0319 (lines ET5529, GT6075, GT6604 and OK0813 were not included because no sequence was available).
- Green cluster 75%, phyA/phyB redundancy (Table 22): ET5555, ET5599 and GT6281 (lines ET6649 and GT6534 were not included because no sequence was available). The green cluster at 95% was not included because the sequence of GT6534 was not available).
- Pink cluster 95% and 75%, stronger response to light than dark (Table 23): CAB and LH21116.

The motifs found with the 'Motif sampler' program were used to search the PlantCARE database (http://sphinx.rug.ac.be:8080/PlantCARE/cgi/index.html): Plant Cis-acting Regulatory Elements. The purpose of this analysis was to determine the presence of similar motifs in plants and their putative function (Tables 20 to 23, columns 2, 3 and 4). The motifs found appear to be elements of well-characterized promoter motifs of which many are involved in light regulation. The shared regulatory motifs were distributed throughout the 500-bp upstream sequence analyzed as shown by the example in Figure 28. Several motifs involved in light regulation were observed along with some motifs for which no information was found either in the PlantCARE database, the PLACE database or the literature (for example, GGCTT and CGAGr, Table 21). These unknown motifs may correspond to cryptic promoters no yet described in the literature. Other types of motifs found included motifs found in genes involved in stress responses (Tables 21 and 23), seed storage protein production and endosperm regulation (Table 22).

Are there any shared regulatory motifs between lines with higher expression in light than dark and those lines with higher expression in dark than light? In order to answer this question, the upstream sequences of the blue, green and pink cluster (Figure 15) at the 75% significance level were analyzed with the 'Motif sampler program. This analysis included lines ET5280, GT5927, GT6647, Ok0319, ET5555, ET5599, GT6281, CAB and LH21116. The focus was on the clusters belonging to the 75% significance because as mentioned before, only one sequence was available for the 95% cluster (phyA/phyB redundancy, green cluster). Only two shared motifs were found (Table 24). One motif, G(t/g)CTT has no similarities to any other motifs found in the database. The other motif was similar to a motif involved in light responses. The three clusters compared have in common a different response in light than dark, but they belong to clusters with different light regulation. The analysis of the upstream sequences of the gene trap and enhancer

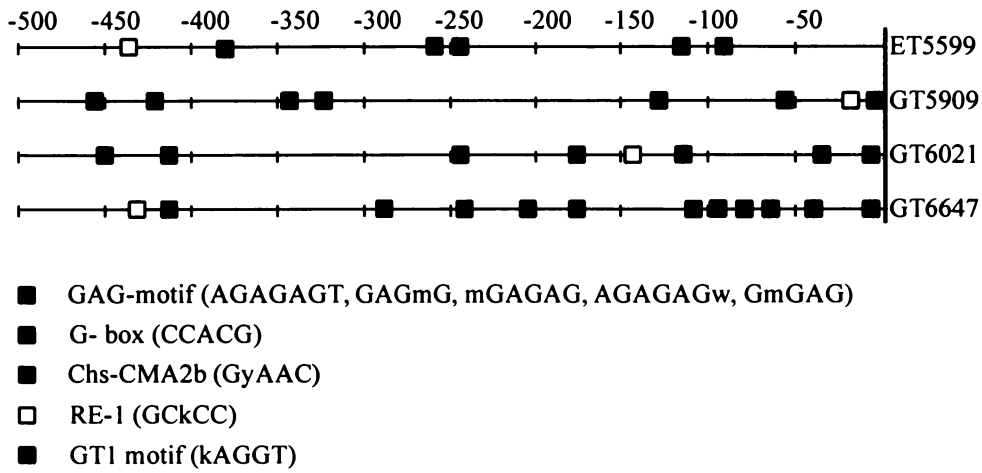


Figure 28. Example of the distribution of the common regulatory motifs found in the blue cluster (Figure 15, 95% significance).

Table 24. Shared regulatory motifs between the blue, green and pink clusters (Figure 15). k = G/T, m= C/A, r = G/A, y= C/T, w= A/T, s= C/G

Shared	**Similarity to other motifs in the database**		
Motif	**Motif**	**Function**	**Sequence**
GkCTT	No information		
GTrTT	ACGT containing element	Cis acting element involved in light responses. *Petroselinum crispum*	CTAAC**GTATT** GACAC**GTGTT**CGATGACGT GGT

5.7 Discussion

The flanking sequences of a subset of Ds:GUS elements were analyzed in order to address a variety of questions. As stated in the introduction, one can search for common elements in the promoters of trapped genes that share common expression characteristic, or one can ask whether the expression profile of a promoter trap line can be understood on the basis of known promoter sequence motifs.

Analysis of flanking sequences of 15 Ds:GUS elements by TAIL PCR established that insertions occurred throughout the genome. The TAIL-PCR reactions usually did not give rise to multiple products. In one case, line GT6407 had 2 bands but after sequencing both bands and doing the BLAST search, the insertion was the same. These data are consistent with the notion that each GT/ET line contains only a single insertion. Given the unambiguous results of the sequence similarity searches, none of the insertions were in repetitive DNA or other transposon-like sequences. There appeared to be a bias towards insertions into the ends of chromosomes rather than the central parts. Chromosome I displayed the highest number of insertions. The reasons for these biases are not clear.

5.7.1. Possible cryptic promoters

Insertions were approximately equally distributed between intergenic regions and 'genes', i.e. putatively transcribed regions. Within the latter group, the insertions were approximately equally likely to reside in the sense or antisense orientation with respect to the GUS ORF and the tagged gene. The approximately random positioning of the Ds:GUS elements with respect to transcribed regions was surprising, given the pre-selection among all insertions for those with moderate-to-strong and light-regulated GUS expression. This finding suggests either that activation of GUS expression is often driven by cryptic promoters; or that the annotation of the genomic transcription units in the Arabidopsis genome database has frequent errors. Both explanations may be true. To address the latter possibility, the annotations from NCBI, TAIR and MIPS databases were compared. In most cases, the annotations were nearly identical. However, in two cases (GT6338 and GT5914) there were differences. For these two lines, the annotation of NCBI and TAIR suggested an insertion within a gene whereas in the MIPS database the insertion was between genes. These two sequences were analyzed by the annotation programs described in materials and methods and the results were compared with the NCBI and TAIR annotations. The NCBI annotation seemed to fit better the predictions of the genes by the various programs so this was the final annotation used for these two genes.

Therefore, the gene expression displayed in many of the GT lines with insertions in non-coding regions or into genes in the antisense orientation could be due to the presence of cryptic promoters. Cryptic promoters are regulatory elements that are inactive at their native locations in the genome but may become functional when positioned adjacent to genes (Wu et al., 2001) and they have been analyzed in different species (Foster et al., 1999, Foster et al., 1999, Irniger et al., 1992). Cryptic promoters or cryptic enhancers may also play a role in the expression of GT insertions in the sense-orientation and in ET lines.

One might argue that the recovery of cryptic promoters in a promoter trapping project compromises the goals of the work. I would argue that, if anything, the opposite is true. First, the goal of this project was not to find endogenous genes that are regulated with a predicted pattern or profile. Instead, the goal was to insert probes into the signal transduction network that report about the entire range of light expression profiles encoded in the light signaling machinery. In other words, if a certain expression profile is artificial in the sense that it can be detected only with a promoter trap but not by surveying the endogenous transcriptome, the decision to use promoter trapping was a good one.

5.7.2. Statistical analysis of promoter motifs

In order to interpret the number of sequence motifs within the GT flanking sequence, statistical analysis was conducted (see methods). One would expect that most promoters, cryptic or not, contain one or more of the most basic transcriptional activation motifs, the TATA box and a CCAAT motif. In fact, the frequency of the CCAAT motif within the first 500 bp of the insertion site was higher than expected by chance among the GT lines. The p(a) value for the observed frequency equals 0.013 and is therefore significant at the standard 0.05 level. However, note that the CAAT motif was not unusually frequent. Considering the TATA box, 10 out of 15 GT lines and 12 out of 14 ET lines have at least one. The frequency of the TATA box motifs was higher than random, although the P(a) value narrowly escaped the definition of significant (p=0.067) in the GT lines. This finding is consistent with the idea that the Ds:GUS elements derive their basal expression from the fortuitous presence of basal promoter motifs. Interestingly, there were no fewer TATA boxes in the flanking sequences of the ET lines where 12 out of 14 lines had one or more TATA boxes. The ET lines do not need another TATA box for expression because in their construction a minimal promoter was included.

In contrast, most other motifs appeared at a frequency near that expected by chance. Motifs implicated in light repression were also rare, which is surprising given that most of the promoter traps analyzed are light repressed. However, we must take into account

that all light repression motifs are rather long, making a perfect match highly unlikely. If the consensus sequences for light repression were defined better, matches among the promoter trap lines might be observed.

5.7.3. Correlations between expression profiles and flanking sequence motifs

The second question concerns the high rate of light-dependent or light-sensitive expression patterns among the promoter trap lines analyzed. It is evident that the Ds:GUS sequence itself contains a substantial number of motifs that have been implicated in light regulation, for example the GATA and related motifs and the GT1 motif. Although GATA and GT-boxes are usually found upstream of a TATA box in native promoters, rather than downstream as is the case here, it remains possible that they contribute to the expression characteristics displayed by the GT/ET lines. However, it is important to realize that additional, line-specific, promoter motifs from the flanking sequence must contribute to the expression profiles observed, because the profiles are all largely different.

The question of which flanking sequence motifs (if any) may correlate with which feature of the light response profile is difficult to address because the lines analyzed for their flanking sequences fall into diverse expression profiles that bear little resemblance apart from the fact that most are induced in darkness. A comprehensive PCR amplification of all flanking sequences was beyond the scope of this work. However, an ongoing effort at CSHL to determine the flanking sequences of all GT/ET lines generated there will eventually address the question of whether lines with similar expression characteristics share certain features in their flanking sequences.

One caveat in this discussion is the implied assumption that the determinants of the Ds:GUS expression patterns reside within the first 500 bp of flanking sequence plus the Ds 5' sequence). Indeed, most native Arabidopsis promoters are compact, such that the expression characteristics of the promoter are often contained within 1000 bp of upstream sequence and many basic features of the expression profile are encoded within 250 bp of upstream sequence (for example, CAB:GUS). However, whether the cryptic promoters emerging from this work are equally compact remains to be shown. It cannot be excluded that the GUS expression of the Ds:GUS elements may be sensitive to much more long-range signals, as they operate in Drosophila and mammals, and that endogenous Arabidopsis genes, but not promoter traps, may be insulated against such effects by yet unknown DNA sequence elements.

Even though the genes analyzed for their promoter sequences were light repressed, only one motif involved in light repression was found (in line ET5267). This suggests that

there exists a yet to be discovered group of motifs that confer induction by darkness or repression by light. The promoter:GUS fusions created for this work may provide a starting point to identify such motifs. For example, it would be interesting to determine whether the potential LREs found in many of the GT promoters are also playing a role in dark induction. If so, it would suggest that there exists a separate 'polarity' switch within the promoter that determines whether an LRE confers light induction or dark induction.

In summary, the isolation of promoter trap flanking sequences has made it very likely that promoter traps may plug into the light signaling machinery in novel ways by revealing cryptic promoters, which may function in conjunction with sequence elements contributed by the Ds:GUS element. Therefore, cryptic promoters may contribute to dissect the light-signaling network, due to their complex expression profiles.

Chapter 6. Microarray analysis

6.1. Introduction

The use of microarrays is based on two principles: first, only nucleic acid strands with complementary sequences can hybridize in a stable molecule; second, an alteration in the abundance of mRNA is the predominant factor underlying changes in gene expression. What distinguishes microarray based mRNA expression profiling from other hybridization bases methods is the scale. In a microarray, typically thousands of distinct DNA samples are "arrayed" on a glass microscope slide in a gridded pattern. cDNAs or oligonucleotides are most commonly used. The complex mixture of mRNA molecules to be analyzed is labeled, usually by conversion into a fluorescently labeled cDNA copy. The abundance of individual mRNA species in the mixture is given by the amount of hybridization signal at each spot of the array.

Microarray analysis is very appropriate for making pair-wise comparisons of, for example, two mRNA samples from two different treatments (Schena et al., 1995). Relative transcript abundance is measured by labeling the two samples with different fluorescent dyes (Cy3 and Cy5) and hybridizing them simultaneously (Cummings and Relman, 2000). The unbound probe is removed and the fluorescent probe that has hybridized to the DNA fragments on the microarray is excited by light. The fluorescence signal from each spot on the array is a reflection of the abundance of the corresponding sequence in the original probe (Kehoe et al., 1999). The fluorescence intensities for each spot on the array are determined with a confocal scanner and a software application is used to capture the image of the array and extract the numerical data for each spot. The data are processed and a table is generated with the fluorescence of each of the two dyes (Brown and Botstein, 1999, Duggan et al., 1999, Shalon et al., 1996). In order to visualize the results, a color code is used in which activated genes (higher hybridization with mRNA sample 1 than sample 2) are colored red and the repressed genes are colored green.

Although transcriptional profiling with microarrays provides an extraordinary amount of information about differentially expressed genes, there are still some problems to be resolved. First, cross-hybridization among closely related gene family members may occur (Kehoe et al., 1999), potentially obscuring differences in their regulation. The potential for cross-hybridization is a clear disadvantage of cDNA microarrays compared to promoter trap reporter gene fusions. Second, the quality of RNA plays an important role in the quality of the data since variations in specimen handling prior to RNA extraction can have significant effects on altering the expression patterns obtained with the microarray (Hoheisel and Vingron, 2000). Even though the software programs available for microarray analysis use the same algorithms, each one of them exports

different signal quality measurements which complicates the comparison between data obtained with different application programs (Bassett et al., 1999). In microarray analysis, the decision as to what constitutes a significant difference in gene expression between test and reference populations is a difficult one. Statistical analysis of variance models have been used to establish the minimum thresholds for significant differences of expression. The problem with this is the assumption that each gene has the same statistical variability which disregards the likelihood that rare messages will indeed show greater variability than highly abundant transcripts (Harmer and Kay, 2000). Multiple independent replicates of the same experiment are important for an accurate analysis of the microarray data.

6.2. Rationale

Compared to gene expression profiling of promoter traps, microarrays offer the assurance that genuine Arabidopsis transcripts are being measured, rather than synthetic transcripts coming from cryptic promoters or from poorly characterized hypothetical genes. Meanwhile, microarrays are not a routine technology at this time and data from the cDNA microarrays currently available for Arabidopsis will be confounded by the problem of cross-hybridization between closely related gene family members. However, given the theoretical power of microarrays, it was decided to take a first step towards building light regulated mRNA expression profiles with this technique. Microarrays carrying approximately 5,000 to 6,000 individual Arabidopsis cDNAs were made available by Monsanto Corporation (Ruan et al., 1998). RNAs were isolated in order to identify genes regulated in response to shifts from dark to light (cD-DL experiment) or from light to dark (cD-DL experiment). As expected, several of the light induced genes were involved in photosynthesis related processes. Unexpectedly, however, many of the dark induced genes were involved in stress responses.

6.3. Materials and methods

6.3.1. RNA preparation

Plant tissue was frozen and ground in liquid nitrogen. The homogenized tissue powder was transferred to a centrifuge tube and 1 ml of Trizol reagent was added for every 50-100 mg of tissue. The mixture was shaken vigorously and incubated for 5 min at 20-22°C and then centrifuged at 10,000 g for 15 min at 4°C. The supernatant was transferred to a fresh tube and 0.2 ml of chloroform per 1 ml of Trizol reagent were added. The tube was vortexed vigorously and then incubated for 5 min. The aqueous phase was transferred to a fresh tube and the RNA was precipitated with 0.5 volumes of isopropyl alcohol per 1 ml Trizol reagent used in the initial homogenization. The sample was centrifuged at 10,000 g for 10 min at 4°C and the RNA pellet was air dried and then resuspended in 1 ml

DEPC-treated water. Samples were shipped on dry ice to Monsanto where the microarray experiments were carried out. The dark-shift experiment (cL-LD) was repeated twice under 'dye-swap' conditions. This means that in experiment 1, cD was labeled with Cy3 dye and DL with Cy5, whereas in experiment 2 the dyes were reversed. The light shift experiment was done only once.

6.4. Results

Microarray experiments were conducted with Landsberg (Ler) Arabidopsis plants in order to compare the expression of up to 5,000 genes under different light conditions. RNA samples were prepared from seedlings grown under either one of the following conditions: 5 days of constant light (cL) or dark (cD), 4 days of light and a shift to darkness for 24 hours (LD), or 4 days in dark and transfer to light for 12 hours (DL). These treatments correspond to the treatments used with the transposant plants in the light-dark shift experiment previously described (Chapter 1).

6.4.1. Hybridization signals of the light shift and dark shift

Per array, the raw hybridization data for each gene consisted of two fluorescence intensity measurements representing the hybridization signal of the red (Cy5: cyanine5) and green (Cy3: cyanine3) channels from the labeled mRNA samples. The raw hybridization signals were corrected for differences in the incorporation of fluorescent dyes during labeling. The balance coefficients were 2.44, 1.1 and 0.80 for experiments cL-LD 1, cL-LD 2 (Figure 29) and cD-DL (figure 30), respectively. An attempt to repeat the cD-DL experiment did not give good results due to the poor incorporation of the Cy5 dye (Monsanto, personal communication).

The individual hybridization signals for all the data points of the two cL-LD experiments (Figure 29) were displayed by plotting the log2 of the intensity of the cL signal (X axis) against the log2 of the intensity of the cL-LD signal (Y axis). Figure 30 shows the corresponding results for the cD-DL experiment. For orientation, a 3 fold difference in expression between the cL and LD is indicated with dashed lines and the regression line by a solid line.

6.4.2. Assessing significance levels in the microarray data

In the first cL-LD experiment 5654 genes were probed whereas in the second experiment 4757 genes were probed in the microarray. 4516 genes had 2 replicates in the cL-LD

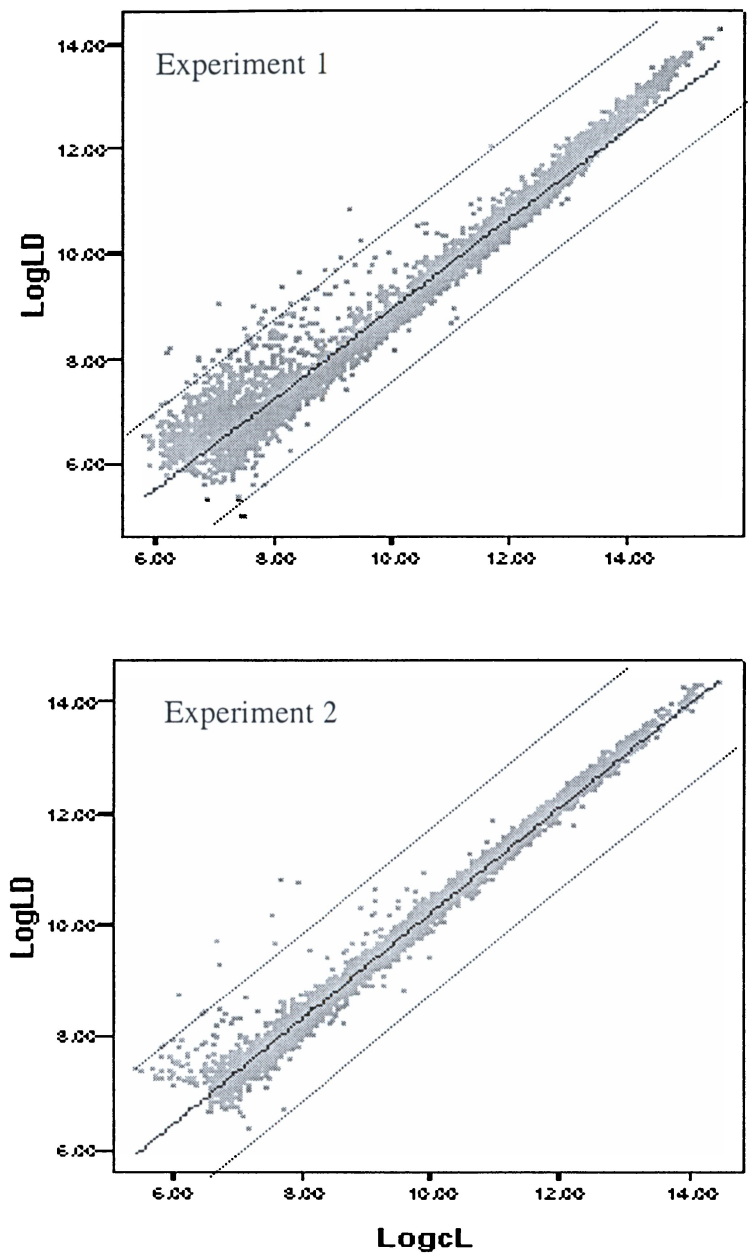


Figure 29. Scatter plots of log2-transformed signals from the two dark-shift microarray experiments. The solid line represents the regression line and the dashed line represents a 3-fold difference in gene expression between the two conditions cL and LD.

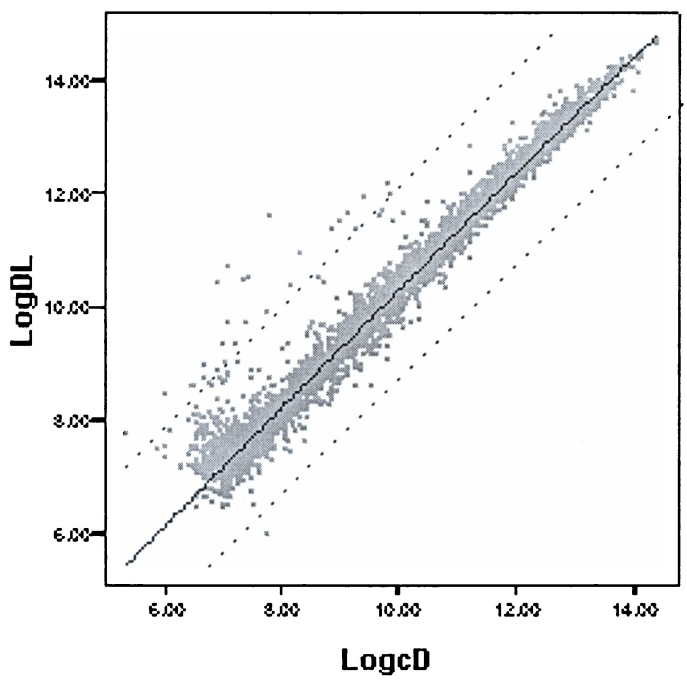


Figure 30. Scatter plot of Log2-transformed signals from the light-shift microarray experiment. The solid line represents the regression line and the dashed line represents a 3-fold difference in gene expression between the two conditions.

experiment. Each of the two experiments had additional datapoints, which were not replicated perhaps due to technical problems such as local high background on the array. 4451 genes out of 4516 genes had less than a two-fold difference, which is typically regarded as insignificant. A total of 349 genes (65 with 2 replicates and 284 with one replicate) displayed a differential gene expression of more than two-fold (Appendix 3A). Surprisingly, all of these genes exhibited an average positive difference, indicating activation of the genes under the shift to dark. For the 65 genes with two replicates, the standard deviation (sd) was calculated. The sd values fluctuated between 0.07 and 2.40. The sd fell in the following categories: 34% of the genes had a sd between 0.07 and 0.50; 32% between 0.51 and 1.00 and 29 % between 1 and 2.00. Only 0.03% of the genes displayed a sd between 2.05 and 2.40. Given the low sd in most of the genes, a total of 274 genes with expression ratios above 2 and below -2 that had only one replicate were added to the pool of genes of 65 genes that had two replicates for a total of 339. It was assumed that the same low sd of the replicated data points might apply to these genes. Under the same assumption, the data from the cD-DL experiment were also examined for evidence of differential expression. Sixty six genes with differential expression greater than 2 and lower than –2 were found (Figure 31). For the cL-LD the most frequent fold induction was between 2.5 and 3 whereas for the cD-DL experiment, the most frequent fold induction was between 2 and 2.5.

6.4.3. Functional analysis of genes responsive to a light shift or dark shift

The genes with differential expression under either the shift to dark (cL-LD) or the shift to light (cD-DL) were classified according to their function (Tables 25 and 26). A more complete list is included in Appendix 3. However, genes without any functional annotation (163 out of 339 for cL-LD and 29 out of 66 for cD-DL) were omitted.

Representative examples of the classification by functional categories are included for both the cL-LD (Table 25) and cD-DL (Table 26). In the shift to dark the most commonly activated genes were those involved in translation closely followed by genes involved in metabolic processes. Genes involved in different stress responses (Tables 27 and 28) including responses to temperature, salt, irradiation, pathogen defense and osmotic stress were upregulated in the shift to dark.

The most commonly induced genes in the shift to light are those involved in metabolism, with a slightly higher percentage for primary metabolism (Table 27) than secondary metabolism. As expected for seedlings undergoing deetiolation in response to light, genes involved in general chloroplast functions and photosynthetic reactions were also activated. In this experiment there were four downregulated genes (Appendix 3), two involved in primary metabolism (genes encoding for aldehyde dehydrogenase and

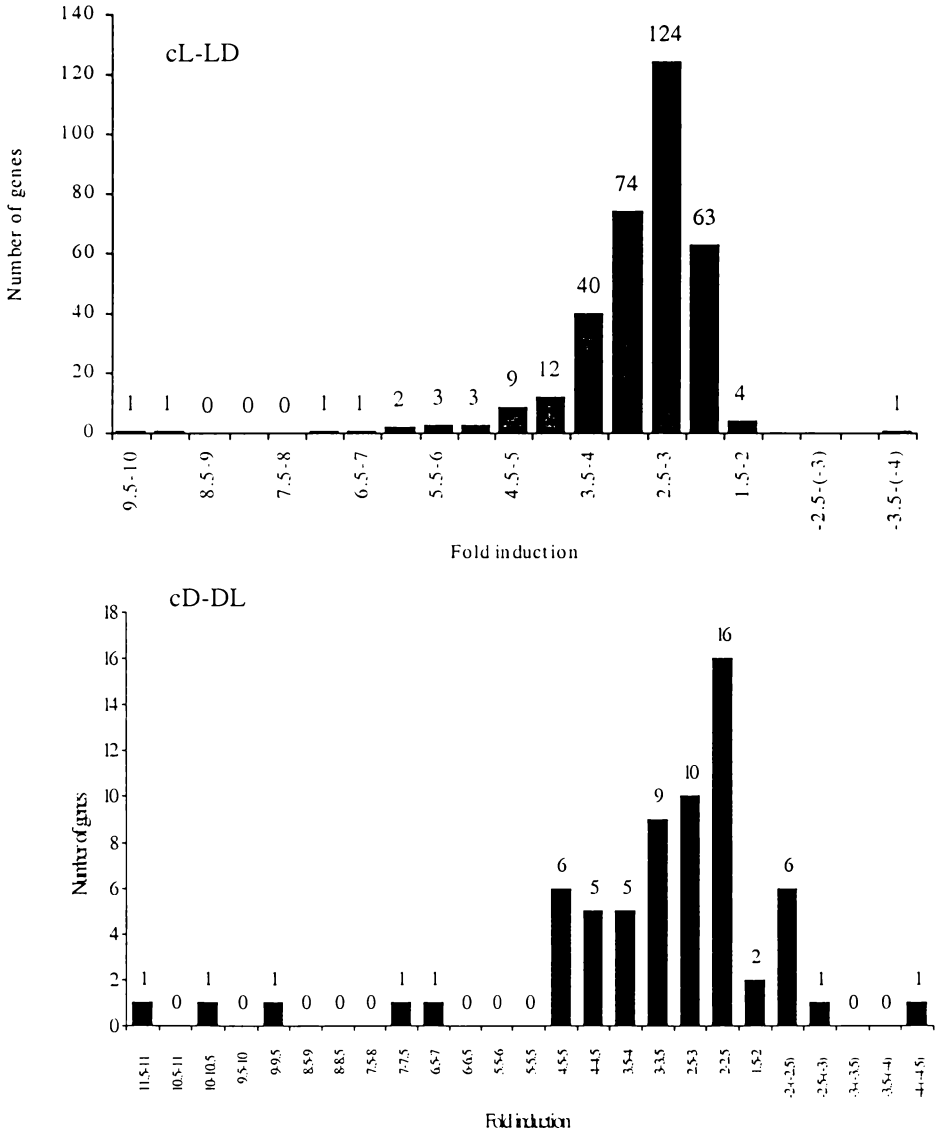


Figure 31. Distribution of the induction in the cL-LD and cD-DL microarray experiment. Only genes with an induction above +2 or below –2-fold are listed.

Table 25. Examples of the classification of the differentially expressed genes in the cL-LD experiment. Only genes that were induced or repressed at least two-fold are included. For the complete list of annotated genes see Appendix 3A.

Gene product	Induction	Standard deviation	Functional Category
NADPH:protochlorophyllide oxidoreductase	4.0		**Chloroplast-**
Chloroplastic outer envelope protein	2.8		**General**
Chloroplast 31KD ribonucleoprotein	2.6		
Phosphoenolpyruvate carboxykinase	9.7		**Chloroplast- dark**
Glyceraldehyde 3-phosphate dehydrogenase	2.3	0.85	**Reactions**
Lhcb6 protein – A. thaliana	3.6		**Chloroplast – light**
Chlorophyll a/b binding protein	2.7		**Reactions**
Cell wall protein	3.7		**Cell division and**
Tubulin alpha chain – A. thaliana	3.5		**Expansion**
Nucleotide sugar epimerases	2.8		
Glucose 6-phosphate/ translocator precursor	2.7		**Metabolism**
Similar to glycosyltransferase	2.7		
Isochorismate synthase	3.0		
Ubiquitin extension protein	2.9		**Protein**
Ubiquitin-specific protease	2.9		**Processing/**
ADP, ATP carrier protein	4.4		
Gamma tonoplast intrinsic protein 2	3.6	0.85	**Transport**
H+ transporting ATPase type 1	3.3		
RNA helicase	3.0		
RNA polymerase II 13.6 kDa chain	2.7		**Transcription**
Similar to MADS box transcription factors	2.0	1.34	
Putative 40S ribosomal protein S25	4.7		
60S ribosomal protein L5	4.6		**Translation**
Putative ribosomal protein L10	3.9		
Translation elongation factor eEF-1	3.5		
Putative serine carboxypeptidase	5.2		**Signal**
Calcium-dependent protein kinase 6	2.7	0.57	**Transduction**
Nucleoside diphosphate kinase type 1	2.4	0.99	
Heat shock cognate protein 70-1	7.5	0.35	
Metallothionein 2a	5.4	0.42	
Drought induced protein Di21	3.9		**Stress response**
Salt stress inducible small GTP binding	2.9		
Stress-induced protein OZI1 precursor	2.8		
Dehydration-induced protein RD22	2.7		

Table 26. Examples of the classification of the differentially expressed genes in the cD-DL experiment. Only genes that were induced or repressed at least two-fold are included. For a complete list of annotated genes see Appendix 3B.

Upregulated genes		
Gene Product	**Induction**	**Functional category**
Tic22-like protein	11.1	
Monogalactosyldiacylglycerol synthase	4.5	
Photosystem I reaction center subunit III	4.4	**Chloroplast general**
Photosystem II oxygen-evolving complex	3.2	
PSI-H subunit – *Brassica rapa*	2.3	
Cytochrome P450	9.2	
Ribulose bisphosphate Carboxylase/oxygenase activase	2.8	**Chloroplast - dark reactions**
Chlorophyll a/b binding protein - like	3.6	**Chloroplast - light**
Lhcb2 protein	2.1	
PSI type III chlorophyll a/b binding protein	3.6	
Succinate CoA ligase	3.6	**Metabolism – primary**
Thiamin biosynthesis protein thi4	3.3	
Aldehyde dehydrogenase like protein	2.5	
Phenylalanine ammonia-lyase	5.0	**Metabolism – secondary**
Putative malonyl-CoA:Acyl transacylase	2.3	
Acyl carrier-like protein	2.1	
Putative proteinase inhibitor II	2.1	**Protein processing-folding**
Selenium-binding protein	2.5	**Transport**
Putative DNA binding protein	3.5	**Transcription**
Homeotic protein Athb-6	2.9	
Plastid ribosomal protein	7.2	**Translation**
Carbonate dehydratase	3.9	**Stress response**
Membrane associated salt-inducible protein	3.0	
Glutathione conjugate transporter	2.1	
Downregulated genes		
Aldehyde dehydrogenase (NAD+)	-2.0	**Metabolism - primary**
Alcohol dehydrogenase-like protein	-4.2	
Ribophorin I-like protein	-2.1	**Transport**
Arginine/serine-rich splicing factor	-2.1	**Transcription**

Table 27. Functional categories for the genes regulated in the shift to dark and the shift to light microarray experiments.

	Functional category	Shift to Dark (cL-LD)		Shift to Light (cD-DL)	
		Number of genes	Percentage	Number of genes	Percentage
1	Chloroplast-general	5	2.8	7	19.4
2	Chloroplast-dark reactions	2	1.1	1	2.7
3	Chloroplast-light reactions	3	1.7	7	19.4
4	Cell division and expansion	13	7.4	0	0
5	Metabolism-primary	25	14.2	6	16.7
	Metabolism-secondary	10	5.7	4	11.1
6	Mitochondria/ Oxidative phosphorylation	3	1.7	0	0
7	Protein processing/folding	12	6.8	1	2.8
8	Transport	19	10.7	3	8.3
9	Transcription	4	2.3	3	8.3
10	Translation	38	21.6	1	2.7
11	Signal transduction	8	4.5	0	0
12	Stress response	25	14.2	3	8.3
13	Other	9	5.1	0	0
	Total	**176**		**36**	

Table 28. Summary of the types of stress responses in which induced stress response genes are involved.

Gene product	Type of stress response
Shift to light	
Cytochrome P450	Defense - insects
Membrane associated salt-inducible protein	Salts
Glutathione conjugate transporter	Defense-chemical toxicity (i.e. herbicide)
Transfer to dark	
Heat shock cognate protein 70-1	Temperature
Metallothionein 2a	Toxicity -Heavy metal
Omega 3-fatty acid desaturase CF3	Temperature – cold response
TMP=A transmembrane protein	Turgor
Germin-like protein	Defense
Drought induced protein Di21	Drought
Putative osr40- *A. thaliana*	Salt
Peroxidase – A. thaliana	Irradiation
Pdr1 A. thaliana	Defense
Nitrilase 1	Defense – pathogens
Monosaccharid transport protein STP4	Defense – pathogen, wounding
Probable glutathione transferase	Defense-pathogens, insects
Transmembrane protein TMP-B	Turgor
Salt stress inducible small GTP binding	Salt stress
Stress-induced protein OZI1 precursor	Defense – ozone, pathogen
Dehydration-induced protein RD22	Drought
Jasmonate inducible protein isolog	Defense
Putative disease resistance protein	Defense - pathogens
AtRab 18- A. thaliana	Drought
Copper homeostasis factor	Toxicity -Heavy metal
Glutathione S-transferase	Defense-chemical toxicity (i.e. herbicide)
Chaperonin-60 beta subunit	Defense - wounding
Metallothionein 2b	Toxicity -Heavy metal

alcohol dehydrogenase-like proteins), one gene involved in transport (encoding for a ribophorin I-like protein) and one gene belonging to the 'other' functional category (encoding for arginine/serine-rich splicing factor-pseudorabies Rsp40). All the other genes were upregulated.

6.5 Discussion

Differentially expressed genes in the shift from cL to cD (cL-LD) and from cD to cL (cD-DL) were found. For reasons that are not entirely clear, the response was in almost every case an induction rather than a repression. Considering just the notorious statistical variation in expression signals in microarray experiments, this finding was surprising and led us to consider technical problems during the array scanning or subsequent steps that were beyond my control. However, given the typical trail of data points showing induction of gene expression, it was worth to continue with the analysis of the results.

For both light shift experiments, the most commonly activated genes coded for proteins involved in metabolism, and among these, primary metabolism outweighed secondary metabolism. This is understandable in light of the metabolic pathways triggered by changes in light conditions. However, apart from the 'metabolism' category, the two shift treatments appeared to affect the functional categories in quite distinct ways (Table 22). As expected, in the shift from constant dark to light a high percentage of the activated genes encoded proteins that play vital roles in the chloroplast, specifically the photosynthetic light reactions. Moreover, genes involved in transport and transcription also became rapidly activated in the shift to light.

Interestingly, the shift to dark elicited the expression of a set of different genes. Genes encoding for proteins involved in cell division or expansion, translation, stress responses and transport were particularly well represented. The dark treatment triggers rapid cell wall elongation, which requires activation of genes for cell wall biosynthesis (Table 20 and Appendix 3 list genes of this type). No activation of genes involved in cell division and expansion was observed after the shift from dark to light. The xyloglucan endotransglycosylase showed a low induction (+1.7) in the cD-DL experiment and therefore it was not reported among the genes thought to have significant activation as described previously. Although cell division is triggered by a shift from darkness to light, the response is confined to the shoot and root apex and is therefore less likely to manifest itself in a total seedling RNA preparation. The striking stimulation of genes encoding for ribosomal proteins and translation factors in the shift to dark suggests that overall levels of translation must increase in order to cope with the dark condition. Alternatively,

specific components of the translation machinery could be used preferentially in the dark rather than in the light.

Also of interest are genes that are known to be light regulated but failed to respond to light or dark shifts. For example, downregulation of the CAB gene in response to dark was not observed. The reason for this is not quite clear and might be related to a technical problem, given that almost no downregulated genes were observed at all. Alternatively, the result might be explained as a circadian phenomenon. CAB gene expression is controlled by the circadian clock and light (Millar et al., 1995, Piechulla, 1999). Rhythmic expression of CAB has been observed for more than 5 days under constant dark and constant light conditions (Sugiyama et al., 2001). Given that CAB genes are clock controlled, it is possible that the response to 24h of darkness was overlaid by the entrained response to the circadian clock, which simply went through one full 24h cycle, i.e. bringing CAB gene expression back close to the same level as that obtained in the '5 day light' control treatment. In an independent microarray experiment, 24h of darkness resulted in only about a two-fold downregulation of CAB (Schaffer et al., 2001). However, the growing conditions in this experiment (12 hours light, 12 hours dark) were different from ours. This might explain the lack of downregulation for CAB and perhaps for other genes analyzed in the microarray.

By selecting 24h as the period for the dark treatment I might have inadvertently corrected for interference by the circadian clock thus focusing more specifically on genes responding to darkness per se. In this regard, it is interesting that 24h of dark treatment caused upregulation of stress-related protein mRNAs. With hindsight, this should not be surprising given that plants in temperate environments never experience 24h of darkness. However, in light of the fact that typical 'dark-adaptation' experiments normally consist of 40 or 48 h or dark treatment, one must suspect that these dark treatments affect the cellular signaling network far beyond the core light signaling pathways.

It is possible that a change to dark triggers in the plant similar processes as those triggered when the days become shorter in the winter, eliciting the expression of genes necessary for the sensed environmental change. This explanation is perhaps too simplistic because the stress related genes stimulated in the shift to dark represented a wide array of processes including responses to cold, dehydration, disease, chemical defense, irradiation, wounding and heavy metal stress (Table 23). An alternative explanation for the upregulation of sets of genes in the experiments described include the connection of these signaling pathways with the light signal transduction pathway as shown with pathogenesis related signaling pathways connected with the light signaling pathways. For example the salicylic acid-mediated signaling pathway crosstalk with the pathway regulated with the phyA, leading to the induction of CAB genes (Schenk et al., 2000). Additionally, high pathogenesis related gene expression in the phyA and phyB signaling

mutant (psi2) also suggested a connection between the light signal transduction and pathogenesis-related gene-signaling pathways (Genoud et al., 1998, Genoud and Métraux, 1999).

In summary, differentially expressed genes were found in the cL-LD and cD-DL shifts. The most common activated genes coded for proteins involved in metabolism in both cases. However, in the cD-DL experiment, genes involved in photosynthetic light reactions, transport and transcription were rapidly activated. Interestingly, in the cL-LD experiment, the most commonly activated genes were those involved in cell division or expansion, stress responses and translation.

Chapter 7. Synopsis

A collection of enhancer and gene trap lines was screened for the GUS expression patterns under different light conditions. Both enhancer and gene trapping methods enable the presence of genes and their patterns of transcriptional regulation to be detected, independently of a mutant phenotype (Meissner et al., 2000, Sundaresan et al., 1995).

Initially, 2098 lines were screened under constant light or constant dark. Among the lines that showed any staining, the most common pattern of gene expression was a tissue-specific pattern of light responsiveness. A total of 365 lines that responded differentially to the cL and cD condition were analyzed for their response to a shift from light to dark (LD) or from dark to light (DL). A set of 286 rapidly responsive genes was identified. In order to characterize the response of the lines to the pathways gated by phyA and phyB, 76 light responsive lines were screened for their response to cR, cFR and pulses of either R or FR. Promoter trapping has the advantage over other approaches of preserving the tissue-specificity of the transcriptional response. The tissue specificity of light responses should one day be incorporated into our models of the light signaling network. The integration between light signaling pathways and pathways that define cell type specific differentiation remain poorly understood.

By measuring the light response profiles of the promoter trap lines under various light conditions it was possible to identify clusters of genes with shared light responses. Each cluster may represent a single output branch of the light-signaling network. Some of the clusters of genes identified showed only a phyA response, other showed phyA/phyB antagonism or phyA/phyB synergism.

Analysis of the genomic flanking sequences established that insertions occurred throughout the genome. Interestingly, there appeared to be a bias towards insertions at the end of the chromosomes. The flanking sequences contained motifs known from common light inducible promoters. Among these, the most frequently observed were the GATABOX, GT1CONSENSUS, and GT1CORE. The GATABOX is required for high level, light regulated and tissue specific expression (Gilmartin et al., 1990; Lam and Chua, 1989). The consensus GT-1 is the binding site of many light regulated genes such as rbcS and bean CHS. Interestingly, the data suggest that the activation of GUS expression may often be driven by cryptic promoters. These are regulatory elements that are inactive at their native locations in the genome but that can become functional when positioned adjacent to genes (Wu et al., 2001). Thus, cryptic promoters may play an important role in the understanding of the light-signaling network.

A preliminary microarray analysis of changes in mRNA expression in response to light and dark shifts resulted in the identification of differentially expressed genes. Interestingly, the shift to dark elicited the expression of genes involved in stress responses. It will be important to perform additional microarray experiments in order to further our understanding of the light-signaling network.

The main purpose and significance of this work was to characterize the types of light response profiles that operate in different tissues of Arabidopsis. Ongoing molecular genetic and specially mutational analyses of the light signaling network may eventually reveal the mechanistic basis for the various light response profiles defined in this work. It may also be worthwhile to compare the GUS expression patterns of the promoter trap lines with patterns of the tagged genes by northern analysis and in situ hybridization. A subset of the tagged genes may itself play a role in mediating light responses; the knockout alleles generated in this work may in the future shed light on the regulatory contributions of the tagged genes to the light signaling network. Understanding the light signaling network will allow us to manipulate crop plants and/or environmental conditions to increase plant productivity.

List of References

Ahmad, M. and Cashmore, A.R. (1993) HY4 gene of *A. thaliana* encodes a protein with characteristics of a blue-light photoreceptor. *Nature* 366: 162-166.

Ahmad, M., Jarillo, J.A., Smirnova, O. and Cashmore, A.R. (1998) The *CRY1* blue light photoreceptor of Arabidopsis interacts with phytochrome A in vitro. *Mol. Cell.* 1(7): 939-948.

Ahmad, M. and Cashmore, A.R. (1996) The pef mutants of *Arabidopsis thaliana* define lesions early in the phytochrome signaling pathway. *Plant J.* 10: 1103-1110.

Altmann, T., Schmidt, R. and Willmitzer, L. (1992) Establishment of a gene tagging system in *Arabidopsis thaliana* based on the maize transposable element. *Theoret. Appl. Genet.* 84, 371-383.

Anderson, S.L. and Kay, S.A. (1995) Functional dissection of circadian clock- and phytochrome-regulated transcription of the Arabidopsis CAB2 gene. *Proc. Natl. Acad. Sci.* 92: 1500-1504.

Anderson, S.L., Teakle, G.R., Martino-Catt, S.J. and Kay, S.A. (1994) Circadian clock and phytochrome-regulated transcription is conferred by a 78 bp cis-acting domain of the Arabidopsis CAB2 promoter. *Plant J.* 6: 457-470.

Ang,L.H. and Deng, X.W. (1994) Regulatory hierarchy of photomorphogenic loci: allele-specific and light-dependent interaction between the HY5 and COP1 loci. *Plant Cell* 6: 613-628.

Ang, L.H., Chattopadhyay, S., Wei, N., Oyama, T., Okada, K., Batschauer, A. and Deng, X.W. (1998) Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of Arabidopsis development. *Mol. Cell* 1(2): 213-222.

Arguello-Astorga, G.R. and Herrera-Estrella, L.R. (1996) Ancestral multiple units in light-responsive plant promoters have structural features correlating with specific phototransduction pathways. *Plant Physiol.* 112: 1151-1166.

Bailey T. L. and Elkan C. (1995) "The value of prior knowledge in discovering motifs with MEME." *Proceedings ISMB* 3: 21-9.

Ballaré, C.L. and Scopel A.L. (1997) Phytochrome signalling in plant canopies: Testing its population-level implications with photoreceptors mutants of Arabidopsis. *Functional Ecology* 11(4): 441-450.

Ballaré, C.L., Scopel, A.L., and Sánchez, R.A. (1990) Far-red radiation reflected from adjacent leaves: an early signal of competition in plant canopies. *Science* 247: 329-332.

Ballaré, C.L., Scopel A.L., Sánchez R.A. and Radosevich S.R. (1992) Photomorphogenic processes in the agricultural environment *Photoch. Photobiol.* 56:777-788.

Barnes, S.A., McGrath, R.B. and Chua, N.H. (1997) Light signal transduction in plants. *Trends Cell Biol.* 7:21-26.

Bassett, D.E., Eisen, M.B., Boguski, M.S. (1999) Gene expression informatics-it's all in your mine. *Nat. Genet.* 21: 51-55.

Batschauer, A. (1998) Photoreceptors of higher plants. *Planta* 206: 479-492.

Batschauer, A., Gilmartin, P.M., Nagy, F. and Schäfer, E. (1994) The molecular biology of photoregulated genes. In: R.E. Kendrick and G.H.M. Kronenberg (eds), Photomorphogenesis in plants. 2nd Edition. Kluwer Academic Publishers. The Netherlands. pp. 559-599.

Bhalla, U.S. and Iyengar, R. (1999) Emergent properties of networks of biological signaling pathways. *Science* 283: 381-387.

Block, A., Dangl, J.L., Hahlbrock, K. and Schulze-Lefert, P. (1990) Functional borders, genetic fine structure and distance requirements of cis elements mediating light responsiveness of the parsley chalcone synthase promoter. *Proc. Natl. Acad. Sci.* 87(14): 5387-5391.

Bolle, C., Koncz, C. and Chua, N.H. (2000) PAT1, a new member of the GRAS family, is involved in phytochrome A signal transduction. *Genes Dev.* 14:1269-1278.

Bowler , C. and Chua, N.H. (1994) Emerging themes of plant signal transduction. *Plant Cell* 6: 1529-1541.

Bowler, C., Neuhaus, G., Yamagata, H. and Chua, N.H. (1994) Cyclic GMP and calcium mediate phytochrome phototransduction. *Cell* 77: 73-81.

Brown, P.O. and Botstein, D. (1999) Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* 21: 33-37.

Bruce,W.B. and Quail, P.H. (1990) cis-acting elements involved in photoregulation of an oat phytochrome promoter in rice. *Plant Cell* 2(11): 1081-1089.

Buche, C., Poppe, C., Schäfer, E. and Kretsch, T. (2000) *eid1*: A new Arabidopsis mutant hypersensitive in phytochrome A-dependent high-irradiance responses: *Plant Cell* 12: 547-558.

Campisi, L., Yang, Y., Yi, Y., Heilig, E., Herman, B., Cassista, A.J., Allen, D.W., Xiang, H. and Jack, T. (1999) Generation of enhancer trap lines in Arabidopsis and characterization of expression patterns in the inflorescence. *Plant J.* 17(6): 699-707.

Carabelli, M., Sessa, G., Baima, S., Morelli, G. and Ruberti, I. (1993) The Arabidopsis Athb-2 and –4 genes are strongly induced by far-red rich light. *Plant J.* 4(3): 469-479.

Carabelli, M., Morelli, G., Whitelam, G. and Ruberti, I. (1996) Twilight zone and canopy shade induction of the Athb-2 homeobox gene in green plants. *Proc. Natl. Acad. Sci.* 93:3530-3535.

Casal, J.J. (2000) Phytochromes, cryptochromes, phototropin: photoreceptor interactions in plants. *Photochem. Photobiol.* 71: 1-11

Casal, J.J. and Boccalandro, H. (1995) Co-action between phytochrome B and HY4 in *Arabidopsis thaliana*. *Planta* 197: 213-218.

Casal, J.J. and Sánchez, R.A. (1998) Phytochromes and seed germination. *Seed Sci. Res.* 8: 317-329.

Casal, J.J., Cerdán, P.D., Staneloni, R.J. and Cattaneo, L. (1998) Different phototransduction kinetics of phytochrome A and phytochrome B in *Arabidopsis thaliana*. *Plant Physiol.* 116: 1533-1538.

Casal, J.J., Sánchez, R.A. and Detegibus,V.A. (1986) Effects of plant density on tillering: the involvement of the R/Fr and the proportion of radiation intercepted per plant. *Exp. environm. Bot.* 26: 365-371.

Casal, J.J. Sánchez, R.A., Benedetto A.H. De Miguel L.C. (1991) Light promotion of seed germination in *Datura ferox* is mediated by a highly stable pool of phytochrome *Photochem. Photobiol.* 53:249-254.

Casal, J.J., Yanovsky, M.J. and Luppi J.P. (2000) Two photobiological pathways of phytochrome A activity, only one of which shows dominant negative suppression of phytochrome B. *Photochem. Photobiol.* 71(4): 481-486.

Castresana, C., García-Luque, I., Alonso, E., Malik, V.S. and Cashmore, A.R. (1988) Both positive and negative regulatory elements mediate expression of photoregulated CAB gene from *Nicotiana plumbaginifolia. EMBO J.* 7(7): 1929-1936

Cerdán, P.D., Staneloni, R.J., Ortega, J., Bunge, M.M., Rodriguez-Batiller, M.J., Sánchez, R.A. and Casal, J.J. (2000) Sustained but not transient phytochrome A signaling targets a region of an Lhcb1*2 promoter not necessary for phytochrome B action. *Plant Cell* 12: 1203-1211.

Cerdán, P.D., Yanovsky, M.J., Reymundo, F.C., Nagatani, A., Staneloni, R.J., Whitelam, G.C. and Casal, J.J. (1999) Regulation of phytochrome B signaling by phytochrome A and FHY1 in *Arabidopsis thaliana. Plant J.* 18(5): 499-507.

Chamovitz, D.A. and Deng, X.W. (1996) Light signaling in plants. *Crit. Rev. Plant Sci.* 15: 455-478.

Chattopadhyay, S., Ang, L.H., Puente, P., Deng, X.W. and Wei, N. (1998a) Arabidopsis bZIP protein HY5 directly interacts with light-responsive promoters in mediating light control of gene expression. *Plant Cell* 10(5): 673-783.

Chattopadhyay, S. Puente, P. Deng, X.W. and Wei, N. (1998b) Combinatorial interaction of light-responsive elements plays a critical role in determining the response characteristics of light-regulated promoters in Arabidopsis. *Plant J.* 15(1): 69-77.

Choi, G., Yi, H., Lee, J. Kwon, Y.K., Soh, M.S., Shin, B., Luka, Z., Hahn, T.R. and Song, P.S. (1999) Phytochrome signaling is mediated through nucleoside diphosphate kinase 2. *Nature* 401(6753): 610-613.

Chory, J. (1993) Out of darkness: mutants reveal pathways controlling light-regulated development in plants. *Trend Genet.* 9, 167-172.

Chory, J. and Peto, C.A. (1990) Mutations in the DET1 gene affect cell-type-specific expression of light regulated genes and chloroplast development in Arabidopsis. *Proc. Natl. Acad. Sci.* 87(22): 8776-8780.

Chory, J. and Wu, D. (2001) Weaving the complex web of signal transduction. *Plant Physiol.* 125: 77-80.

Chory, J., Chatterjee, M., Cook R.K., Elich, T., Fankhauser, C., Li, J. Nagpul, P., Neff, M., Pepper, A., Poole, D., Reed, J. and Vitart, V. (1996) From seed germination to flowering, light controls plant development via the pigment phytochrome. *Proc. Natl. Acad. Sci.* 93(22):12066-12071.

Chory, J., Cook, R.K., Dixon, R., Elich, T., Li, H.M., López, E., Mochizuki, N., Nagpal, P, Pepper, A., Poole, D. and Reed, J. (1995) Signal-transduction pathways controlling light-regulated development in Arabidopsis. *Phil. Trans. R. Soc. Lond.* 350: 59-65.

Christie, J.M., and Jenkins, G.I. (1996) Distinct UV-B and UV-A blue light signal transduction pathways induce chalcone synthase gene expression in Arabidopsis cells. *Plant Cell* 8:1555-1567.

Clack, T., Mathews, S. and Sharrock, R.A. (1994) The phytochrome apoprotein family in Arabidopsis is encoded by five genes: the sequences and expression of PHYD and PHYE. *Plant Mol. Biol.* 25:413-427

Cone, J.W. Jaspers, P.A. and Kendrick R.E. (1985) Biphasic fluence response curves for light induced germination of *Arabidopsis thaliana* seeds. *Plant Cell and environm.* 8: 605-612.

Coruzzi, G., Broglie, R., Edwards, C. and Chua, N.H. (1984) Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5-biphosphate carboxylase. *EMBO J.* 3(8): 1671-1679.

Cummings, C.A. and Relman, D.A. (2000) Using DNA microarrays to study host-microbe interactions. *Genomics* 6(5): 513-525

Dehesh, K., Bruce, W.B. and Quail, P.H. (1990) A trans-acting factor that binds to a GT-motif in a gene promoter. *Science* 250(4986): 1397-1399.

Deng, X.W. and Quail, P.H. (1999) Signaling in light-controlled development. *Sem. Cell Dev. Biol.* 10: 121-129.

Devlin, P.F., Robson, P.R.H., Patel, S.R., Goosey, L. Sharrock, R.A. and Whitelam, G.C. (1999) Phytochrome D acts in the shade-avoidance syndrome in Arabidopsis by controlling elongation growth and flowering time. *Plant Physiol.* 119: 909-915.

Dickey, L.F., Petracek, M.E., Nguyen, T.T., Hansen, E.R. and Thompson, W.F. (1998) Light regulation of Fed-1 mRNA requires an element in the 5' untranslated region and correlates with differential polyribosome association. *Plant Cell* 10: 475-484.

Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. and Trent, J. (1999) Expression profiling using cDNA microarrays. *Nat. Genet.* 21: 10-14.

El-Hamdouchi, A. and Willett, P. (1986) Hierarchic document classification using Ward's clustering method. Proceedings of 1986 ACM conference on research and development in information retrieval 149-156.

Enoki, H., Izawa, T., Kawahara, M., Komatsu, M., Koh, S., Kyozuka, J. and Shimamoto, K. (1999) Ac as a tool for the functional genomics of rice. *Plant J.* 19(5): 605-613.

Fairchild, C.D., Schumaker, M.A. and Quail, P.H. (2000) HFR1 encodes an atypical bHLH protein that acts in phytochrome A signal transduction. *Genes Dev.* 14(18): 2377-2391.

Fankhauser, C. (2001) The phytochromes, a family of red/far-red absorbing photoreceptors. *J. Biol. Chem.* 276(15): 11453-11456.

Fankhauser, C. and Chory, J. (1997) Light control of plant development. *Annu. Rev. Cell Dev. Biol.* 13:203-229.

Fankhauser, C., Yeh, K.C., Lagarias, J.C., Zhang, H., Elich, T.D. and Chory, J. (1999) PKS1, a substrate phosphorylated by phytochrome that modulates light signaling in Arabidopsis. *Science* 284(5419): 1539-1541.

Fluhr, R., Kuhlemeier, C., Nagy, F. and Chua, N.H. (1986) Organ-specific and light-induced expression of plant genes. *Science* 232: 1106-1112.

Foster, E., Hattori, J., Labbe, H., Ouellet, T., Fobert, P.R., Jamer, L.E., Iyer, V.N. and Miki, B.L. (1999) A tobacco cryptic constitutive promoter, tCUP, revealed by T-DNA tagging. *Plant Mol. Biol.* 41(1): 45-55.

Gallo-Meagher, M., Sowinski, D.A., Elliot, R.C. and Thompson, W.F. (1992) Both internal and external regulatory elements control expression of the pea Fed-1 gene in transgenic tobacco seedlings. *Plant Cell* 4: 389-395.

Genoud, T. and Métraux. (1999) Crosstalk in plant cell signaling: structure and function of the genetic network. *Trends Pl. Sci.* 4(12): 503-507.

Genoud, T., Millar, A.J., Nishizawa, N., Kay, S.A., Schafer, E., Nagatani, A. and Chua, N.H. (1998) An Arabidopsis mutant hypersensitive to red and far-red light signals. *Plant Cell* 10: 889-904 .

Gil, P., Dewey, E., Friml, J., Zhao, Y., Snowden, K.C., Putterill, J., Palme, K., Estelle, M. and Chory, J. (2001) BIG: a calossin-like protein required for polar auxin transport in Arabidopsis. *Genes and Dev.* 15: 1985-1997.

Gilmartin, P.M., Sarokin, L., Memelink, J. and Chua, N.H. (1990) Molecular light switches for plant genes. *Plant Cell* 2(5): 369-378.

Ha, S.B. and An, G. (1988) Identification of upstream regulatory elements involved in the developmental expression of the *Arabidopsis thaliana* cab1 gene. *Proc. Natl. Acad. Sci.* 85 (21): 8017-8021.

Halliday, K.J., Hudson, M., Ni, M., Qin, M. and Quail, P.H. (1999) poc1: an Arabidopsis mutant perturbed in phytochrome signaling because of a T-DNA insertion in the promoter of PIF3, a gene encoding a phytochrome-interacting bHLH protein. *Proc. Natl. Acad. Sci.* 96(10): 5832-5837.

Hansen, E.R., Petracek, M.E., Dickey, L.F. and Thompson, W.F. (2001) The 5' end of the pea ferredoxin-1 mRNA mediates rapid and reversible light-directed changes in translation in tobacco. *Plant Physiol.* 125: 770-778.

Hardtke, C.S. and Deng, X.W. (2000) The cell biology of the COP/DET/FUS proteins: Regulating proteolysis in photomorphogenesis and beyond?. *Plant Physiol.* 124: 1548-1557.

Hardtke, C.S., Gohda, K., Osterlund, M.T., Oyama, T., Okada, K. and Deng, X.W. (2000) HY5 stability and activity in Arabidopsis is regulated by phosphorylation in its COP1 binding domain. *Embo J.* 19(18): 4997-5006.

Harmer, S.L. and Kay, S.A. (2000) Microarrays: Determining the balance of cellular transcription. *Plant Cell* 12: 613-615.

Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A. and Kay, S.A. (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science* 290: 2110-2113.

Hennig, L., Poppe, C., Sweere, U., Martin, A. and Schäfer, E. (2001) Negative interference of endogenous phytochrome B with phytochrome A function in Arabidopsis. *Plant Phys.* 125:1036-1044.

Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acid. Res.* 27(1): 297-300.

Hoecker, U., Tepperman, J.M. and Quail, P.H. (1999) SPA1, a WD-repeat protein specific to phytochrome A signal transduction. *Science* 284, 496-499.

Hoecker, U., Xu, Y. and Quail, P.H. (1998) SPA1: a new genetic locus involved in phytochrome A-specific signal transduction. *Plant Cell* 10(1): 19-33.

Hoheisel, J.D. and Vingron, M. (2000) Transcriptional profiling: is it worth the money? *Res. Microbiol.* 151: 113-119.

Horwitz, B.A., Thompson, W.F. and Briggs, W.R. (1988) Phytochrome regulation of greening in *Pisum*. *Plant Physiol.* 86: 299-305.

Hou, Y., von Arnim, A.G. and Deng, X.W. (1993) A new class of Arabidopsis constitutive photomorphogenic genes involved in regulating cotyledon development. *Plant Cell* 5: 329-339.

Hudson, M., Ringli, C., Boylan, M.T. and Quail. P.H. (1999) The FAR1 locus encodes a novel nuclear protein specific to phytochrome A signaling. *Genes Dev.* 13, 2017-2027.

Hui, E.K., Wang, P.C. and Lo, S.J. (1998) Strategies for cloning unknown cellular flanking DNA sequences from foreign integrants. *Cell. Mol. Life Sci.* 54: 1403-1411.

Huq, E., Kang, Y., Halliday, K.J., Qin, M. and Quail, P.H. (2000) SRL1: a new locus specific to the phyB-signaling pathway in Arabidopsis. *Plant J.* 23(4): 461-470.

Ideker, T., Thorsson, V., Ranish, J.A., Christmas, R., Buhler, J., Eng, J.K., Bumgarner, R., Goodlett, D.R., Aebersold, R. and Hood, L. (2001) Integrated genomic and proteomic analyses of a systemically perturbed metabolic network. *Science* 292: 929-934.

Inaba, T., Nagano, Y., Reid, J.B. and Sasaki, Y. (2000) DE1, a 12-base pair cis-regulatory element sufficient to confer dark-inducible and light down-regulated expression to a minimal promoter in pea. *J. Biol. Chem.* 275(26): 19723-19727.

Inaba, T., Nagano, Y., Sakakibara, T. and Sasaki, Y. (1999) Identification of a cis-regulatory element involved in phytochrome down-regulated expression of the pea small GTPase gene *pra2*. *Plant Phys.* 120: 491-499.

Irniger, S., Egli, C.M., Kuenzler, M., and Braus, G.H. (1992) The yeast actin intron contains a cryptic promoter that can be switched on by preventing transcriptional interference. *Nucleic Acids. Res.* 20(18): 4733-4739

Ito, T., Kito, K., Adati, N., Mitsui, Y., Hagiwara, H. and Sakaki, Y. (1994) Fluorescent differential display: arbitrarily primed RT-PCR fingerprinting on automated DNA sequencer. *FEBS Lett.* 351: 231-236.

Jackson, J.A., Fuglevand, G., Brown, B.A., Shaw, M.J. and Jenkins, G.I. (1995) Isolation of Arabidopsis mutants altered in the light-regulation of chalcone synthase gene expression using a transgenic screening approach. *Plant J.* 8(3): 369-380.

Jensen, P.J., Hangarter, R.P., and Estelle, M. (1998) Auxin transport is required for hypocotyl elongation in light-grown but not dark-grown Arabidopsis. *Plant Physiol.* 116: 455-462.

Jordan, J.D., Landau, E.M. and Iyengar, R. (2000) Signaling networks: The origins of cellular multitasking. *Cell* 103: 192-200

Karlin-Neumann, G.A., Brusslan, J. and Tobin, E. (1991) Phytochrome control of the *tms2* gene in transgenic Arabidopsis: A strategy for selecting mutants in the signal transduction pathway. *Plant Cell* 3(6): 573-582.

Kaufman, L.S., Briggs, W.R. and Thompson, W.F. (1985) Phytochrome control of specific mRNA levels in developing pea buds. *Plant Phys.* 78: 388-393.

Kaufman, L.S., Roberts, L.L., Briggs, W.R. and Thompson, W.F. (1986) Phytochrome control of specific mRNA levels in developing pea buds. *Plant Physiol.* 81: 1033-1038.

Kaufman, L.S., Thompson, W.F. and Briggs W.R. (1984) Different red light requirements for phytochrome induced accumulation of Cab RNA and rbcS RNA. *Science* 226: 1447-1449.

Kehoe, D.M., Degenhardt, J., Winicov, I. and Tobin, E.M. (1994) Two 10-bp regions are critical for phytochrome regulation of a *Lemna gibba Lhcb* gene promoter. *Plant Cell* 6: 1123-1134.

Kehoe, D.M., Villand, P. and Somerville, S. (1999) DNA microarrays for studies of higher plants and other photosynthetic organisms. *Trends Plant Sci.* 4(1): 38-41

Khurana, J.P., Kochhar, A. and Tyagi, A.K. (1998) Photosensory perception and signal transduction in higher plants- molecular genetic analysis. *Crit. Rev. Plant Sci.* 17(5): 465-539.

Kircher, S., Kozma-Bognar, L., Kim, L., Adam, E., Harter, K., Schäfer and Nagy, F. (1999) Light quality-dependent nuclear import of the plant photoreceptors phytochrome A and B. *Plant Cell* 11: 1445-1456.

Kircher, S., Wellmer, F., Nick, P., Rügner, A., Schäfer, E. and Harter, K. (1999a) Nuclear import of the parsley bZIP transcription factor CPRF2 is regulated by phytochrome photoreceptors. *J. Cell Biol.* 144: 201-211.

Krall, L. and Reed, J.W. (2000) The histidine kinase-related domain participates in phytochrome B function but is dispensable. *Proc. Natl. Acad. Sci.* 97(14): 8169-8174.

Krogh, A., Mian, I.S., and Haussler, D. (1994) A hidden Markov model that finds genes in *E. coli* DNA. *Nucl. Acids Res.* 22, 4768-4778.

Kuno, N., Maramatsu, T., Hamazato, F. and Furuya, M. (2000) Identification by large-scale screening of phytochrome regulated genes in etiolated seedlings of Arabidopsis using a fluorescent differential display technique. *Plant Physiol.* 122: 15-24.

Lam, E. and Chua, N.H. (1989) ASF-2: A factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in cab promoter. *Plant Cell* 1:1147-1156.

Lam, E. and Chua, N.H. (1990) GT-1 binding site confers light responsive expression in transgenic tobacco. *Science* 248(4954): 471-474.

Lawrence C. E., Altschul, S.F., Boguski, M.S., Liu, J.S., Neuwald, A.F. and Wootton, J.C. (1993) Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science* 262 (5131): 208-14.

Leu, W.M., Cao, X.L., Wilson, T.J., Snustad, D.P. and Chua, N.H. (1995) Phytochrome A and phytochrome B mediate the hypocotyl-specific downregulation of TUB1 by light in Arabidopsis. *Plant Cell* 7: 2187-2196.

Leutwiler, L.S., Hough-Evans, B.R. and Meyerowitz, E.M. (1984) The DNA of *Arabidopsis thaliana*. *Mol. Gen. Genet.* 194: 15-23.

Li, H.M., Altschmied, L. and Chory, J. (1994) Arabidopsis mutants define downstream branches in the phototransduction pathway. *Genes Dev.* 8(3): 339-349.

Li, H.M., Washburn, T. and Chory, J. (1993) Regulation of gene expression by light. *Curr. Opin. Cell Biol.* 5: 455-460

Liang, P. and Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257: 967-971

Liebman, S.W. and Newman, G. (1993) A ubiquitin-conjugating enzyme, RAD6, affects the distribution of Ty1 retrotransposon integration positions. *Genetics* 133: 499-508.

Lin, C. (2000) Photoreceptors and regulation of flowering time. *Plant Phys.* 123: 39-50

Liscum, E. and Briggs, W.R. (1995) Mutations in the NPH1 locus of Arabidopsis disrupt the perception of phototropic stimuli. *Plant Cell* 7: 473-485.

Liu, Y.G. and Whittier, R.F. (1995) Thermal asymmetric interlaced PCR: Automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 25: 674-681

Liu, X.L., Covington, M.F., Fankhauser, C., Chory, J. and Wagner, D.R. (2001) ELF3 encodes a circadian clock-regulated nuclear protein that functions in an Arabidopsis PHYB signal transduction pathway. *Plant Cell* 13(6): 1293-1304.

Long, D., Martin, M., Sundberg, E., Swinburne, J., Puangsomlee, P. and Coupland, G. (1993) The maize transposable element system Ac/Ds as a mutagen in Arabidopsis: Identification of an albino mutation induced by Ds insertion. *Proc. Natl. Ac. Sci.* 21: 10370-10374.

Lubberstedt, T., Bolle, C.E., Sopory, S., Flueger, K., Herrmann, R.G. and Oelmuller, R. (1994) Promoters from genes for plastid proteins possess regions with different sensitivities toward red and blue light. *Plant Physiol.* 104(3): 997-1006.

Lyapina, S., Cope, G., Shevchenko, A., Serino, G., Tsuge, T., Zhou, C., Wolf, D.A., Wei, N., Shevchenko, A. and Deshaies, R.J. (2001) Promotion of NEDD-CUL1 conjugate cleavage by COP9 signalosome. *Science* 292(5520): 1382-1385.

Mancinelli, A.L. (1994) The physiology of phytochrome action. In: Kendrick, R.E. and Kronenberg, G.H.M. (eds), Photomorphogenesis in Plants. 2nd Edition. Kluwer Academic Publishers. The Netherlands. Pp.211-269.

Martínez-García, J.F., Huq, E. and Quail, P.H. (2000) Direct targeting of light signals to a promoter element-bound transcription factor. *Science* 288(5467): 821-822.

Mas, P., Devlin, P.F., Panda, S. and Kay, S.A. (2000) Functional interaction of phytochrome B and cryptochrome 2. *Nature* 408(76809): 207-211.

Mazzella, M.A., Alconoda-Magliano, T.M. and Casal, J.J. (1997) Dual effect of phytochrome A on hypocotyl growth under continuous red light. *Plant Cell Environ.* 20, 261-267.

McNellis, T.W. and Deng, X.W. (1995) Light control of seedling morphogenetic pattern. *Plant Cell* 7: 1749-1761.

Meagher, R.B., McKinney, E.C. and Kandasamy, M.K. (1999) Isovariant dynamics expand and buffer the responses of complex systems: the diverse plant actin gene family. *Plant Cell* 11, 995-1006.

Meier, I., Callan, K.L., Fleming, A.J. and Gruissem, W. (1995) Organ-specific differential regulation of a promoter subfamily for the ribulose-1,5-biphosphate carboxylase/oxygenase small subunit genes in tomato. *Plant Phys.* 107(4): 1105-1118.

Meissner, R., Chague, V., Zhu, Q., Emmanuel, E., Elkind, Y. and Levy, A.A. (2000) Technical advance: a high throughput system for transposon tagging and promoter trapping in tomato. *Plant J.* 22(3): 265-274.

Millar, A. and Kay, S.A. (1996) Integration of circadian and phototransduction pathways in the network controlling CAB gene transcription in Arabidopsis. *Proc. Natl. Acad. Sci.* 93: 15491-15496.

Millar, A.J., Carré, I.A., Strayer, C.A., Chua, N.H. and Kay, S.A. (1995) Circadian clock mutants in Arabidopsis identified by luciferase imaging. *Science* 267: 1161-1163.

Millar, A.J., McGrath, R.B. and Chua, N.H. (1994) Phytochrome phototransduction pathways. *Ann. Rev. Gen.* 28: 325-349.

Mockler, T.C., Guo, H., Yang, H., Duong, H. and Lin, C. (1999) Antagonistic actions of Arabidopsis chryptochromes and phytochrome B in the regulation of floral induction. *Development* 126(10): 2073-2081.

Møller, S.G. and Chua, N.H. (1999) Interactions and intersections of plant signaling pathways. *J. Mol. Biol.* 293: 219-234.

Mulligan, R.M., Chory, J. and Ecker, J.R. (1997) Signaling in plants. *Proc, Natl. Acad.* Sci. 94(7): 2793-2795

Nagatani, A., Reed, J.W. and Chory, J. (1993) Isolation and initial characterization of Arabidopsis mutants that are deficient in functional phytochrome A. *Plant Phys.* 102: 269-277.

Nagy, F., Fluhr, R., Kuhlemeier, C., Kay, S., Boutry, M., Green, P., Poulsen, C. and Chua, N.H. (1986) Cis-acting elements for selective expression of two photosynthetic genes in transgenic plants. *Phil. Trans. R. Soc. Lond.* 314 (1166): 493-500.

Nakagawa, Y., Machida, C., Machida, Y. and Toriyama, K. (2000) Frequency and pattern of transposition of the maize transposable element Ds in transgenic rice plants. *Plant Cell Physiol.* 41(6): 733-742.

Neff, M.M. and Chory, J. (1998) Genetic interactions between phytochrome A, phytochrome B and cryptochrome 1 during Arabidopsis development. *Plant Phys.* 118: 27-36.

Neff, M.M., Fankhauser, C. and Chory, J. (2000) Light: an indicator of time and place. *Genes Dev.* 14: 257-271.

Neuhaus, G., Bowler, C., Hiratsuka, K., Yamagata, H. and Chua, N.H. (1997) Phytochrome-regulated repression of gene expression requires calcium and cGMP. *EMBO J.* 16: 2254-2264.

Neuhaus, G., Bowler, C., Kern, R. and Chua, N.H. (1993) Calcium/calmodulin-dependent and –independent phytochrome signal transduction pathways. *Cell* 73(5): 937-952.

Ngai , N., Tsai, F.Y. and Coruzzi, G. (1997) Light-induced transcriptional repression of the pea AS1 gene: identification of cis-elements and transfactors. *Plant J.* 12(5): 1021-1034

Ni, M., Tepperman, J.M. and Quail, P.H. (1998) PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein. *Cell* 95: 657-667.

Ni, M., Tepperman, J.M. and Quail, P.H. (1999) Binding of phytochrome B to its nuclear signaling partner PIF3 is reversibly induced by light. *Nature* 400(6746): 710-711.

Nussaume, L., Harrison, K., Klimyuk, V., Martienssen, R., Sundaresan, V. and Jones, J.D. (1995) Analysis of splice donor and acceptor site function in a transposable gene trap derived from the maize element Activator. *Mol. Gen. Genet.* 249(1): 91-101.

Ohgishi, M., Oka, A., Morelli, G., Ruberti, I. and Aoyama, T. (2001) Negative autoregulation of the Arabidopsis homeobox gene ATHB-2. *Plant J.* 25(4): 389-398.

O'Kane, C. and Gehring, W. J. (1987) Detection in situ of genomic regulatory elements in Drosophila. *Proc. Natl. Acad. Sci.* 84, 9123-9127.

Okamoto, H., Matsui, M. and Deng, X.W. (2001) Overexpression of the heterotrimeric G-protein α-subunit enhances phytochrome-mediated inhibition of hypocotyl elongation in Arabidopsis. *Plant Cell* 13: 1639-1651.

Orozco, B.M. and Ogre, W.L. (1993) Localization of light-inducible and tissue-specific regions of the spinach ribulose biphosphate carboxylase/oxygenase (rubisco) activase promoter in transgenic tobacco plants. *Plant Mol. Biol.* 23(6): 1129-1138.

Osterlund, M.T., Hardtke, C.S., Wei, N. and Deng, X.W. (2000) Targeted destabilization of HY5 during light regulated development of Arabidopsis. *Nature* 405: 462-466.

Parinov, S., Sevugan, M., Ye, D., Yang, Y.W., Kumaran, M. and Sundaresan, V. (1999) Analysis of flanking sequences from Dissociation insertion lines: A database for reverse genetics in Arabidopsis. *Plant Cell* 11: 2263-2270.

Parks, B.M. and Quail, P.H. (1991) Phytochrome-deficient hy1 and hy2 long hypocotyl mutants of Arabidopsis are defective in phytochrome chromophore biosynthesis. *Plant Cell* 3: 1177-1186.

Parks, B.M., Shanklin, J., Koonneef, M., Kendrick, R.E. and Quail, P.H. (1989) Immunochemically detectable phytochrome is present at normal levels but is photochemically nonfunctional in the hy1 and hy2 hypocotyl mutants of Arabidopsis. *Plant Mol. Biol.* 12: 425-437.

Petracek, M.E., Dickey, L.F., Nguyen, T.T., Gatz, C., Sowinski, D.A., Allen, G.C. and Thompson, W.F. (1998) Ferredoxin-1 mRNA is destabilized by changes in photosynthetic electron transport. *Proc. Natl. Acad. Sci.* 95: 9009-9013.

Piechulla, B. (1999) Circadian expression of the light-harvesting complex protein genes in plants. *Chronobiol. Int.* 16: 115-128.

Poppe, C., Sweere, U., Drumm-Herrel, H. and Schafer, E. (1998) The blue light receptor cryptochrome 1 can act independently of phytochrome A and B in *Arabidopsis thaliana*. *Plant J.* 16(4): 465-471.

Puente, P., Wei, N. and Deng, X.W. (1996) Combinatorial interplay of promoter elements constitutes the minimal determinants for light and developmental control of gene expression in Arabidopsis. *EMBO J.* 15(14): 3732-3743.

Quail, P.H., Boylan, M.T., Parks, B.M., Short, T.W., Xu, Y. and Wager, D. (1995) Phytochromes: photosensory perception and signal transduction. *Science* 268, 675-680

Reed, J.W., Nagpal, P., Poole, D.S., Furuya, M. and Chory, J. (1993) Mutations in the gene for the red/far-red light photoreceptor phytochrome B alter cell elongation and physiological responses throughout Arabidopsis development. *Plant Cell* 5: 147-157.

Ruan, Y, Gilmore, J. and Conner, T. (1998) Towards Arabidopsis genome analysis: monitoring expression profiles of 1400 genes using cDNA microarrays. *Plant J.* 15(6): 821-833.

Salomon, M., J.M. Christie, E. Knieb, U. Lempert and W.R. Briggs. (2000) Photochemical and mutational analysis of FMN-binding domains of the plant blue light receptor, phototropin. *Biochemistry* 39: 9401-9410.

Schaffer, R., Landgraf, J., Accerbi, M., Simon, V., Larson, M. and Wisman, E. (2001) Microarray analysis of diurnal and circadian-regulated genes in Arabidopsis. *Plant Cell* 13: 113-123.

Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270: 467-470.

Schenk, P.M>, Kazan, K., Wilson, I., Anderson, J.P., Richmond, T., Somerville, S.C. and Manners, J.M. (2000) Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. *Proc. Nat. Acad. Sci.* 97(21): 11655-11660.

Schmidt, R. and Willmitzer, L. (1989) The maize autonomous element activator (Ac) shows a minimal germinal excision frequency of 0.2%-0.5% in transgenic *Arabidopsis thaliana* plants. *Mol. Gen. Genet.* 220: 17-24.

Schmitt, J., McCormac, A.C. and Smith, H. (1995) A test of the adaptive plasticity hypothesis using transgenic and mutant plants disabled in phytochrome-mediated elongation responses to neighbors. *Am. Nat.* 146(6): 937-953.

Schwechheimer, C., Serino, G., Callis, J., Crosby, W.L., Lyapina, S., Deshaies, R.J., Gray, W.M., Estelle, M. and Deng. X.W. (2001) Interactions of the COP9 signalosome with the E3 ubiquitin ligase SCFTIRI in mediating auxin response. *Science.* 292(5520): 1379-1382.

Scopel, A.L., Ballaré, C.L. and Sánchez, R.A. (1991) Induction of extreme light sensitivity in buried weed seeds and its role in the perception of soil cultivations. *Plant Cell Environ.* 14:501-508.

Serikawa, K.A., Martinez-Laborda, A., Kim, H.S. and Zambryski, P.C. (1997) Localization of expression of KNAT3, a class of knotted1-like gene. *Plant J.* 11(4): 853-861.

Shalon, D., Smith, S.J. and Brown, P.O. (1996) A DNA micro-array system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* 6: 639-645.

Sheng, J., Jeon, J. and Mehdy, M.C. (1993) Developmental regulation and phytochrome-mediated induction of mRNAs encoding a proline-rich protein, glycine-rich proteins and hydroxyproline-rich glycoproteins in *Phaseolus vulgaris* L. *Proc. Natl. Acad. Sci.* 90(3): 828-832.

Shinomura, T., Nagatani, A., Hanzawa, H., Kubota, M., Watanabe, M. and Furuya, M. (1996) Action spectra for phytochrome A- and B-specific photoinduction of seed germination in *Arabidopsis thaliana. Proc. Natl. Acad. Sci.* 93(15): 8129-8133.

Skarnes, W.C. (1990) Entrapment vectors: a new tool for mammalian genetics. *Biotech.* 8(9): 827-831.

Smith, H. (2000) Phytochromes and light signal perception by plants - an emerging synthesis. *Nature* 407: 585-591.

Smith, H. (1982) Light quality, photoreception and plant strategy. Ann.Rev. *Plant Phys.* 33:481-518.

Smith, H. (1994) Sensing the light environment: the functions of the phytochrome family. In: R.E. Kendrick & G.H.M. Kronenberg (eds), Photomorphogenesis in Plants. 2nd Edition. Kluwer Academic Publishers. The Netherlands, pp.377-416.

Smith, H., Xu, Y. and Quail, P.H. (1997) Antagonistic but complementary actions of phytochromes A and B allow seedling de-etiolation. *Plant Physiol.* 114(2):637-641.

Soh, M.S., Hong, S.H., Hanzawa, H., Furuya, M. and Nam, H.G. (1998) Genetic identification of FIN2, a far red light-specific signaling component of *Arabidopsis thaliana. Plant J.* 16, 411-419.

Soh, M.S., Kim, Y.M., Han S.J. and Song P.S. (2000) REP1, a basic helix-loop-helix protein, is required for a branch pathway of phytochrome A signaling in Arabidopsis. *Plant Cell* 12(11): 2061-2074.

Solis, R., Takumi, S., Mori, N. and Nakamura, C. (1999) Ac-mediated trans-activation of the Ds element in rice (*Oryza sative* L.) cells as revealed by GUS assay. *Hereditas* 131(1): 23-31.

Somers, D.E. and Quail, P.H. (1995) Temporal and spatial expression patterns of PHYA and PHYB genes in Arabidopsis. *Plant J.* 7(3): 413-427.

Somers, D.E., Sharrock, R.A., Tepperman, J.M. and Quail, P.H. (1991) They hy3 long hypocotyl mutant of Arabidopsis is deficient in phytochrome B. *Plant Cell* 3:1263-1274.

Spiegelman, J.L., Mindrinos, M.N., Fankhauser, C., Richards, D., Lutes, J., Chory, J. and Oefner, P.J. (2000) Cloning of the Arabidopsis RSF1 gene by using a mapping strategy based on high-density DNA arrays and denaturing high-performance liquid chromatography. *Plant Cell* 12: 2485-2498.

Spradling, A.C., Stern, D.M., Kiss, I., Roote, J., Laverty, T. and Rubin, G.M. (1995) Gene disruptions using P transposable elements: An integral component of the Drosophila genome project. *Proc. Natl. Acad. Sci.* 92: 10824-10830.

SPSS®. (1999) SPSS® base 10.0 applications guide.

Sugiyama, N., Izawa, T., Oikawa, T. and Shimamoto, K. (2001) Light regulation of circadian clock-controlled gene expression in rice. *Plant J.* 26(6): 607-615.

Sun, L. and Tobin, E.M. (1990) Phytochrome-regulated expression of genes encoding light-harvesting chlorophyll a/b -protein in two long hypocotyl mutants and wild type plants of *Arabidopsis thaliana*. *Photochem. Photobiol.* 52: 51-56.

Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J.D.G., Dean, C., Ma, H. and Martienssen, R. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements.. *Genes and Dev.* 9, 1797-1810.

Takayuki, K., Chiba, T., Suzuki, T., Iwai, K., Yamanaka, K., Minato, N., Suzuki, H., Shimbara, N., Hidaka, Y., Osaka, F., Omata, M. and Tanaka, K. (2001) NEDD8 recruits E2-ubiquitin to SCF E3 ligase. *Embo J.* 20(15): 4003-4012.

Teakle, G.R. and Kay, S.A. (1995) The GATA-binding protein CGF-1 is closely related to GT-1. *Plant mol. Biol.* 29(6): 1253-1266.

Terzaghi, W.B. and Cashmore, A.R. (1995) Light-regulated transcription. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 46: 445-474.

Tepperman, J.M., Zhu, T., Chang, H.S., Wang, X. and Quail, P.H. (2001) Multiple transcription-factor genes are early targets of phytochrome A signaling. *Proc. Natl. Acad. Sci.* 98(16): 9437-9442.

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

Thompson, W.F., Everett, M., Polans, N.O., Jorgensen, R.A. and Palmer, J.D. (1983) Phytochrome control of RNA levels in developing pea and mung bean leaves. *Planta* 158: 487-500.

Tobin, E.M. and Silverthorne, J. (1985) Light regulation of gene expression in higher plants. Annu. Rev. *Plant Physiol.* 36: 569-593.

Tucker, C.L., Gera, J.F. and Uetz, P. (2001) Towards an understanding of complex protein networks. *Trends Cell Biol.* 11(3): 102-106.

Ulla, H., Chen, J.G., Young, J.C., Im, K.H., Sussman, M.R. and Jones, A.M. (2001) Modulation of cell proliferation by heterotrimeric G protein in Arabidopsis. *Science* 292: 2066-2069.

Velculescu, V.E. , Zhang, L., Vogelstein, B. and Kinzler, K.W. (1995) Serial analysis of gene expression. *Science* 270: 484-487.

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Hieter, P., Vogelstein, B. and Kinzler, K.W. (1997) Characterization of the yeast transcriptome. *Cell,* 88: 243-251.

Vierstra, R.D. and Quail, P.H. (1983) Purification and initial characterization of 124-kilodalton phytochrome from *Avena*. *Biochemistry* 22: 2498-2505.

Vohradsky, J. (2001) Neural network model of gene expression. *FASEB J.* 15(3): 846-854.

Wagner , D., Hoecker, U. and Quail, P.H. (1997) RED1 is necessary for phytochrome B-mediated red light specific signal transduction in Arabidopsis. *Plant Cell* 9(5): 731-743

Walter, M.H., Liu, J.W., Wunn, J. and Hess, D. (1996) Bean ribonuclease-like pathogenesis-related protein genes (Ypr10) display complex patterns of developmental, dark-induced and exogenous-stimulus-dependent expression. *Eur. J. Biochem.* 239(2): 281-293.

Wang, Z.Y., and Tobin, E.M. (1998) Constitutive expression of the circadian clock associated 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell* 93: 1207-1217

Wang, H., Ma, L.G., Li, J.M., Zhao, H.Y. and Deng, X.W. (2001a) Direct interaction of Arabidopsis cryptochromes with COP1 in mediation of photomorphogenic development. Published online August 16, 2001; 10.1126/Science. 1063630 (Science Express Reports).

Wang, X.Q., Ulla, H., Jones, A.M. and Assmann, S.M. (2001b) G protein regulation of ion channels and abscisic acid signaling in Arabidopsis guard cells. *Science* 292: 2070-2072.

Wang, Z.Y., Kenibuch, D., Sun, L., Harel, E., Ong, M.S. and Tobin, E.M. (1997) A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. *Plant Cell* 9: 491-507.

Wehmeyer, B., Cashmore, A.R. and Schäfer, E. (1990) Photocontrol of the expression of genes encoding chlorophyll a/b binding proteins and small subunit of ribulose-1,5-bisphosphate carboxylase in etiolated seedlings of *Lycopersicum esculentum* (L) and *Nicotiana tabacum* (L). *Plant Physiol.* 93: 990-997.

Weisshaar, B., Armstrong, G.A., Block, A., de Costa e Silva, O., and Hahlbrock, K. (1991) Light-inducible and constitutively expressed DNA-binding proteins recognizing a plant promoter with functional relevance in light responsiveness. *EMBO J.* 10: 17777-1786.

Weng, G., Bhalla, U. and Iyengar, R. (1999) Complexity in biological signaling systems. *Science* 284: 92-96.

Whitelam, G.C. and Devlin, P.F. (1997) Roles of different phytochromes in Arabidopsis photomorphogenesis. *Plant Cell Environ.* 20, 752-758.

Wisman, E., Cardon, G.H., Fransz, P. and Saedler, H. (1998) The behavior of the autonomous maize transposable element En/Spm in *Arabidopsis thaliana* allows efficient mutagenesis. *Plant Mol. Biol.* 37: 989-999.

Wu, K., Malik, K., Tian, L., Hu, M., Martin, T., Foster, E., Brown, D. and Miki, B. (2001) Enhancers and core promoter elements are essential for the activity of a cryptic gene activation sequence from tobacco, tCUP. Mol. Genet. *Genomics* 265(5): 763-770.

Yamamoto, Y.Y., Deng, X.W. and Matsui, M. (2001) CIP4, a new COP1 target, is a nucleus-localized positive regulator of Arabidopsis photomorphogenesis. *Plant Cell* 13: 399-411.

Yamamoto, Y.Y., Matsui, M., Ang, L.H. and Deng, X.W. (1998) Role of a COP1 interactive protein in mediating light-regulated gene expression in Arabidopsis. *Plant Cell* 10(7): 1083-1094.

Yanagisawa, S. (2000) Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. *Plant J.* 21(3): 281-288.

Yang, H.Q., Wu, Y.J., Tang, R.H., Liu, D, Liu, Y. and Cashmore, A.R. (2000) The C termini of Arabidopsis cryptochromes mediate a constitutive light response. *Cell* 103: 815-827.

Yanovsky, M.J., Casal, J.J. and Luppi, J.P. (1997) The VLF loci, polymorphic between ecotypes *Landsberg erecta* and Columbia, dissect two branches of phytochrome A signal transduction that correspond to very-low-fluence and high irradiance responses. *Plant J.* 12(3): 659-667.

Yanovsky, M.J., Casal, J.J. and Whitelam, G.C. (1995) Phytochrome A, Phytochrome B and HY4 are involved in hypocotyl growth responses to natural irradiation in Arabidopsis: weak de-etiolation of the phyA mutant under dense canopies. *Plant Cell Env.* 18: 788-794

Yeh, K.C. and Lagarias, J.C. (1998) Eukaryotic phytochromes: light-regulated serine/threonine protein kinases with histidine kinase ancestry. *Proc. Natl. Acad. Sci.* 95(23): 13976-13981.

Yuh, C.H., Bolouri, H. and Davidson, E.H. (1998) Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science* 279: 1896-1902

Zhou, D.X. (1999) Regulatory mechanism of plant gene transcription by GT elements and GT factors. *Trends Plant Sci.* 4(6): 1380-1385.

Appendices

Appendix 1

Tissue specific patterns of expression of gene and enhancer trap lines. Five day old seedlings were stained for GUS and their patterns of expression analyzed under the stereoscope. Included are the patterns of the following lines:

AJ146.55	ET6633	GT6279
ET5158	ET6649	GT6281
ET5203	GT5874	GT6325
ET5267	GT5905	GT6338
ET5280	GT5909	GT6341
ET5359	GT5914	GT6353
ET5403	GT5927	GT6372
ET5443	GT5929	GT6407
ET5487	GT5939	GT6494
ET5491	GT5957	GT6508
ET5529	GT5964	GT6534
ET5555	GT5971	GT6545
ET5580	GT6021	GT6604
ET5599	GT6027	GT6634
ET5627	GT6039	GT6647
ET5642	GT6052	GT6670
ET5653	GT6064	GT6671
ET6375	GT6067	GT6675
ET6417	GT6075	GT6688
ET6426	GT6112	LH211.16
ET6428	GT6113	OK001.15
ET6435	GT6123	OK001.23
ET6537	GT6227	OK003.19
ET6561	GT6228	OK008.13
ET6566	GT6236	OK011.22
ET6611	GT6240	

Symbols:

C = colet	**CE** = cortex-epidermis	**H** = hydathodes
LP = Lateral root primordia	**ME** = mesophyll-epidermis	**P** = petiole
RH = root hairs	**RT** = root tip	**SA** = shoot apex
T = trichomes	**V** = vasculature	

AJ146.55					
Experiment 1					
Treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME3	-	-	-
D	-	-	-	-	-
LD	-	ME3	-	-	-
DL	-	ME3	-	-	-
Experiment 2					
L	-	ME3	-	-	-
D	-	-	-	-	-
R	-	ME3	-	-	-
FR	-	ME3	-	-	-
Rp	-	ME3	-	-	-
R-FR	-	ME3	-	-	-
FRp	-	ME3	-	-	-
FR-R	-	ME3	-	-	-

ET5158					
Experiment 1					
Treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	-	CE3 V3	CE2 C3	-
LD	-	-	-	C2	-
DL	-	-	CE3 V3	CE2 C3	-
Experiment 2					
L	-	ME3	-	-	-
D	-	-	-	-	-
R	-	ME3	-	-	-
FR	-	ME3	-	-	-
Experiment 3					
L	-	-	-	-	-
D	-	ME3 V3	-	-	-
R	-	ME3 V3	-	-	-
FR	-	ME1 V1	-	-	-

ET5203					
Experiment 1					
Treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA3	P3	-	V3	-
D	-	-	-	-	-
LD	-	P1	-	V2	-
DL	-	ME1 P1	CE1 V2	V3	-
Experiment 2					
L	SA3	P1	-	-	-
D	-	-	-	V3 LRP1	-
R	V3	ME1	CE2 V3	V3	-
FR	V2	-	CE1 V2	V3 LRP2	-
Experiment 3					
L	SA3	P3	-	-	-
D	-	-	-	V3	-
R	V3	-	CE2 V3	V3	-
FR	V2	-	CE1 V2	V3	-
Rp	-	-	CE1 V2	V3	-
R-FR	-	-	CE1	V3	-
FRp	-	-	CE1 V3	V3	-
FR-R	V2	-	CE2 V3	V3	-

ET5267					
Experiment 1					
Treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	CE1 RH1	-
D	SA1	ME2 V2	-	-	-
LD	-	ME1	-	-	-
DL	SA1	ME2 V2	CE2 V3	CE1 RH1	-
Experiment 2					
L	-	ME1	-	CE1 RH1	-
D	-	ME2 V2	CE2 V2	-	-
R	-	ME2	-	-	-
FR	-	ME2 V2	-	RH1	-

ET5267					
Experiment 3					
Treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	ME1	-	RH1	-
D	-	ME2 V2	-	-	-
R	-	ME2	-	-	-
FR	-	ME2 V2	-	RH1	-
Rp	-	ME2 V2	CE1	-	-
R-FR	-	ME2 V2	CE1	-	-
FRp	-	ME2 V2	CE1	-	-
FR-R	-	ME2 V2	CE1	-	-

ET5280					
Experiment 1					
Treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	P1	CE2 V3	-	-
LD	-	-	-	-	-
DL	-	-	CE1 V2	-	-
Experiment 2					
L	-	-	CE1	-	-
D	-	P1	CE2 V3	-	-
R	-	-	CE1 V1	-	-
FR	-	-	CE1	-	-
Experiment 3					
L	-	-	CE1	-	-
D	-	P1	CE2 V3	-	-
R	-	-	CE1 V1	-	-
FR	-	-	CE1	-	-
Rp	-	P1	CE2 V3	-	-
R-FR	-	P1	CE2 V3	-	-
FRp	-	P1	CE2 V3	-	-
FR-R	-	-	CE2 V3	-	-

ET5280					
Experiment 4					
Treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	CE1	-	-
D	-	P1	CE2 V3	-	-
R	-	-	CE1 V1	-	-
FR	-	-	CE1	-	-
Rp	-	P1	CE2 V3	-	-
R-FR	-	P1	CE2 V3	-	-
FRp	-	P1	CE2 V3	-	-
FR-R	-	-	CE2 V3	-	-

ET5359					
Experiment 1					
Treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	SA3 V3	P3	CE3 V3	C3	-
D	-	-	CE2 V2	-	-
LD	SA2 V2	P2	CE3 V3	C3	ME1
DL	SA2	-	CE3 V3	-	-
Experiment 2					
L	SA3 V3	P3	CE3 V3	C3	-
D	-	-	CE1 V2	-	-
R	SA3 V3	P3	CE3 V3	C3	-
FR	SA3 V3	P3	CE3 V3	C3	-
Experiment 3					
L	SA3 V3	P3	CE3 V3	C3	-
D	-	-	CE1 V2	-	-
R	SA3 V3	P3	CE3 V3	C3	-
FR	SA3 V3	P3	CE3 V3	C3	-
Rp	-	-	CE1 V2	C2	-
R-FR	-	-	CE1 V2	C2	-
FRp	SA1	-	CE1 V2	C2	-
FR-R	SA1	-	CE1 V2	C2	-

ET5359					
Experiment 4					
Treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA3 V3	P3	CE3 V3	C3	-
D	-	-	CE1 V2	-	-
R	SA3 V3	P3	CE3 V3	C3	-
FR	SA3 V3	P3	CE3 V3	C3	-
Rp	-	-	CE1 V2	C2	-
R-FR	-	-	CE1 V2	C2	-
FRp	SA1	-	CE1 V2	C2	-
FR-R	SA1	-	CE1 V2	C2	-

ET5403					
Experiment 1					
Treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	CE3 C3 RH3	-
D	-	-	CE3 V3	CE2 C3 RH3	-
LD	-	-	CE2 V3	CE3 C2 RH2	-
DL	-	-	CE2 V2	CE1 C3 RH2	-
Experiment 2					
L	-	-	-	CE2 C3 RH3	-
D	-	-	CE3 V3	CE1 C3 RH3	-
R	-	-	CE3 V3	CE1 C3 RH3	-
FR	-	-	CE3 V3	CE1 C3 RH3	-
Experiment 3					
L	-	-	-	CE2 C3 RH3	-
D	-	-	CE3 V3	CE1 C3 RH3	-
R	-	-	CE3 V3	C3 RH3	-
FR	-	-	CE3 V3	C3 RH3	-
Rp	-	-	CE3 V3	C3 RH3	-
R-FR	-	-	CE3 V3	CE1 C3 RH3	-
FRp	-	-	CE3 V3	CE1 C3 RH3	-
FR-R	-	-	CE3 V3	CE1 C3 RH3	-

ET5443					
Experiment 1					
Treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	V1	-
D	-	-	-	C2 CE2	-
LD	-	ME1	-	V1	-
DL	-	ME1	-	V1 C1	-
Experiment 2					
L	-	ME1	-		-
D	-	-	-	C2 CE2	-
R	-	ME1	-	CE2	-
FR	-	ME1	-	CE1	-
Experiment 3					
L	-	-	-	CE2	-
D	-	-	-	C2 CE2	-
R	-	ME1	-	CE1	-
FR	-	ME1	-	CE1	-
Rp	-	-	-	C2 CE2	-
R-FR	-	-	-	C2 CE2	-
FRp	-	-	-	C2 CE2	-
FR-R	-	-	-	C2 CE2	-
Experiment 4					
L	-	ME1	-	CE2	-
D	-	-	-	C2 CE2	-
R	-	ME1	-	CE1	-
FR	-	ME1	-	CE1	-
Rp	-	-	-	C2 CE2	-
R-FR	-	-	-	C2 CE2	-
FRp	-	-	-	C2 CE2	-
FR-R	-	-	-	C2 CE2	-
Experiment 5					
L	-	ME1	-	CE2	-
D	-	-	-	C2 CE2	-
R	-	ME1	-	CE1	-
FR	-	ME1	-	CE1	-
Rp	-	-	-	C2 CE2	-
R-FR	-	-	-	C2 CE2	-
FRp	-	-	-	C2 CE2	-
FR-R	-	-	-	C2 CE2	-

ET5443					
Experiment 6					
Treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME1	-	CE2	-
D	-	-	-	C2 CE2	-
R	-	ME1	CE1	C2 CE2	-
FR	-	ME1	-	C2 CE2	-
Rp	-	-	-	C2 CE2	-
R-FR	-	-	-	C2 CE2	-
FRp	-	-	-	C2 CE2	-
FR-R	-	-	-	C2 CE2	-

ET5487					
Experiment 1					
Treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	V3	-	-	-
D	V2	V2	CE1 V2	V2	-
LD	-	V2	-	-	-
DL	-	ME1 V2	V2	-	-
Experiment 2					
L	-	V2	-	-	-
D	V2	V2	CE1 V2	V2	-
R	V2	ME1 V2	V2	V2	-
FR	-	ME1 V3	-	V1	-
Experiment 3					
L	NO DATA				
D	V2	ME1 V2	CE1 V2	-	-
R	V2	ME1 V2	V2	V2	-
FR	-	ME1 V2	-	V1	-
Rp	V2	ME1 V2	CE1 V2	-	-
R-FR	V2	ME1 V2	CE1 V2	-	-
FRp	V2	ME1 V2	CE1 V2	-	-
FR-R	V2	ME1 V2	CE1 V2	V1	-

ET5491					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME1 V1	-	V2 RH1	-
D	-	ME1 V1	CE2 V2	V2 RH1 RT1	-
LD	-	ME1 V1	CE2 V2	V2 RH1 RT1	-
DL	-	ME2 V1	CE1 V1	V2 RH1 RT1	-
Experiment 2					
L	-	ME1 V1	-	V2 RT1	-
D	-	ME1 V1	CE2 V1	V2 RT1	-
R	-	ME1 V2	CE1	V2 RT1	-
FR	-	ME2 V2	-	V2 RT1	-
Experiment 3					
L	-	ME1 V2	V1	V2 RT1	-
D	-	ME1 V2	CE2 V2	V2 RT1	-
R	-	ME1 V2	-	V2 RT1	-
FR	-	ME1 V2	-	RT1	-
Rp	-	ME1 V1	CE1 V2	V1 RT1	-
R-FR	-	ME1	-	RT1	-
Experiment 4					
L	-	ME1 V2	-	V2 RT1	
D	-	ME1 V2	CE2 V2	V2 RT1	
R	-	ME1 V2	-	V2 RT1	
FR	-	ME2 V2	-	V2 RT1	
Rp	-	ME1 V1	CE1 V2	V1 RT1	
R-FR	-	ME1	CE1 V2	V1 RT1	
FRp	-	ME1	CE1 V2	V1 RT1	
FR-R	-	ME1 V2	CE1 V2	V2 RT1	
Experiment 5					
L	-	ME1 V2	-	V2 RT1	
D	-	ME1 V2	CE2 V2	V2 RT1	
R	-	ME1	-	V2 RT1	
FR	-	ME2 V2	-	V2 RT1	
Rp	-	ME1 V1	-	V1 RT1	
R-FR	-	ME1	CE1 V1	V1 RT1	
FRp	-	ME1	CE1 V2	V1 RT1	
FR-R	-	ME1 V2	CE1 V2	RT1	

ET5491					
Experiment 6					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME1 V2	-	V2 RT1	-
D	-	ME1 V2	CE2 V2	V2 RT1	-
R	-	ME1 V2	-	V2 RT1	-
FR	-	ME2 V2	-	V2 RT1	-
Rp	-	ME1 V1	CE1 V2	V1 RT1	-
R-FR	-	ME1	CE1 V2	V1 RT1	-
FRp	-	ME1	CE1 V2	V1 RT1	-
FR-R	-	ME1 V2	CE1 V2	V2 RT1	-

ET5529					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	CE3 V3 RH1	-
D	-	-	CE2 V2	CE3 V3 RH1	-
LD	-	-	-	CE3 V3 RH3	-
DL	-	-	CE2 V2	CE3 V3 RH1	-
Experiment 2					
L	SA3 V3	ME3 V3	CE3	CE2 RT2 RH1 LP2	ME2 T2
D	SA2 V3	ME2 V3	CE3 V3	CE2 RH1	ME1
R	SA2 V3	ME2 V3	CE3 V3	CE1	ME1 T1
FR	SA3 V3	ME2 V3	CE3 V3	CE2 RH1	ME1
Experiment 3					
L	SA1 S2	ME1 V2	CE2 V3	CE1 RH1 LP2	T2
D	SA2	ME2 V3	CE2 V3	CE1 RT2 RH1	-
R	SA2 S3	ME2 V3	CE2 V3	CE1 RT2 RH2	ME1 T2
FR	SA2 S3	ME1 V3	CE2 V3	CE1 RH1	ME1 T2

ET5529					
Experiment 4					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA2 S2	ME1 V2	CE2 V3	CE1 RH1 LP2	-
D	SA2	ME2 V3	CE2 V3	CE1 RT2 RH1	-
R	SA2 S3	ME2 V3	CE2 V3	CE1 RT2	-
FR	SA2 S3	ME1 V2	CE2 V3	CE1 RH1	-
Rp	NO DATA				
R-FR	SA2	ME2 V3	CE2 V3	CE2 RT2 LP2 RH1	-
FRp	SA2	ME2 V3	CE2 V3	CE2 RT2 LP2 RH1	-
FR-R	SA2	ME2 V3	CE2 V3	CE2 RT2 LP2 RH1	-
Experiment 5					
L	SA2 S2	ME1 V2	CE2 V3	CE2 RT1 RH1	-
D	SA2	ME2 V3	CE2 V3	CE2 RT2 RH2	-
R	SA2	ME2 V3	CE2 V3	CE2 RT2 RH2 LP2	-
FR	SA2	ME1 V2	CE2 V3	CE1 RH1	-
Rp	SA2	ME2 V3	CE2 V3	CE2 RT2 LP2 RH2	-
R-FR	SA2	ME2 V3	CE2 V3	CE2 RT2 LP2 RH2	-
FRp	SA2	ME2 V3	CE2 V3	CE2 RT2 LP2 RH2	-
FR-R	SA2	ME2 V3	CE2 V3	CE2 RT2 LP2 RH1	-
Experiment 6					
L	SA1 S2	ME2 V2	CE2 V3	CE2 RT2 RH1	-
D	SA2	ME2 V3	CE2 V3	CE2 RT2 RH2	-
R	SA2	ME2 V3	CE2 V3	CE2 RH2	-
FR	SA3	ME2 V2	CE2 V3	CE2 RT2 LP2 RH2	-
Rp	SA2			CE2 RT2 LP2 RH2	
R-FR	SA1	ME2 V3	CE2 V3	CE2 RT2 LP2 RH2	-
FRp	SA2	ME2 V3	CE2 V3	CE2 RT2 LP2 RH2	-
FR-R	SA2	ME2 V3	CE2 V3	CE2 RT2 LP2 RH2	-

ET5555					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	CE2 C2 RH2	-
D	-	-	CE1 V1	CE2 V3 C3 RH2 LP3	-
LD	-	-	-	CE2 V2 C2 RH2 LP2	-
DL	-	-	-	CE1 V2 C2 RH2 LP2	-
Experiment 2					
L	-	-	-	CE2 C3 RH2	-
D	-	-	CE1 V1	CE2 V2 C3 RH2 LP2	-
R	-	-	-	CE2 V2 C3 RH2 LP3	-
FR	-	-	-	CE1 V1 RH2 LP2	-
Experiment 3					
L	-	-	-	CE2 V2 C3 RH2	-
D	-	-	CE1 V1	CE2 V2 C3 RH2 LP3	-
R	-	-	-	CE2 V2 C3 RH2 LP2	-
FR	-	-	-	CE2 V2 C3 RH2 LP3	-
R	-	-	-	CE2 V2 C3 RH2 LP3	
Rp	-	-	-	CE2 V2 C3 RH2 LP3	
R-FR	-	-	-	CE2 V2 C3 RH2 LP3	
FRp	-	-	-	CE2 V2 C3 RH2 LP3	
FR-R	-	-	-	CE2 V2 C3 RH2 LP3	-

ET5580					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	S2	ME1	-	-	-
D	-	-	-	-	-
LD	S2	ME1 P1	-	-	-
DL	S2	ME1 P1	-	-	-
Experiment 2					
L	S2	ME1	-	-	-
D	-	-	-	-	-
R	S2	ME1	-	-	-
FR	-	-	-	-	-

ET5580					
Experiment 3					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	S2	ME1	-	-	-
D	-	-	-	-	-
R	S2	ME1	-	-	-
FR	-	-	-	-	-
Rp	-	-	-	-	-
R-FR	-	-	-	-	-
FRp	-	-	-	-	-
FR-R	-	-	-	-	-

ET5599					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	CE1	-
D	-	V1	CE2 V2	V2 C3	-
LD	-	-	-	V2 RT2	-
DL	-	V1	CE2 V2	RT2	-
Experiment 2					
L	-	ME1 P1	-	CE1 C1	-
D	-	ME1	CE1 V1	CE1 V2 C3 LP3	-
R	S2	ME1 V2	CE1 V3	CE2 V2 C3 LP3	-
FR	-	-	-	CE1 V2 C2	-
Experiment 3					
L	S1	-	-	CE1 C1 V2	-
D	-	ME1	CE1 V1	CE1 V2 C3 LP3	-
R	S2	ME1 V1	CE1 V1	CE1 V2 C3 LP3	-
FR	S2	-	-	CE1 V2 C2	-
Rp	-	ME1	CE1 V1	CE1 V2 C2 LP3	-
R-FR	-	-	CE1 V1	CE1 V1 C2 LP2	-

ET5599					
Experiment 4					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	S2	-	-	CE1 C1 V2	-
D	-	ME1	CE1 V1	CE1 C3 LP3	-
R	S2	ME1 V1	CE1 V1	CE1 C3 LP3	-
FR	-	-	-	CE1 V2 C2	-
Rp	-	ME1	CE1 V1	CE1 V2 C2 LP3	-
R-FR	-	-	CE1 V1	CE1 V1 C2 LP2	-
FRp	-	-	CE1 V2	CE1 V3 C3 LP3	-
FR-R	-	-	CE1 V2	CE1 V3 C3 LP3	-

ET5627					
Experiment 1					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	-	M3 P3 V3	CE2	-	-
D	-	-	CE3 V3	V3	-
LD	-	M3 V3 P3	CE3 V3	CE1	-
DL	-	ME1	CE3 V3	V3	-
Experiment 2					
L	-	ME3 P3 V3	CE2	-	-
D	-	-	CE2 V3	CE1 V3	-
R	-	ME2 V3	CE2 V3	CE1	-
FR	-	ME2 V3	CE2 V3	CE1 V3	-
Experiment 3					
L	-	ME3 P3 V3	CE2	CE1	-
D	-	-	CE2 V3	CE	-
R	-	ME1 V2 P1	CE2 V3	CE1	ME2
FR	-	ME1 V2 P1	CE2 V2	CE1	-
Rp	-	-	CE2 V3	CE1 RH2	-
R-FR	-	-	CE2 V3	CE1	-

ET5627					
Experiment 4					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME3 P3 V3	CE2	CE1 RH1	-
D	-	-	CE2 V3	CE1	-
R	-	ME1 V2 P1	CE2 V3	CE1	-
FR	-	ME1 V2 P1	CE1 V2	CE1	-
Rp	-	ME1 V1	CE2 V3	CE1 RH2	-
R-FR	-	-	CE2 V3	CE1	-
FRp	-	-	CE1 V3	CE1 V1	-
FR-R	-	-	CE2 V3	CE1 V1	-
Experiment 5					
L	-	ME3 P3 V3	CE2	CE1 RH1	-
D	-	-	CE2 V3	CE1	-
R	-	ME1 V2 P1	CE2 V3	CE1	-
Fr	-	ME1 V1 P1	-	CE1	-
Rp	-	-	CE2 V3	CE1 RH2	-
R-Fr	-	-	CE2 V3	CE1	-
Frp	-	-	CE1 V3	CE1 V2	-
Fr-r	-	-	CE2 V3	CE1 V2	-
Experiment 6					
L	-	ME3 P3 V3	-	CE1 RH1	-
D	-	-	CE2 V3	CE1	-
R	-	ME1 V2 P1	CE2 V3	CE1	-
Fr	-	-	-	CE1	-
Rp	-	-	CE2 V3	CE1 RH2	-
R-Fr	-	-	CE2 V3	CE1	-
Frp	-	-	CE1 V3	CE1 V2	-
Fr-r	-	-	CE2 V3	CE1 V2	-
Experiment 7					
L	-	ME2 P2	-	CE1 RH1	-
D	-	-	CE2 V3	CE1	-
R	-	ME1 V2 P1	CE2 V3	CE1	-
Fr	-	-	-	-	-
Rp	-	-	CE2 V3	CE1 RH2	-
R-Fr	-	-	CE2 V3	CE1	-
Frp	-	-	CE1 V3	CE1 V2	-
Fr-r	-	-	CE2 V3	CE1 V2	-

ET5627					
Experiment 8					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	ME3 P3 V3	-	CE1	-
D	-	-	CE2 V3	CE1	-
R	-	ME1 V2 P1	CE2 V3	CE1	-
Fr	-	-	-	-	-
Rp	-	-	CE2 V3	CE1	-
R-Fr	-	-	CE2 V3	CE1	-
Frp	-	-	CE1 V3	CE1 V1	-
Fr-r	-	ME1	CE2 V3	CE1 V1	-

ET5642					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	SA3 V3	ME3 P3	CE3 V3	CE2 RH3	ME3
D	SA3 V3	ME3 V3	CE2V3	CE1 RH2	-
LD	SA3 V3	ME3 P3	CE3 V3	CE1 RH3	ME3
DL	SA3 V3	ME3 P3	CE2 V3	CE2 RH2	-
Experiment 2					
L	SA3 V3	ME3 P3	CE3 V3	CE2 RH3	ME3
D	-	ME3 V3	-	-	-
R	SA3 V3	ME3 P3	CE2 V2	RH3	-
Fr	SA3 V3	ME3 P3	CE2	RH3	-
Rp	SA3 V3	ME3 P3	CE2	RH2	-
R-Fr	SA2 V2	ME2 P2	CE2	RH2	-
Frp	SA2 V2	ME2 P2	CE2	RH2	-
Fr-r	SA2 V2	ME2 P2	CE2	RH2	-

ET5653					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	S3	ME3 V3 P3	CE1 V1	CE1 RH2	ME2 T1
D	-	ME3 V3 P3	CE3 V3	CE2 RH2	-
LD	S3	ME3 V3 P3	CE3 V3	CE2 RH2	ME3 T1
DL	-	ME3 V3 P3	CE3 V3	CE2 RH2	-
Experiment 2					
L	S3	ME3 V3	CE2	CE1	ME1 T1
D	-	ME3 V3 P3	CE2 V3	CE2	
R	S3	ME3 V3	CE3 V3	CE2	ME1 V2 T2
FR	-	ME3 V3	CE1 V2	-	-

ET6375					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	SA2	-	-	-	ME2
D	-	ME1 V1	CE1 V2	-	ME1
LD	SA1	-	CE1 V2	-	ME1
DL	SA2	ME1 V1	CE1 V2	-	ME2
Experiment 2					
L	S3	ME3 V3 P2	-	CE1 RH2 C2	-
D	S3	ME3 V3	CE2 V3	C2	-
R	S3	ME3 V3	CE1 V2	C2	-
Fr	S3	ME3 V3	CE1	-	-
Experiment 3					
L	S3	ME2 V3 P2	-	CE1 RH2 C2	-
D	-	ME3 V3	CE1 V1	CE1 RH2 C2	-
R	S3	ME3 V3	CE1	C2	-
Fr	S3	ME3 V3	CE1	-	-
Rp	-	ME2 V3	CE1 V2	C2	-
R-Fr	-	ME2	CE2	C2	-

ET6375					
Experiment 4					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	S3	ME2 V3 P2	-	CE1 RH2 C2	-
D	-	ME3 V3	CE2 V3	CE1 RH2 C2	-
R	S3	ME3 V3	CE2 V3	C2	-
Fr	S3	ME3 V3	CE1	-	-
Rp	-	ME2 V3	CE2 V3	C2	-
R-Fr	-	ME2	CE2 V3	C2	-
Frp	-	ME2	CE2 V3	C2 RH2	-
Fr-r	-	ME2 V3	CE2 V3	C2	-
Experiment 5					
L	S3	ME2 V3 P2	-	CE1 RH2 C2	-
D	-	ME3 V3	CE2 V3	CE1 RH2 C2	-
R	S3	ME3 V3	CE2 V3	C2	-
Fr	S3	ME3 V3	CE1	-	-
Rp	-	ME2 V3	CE2 V3	C1	-
R-Fr	-	ME2	CE2 V3	C1	-
Frp	-	ME2	CE2 V3	C2 RH2	-
Fr-r	-	ME2 V3	CE2 V3	C2	-

ET6417					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	H1	-	CE1	-
D	-	-	-	CE2	-
R	-	H2	-	CE2	-
Fr	-	-	-	-	-
Experiment 2					
L	-	H1	-	CE1	T1
D	-	-	-	CE2	-
R	-	H3	-	-	-
Fr	-	-	-	-	-
Experiment 3					
L	-	H2	-	-	H2
D	-	-	-	CE1	-
R	-	H2	-	-	-
Fr	-	-	-	-	-

ET6417					
Experiment 4					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	H1	-	CE1	-
D	-	-	-	CE2 RH2	-
R	-	-	-	CE2 RH1	-
Fr	-	-	-	CE1	-
Rp	-	H2	-	CE2 RH1	-
R-Fr	-	-	-	CE2 RH1	-
Frp	-	-	-	CE2 RH1	-
Fr-r	-	-	-	CE1 RH1	-

ET6426					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	V2	V2	-
D	-	-	CE2 V3	V3	-
LD	-	-	V2	V2	-
DL	-	-	CE1 V2	-	-
Experiment 2					
L	-	-	-	-	-
D	V2	-	V2	-	-
R	V3	-	V3	-	-
Fr	-	-	-	-	-
Experiment 3					
L	-	-	V2	-	-
D	V2	-	V2	-	-
R	V2	-	V2	-	-
Fr	-	-	V2	-	-

ET6428					
Experiment 1					
treatment	Shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	-	CE2 V3	-	-
LD	-	-	-	-	-
DL	-	-	CE1 V2	-	-
Experiment 2					
L	-	-	-	-	-
D	-	-	CE1 V2	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Experiment 3					
L	-	-	-	-	-
D	-	-	CE1	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Experiment 4					
L	-	-	-	-	-
D	-	-	CE1 V2	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Rp	-	-	CE1 V2	-	-
R-Fr	-	-	CE1 V2	-	-
Frp	-	-	CE1 V2	-	-
Fr-r	-	-	CE1 V2	-	-
Experiment 5					
L	-	-	-	-	-
D	-	-	CE1 V1	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Rp	-	-	CE1 V2	-	-
R-Fr	-	-	CE1 V2	-	-
Frp	-	-	CE1 V2	-	-
Fr-r	-	-	CE1 V2	-	-

ET6428					
Experiment 6					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	-	CE1 V2	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Rp	-	-	CE1 V1	-	-
R-Fr	-	-	CE1 V1	-	-
Frp	-	-	CE1 V1	-	-
Fr-r	-	-	CE1 V2	-	-

ET6435					
Experiment 1					
treatment	Shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	SA1 V1	ME3 V3 P2	V2	-	-
D	-	-	-	-	-
LD	SA1 V1	ME2 V2 P1	V2	-	-
DL	SA1 V1	ME1	V2	-	-
Experiment 2					
L	-	ME3 V3	V1	-	-
D	-	-	-	-	-
R	SA3	ME3 V3	V3	-	-
Fr	-	ME2	-	-	-
Experiment 3					
L	-	ME2	-	-	-
D	-	-	-	-	-
R	SA 2 V3	ME2 V3	CE1 V3	-	-
Fr	SA2 V3	ME1 V2	-	-	-
Experiment 4					
L	-	ME2 V3	-	-	ME2 V2 T2
D	No data		-	-	-
R	-	ME2 V3 P1	CE1 V1	-	ME2 V2 T2
Fr	-	ME2 V3	CE1 V2	-	-
Rp	-	ME1	-	-	-
R-Fr	-	-	CE1	-	-

ET6435					
Experiment 5					
treatment	Shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	ME2 V3	-	-	-
D	-	-	-	-	-
R	-	ME2 P1	CE1 V1	-	-
Fr	-	ME1 V2	CE1 V1	-	-
Rp	-	ME1	V1	-	-
R-Fr	-	-	-	-	-
Frp	-	-	-	-	-
Fr-r	-	ME2	V1	-	-
Experiment 6					
L	-	ME2 V3	-	-	-
D	-	-	-	-	-
R	-	ME2 P1	CE1 V1	-	-
Fr	-	ME1 V2	CE1 V1	-	-
Rp	-	ME1	V1	-	-
R-Fr	-	-	-	-	-
Frp	-	-	-	-	-
Fr-r	-	ME2	V1	-	-

ET6537					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	ME2	CE2 V2	-	-
LD	-	ME1	-	-	-
DL	-	ME1	-	-	-
Experiment 2					
L	-	-	-	-	-
D	-	ME1	-	-	-
R	-	ME1	-	-	-
Fr	-	ME1	-	-	-

ET6537					
Experiment 3					
Treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	ME1	-	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Rp	-	ME1	-	-	-
R-Fr	-	ME1	-	-	-
Frp	-	ME1	-	-	-
Fr-r	-	ME1	-	-	-

ET6561					
Experiment 1					
Treatment	Shoot apex	Cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	ME2 V2	CE1	-	-
LD	-	ME1 V2	CE1	-	-
DL	-	ME1	CE1 V1	-	-
Experiment 2					
L	-	-	-	-	-
D	-	-	CE1 V1	CE1	-
R	-	-	CE1 V1	CE1	-
Fr	-	-	CE1 V1	-	-
Experiment 3					
L	-	-	-	-	-
D	-	-	CE1	-	-
R	-	-	CE1	-	-
Fr	-	-	-	-	-
Rp	-	-	CE1	-	-
R-Fr	-	-	CE1	-	-
Frp	-	-	CE1	-	-
Fr-r	-	-	CE1	-	-

ET6561					
Experiment 4					
treatment	Shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	ME1 V1	CE1	-	-
R	-	ME1 V1	CE1	-	-
Fr	-	ME1 V1	-	-	-
Rp	-	-	CE1	-	-
R-Fr	-	-	CE1	-	-
Frp	-	-	CE1	-	-
Fr-r	-	-	CE1	-	-

ET6566					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	CE1	-	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Experiment 2					
L	-	-	-	-	-
D	-	CE1	-	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Experiment 3					
L	-	-	-	-	--
D	-	CE1	-	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Rp	-	-	-	-	-
R-Fr	-	-	-	-	-
Frp	-	-	-	-	-
Fr-r	-	-	-	-	-

ET6611					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	ME1 V2	-	-	-
LD	-	-	-	-	-
DL	-	-	-	-	-
Experiment 2					
L	-	-	-	-	-
D	-	ME1	-	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Experiment 3					
L	-	-	-	-	-
D	-	ME1	-	-	-
R	-	ME1	-	-	-
Fr	-	-	-	-	-
Rp	-	-	-	-	-
R-Fr	-	-	-	-	-
Frp	-	-	-	-	-
Fr-r	-	-	-	-	-
Experiment 4					
L	-	-	-	-	-
D	-	ME2	-	-	-
R	-	-	-	-	-
Fr	-	-	-	-	-
Rp	-	-	-	-	-
R-Fr	-	ME1	-	-	-
Frp	-	ME1	-	-	-
Fr-r	-	-	-	-	-

ET6633					
Experiment 1					
treatment	Shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	SA3	H2 P3	CE3	CE1	-
D	SA3	ME2 V2 P3	CE3 V3	CE2 RH1	ME2
LD	SA3	ME1 P3	CE3 V2	CE1 RH1	ME2
DL	SA3	ME1 V2 H1 P3	CE3 V3	CE1 RH1	ME2

ET6633					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	CE2 V3	-	-
D	-	-	CE2 V3	-	-
R	-	-	CE2 V3	-	-
Fr	-	-	CE1	-	-
Experiment 3					
L	-	-	CE2 V3	-	-
D	-	-	CE2 V3	-	-
R	-	-	CE2 V3	-	-
Fr	-	-	CE1	-	-
Experiment 4					
L	-	-	CE3 V3	-	-
D	-	ME3	CE3 V3	CE2 RH2	-
R	No data				
Fr	-	ME1	-	-	-
Rp	-	ME1	CE3 V3	-	-
R-Fr	No data				
Experiment 5					
L	-	-	CE2	-	--
D	-	-	CE2 V3	-	-
R	-	-	CE2	-	-
Fr	-	-	CE1	-	-
Rp	-	-	CE2 V3	-	-
R-Fr	-	-	CE2 V3	-	-
Frp	-	-	CE2 V3	-	-
Fr-r	-	-	CE2 V3	-	-

ET6649					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	ME2	CE2	-	-
D	-	ME2 V3	CE2 V3	CE2 RH2	-
LD	-	ME2 V3	CE2 V2	-	-
DL	-	ME2 V3	CE1 V2	-	-

ET6649					
Experiment 2					
treatment	Shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	SA2 V2	ME2 V2	CE2 V3	CE1	-
D	SA2 V2	ME1	CE2 V3	V2 RH1	-
R	SA2	ME1 V2	CE2 V2	V2	-
Fr	SA3	ME1 V1	CE1 V2	V2	-
Experiment 3					
L	SA2 V2	ME2 V2	CE2	-	-
D	SA2 V2	ME1 V2	CE2 V3	RH1	-
R	SA2	ME2 V3	CE2 V2	CE1 RH1	-
Fr	SA3	ME1 V1	CE1 V1	RH1	-
Experiment 4					
L	SA2 V2	ME2 V3	CE2	-	-
D	SA2 V2	ME1 V3	CE2 V3	RH1	-
R	SA2	ME1 V3	CE2 V3	CE1 RH1	-
Fr	SA3	ME1 V2	CE1 V1	RH1	-
Experiment 5					
L	SA2	ME3	CE2 V3	CE1	-
D	SA2 V2	ME3 V3	CE2 V3	CE1 RH1	-
R	SA2	ME3 V3	CE3 V3	CE1 RH1	-
Fr	SA2	ME3 V3	CE3 V3	RH1	-
Experiment 6					
L	SA2	ME2 V2	CE2 V2	CE1	-
D	SA2 V2	ME1 V2	CE2 V3	CE1 RH1	-
R	SA2	ME1 V2	CE2 V2	CE1	-
Fr	SA2	ME1 V1	CE1 V2	-	-
Rp	SA2	ME1 V2	CE2 V3	CE1	-
R-Fr	SA2	ME1 V2	CE2 V3	CE1 RH1	-
Frp	SA2	ME1 V2	CE2 V3	CE1 RH1	-
Fr-r	SA2	ME1 V2	CE2 V3	RH1	-
Experiment 7					
L	SA1	ME1 V1	CE1 V1	-	-
D	SA2 V2	ME1 V2	CE2 V3	-	-
R	SA1	ME1 V2	CE2 V3	CE1	-
Fr	No data				
Rp	SA2	ME1 V2	CE2 V3	CE1	-
R-Fr	SA2	ME1 V2	CE2 V3	CE1 RH1	-
Frp	SA2	ME1 V2	CE2 V3	CE1 RH1	-
Fr-r	SA2	ME1 V2	CE2 V3	RH1	-

GT5874					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	ME2 V2 P2	-	CE1	-
D	-	ME2 V3 P3	CE2	-	-
R	-	ME3 V3 P3	CE2	CE1	-
FR	-	ME3 V3 P2	CE2	CE1	-
Experiment 2					
L	S2	ME2 V2 P2	-	-	H2
D	-	ME2 V3 P3	CE2	-	-
R	S3 SA2	ME2 V3 P3	CE1	-	-
FR	S3	ME3 V3 P2	-	-	-
Experiment 3					
L	S2	ME2 V2 P2	-	-	H2
D	-	ME3 V3 P3	CE2	-	-
R	S3	ME2 V3 P3	CE1	-	-
FR	S3	ME3 V3 P2	-	-	-
Rp	-	ME2 V3 P2	CE1 V1	-	-
R-FR	-	ME2 V3 P2	CE1 V2	-	-
FRp	-	ME2 V3 P2	CE1	-	-
FR-R	-	ME2 V3 P2	CE2	-	-
Experiment 4					
L	-	ME2 V2 P2	-	-	H2
D	-	ME2 V3 P2	CE2	-	-
R	S2	ME2 V3 P2	-	-	-
FR	S3	ME3 V3 P2	-	-	-
Rp	-	ME2 V3 P2	CE1 V1	-	-
R-FR	-	ME2 V3 P2	CE1 V2	-	-
FRp	-	ME2 V2 P2	CE1	-	-
FR-R	-	ME2 V3 P2	CE1	-	-
Experiment 5					
L	S2	ME2 V2 P2	-	-	H2
D	-	ME3 V3 P3	CE2	-	-
R	S3	ME2 V3 P3	CE1	-	-
FR	S3	ME3 V3 P2	-	-	-
Rp	-	ME2 V3 P2	CE1 V1	-	-
R-FR	-	ME2 V3 P2	CE1 V2	-	-
FRp	-	ME2 V3 P2	CE1	-	-
FR-R	-	ME2 V3 P2	CE2	-	-

GT5874					
Experiment 6					
treatment	shoot apex	Cotyledon	Hypocotyl	Root	immature leaf
L	S2	ME2 V2	-	-	H2
D	-	ME3 V3 P3	CE2	-	-
R	S3	ME2 V3 P3	CE1	-	-
FR	S3	ME3 V3 P2	-	-	-
Rp	-	ME2 V3 P2	CE1 V1	-	-
R-FR	-	ME2 V3 P2	CE1 V2	-	-
FRp	-	ME2 V3 P2	CE1	-	-
FR-R	-	ME2 V3 P2	CE2	-	-

GT5905					
Experiment 1					
treatment	Shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	ME2 V2	CE1	V3 LP3	-
LD	-	ME1 V2	-	LP2	-
DL	-	ME1 V2	-	-	-
Experiment 2					
L	-	-	-	-	-
D	-	ME1 V1	CE1	-	-
R	-	ME1	-	-	-
FR	-	ME1 V1	-	-	-
Experiment 3					
L	-	-	-	-	-
D	-	ME1 V1	CE1	-	-
R	S1	ME1	-	-	-
FR	-	ME1	-	-	-
Rp	-	ME1	CE1	LP2	-
R-FR	-	ME1	CE1	LP2	-
FRp	-	ME1	CE1	LP2	-
FR-R	-	ME1	CE1	LP2	-

GT5909					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME1 V1	CE3	CE1 RH1	-
D	-	ME3 V3	CE1 V2	CE1 RH1	-
LD	-	ME3 V3	CE3 V2	-	-
DL	-	ME2 V2	CE1	-	-
Experiment 2					
L	-	ME2 V3	CE2 V2	-	-
D	-	ME3 V3	CE2 V3	-	-
R	-	ME3 V3	CE2 V3	-	-
FR	-	ME2 V3	CE2 V2	-	-
Experiment 3					
L	-	ME1 V1	CE2 V2	-	-
D	-	ME3 V3	CE1 V2	-	-
R	-	ME3 V3	CE2 V3	-	-
FR	-	ME3 V3	CE2 V3	-	-
Experiment 4					
L	-	E1 V1	-	-	-
D	-	ME3 V3	CE1	-	-
R	-	ME3 V3	CE1	-	-
FR	-	ME2 V3	CE1	-	-
Rp	-	ME3 V3	CE1	-	-
R-FR	-	ME3 V3	CE1	-	-
Experiment 5					
L	-	ME1	-	-	-
D	-	ME3 V3	CE1	-	-
R	-	ME3 V3	CE1	-	-
FR	-	ME2 V3	CE1	-	-
Rp	-	ME3 V3	CE1	-	-
R-FR	-	ME3 V3	CE1	-	-
FRp	-	ME3 V3	CE1	-	-
FR-R	-	ME3 V3	CE1	-	-

GT5914					
Experiment 1					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	S3	-	-	CE1 V2 C2 RH2	-
D	-	-	V2	C3	-
LD	S3	-	-	CE1 V2 C3 RH2	-
DL	-	-	V2	CE1 V2 C3	-
Experiment 2					
L	S3	-	-	-	-
D	-	-	V2	CE3 C3 RH2	-
R	S2	-	V2	CE2 C3 RH1	-
FR	-	-	-	CE1 C1	-
Experiment 3					
L	S3	-	-	-	-
D	-	-	-	CE2 C3	-
R	S3	-	-	CE1 C3 RH2	-
FR	-	-	-	CE1 C2	-
Experiment 4					
L	S3	-	-	-	-
D	-	-	-	V2 C3	-
R	S3	-	-	-	-
FR	S3	-	-	-	-
Rp	-	-	-	V2 C2	-
R-FR	-	-	-	V1 C2	-
FRp	-	-	-	V1 C2	-
FR-R	-	-	-	V1 C2	-

GT5927					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA2	V2 H2	-	V3 C3 LP3	ME2
D	SA3	ME3 V3 H2	CE1 V3	CE3 V3 C3 RH1	ME2
LD	SA2	V2	-	V2 C3	ME2
DL		V3	CE1V3	V3 C3	-

GT5927					
Experiment 2					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	S3 SA2	V2 H2	-	CE2 V2 C3	ME2
D	SA2	ME2 V2 H2	CE1 V3	CE2 V2 C3	ME2
R	S3 SA3	V2 H2	CE1 V2	CE1 V3 C3 RH1	ME2
FR	SA1	ME1 V2 H2	-	CE2 V3 C3 RH2	-
Experiment 3					
L	S3 SA2	ME1 V2 H2	-	CE3 V3 C3	ME2
D	SA2	ME2 V2 H2	CE1 V3	CE2 V3 C3	ME2
R	S3 SA2	V2 H2	CE1 V2	CE1 V3 C3	ME2
FR	S3 SA3	ME1 V2 H2	-	CE2 V2 C2	-
Experiment 4					
L	S3 SA2	V2 H2	-	CE2 V2 C3 LP3	ME2
D	SA2	ME2 V2 H2	CE1 V3	CE2 V2 C3 LP3	ME2
R	S3 SA3	V2 H2	CE1 V2	CE1 V3 C3 LP3 RH1	ME2
FR	S3 SA3	ME1 V2 H2	-	CE2 V3 C3 LP3	-
Rp	SA2	ME1	CE1 V3	V3 C3 LP3	-
Experiment 5					
L	S3 SA2	V2 H2	-	CE2 V3 C3 LP3	ME
D	SA3	ME2 V2 H2	CE1 V3	CE2 V3 C3 LP3	ME2
R	S3 SA3	V2 H2	CE2 V2	CE1 V3 C3 RH1 LP3	ME2
FR	S3 SA3	ME1 V2 H2	-	CE2 V2 C2 LP3	-
Rp	SA2	ME1	CE2 V3	V3 C3 LP3	ME2
R-FR	SA2	ME1	CE2 V3	V3 C3 LP3	-
FRp	SA2	ME1	CE2 V3	V3 C3 LP3	-
FR-R	SA2	ME1	CE2 V3	V3 C3 LP3	-
Experiment 6					
L	S3 SA2	V2 H2	-	CE2 V2 C3 LP3	ME2
D	SA3	ME2 V2 H2	CE1 V3	CE2 V3 C3 LP3	ME2
R	S3 SA3	V2 H2	CE2 V2	CE1 V3 C3 LP3 RH1	ME2
FR	SA1	ME1 V2 H2	-	CE2 V2 C2 LP3	-
Rp	SA2	ME1	CE2 V3	V3 C3 LRP3	-
R-FR	SA2	ME1	CE2 V3	V3 C3 LP3	-
FRp	SA2	ME1	CE2 V3	V3 C3 LRP3	-
FR-R	SA2	ME1	CE2 V3	V3 C3 LP3	-

GT5929					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	Immature leaf
L	SA3	-	-	-	-
D	-	V1	-	-	-
LD	SA1	ME1 V1	-	-	-
DL	SA1	ME1 V2	-	-	-
Experiment 2					
L	S3	-	-	-	ME3 T3
D	S3 SA3	-	CE2 V3	-	ME3
R	S3 SA3	-	CE2 V3	-	ME3
FR	S3	-	CE1	-	ME3 T3

GT5939					
Experiment 1					
treatment	Shoot apex	cotyledon	Hypocotyl	Root	Immature leaf
L	-	ME2 V2	-	C1	-
D	-	ME2 V2	CE1 V2	-	-
LD	-	ME2 V2	CE1	C1	-
DL	-	ME2 V2	CE1 V2	-	-
Experiment 2					
L	-	ME2 V2	-	RH1	-
D	-	ME2 V2	CE1 V2	-	-
R	-	ME2 V2	-	-	-
FR	-	ME1 V2	-	-	-
Experiment 3					
L	-	ME2 V2	-	RH1	-
D	-	ME2 V2	CE1 V1	-	-
R	-	ME2 V2	-	-	-
FR	-	ME1 V1	-	-	-
Experiment 4					
L	-	ME1	-	-	-
D	-	ME2 V2	CE1	-	-
R	-	ME2 V2	-	C1	-
FR	-	ME1	-	-	-
Rp	-	ME2 V2	CE1		-
R-FR	-	ME2 V2	CE1	-	-
FRp	-	ME2 V2	CE1	-	-
FR-R	-	ME2 V2	CE1	-	-

GT5939					
Experiment 5					
treatment	shoot apex	cotyledon	hypocotyl	Root	Immature leaf
L	-	ME1	-	-	-
D	-	ME2 V2	CE1	-	-
R	-	ME1 V2	CE1	-	-
FR	-	ME1	-	-	-
Rp	-	ME1 V2	CE1	-	-
R-FR	-	ME1 V2	-	-	-
FRp	-	ME2 V2	CE1	-	-
FR-R	-	ME1 V2	CE1	-	-
R-FR	-	ME2 V2	CE1	-	-

GT5957					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	Immature leaf
L	-	-	-	-	-
D	-	ME2	-	-	-
LD	-	-	-	-	-
DL	-	ME1	-	-	-
Experiment 2					
L	-	-	-	-	-
D	-	ME2	-	-	-
R	-	-	-	-	-
FR	-	ME1	-	-	-

GT5964					
Experiment 1					
treatment	Shoot apex	Cotyledon	hypocotyl	Root	Immature leaf
L	SA3	ME2 V3	-	C1	-
D	SA2	ME3 V3	CE1 V2	-	-
LD	SA1	ME3 V3	CE1	C1	-
DL		ME3	CE1 V2	-	

GT5964					
Experiment 2					
treatment	Shoot apex	cotyledon	hypocotyl	Root	Immature leaf
L	SA3 V3	ME2 V3	-	-	-
D	SA3	ME2 V3	CE1 V1	-	-
R	SA3	ME2 V3	CE2 V2	-	-
FR	SA3	ME2 V3	CE2 V2	-	-
Experiment 3					
L	SA3 V3	ME2 V3	CE2 V3	-	-
D	SA2	ME3 V3	CE1 V1	-	-
R	SA2	ME2 V3	CE1 V3	-	-
FR	SA2	ME3 V3	CE1 V3	-	-
Rp	SA2	ME3 V3	CE1 V2	-	-
R-FR	SA2	ME3 V3	CE1 V2	-	-
Experiment 4					
L	SA2	ME2 V3	CE1 V2	-	-
D	SA1	ME2 V3	CE1 V2	-	-
R	SA2	ME2 V3	CE1 V2	-	-
FR	SA2	ME1 V2	CE1	-	-
Rp	SA2	ME3 V3	CE1 V3	-	-
R-FR	SA2	ME3 V3	CE1 V3	-	-
FRp	SA2	ME2 V2	CE1 V3	-	-
FR-R	SA2	ME3 V3	CE1 V3	-	-
Experiment 5					
L	SA3	ME1 V2	CE1	-	-
D	SA2	ME2 V3	CE1 V2	-	-
R	SA2	ME2 V3	CE1 V2	-	-
FR	SA2	ME1 V2	CE1	-	-
Rp	SA3	ME3 V3	CE1 V2	-	-
R-FR	SA2	ME3 V3	CE1 V2	-	-
FRp	SA1	ME2 V2	-	-	-
FR-R	SA2	ME3 V3	CE1 V2	-	-

GT5964					
Experiment 6					
treatment	shoot apex	Cotyledon	hypocotyl	Root	Immature leaf
L	SA3	ME1 V2	CE1	-	-
D	SA1	ME2 V3	CE1 V3	-	-
R	SA2	ME2 V3	CE1 V2	-	-
FR	SA2	ME1 V2	CE1	-	-
Rp	SA3	ME3 V3	CE1 V2	-	-
R-FR	SA2	ME3 V3	CE1 V2	-	-
FRp	SA1	ME2 V2		-	-
FR-R	SA2	ME3 V3	CE1 V2	-	-

GT5971					
Experiment 1					
treatment	shoot apex	Cotyledon	hypocotyl	Root	Immature leaf
L	-	ME2 V3	CE2	RT2 RH1	-
D	-	ME3 V3	CE3 V3	CE3 RT2 RH1	-
LD	-	ME3 V3	CE2	RT2 RH2	-
DL	-	ME3 V3	CE2V3	RH1	-
Experiment 2					
L	SA2	ME2 V3	CE2 V3	CE1 RT2 RH1	-
D	SA3	ME3 V3	CE2 V3	RT1	-
R	SA3	ME3 V3	CE2 V3	CE1 RT2 RH1	-
FR	SA2	ME2 V2	CE2 V3	RT1 RH1	-
Experiment 3					
L	SA1	ME2 V3	CE1 V2	-	-
D	SA3	ME3 V3	CE2 V3	-	-
R	SA3	ME3 V3	CE2 V2	-	-
FR	SA2	ME2 V3	CE1 V2	-	-
Rp	SA3	ME3	CE2 V3	-	-
R-FR	No data			-	

GT5971					
Experiment 4					
treatment	shoot apex	Cotyledon	hypocotyl	Root	Immature leaf
L	SA1	ME1	CE1	-	-
D	SA3	ME3 V3	CE2 V3	-	-
R	SA3	ME2 V3	CE2 V3	-	-
FR	SA1	ME2 V3	CE1 V2	-	-
Rp	SA3	ME3 V3	CE2 V3	-	-
R-FR	SA3	ME2 V3	CE2 V3	-	-
FRp	SA3	ME2 V3	CE2 V3	-	-
FR-R	SA3	ME2 V3	CE2 V3	-	-
Experiment 5					
L	SA1	ME1	CE1	-	-
D	SA2	ME2 V3	CE2 V3	-	-
R	SA2	ME2 V3	CE2 V3	-	-
FR	SA2	ME1	-	-	-
Rp	SA2	ME2 V3	CE2 V3	-	-
R-FR	SA2	ME2 V3	CE2 V3	-	-
FRp	SA2	ME2 V3	CE2 V3	-	-
FR-R	SA2	ME2 V3	CE2 V3	-	-

GT6021					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA2	ME1 V3	CE2 V3	V2 RT2 C3	-
D	SA3	ME3 V3	CE2 V3	V3 LP3 RT2 C3	-
LD	SA3	ME2 V3	CE2 V3	V3 LP3 RT2 C3	-
DL	SA3	ME2 V3	CE2 V3	V2 LP3 RT2 C3	-
Experiment 2					
L	SA2	ME1 V3	-	-	-
D	SA3	ME3 V3	CE3 V3	V3 LP3 RT2 C3	-
R	SA3	ME2 V3	CE2 V3	V3 LP3 RT2 C3	-
FR	SA2	ME1 V3	CE2 V3	V2 LP3 RT1 C3	-

GT6021					
Experiment 3					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	SA3	ME3 V3	CE2 V3	C2	ME1 T2
D	SA3	ME3 V3	CE3 V3	CE2 LP3 RH2 C3	-
R	SA3	ME2 V3	CE2 V3	CE2 LP3 RH2 C3	ME1 T2
FR	SA2	ME1 V2	CE1 V2	CE2 LP2 RH2 C3	ME1 T2
Experiment 4					
L	SA2	ME1 V2	-	-	-
D	SA3	ME3 V3	CE3 V3	V3 LP3 RT2 C3	-
R	SA3	ME2 V3	CE2 V3	V3 LP3 RT2 C3	-
FR	SA2	ME1 V2	CE1 V2	V2 LP3 RT1 C3	-
Rp	SA3	ME3 V3	CE1 V3	V2 LP3 RT2 C3	-
R-FR	SA3	ME3 V3	CE2 V3	V2 LP3 RT2 C3	-
FRp	SA3	ME3 V3	CE2 V3	V2 LP3 RT2 C3	-
Experiment 5					
L	SA1	ME1	-	LP3 C3	-
D	SA3	ME3 V3	CE3 V3	V3 LP3 RT2 C3	-
R	SA3	ME2 V3	CE2 V3	V3 LP3 RT2 C3	-
FR	SA1	ME2 V2	CE2 V3	V2 LP3 RT1 C3	-
Rp	SA3	ME3 V3	CE2 V3	V2 LP3 RT1 C3	-
R-FR	SA3	ME3 V3	CE2 V3	V2 LP3 RT3 C3	-
FRp	SA3	ME3 V3	CE2 V3	V2 LP3 RT2 C3	-
FR-R	SA3	ME3 V3	CE2 V3	V2 LP3 RT2 C3	-
Experiment 6					
L	SA1	ME1	-	LP3	-
D	SA3	ME3 V3	CE3 V3	V3 LP3 RT2 C3	-
R	SA3	ME2 V3	CE2 V3	V3 LP3 RT2 C3	-
FR	No data				
Rp	SA3	ME3 V3	CE2 V3	V2 LP3 RT1 C3	-
R-FR	SA3	ME3 V3	CE2 V3	V2 LP3 RT3 C3	-
FRp	SA3	ME3 V3	CE2 V3	V2 LP3 RT3 C3	-
FR-R	SA3	ME3 V3	CE2 V3	V2 LP3 RT2 C3	-

GT6027					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME3 V3	-	RH1	T2
D	SA2	ME3 V3	-	CE2 RH1	ME1
LD	SA1	ME3 V3	CE1	CE2 RH1	ME1
DL	-	ME3 V3	CE1	-	ME1
Experiment 2					
L	S2	ME3 V3	CE1	RH2	ME1 T1
D	-	ME3 V3	-	CE1 RH2	-
R	S2	ME3 V3	CE1	CE2 RH2	
FR	-	ME3 V3	CE1	RH1	ME1

GT6039					
Experiment 1					
treatment	Shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	ME2	-	-	-
D	-	-	-	-	-
LD	-	-	-	-	-
DL	-	-	-	-	-
Experiment 2					
L	-	ME3	-	-	-
D	-	-	-	-	-
R	SA3	ME3 V3	CE2 V3	-	-
FR	-	-	-	-	-

GT6052					
Experiment 1					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	S3	-	-	-	-
D	-	-	-	-	-
LD	S3	-	-	-	-
DL	S3	-	-	-	-

GT6052					
Experiment 2					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	S3	-	-	-	-
D	S1	-	-	-	-
R	S3	-	-	-	-
FR	S3	-	-	-	-
Experiment 3					
L	S3	-	-	-	-
D	-	-	-	-	-
R	S3 SA1	-	-	-	-
FR	S3 SA1	-	-	-	-
Rp	-	-	-	-	-
R-FR	-	-	-	-	-
Experiment 4					
L	S3	-	-	-	-
D	-	-	-	-	-
R	S3 SA1	-	-	-	-
FR	S3	-	-	-	-
Rp	S2	-	-	-	-
R-FR	S3	-	-	-	-
FRp	S3	-	-	-	-
FR-R	-	-	-	-	-
Experiment 5					
L	S3	-	-	-	-
D	-	-	-	-	-
R	S3	-	-	-	-
FR	S3	-	-	-	-
Rp	-	-	-	-	-
Experiment 6					
L	S3	-	-	-	-
D	-	-	-	-	-
R	S3	-	-	-	-
FR	S3	-	-	-	-
Rp	-	-	-	-	-
R-FR	-	-	-	-	-
FRp	-	-	-	-	-
FR-R	-	-	-	-	-

GT6064					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	SA3	ME3	CE2 V3	-	-
LD	-	-	-	-	-
DL	-	-	-	-	-
Experiment 2					
L	S3	ME3 V3 P1	-	-	-
D	SA3	ME3 V3 P3	CE2	CE1	-
R	SA3 S3	ME3 V3 P3	CE2	CE1	
FR	SA2 S2	ME3 V3 P1	-	CE1	-
Experiment 3					
L	-	ME1	-	-	-
D	-	ME2 V3 P2	-	-	-
R	-	ME1	-	-	-
FR	-	ME1	-	-	-
Rp	-	ME1	-	-	-
R-FR	-	ME1	-	-	-
FRp	-	ME1	-	-	-
FR-R	-	ME1	-	-	-
Experiment 4					
L	-	ME1	-	-	-
D	SA2	ME3 V3 P3	CE2	CE1	-
R	SA3 S3	ME3 V3 P3	CE2	-	-
FR	SA3	ME3 V3 P3	CE1	-	-
Rp	SA2	ME3 V3 P3	CE2 V2	CE1	-
R-FR	No data				
FRp	No data				
FR-R	-	ME2 V2 P1	CE2 V3	V1	-

GT6067					
Experiment 1					
treatment	Shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	V3	ME1	CE1	CE1 C2	-
D	SA3	ME3 V3	CE2 V3	V3 C3	-
LD	SA3 V3	ME3 V3	CE2 V3	V3 C2	-
DL	SA3 V3	ME3 V3	CE3 V3	CE3 V3 C3	-
Experiment 2					
L	V2	ME1	CE1 V3	C2	E1 T1
D	SA3 V3	ME3	CE2 V2	V3 C2	-
R	SA3 V3	ME2	CE1	V3 C2	-
FR	SA3 V3	ME3	-	-	-
Rp	SA2 V3	ME3	CE1	V3 C3	-
R-FR	SA2 V2	ME3	CE1	V3 C3	-
FRp	SA3 V3	ME3 V3	CE1	V3 C3	-
FR-R	SA3 V3	ME3 V3	CE1 V2	V3 C3	-
Experiment 3					
L	SA3 V3	ME2	CE1	C2	-
D	SA3 V3	ME3	CE1 V1	C1	-
R	SA3 V3	ME2	-	-	-
FR	SA3 V3	-	-	-	-
Rp	SA3 V3	ME3 V3	CE1 V2	V2 C3	-
R-FR	SA3 V3	ME3 V3	CE1 V2	V2 C3	-
FRp	SA2 V3	ME3 V3	CE1 V2	V2 C3	-
FR-R	SA2 V3	ME3 V3	CE1 V2	V2 C3	-
FR-R	SA3 V3	ME3 V3	CE1 V2	V3 C3	

GT6075					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA3	ME1	CE3		ME3 T1
D	SA3	ME3 V3	CE1 V1		ME2
LD	SA3	ME3	CE1		ME3 T1
DL	SA3	ME3 V3	CE3 V3		ME3

GT6075					
Experiment 2					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA3	ME3	CE2 V2	LP2	-
D	SA3	ME3 V3	CE2 V2	-	-
R	SA3	ME3 V3	CE2 V3	LP2	-
FR	SA3	ME1 V1	CE2 V2	-	-
Experiment 3					
L	SA3	ME3	CE2 V3	LP2 RT2	-
D	SA3	ME3 V3	CE2 V3	LP2 RT2	-
R	SA3	ME3 V3	CE2 V3	LP2 RT2	-
FR	SA3	ME2 V3	CE2 V2	LP2 RT2	-
Experiment 4					
L	SA2 V3	ME2 V2	CE2 V2	V2 LP2	ME2 T2
D	SA2 V3	ME2 V2	CE2 V2	LP2	-
R	SA2 V3	ME2 V2	CE2 V3	LP2	-
Fr	SA2 V3	ME2 V2	CE2 V3	LP2	ME2
Rp	SA2 V3	ME2 V3	CE2 V3	LP2	-
R-Fr	SA2 V3	ME2 V3	CE2 V3	LP2	-
Experiment 5					
L	SA2 V3	ME2 V2	CE2 V2	V2 LP2	ME2 T2
D	SA2 V3	ME2 V3	CE2 V3	C2 LP2	-
R	SA2 V3	ME2 V3	CE2 V3	C2 LP2	-
Fr	SA2 V3	ME2 V3	CE2 V3	C2LP2	ME2
Rp	SA2 V3	ME2 V3	CE2 V3	C2 LP2	-
R-Fr	SA2 V3	ME2 V3	CE2 V3	C2 LP2	-
Frp	SA2 V3	ME2 V3	CE2 V3	C2 LP2	-
Fr-r	SA2 V3	ME2 V3	CE2 V3	C2 LP2	-
Experiment 6					
L	SA2 V2	ME2 V2	CE2 V2	LP2	ME2 T2
D	SA2 V3	ME2 V2	CE2 V2	LP2	-
R	SA2 V3	ME2 V2	CE2 V3	LP2	-
Fr	SA2 V3	ME2 V2	CE2 V3	LP2	-
Rp	SA2 V3	ME2 V3	CE2 V3	LP2	-
R-Fr	SA2 V3	ME2 V3	CE2 V3	LP2	-
Frp	SA2 V3	ME2 V3	CE2 V3	-	-
Fr-r	SA2 V3	ME2 V3	CE2 V3	-	-

GT6112					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA3	-	-	LP3 RT3	ME3
D	SA3	ME3 V3	CE1	LP3 RT3	ME3
LD	SA3	ME2 V2	CE1	RT3	ME3
DL	SA3	ME3 V3	-	LP3 C3 RT3	ME3
Experiment 2					
L	SA3	-	-	LP3 RT3	
D	SA3	ME3 V3	CE1	LP3 RT3	
R	SA3	ME1 V1	CE1	LP3 RT3	
FR	SA2	-	-	-	
Experiment 3					
L	SA2	-	-	LP3 RT3	E1
D	SA3	ME3 V3	CE1	LP3 RT3	
R	SA3	ME1 V1	CE1	LP3 RT3	
FR	SA3	-	-	-	
Rp	SA3	ME1 V2	CE1	RT3	
R-FR	SA3	ME2 V3	CE1	RT3	
FRp	SA3	ME1	CE1	LP3	
FR-R	SA2	-	CE1	RT3	
Experiment 4					
L	SA1	-	-	RT3	
D	SA3	ME3 V3	-	LP3 RT3	
R	SA2	ME1 V1	-	LP3 RT3	
FR	SA2	-	-	-	
Rp	SA3	ME2 V3	-	RT3	
R-FR	SA3	ME2 V3	-	RT3	
FRp	SA3	ME2 V3	-	LP3	
FR-R	SA3	ME2 V2	-	RT3	

GT6113					
Experiment 1					
treatment	shoot apex	Cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	RT3	-
D	-	ME1 V1	-	RT1	-
LD	-	-	-	RT3	-
DL	-	V1	-	RT3	-

GT6113					
Experiment 2					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	RT2	-
D	-	ME1	-	CE1 RH3 RT3	-
R	-	ME1 H1	-	CE1 RT2	-
FR	-	-	-	RT3	-
Experiment 3					
L	-	-	-	RT3	-
D	-	ME1	-	CE2 RT3	-
R	-	ME1 H1	-	CE2 RT2	-
FR	-	-	-	RT3	-
Rp	-	ME1	CE1	CE2 RT3	-
R-FR	-	ME1	-	CE3 RH1 RT3	-
FRp	-	ME1	-	CE3 RH1 RT3	-
FR-R	-	ME1	CE1	CE3 RH1 RT3	-
Experiment 4					
L	-	-	-	CE2 RT3	-
D	-	ME1	-	CE2 RT3	-
R	-	ME1 H1	-	CE2 RT3	-
FR	-	H1	-	CE2 RT2	-
Rp	-	ME1	CE1	CE2 RT3	-
R-FR	-	ME1	CE1	CE3 RT3	-
FRp	-	ME1	-	CE3 RT3	-
FR-R	-	ME1	-	CE3 RT3	-
Experiment 5					
L	-	-	-	RT3	-
D	-	ME1 V1	-	CE2 RT3	-
R	-	ME1 V1 H1	-	CE2 RT3	-
FR	-	-	-	RT3	-
Rp	-	ME1 V1	-	CE2 RT3	-
R-FR	-	ME1 V1	-	CE3 RT3	-
FRp	-	ME1 V1	-	CE2 RT3	-
FR-R	-	ME1 V1	-	CE3 RT3	-

GT6123					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	SA 2	ME1 V2	-	-	ME2
LD	SA2	ME1	-	-	ME2
DL	SA2	ME1 V2	-	-	ME2
Experiment 2					
L	-	-	-	-	-
D	SA 1	ME1 V1	-	-	ME1
R	S1	V1	-	-	-
FR	-	-	-	-	-

GT6227					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	SA2	ME2 V2	CE1	-	-
LD	SA2	-	-	-	-
DL	-	-	-	-	-
Experiment 2					
L	-	-	-	-	-
D	SA2	ME2 V2	CE1 V1	-	-
R	SA2	-	-	-	-
FR	-	-	-	-	-

GT6228					
Experiment 1					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	-	-	-	-	ME1
D	SA3 V3	ME3 V3	CE2 V1	LP3	ME3
LD	SA3	ME3 V3	-	-	ME2
DL	SA3 V3	ME3 V3	-	-	ME2

GT6228					
Experiment 2					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	ME1
D	SA3 V3	ME3 V3	CE2	C3 LP3	ME3
R	SA3 V3	ME2 V3	CE1	C3 LP3	ME3
FR	SA3 V3	ME2 V3	CE1	C3 LP3	ME2
Experiment 3					
L	-	-	-	-	ME1
D	SA2	ME3	-	-	ME3
R	-	ME1	-	-	ME3
FR	SA2	ME2	-	-	ME2
Rp	-	-	-	-	ME3
R-FR	-	ME2	-	-	ME2
FRp	-	ME2	-	-	ME3
FR-R	-	ME2	-	-	ME2
Experiment 4					
L	-	-	-	-	ME1
D	SA3 V3	ME3 V3	CE2	C3 LP3	ME3
R	SA3 V3	ME2 V3	CE1	C3 LP3	ME3
FR	SA3 V3	ME2 V3	-	C3 LP3	ME2
Rp	SA3 V3	ME3 V3	CE1	C3 LP3	ME3
R-FR	SA3 V3	ME2 V3	CE1	-	ME2
FRp	SA3 V3	ME2 V3	CE1	LP3	ME3
FR-R	-	ME2 V3	-	-	ME2

GT6236					
Experiment 1					
treatment	Shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	SA2 V2	ME3 V3	-	-	-
LD	-	ME1 V1	-	-	-
DL	-	ME1 V1	-	-	-
Experiment 2					
L	-	-	-	-	-
D	SA2 V2	ME3 V3	-	-	-
R	-	-	-	-	-
FR	-	-	-	-	-

GT6236					
Experiment 1					
treatment	Shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	SA1	ME1	-	-	-
D	-	ME1	-	-	-
R	-	ME1	-	-	-
FR	-	ME1	-	-	-
Rp	SA1	ME1	-	-	-
R-FR	SA1	ME1	-	-	-
FRp	SA1	ME1	-	-	-
FR-R	-	ME1	-	-	-

GT6240					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	-	CE1	RT3	-
LD	-	-	-	-	-
DL	-	-	CE1	RT1	-
Experiment 2					
L	-	-	-	-	-
D	-	-	CE1	RT2	-
R	-	-	-	-	-
FR	-	-	-	-	-

GT6279					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	CE1	-	-
D	SA2	-	CE1 V3	-	-
LD	SA2	-	CE1 V1	-	-
DL	SA2	-	CE1 V3	-	-

GT6279					
Experiment 2					
treatment	shoot apex	cotyledon	Hypocotyl	Root	immature leaf
L	S2	-	-	-	-
D	SA1	-	CE1	-	-
R	S2 SA1	-	CE1 V2	-	-
FR	-	-	-	-	-
Experiment 3					
L	S2	-	-	-	-
D	SA1	-	CE1	-	-
R	S2 SA1	-	CE1 V2	-	-
FR	-	-	-	-	-

GT6281					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	SA3	ME3 V3	CE1 V2	-	-
LD	SA3	ME1	-	-	-
DL	SA1	ME1 V3	-	-	-
Experiment 2					
L	SA1	-	-	-	-
D	SA2	ME2 V2	CE1 V1	-	-
R	SA2	ME1	-	-	-
FR	SA1	ME1	-	-	-
Experiment 3					
L	SA1	-	-	-	-
D	SA2	ME2 V2	CE1 V1	-	-
R	SA2	ME1	-	-	-
FR	SA2	ME1	-	-	-
Rp	SA2	ME1 V2	CE1 V2	-	-
R-FR	SA2	ME1 V2	CE1 V2	-	-
FRp	SA2	ME1 V1	CE1 V1	-	-
FR-R	SA1	ME1 V1	CE1 V1	-	-

GT6281					
Experiment 4					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	SA1	ME1	CE1	-	-
R	SA1	-		-	-
FR		-	-	-	-
Rp	SA1	ME1	CE1 V1	-	-
R-FR	SA1	ME1	CE1	-	-
FRp	SA1	ME1	CE1	-	-
FR-R	SA1	ME1	CE1	-	-
Experiment 5					
L	SA1	-	-	-	-
D	SA1	ME1	CE1	-	-
R	SA1	-		-	-
FR	-	-	-	-	-
Rp	SA1	ME1	CE1 V1	-	-
R-FR	SA1	ME1	CE1	-	-
FRp	SA1	ME1	CE1	-	-
FR-R	SA1	ME1	CE1	-	-
Experiment 6					
L	SA1	-	-	-	-
D	SA1	ME1	CE1	-	-
R	SA1	-		-	-
FR	-	-	-	-	-
Rp	SA1	ME1	CE1 V1	-	-
R-FR	SA1	ME1	CE1	-	-
FRp	SA1	ME1	CE1	-	-
FR-R	SA1	ME1	CE1	-	-

GT6325					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	S3	-	V2	C2 LP2	-
D	SA1 S3	-	CE1 V1	C2 LP1	-
LD	SA1 S3	-	CE1	C2 LP1	-
DL	-	-	-	-	-

GT6325					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	S3	-	V2	C2 LP2	-
D	-	-	CE1 V2	C2 LP1	-
R	S3 SA1	-	CE1 V3	C2	-
FR	S3	-	-	-	-
Experiment 3					
L	S3	-	V2	C2 LP2	-
D	-	-	CE1 V2	C2 LP1	-
R	S3 SA1	-	CE1 V3	C2	-
FR	S3	-	-	-	-
Rp	S3	-	CE1 V2	C2 LP2	-
R-FR	S3	-	CE1 V2	C2 LP2	-
FRp	S3	-	-	-	-
FR-R	S3 SA1	-	CE2 V3	CE3 LP3	-

GT6338					
Experiment 1					
Treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA2	ME2 V3 H2	-	CE3 V3 C3 RH1	-
D	SA2	ME2 V3	CE1 V3	C3 LP3	ME2
LD	SA2	ME2 V2	V1	CE3 C3 LP3	ME2
DL	-	-	V1	C3 LP3	-
Experiment 2					
L	SA2	ME2 V3 H2	-	CE3 V3 C3 RH1	-
D	SA2	ME2 V3	CE1 V3	V3 C3 LP3	-
R	SA2	ME2 V2	-	V3 C3 LP3	-
FR	SA1	ME2	-	-	-
Experiment 3					
L	SA2	ME2 V3 H2	-	CE3 V3 C3 RH1	
D	SA2	ME2 V3	CE1 V3	V3 C3 LP3	
R	SA2	ME2 V2	-	V3 C3 LP3	
FR	SA1	ME2 H1	-	-	
Rp	-	-	V3	LP3	
R-FR	SA2	ME2 V3	CE1 V3	CE2 V3 C3 LP3	
FRp	SA1	ME1	CE1 V3	V2 C3 LP3	
FR-R	SA1	ME1	CE1 V3	V3 C3 LP3	

GT6341					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME1 V1	-	CE3	-
D	SA3	ME3 V3	CE3 V3	CE3 RH3	ME3
LD	-	ME3 V3	CE3 V3	-	ME3 T3
DL	SA3	ME3 V3	CE3 V3	-	ME3
Experiment 2					
L	-	ME1 V2	-	-	-
D	SA3	ME3 V3	CE3 V3	CE3 RH3	-
R	SA3	ME3 V3	CE3 V3	CE3	-
FR	SA3	ME3 V3	CE3 V3	CE2	-
Experiment 3					
L	-	ME2 V2	CE2 V2	CE1	-
D	SA3	ME3 V3	CE3 V3	CE2 RH2	-
R	SA3	ME3 V3	CE2 V3	RH2	-
FR	SA3	ME3 V3	CE2 V3	RH2	-

GT6353					
Experiment 1					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	CE2 C2	-
D	-	-	-	C2	-
LD	-	-	-	C2	-
DL	-	-	-	CE2 C2	
Experiment 2					
L	-	-	-	CE2 C1 RH2	-
D	-	-	-	CE2 C2 RH3	-
R	-	-	-	CE2 C2 RH3	-
FR	-	-	-	CE2 C2 RH3	-
Rp	-	-	-	CE2 C1 RH1	-
R-FR	-	-	-	CE3 C3 RH3	-
FRp	-	-	-	CE3 C3 RH3	-
FR-R	-	-	-	CE2 C2 RH3	-

GT6372					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME2	V1	CE3C3 RH3	-
D	SA1	ME2 V2	CE2 V3	CE2 C2 RH2	-
LD	SA1	ME2 V2	E1 V1	CE3 C3 RH3	-
DL	-	-	-	CE3 C3 RH3	-
Experiment 2					
L	-	-	-	CE3V3 RH3	-
D	SA2	ME2	V3	V3 RH2	-
R	SA1 S3	ME2	CE1 V3	CE1 V1 C2 RH1	-
FR	S3	ME1 V2	-	-	-
Rp	S3	ME1 V3	V1	RH2	
Experiment 3					
L	S2	ME2	V1	CE3C3 RH3	-
D	SA1	ME2 V3	CE2 V3	CE3 C3 RH2	-
R	SA1	ME1 V3	CE2 V3	CE3 C3 RH3	-
FR	-	ME1 V3	CE2 V3	CE3 C3 RH3	-
Rp	SA2	ME1 V2	CE1 V3	CE3 C3 RH3	-
R-FR	SA2	ME2 V3	CE2 V3	CE3 C3 RH3	-
FRp	SA1	ME2 V3	CE1 V3	CE3 CE RH3	-
FR-R	SA1	ME2 V3	CE1 V2	CE3 C3 RH3	-
Experiment 4					
L	-	-	-	CE2 C2 RH2	-
D	SA1	ME1	CE2 V3	CE3 C3 RH2	-
R	-	ME1	CE2 V3	CE3 C3 RH3	-
FR	-	ME1	-	-	-
Experiment 5					
L	-	ME1	-	CE2C2 RH2	-
D	SA1	ME1	CE2 V3	CE3 C3 RH2	-
R	SA1	ME2	CE2 V3	CE3 C3 RH3	-
FR	-	ME2	CE1	CE3 C3 RH3	-
Rp	-	-	CE1 V3	CE3 C3 RH2	-
R-FR	SA1	-	CE2 V3	CE1 C3 RH3	-
FRp	-	-	CE1 V3	CE3 C3 RH3	-
FR-R	SA1	-	CE1 V2	CE3 C3 RH3	-

GT6372					
Experiment 6					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME1	-	CE2 C2 RH2	-
D	SA1	ME1	CE2 V3	CE3 C3 RH2	-
R	SA1	ME2	CE2 V3	CE3 C3 RH3	-
FR	-	ME2	CE1	CE3 C3 RH3	-
Rp	-	-	CE1 V3	CE3 C3 RH2	-
R-FR	SA1	-	CE2 V3	CE1 C3 RH3	-
FRp	-	-	CE1 V3	CE3 C3 RH3	-
FR-R	SA1	-	CE1 V2	CE3 C3 RH3	-

GT6407					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	SA2	ME2	-	-	-
LD	-	ME2	-	-	-
DL	-	-	-	-	-
Experiment 2					
L	-	-	-	-	-
D	-	ME1	-	-	-
R	-	ME1	-	-	-
FR	-	ME1	-	-	-
Experiment 3					
L	-	-	-	-	-
D	-	ME1	CE1	-	-
R	-	ME1	-	-	-
FR	-	-	-	-	-
Rp	-	ME1	CE1	-	-
R-FR	-	ME1	CE1	-	-
FRp	-	ME1	CE1	-	-
FR-R	-	ME1	CE1	-	-

GT6407					
Experiment 4					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	SA1	ME1	-	-	-
R	-	ME1	-	-	-
FR	-	-	-	-	-
Rp	SA1	ME1	-	-	-
R-FR	SA1	ME1	CE1	-	-
FRp	SA1	ME1	CE1	-	-
FR-R	SA1	ME1	CE1	-	-

GT6494					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA3 S3	-	CE2 V3	V3 LP3	ME1
D	SA1	-	CE2 V3	V3 LP3 C3	-
LD	SA2	-	CE2 V3	V3 LP3 C3	ME1
DL	SA1	-	CE2 V3	V3 LP3 C3	ME1
Experiment 2					
L	SA2 S2	-	CE2 V3	V3	-
D	SA1	-	CE2 V3	V3 LP3	-
R	SA1 S2	-	V2	-	-
FR	SA3 S3	-	CE1 V3	V3	-
Experiment 3					
L	SA1	-	CE2 V3	V3	-
D	SA1	-	CE2 V3	V3 LP3	-
R	No data				
FR	SA1	-	CE2 V3	V3 LP3	-
Rp	SA1	-	CE2 V3	-	-
R-FR	SA1	-	CE2 V3	V3 LP3	-
FRp	SA1	-	CE2 V3	V3 LP3	-
FR-R	No data				

GT6508					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	CE2 C2	-
D	-	ME1 V1	CE3 V3	CE3 RH2	-
LD	-	ME1	CE3	CE3 C3	-
DL	-	ME1 V2	CE3 V3	CE3 RH2	-
Experiment 2					
L	-	-	-	CE2	-
D	-	ME1 V1	CE3 V3	CE3 RH2	-
R	-	-	CE2 V2	CE2 RH2	-
FR	-	-	CE1	CE3 C3 RH2	-
Experiment 3					
L	-	-	CE3	E2	-
D	-	ME1 V1	CE3 V3	E3 RH2	-
R	-	-	CE2 V2	E2 RH2	-
FR	-	-	CE1	E3 C3 RH2	-
Experiment 4					
L	-	-	-	E2	-
D	-	ME1 V1	CE3 V3	E3 RH2	-
R	-	-	CE2 V2	E2 RH2	-
FR	-	-	CE1	E1 C2	-
Rp	-	ME1 V1	CE3 V3	E3 RH3	-
R-FR	-	-	CE3 V3	E3	-
FRp	-	ME1	CE3 V3	E2	-
FR-R	-	-	CE3 V3	E2	-

GT6534					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	V1	CE1	-
D	-	-	CE2 V3	CE3 LP3	-
LD	SA1	-	CE2 V3	CE3 LP3	-
DL	SA1	-	CE2 V2	CE3 LP3	-
Experiment 2					
L	-	-	-	V2	-
D	-	-	CE2 V3	LP3 V3	-
R	-	-	CE2 V3	LP3	-
FR	-	-	V2	-	-

GT6534					
Experiment 3					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	No data				
D	-	-	CE2 V3	CE3 V3 LP3 RT1	-
R	-	-	CE1 V2	CE2 V3 LP3	-
FR	-	-	CE1 V3	CE1 V2 LP3	-
Rp	-	-	CE2 V2	CE3 V3 LP3	-
R-FR	-	ME3	CE2 V2	CE3 V3 LP3	-
FRp	-	-	CE2 V3	CE3 V3 LP3	-
FR-R	-	-	CE2 V3	CE3 V3 LP3	-
Experiment 4					
L	-	-	-	CE3 V3 LP3	-
D	-	-	CE2 V3	CE2 V3 LP3	-
R	-	-	CE1 V3	CE3 V3 LP3	-
FR	-	-	CE1 V2	CE1 V2	-
Rp	-	-	CE2 V3	CE3 V3 LP2	-
R-FR	-	-	CE2 V3	CE3 V3 LP3	-
FRp	-	-	CE2 V3	CE3 V3 LP3	-
FR-R	-	-	CE2 V3	CE3 V3 LP3	-
Experiment 5					
L	No data				
D	-	-	CE1 V3	CE3 V3 LRP3	-
R	-	-	CE1 V2	CE3 LP3	-
FR	-	-	CE1	V2	-
Rp	-	-	CE2 V3	CE3 V3 LP3	-
R-FR	-	-	CE2 V3	CE3 V3 LP3	-
FRp	-	-	CE2 V3	CE3 V3 LP3	-
FR-R	-	-	CE2 V3	CE3 V3 LP3	-

GT6545					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	ME2 V2	-	-	-
LD	-	ME1	-	-	-
DL	-	ME1	-	-	-

GT6545					
Experiment 2					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	ME1	-	-	-
R	-	ME1	-	-	-
FR	-	-	-	-	-
Experiment 3					
L	-	-	-	-	-
D	-	-	-	-	-
R	-	ME1	-	-	-
FR	-	-	-	-	-
Rp	-	-	-	-	-
R-FR	-	-	-	-	-
FRp	-	-	-	-	-
FR-R	-	-	-	-	-

GT6604					
Experiment 1					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	-	V2	-	CE2 C3 RH3	-
D	-	ME1 V2	CE1 V2	CE2 C1 RH1	-
LD	-	ME3 V3	CE1	CE2 C3 RH3	-
DL	-	ME1 V2	-	CE2 C3 RH3	-
Experiment 2					
L	-	ME1 V2	-	CE2 C3 RH3	-
D	-	ME1 V2	CE1 V2	CE3 C3 RH3	-
R	-	ME1	CE1 V1	CE2 C1 RH3	-
FR	-	ME1 V2	-	CE1 C1 RH3	-
Experiment 3					
L	-	ME1	-	CE2 C3 RH3	-
D	-	ME1 V2	CE1 V2	CE3 C3 RH3	-
R	-	ME1	CE1 V2	CE2 C1 RH3	-
FR	-	ME1 V1	-	CE2 C1 RH3	-
Rp	-	ME1 V2	CE1 V2	CE1 C1 RH3	-
R-FR	-	ME1 V2	CE1 V2	CE1 C2 RH3	-
FRp	-	ME1 V2	CE1 V2	CE1 C3 RH3	-
FR-R	-	ME1	CE1 V2	CE1 C2 RH3	-

GT6604					
Experiment 4					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	-	ME1 V2	-	CE3 C3 RH3	-
D	-	ME1 V2	CE1 V2	CE3 C3 RH3	-
R	-	ME1	CE1 V2	CE3 C3 RH3	-
FR	-	ME1 V2	-	CE1 C1 RH3	-
Rp	-	ME1 V2	CE1 V2	CE1 C1 RH3	-
R-FR	-	ME1 V2	CE1 V2	CE1 C2 RH3	-
FRp	-	ME1 V2	CE1 V2	CE1 C3 RH2	-
FR-R	-	ME1	CE1 V2	CE1 C2 RH3	-
Experiment 4					
L	-	-	-	CE3 C3 RH2	-
D	-	-	V2	CE3 CE RH3	-
R	-	-	-	CE2 C2 RH2	-
FR	-	-	-	CE2 C2 RH2	-
Rp	-	ME1	V2	CE2 C2 RH2	-
R-FR	-	ME1	V2	CE2 C2 RH2	-
FRp	-	ME1	V1	CE2 C2 RH2	-
FR-R	-	-	-	CE2 C2 RH2	-

GT6634					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA3	V3	-	C3 LP3	CE3
D	SA3	ME3 V3	V3	C3 LP3	CE3
LD	SA3	ME3 V3	V3	C3 LP3	CE3
DL	SA3	ME1V3	-	C3 LP3	CE3
Experiment 2					
L	SA3	V3	-	C1 LP1	-
D	SA3	ME3 V3	CE1V3	C3 LP3	-
R	SA3	ME1V2	CE1V2	C2 LP2	-
FR	SA3	ME1V1	CE1 V1	C1 LP1	-
Experiment 3					
L	SA3	ME1V2	-	LP2 C1	CE1 T2
D	SA3	ME3 V3	CE1 V3	LP3 C3	CE1
R	SA2	ME3 V3	CE1 V3	LP3 C3	CE1 V2 T2
FR	SA3	ME2 V2	CE1	LP3 C3	-

GT6634					
Experiment 4					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA2 V2	-	-	LP3	CE2 V3
D	SA2 V2	ME3	-	LP3	-
R	SA2 V2	ME1	CE1	LP3	CE3 V3
FR	SA3 V2	ME1	CE1	LP3	CE3 V3
Rp	SA1 V2	ME1	-	LP3	-
R-FR	SA2 V2	ME1	-	-	-
Experiment 5					
L	SA2 V2	ME1 V1	-	LP2 C2	
D	SA1 V1	ME3 V3	V1	LP2 C3	
R	SA2 V2	ME1 V2	V2	LP3 C3	
FR	SA3 V2	ME1 V1	-	LP3 C3	
Rp	SA2 V2	ME2 V2	CE1 V2	LP3 C3	
R-FR	SA2 V2	ME2 V2	CE1 V2	LP3 C3	
FRp	SA1 V1	ME2 V2	CE1 V2	LP3 C3	
FR-R	SA1 V1	ME1 V1	V2	LP2 C3	
Experiment 6					
L	SA2 V2	ME1 V1	-	LP2 C2	CE2 V3
D	SA2 V3	ME3 V3	V1	LP2 C3	-
R	SA2 V2	ME1 V2	V2	LP3 C3	CE3 V3
FR	SA2 V2	ME1 V1	-	LP3 C3	CE3 V3
Rp	SA1 V1	ME2 V2	CE1 V2	LP3 C3	-
R-FR	SA1 V1	ME2 V2	CE1 V2	LP3 C3	-
FRp	SA2 V2	ME2 V2	CE1 V2	LP3 C3	-
FR-R	SA1 V1	ME1 V1	V2	LP2 C3	-
Experiment 6					
L	SA2 V2	ME1	-	LP3	CE2 V3
D	SA3 V3	ME1	-	LP3	-
R	SA2 V2	ME1	CE1	LP3	CE3 V3
FR	SA2 V2	ME1	CE1	LP3	CE3 V3
Rp	SA2 V2	ME1	CE1	LP3	-
R-FR	SA2 V2	ME1	CE1 V2	LP3	-
FRp	SA1 V1	ME1	CE1 V2	LP3	-
FR-R	SA1 V1	ME1	CE1 V2	LP3	-

GT6647					
Experiment 1					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	-	V3	-	CE3 C3 RH3	-
D	-	ME3 V3	CE3 V3	CE3 C3 RH3	-
LD	-	ME3 V3	CE1 V1	CE3 C3 RH3	-
DL	-	ME3 V3	CE3 V3	CE1 C3 RH3	-
Experiment 2					
L	-	H1	-	C2 RH2	-
D	-	P2	V3	C2	-
R	-	ME2	V2	C2	-
FR	-	ME2	-	C1	-
Experiment 3					
L	-	ME1 V2	-	CE3 C3 RH3	-
D	-	ME1 V2	CE1 V2	CE3 C3 RH3	-
R	-	ME1 V2	CE1 V2	CE3 C3 RH3	-
FR	-	ME2	CE1 V1	CE1 C3 RH2	-
Rp	-	ME1	CE1 V1	CE2 C3 RH3	-
R-FR	-	ME2 V3	CE2 V3	CE2 C3	-
FRp	-	ME2 V1	CE2 V3	CE3 C3 RH3	-
FR-R	-	ME1 V1	CE1 V3	CE3 C3 RH3	-
Experiment 4					
L	-	ME1 V2	V2	CE3 C3 RH3	-
D	-	ME1 V2	CE1 V2	CE3 C3 RH3	-
R	-	ME1 V2	CE1 V2	CE3 C3 RH3	-
FR	-	ME1 V2	V1	CE1 C3 RH2	-
Rp	-	ME1 V2	CE1 V1	CE3 C3 RH2	-
R-FR	-	ME2 V3	CE2 V3	CE3 C3 RH2	-
FRp	-	ME2	CE1 V3	CE2 C3 RH3	-
FR-R	-	ME2 V3	CE1 V3	CE2 C3 RH2	-
Experiment 5					
L	-	ME1 V3	V2	CE3 C3 RH3	-
D	-	ME1 V1	CE1 V2	CE3 C3 RH3	-
R	-	ME1 V1	CE1 V2	CE3 C3 RH3	-
FR	-	ME2 V3	CE1 V1	CE1 C3 RH2	-
Rp	-	ME1	CE1 V1	CE2 C2 RH2	-
R-FR	-	ME2 V3	CE2 V3	CE2 C3 RH2	-
FRp	-	ME1 V1	CE1 V3	CE3 C3 RH3	-
FR-R	-	ME1 V1	CE1 V3	CE2 C2 RH3	-

GT6670					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA1 S1	-	CE1	-	-
D	SA1	-	CE1 V2	V2 LP2	-
LD	-	-	-	-	-
DL	SA1	-	CE1 V2	-	-
Experiment 2					
L	-	-	-	-	-
D	SA1	-	CE2 V2	C2 RT1 LP2	-
R	SA1	-	CE1 V2	C1 RT1 LP2	-
FR	-	-	CE1 V2	-	-
Experiment 3					
L	-	-	-	-	-
D	SA1	-	CE2 V2	C1 RT1 LP2	-
R	-	-	-	-	-
FR	-	-	-	-	-
Rp	SA1	-	CE1 V2	C1	-
R-FR	SA1	-	CE1 V2	-	-
FRp	SA1	-	CE1 V2	C1 RT1	-
FR-R	SA1	-	CE1 V1	-	-
Experiment 4					
L	-	-	CE1	-	-
D	SA1	-	CE1 V2	C2 RT1 LP2	-
R	SA1	ME1	CE1 V2	C3 RT1 LP2	-
FR	-	-	CE1	C1	-
Rp	SA1	-	CE1 V2	C1	-
R-FR	SA1	-	CE1 V2	LP2 C1	-
FRp	SA1	-	CE1 V2	LP3 C2 RT2	-
FR-R	-	-	CE1 V2	C1 RT1	-

GT6671					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA2	P1	CE2 V3	-	-
D	SA1	ME2 V3 P1	CE1 V3	-	-
LD	SA1	ME2 V3 P1	CE2 V3	-	-
DL	SA1	P1	CE1 V3	-	-

GT6671					
Experiment 2					
treatment	shoot apex	Cotyledon	hypocotyl	Root	immature leaf
L	SA1	P1	CE1 V1	C1	-
D	SA2	ME1 V2 P1	CE1 V3	V2	-
R	SA2	ME2 V2 P1	CE1 V3	V2	-
FR	-	-	CE1 V1	-	-
Experiment 3					
L	-	ME1 P1	CE1	-	-
D	SA1	ME1 P1	CE1 V2	-	-
R	SA1	ME1 P1	CE1 V2	-	-
FR	-	-	CE1 V1	-	-
Rp	-	ME1	CE1 V2	-	-
R-FR	SA1	ME1	CE1 V3	-	-
FRp	SA1	P1	CE1 V3	-	-
FR-R	SA1	P1	CE1 V3	-	-
Experiment 4					
L	-	-	CE1	-	-
D	-	ME1 P1	CE1 V2	-	-
R	-	ME1 P1	CE1 V2	-	-
FR	-	-	CE1	-	-
Rp	SA1	ME1 P1	CE1 V2	-	-
R-FR	SA1	ME1 P1	CE2 V3	-	-
FRp	SA1	P1	CE1 V3	-	-
FR-R	SA1	ME1 P1	CE1 V2	-	-

GT6675					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME1	-	-	-
D	-	ME2 V2	-	-	-
LD	-	ME1	-	-	-
DL	-	ME1	-	-	-
Experiment 2					
L	S2	ME1	-	-	-
D	-	ME2	-	-	-
R	S2	ME1	-	-	-
FR	S2	ME1	-	-	-

GT6675					
Experiment 3					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	S2	ME1	-	-	-
D	-	ME1	CE1 V2	-	-
R	S2	ME1	-	-	-
FR	S2	ME1	CE1	-	-
Rp	-	ME1	CE1 V1	-	-
R-FR	-	ME1	CE1	-	-
Experiment 4					
L	-	ME1	-	-	-
D	-	ME1	CE1 V1	C1	-
R	-	ME1	-	C1	-
FR	-	ME1	-	-	-
Rp	-	ME1	CE1	-	-
R-FR	-	ME1	CE1 V1	C2	-
FRp	-	ME1	CE1 V1	C2	-
FR-R	-	ME1	CE1	C1	-
Experiment 5					
L	-	-	-	-	-
D	-	ME2 V2	CE1 V2	-	-
R	-	ME1 V1	-	-	-
FR	-	ME1	-	-	-
Rp	-	ME2 V2	CE1 V2	-	-
R-FR	-	ME2 V2	CE1 V2	-	-
FRp	-	ME2 V2	CE1 V2	-	-
FR-R	-	ME2 V2	CE1 V2	-	-
Experiment 6					
L	-	ME1	-	-	-
D	-	ME2 V2	CE1 V2	-	-
R	-	ME1 V1	-	-	-
FR	-	ME1	-	-	-
Rp	-	ME2 V2	CE1 V2	-	-
R-FR	-	ME2 V2	CE1 V2	-	-
FRp	-	ME2 V2	CE1 V2	-	-
FR-R	-	ME2 V2	CE1 V2	-	-

GT6675					
Experiment 7					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME1	-	-	-
D	-	ME1	-	-	-
R	-	ME1	-	-	-
FR	-	-	-	-	-
Rp	-	ME1	-	-	-
R-FR	-	ME1	CE1	-	-
FRp	-	ME1	CE1 V1	-	-
FR-R	-	-	CE1 V1	-	-

GT6688					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	ME1 V1	-	-	-
D	-	ME3 V3	CE1	-	-
LD	-	ME2 V3	-	-	-
DL	No data		-	-	-
Experiment 2					
L	-	ME1	-	-	-
D	-	ME2	-	-	-
R	-	ME2 V2	-	-	-
FR	-	ME1	-	-	-

LH211.16					
Experiment 1					
Treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	P3		LP3	T3
D	-	-	-	-	-
LD	-	P1	V1	-	T1
DL	-	P3	CE3 V3	CE2	-

LH211.16					
Experiment 2					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA2	-	CE1	CE2 V2 RH2	T3
D	-	ME2	CE1	-	-
R	-	-	V2	-	T3
FR	-	-	CE1	-	-
Rp	SA2	ME2	CE1 V2	-	-
R-FR	SA1	ME1	CE1	-	-
FRp	SA1	-	CE1	-	-
FR-R	SA2	ME2	CE1 V1	-	-
Experiment 3					
L	SA1	-	CE1	CE2 V2 RH2	T3
D	SA2	ME2 V2	CE2 V2	CE2 V2 RH2	-
R	SA1	-	V2	CE2 V2	T3
FR	SA2	ME2 V2	CE2 V2	CE2 V2 RH2	-
Rp	SA2	ME2	CE1 V2	-	-
R-FR	SA1	ME1	CE1	-	-

OK001.15					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	CE3 C3 RH3	-
D	-	-	-	-	-
LD				CE2 C1 RH2	-
DL	SA1	-	CE2	CE1 C1	-
Experiment 2					
L	-	-	-	CE3 C3 RH3	
D	-	-	CE1	-	-
R	-	-	-	-	-
FR	-	-	-	-	-
Rp	-	-	-	-	-
R-FR	-	-	-	-	-
FRp	-	-	-	-	-
FR-R	-	-	-	-	-

OK001.23					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	-	-
D	-	ME2 V3	CE2 V3	-	-
LD	-	ME1 V1	CE1	-	-
DL	-	ME3 V3	CE2V3	-	-
Experiment 2					
L	S1	ME1	-	-	-
D	-	ME2 V3	CE2 V3	-	-
R	-	ME2V2	CE1	-	-
FR	-	ME2	-	-	-
Experiment 3					
L	-	-	-	-	-
D	-	ME2 V3	CE2 V3	-	-
R	-	-	-	-	-
FR	-	-	-	-	-
Rp	-	-	CE2 V3	-	-
R-FR	-	-	CE1 V2	-	-
FRp	-	-	CE1 V2	-	-
FR-R	-	-	CE1 V2	-	-

OK003.19					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA1	ME3 V3 P3	CE1	CE2 RH1	ME1 T2
D	SA3	ME3 V3 P3	CE3 V3	CE2 RH1	ME3
LD	SA1	ME3 V3 P3	CE1	CE2 RH1	ME1 T2
DL	SA3	ME3 V3 P3	CE3 V3	CE2 RH1	ME3 T2
Experiment 2					
L	SA1	ME3 V3 P3	CE1	CE1 RH1	ME1
D	SA3	ME3 V3 P3	CE3 V3	CE1 RH1	ME3
R	SA1	ME3 V3 P3	CE1	CE1	-
FR	SA3	ME3 V3 P3	CE3 V3	CE1 RH1	-

OK003.19					
Experiment 3					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA1	ME2 V2	-		-
D	SA3	ME2 V2 P3	CE3 V3	-	-
R	SA1	ME2 V2 P3	-	-	-
FR	SA1	ME2 V2 P2	CE3 V3	-	-
Rp	SA1	ME1	CE2 V3	-	-
R-FR	SA3	ME3 V3 P3	CE3 V3	CE1	-
FRp	SA3	ME3 V3 P3	CE2 V3	-	-
FR-R	SA3	ME3 V3 P3	CE2 V3	-	-

OK008.13					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA3	ME1	-	-	CE3 V3 T2
D	SA3	ME2 V2	CE1 V2	-	-
LD	SA3	ME1 V1	CE1 V1	-	CE3 V3 T2
DL	SA3	ME2 V2	CE2 V2	-	-
Experiment 2					
L	SA3	ME1 V1	-	-	CE3 V3 T2
D	SA3	ME2 V2	CE1 V1	LP2	-
R	SA3	ME1 V1	CE1 V1	-	CE3 V3 T2
FR	SA3	ME2 V1	CE1 V1	-	CE3 V3 T2
Experiment 3					
L	SA3	ME1 V1	-	-	CE3 V3 T2
D	SA3	ME2 V2	CE1 V2	LP2	-
R	SA3	ME1 V1	CE1 V2	-	CE3 V3 T2
FR	SA3	ME1 V1	CE1	-	CE3 V3 T2
Rp	SA3	ME2 V2	CE1 V2	LP2	-
R-FR	SA3	ME2 V2	CE2 V2	-	-

OK008.13					
Experiment 4					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	SA3	ME1 V1	-	-	ME2 V2 T2
D	SA3	ME2 V2	CE2 V2	LP2	-
R	SA3	ME1 V1	CE1 V2	-	ME2 V2 T2
FR	SA3	ME1 V1	CE1	-	ME2 V2 T2
Rp	SA3	ME2 V2	CE2 V2	-	-
R-FR	SA3	ME2 V2	CE2 V2	-	-
Experiment 5					
L	SA3	ME1 V1	-	-	ME3 T2
D	SA3	ME2 V2	CE1 V1	-	ME3 T2
R	SA3	ME1 V1	CE1 V1	-	ME3 T2
FR	SA3	ME1 V1	CE1 V1	-	ME3 T2
Rp	SA3	ME2 V2	CE1 V2	LP2	ME2
R-FR	SA3	ME2 V2	CE1 V2	LP2	ME3
FRp	SA3	ME2 V2	CE1 V2	LP2	ME2
FR-R	SA3	ME2 V2	CE1 V2	LP2	ME2
Experiment 7					
L	SA3	ME1 V1	-	-	ME3 V3 T2
D	SA3	ME2 V2	CE1 V1	LP2	-
R	SA3	ME1 V1	CE1 V1	-	ME3 V3 T2
FR	SA3	ME1 V1	CE1 V1	-	ME3 V3 T2
Rp	SA3	ME2 V2	CE1 V2	LP2	-
R-FR	SA3	ME2 V2	CE1 V2	LP2	-
FRp	SA3	ME2 V2	CE1 V2	LP2	-
FR-R	SA3	ME2 V2	CE1 V2	-	-
Experiment 7					
L	SA3	-	-	-	ME3 V3 T2
D	SA3	ME1 V2	CE1 V1	-	ME3 V3
R	SA3	ME1 V1	CE1 V1	-	ME3 V3
FR	SA3	-	CE1 V1	-	ME3 V3
Rp	SA3	ME2 V2	CE1 V1	-	ME3 V3
R-FR	SA3	ME1 V2	CE1 V2	-	ME3 V3
FRp	SA3	ME1 V2	CE1 V2	-	ME3 V3
FR-R	SA3	ME1 V2	CE1 V2	-	ME3 V3

OK011.22					
Experiment 1					
treatment	Shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	C1	-
D	-	-	-	C3	-
LD	-	-	-	C1	-
DL	-	-	-	C1	-
Experiment 2					
L	S3	-	-	-	-
D	-	-	CE2	C1	-
R	-	-	-	C2	-
FR	-	-	-	C1	-
Experiment 3					
L	S3	-	-	-	T3
D	-	-	CE2	C1	-
R	-	-	CE2	C2	-
FR	-	-	CE2	C1	-
Experiment 4					
L	S3	-	-	-	T3
D	-	-	CE2	C1	-
R	-	-	-	C1	T3
FR	-	-	-	C1	-
Experiment 5					
L	S3	-	-	-	T3
D	-	-	CE2	C2	-
R	-	-	-	C1	T3
FR	-	-	-	C1	-
Experiment 6					
L	-	-	-	-	T3
D	-	-	CE2	C2	-
R	-	-	-	C1	-
FR	-	-	-	C1	-
Rp	-	-	CE1	C1	-
R-FR	-	-	CE1	C1	-
FRp	-	-	CE1	C1	-
FR-R	-	-	CE1	C1	-

OK011.22					
Experiment 1					
treatment	shoot apex	cotyledon	hypocotyl	Root	immature leaf
L	-	-	-	C1	-
D	-	-	CE2	-	-
R	-	-	CE1	C1	-
FR	-	-	-	-	-
Rp	-	-	CE1	C1	-
R-FR	-	-	CE1	C1	-
FRp	-	-	CE1	C1	-
FR-R	-	-	CE1	C1	-

Appendix 2

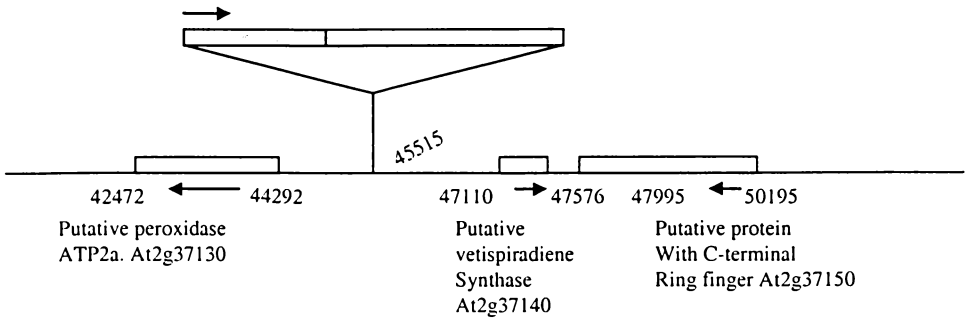
Graphical representation of gene insertions in promoter trap lines:

ET5158
ET5267
ET5280
ET5359
ET5599
ET5627
ET5653
ET6375
ET6566
ET6633

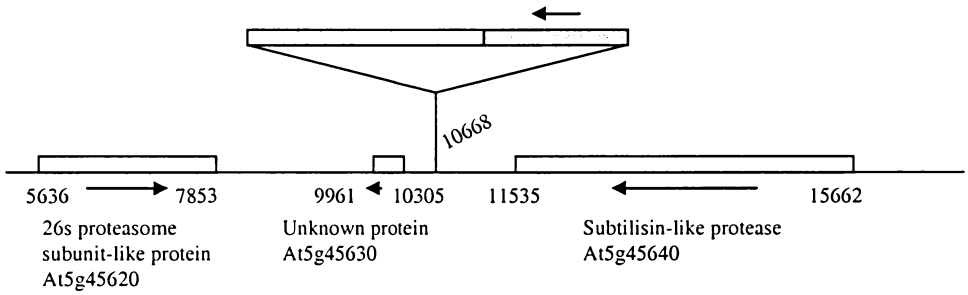
GT5909
GT5927
GT5914
GT5957
GT5964
GT6021
GT6039
GT6227
GT6228
GT6241
GT6281
GT6341
GT6372
GT6407
GT6634
GT6671

OK001.23
OK003.19
OK003.34
LH211.16
AJ146.55
OK011.22

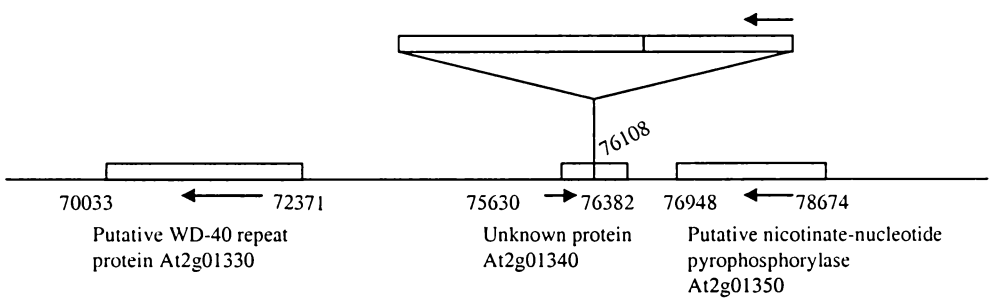
ET5158: *A.thaliana* chromosome 2. Sequence from clones T1j8, T2N18, F3G (Ds5-3 primer)



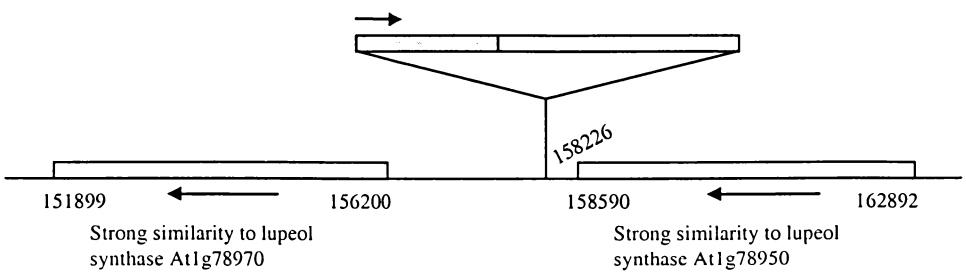
ET5267: *A. thaliana* genomic DNA, chromosome 5, P1 clone:MRA19 (Ds5-3 primer)



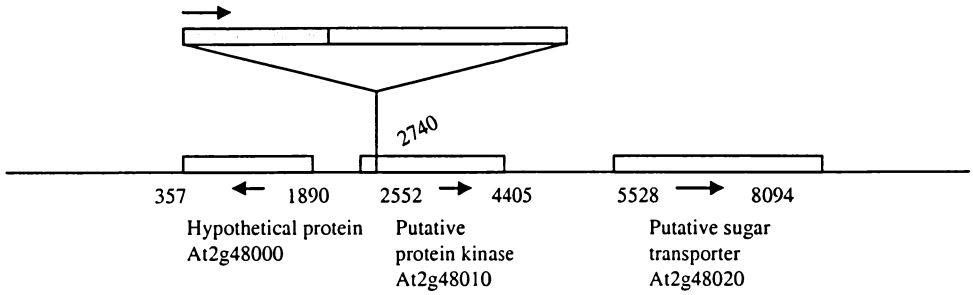
ET5280: *A. thaliana* chromosome 2 section 2. Sequence from clones F23H14,F10A8. (Ds5-4 primer)



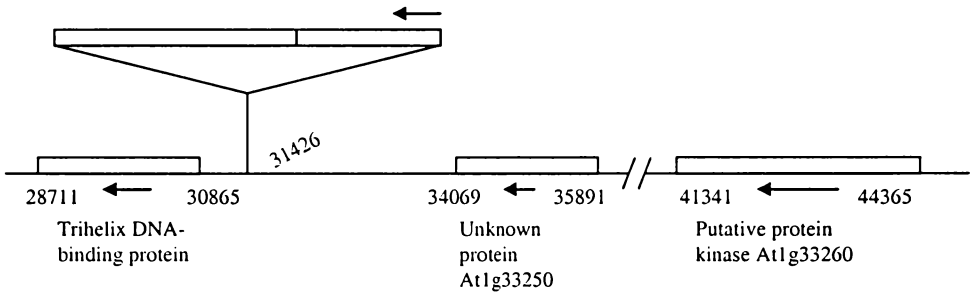
ET5359: *A.thaliana* chromosome 1 YAC YUP8H12R (Ds5-4 primer)



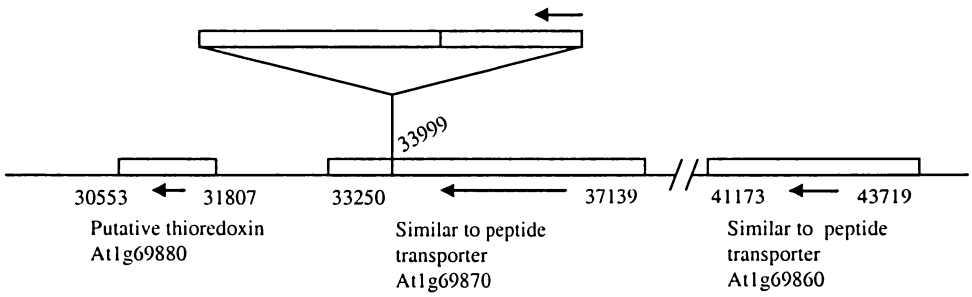
ET5599: *A. thaliana* chromosome 2 clone T9J23 (Ds5-4 primer)



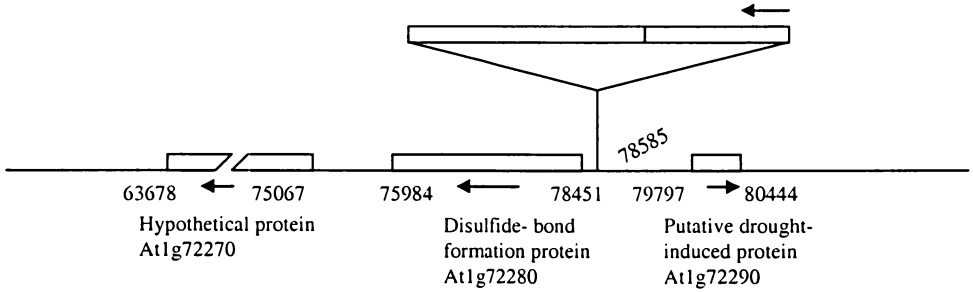
ET5627: *A. thaliana* chromosome 1 BAC T1609(Ds5-4 primer)



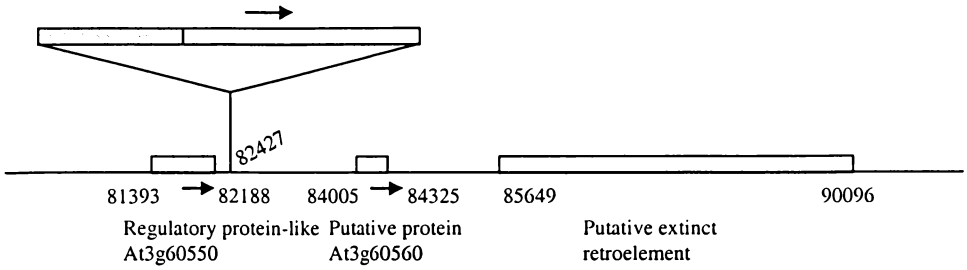
ET5653: *A. thaliana* chromosome 1 BAC T17F3 (Ds5-4 primer)



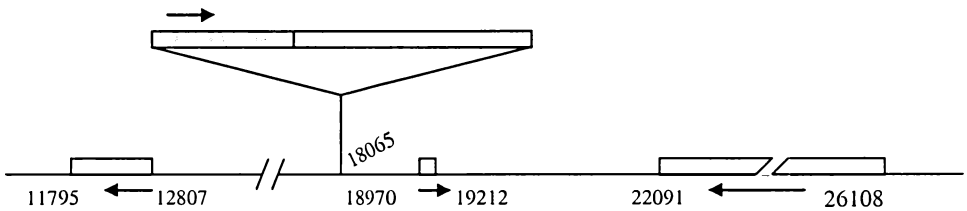
ET6375: *A. thaliana* chromosome 1 BAC T9N14 (Ds5-3 primer)



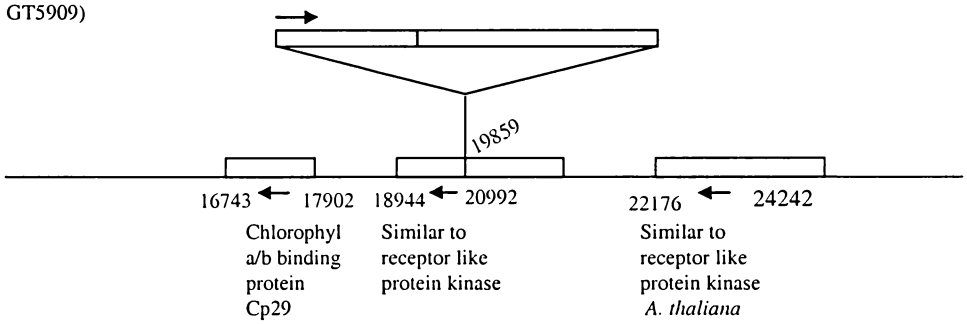
ET6566: *A. thaliana* Chromosome 3 BAC clone T8B10 (Ds5-3)



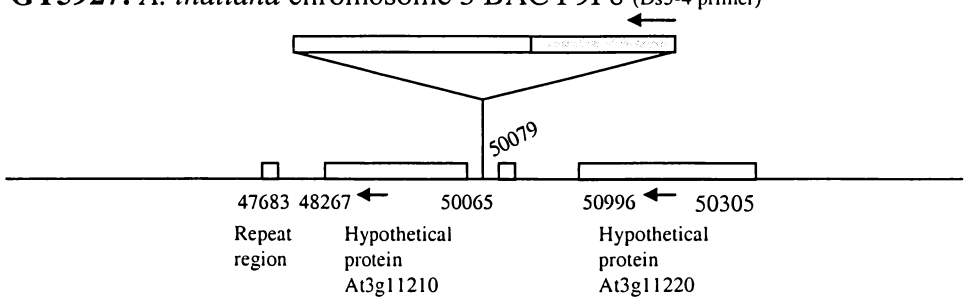
ET6633: *A. thaliana* chromosome 2 section 27 of 255 clone F16J10 (Ds5-4)



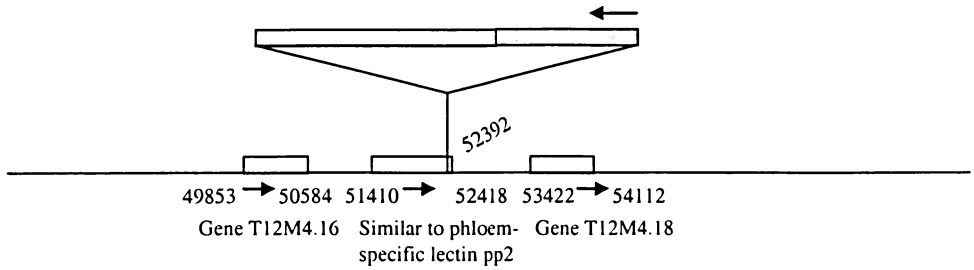
GT5909: *A. thaliana* chromosome 5 BAC F7A7 (3' flanking sequence of GT5909)



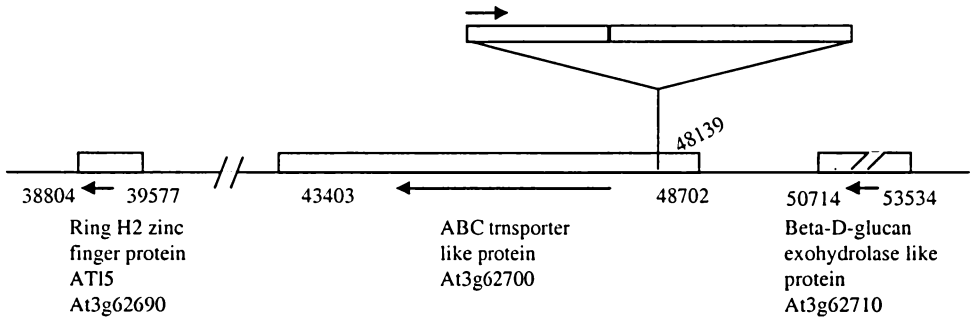
GT5927: *A. thaliana* chromosome 3 BAC F9F8 (Ds5-4 primer)



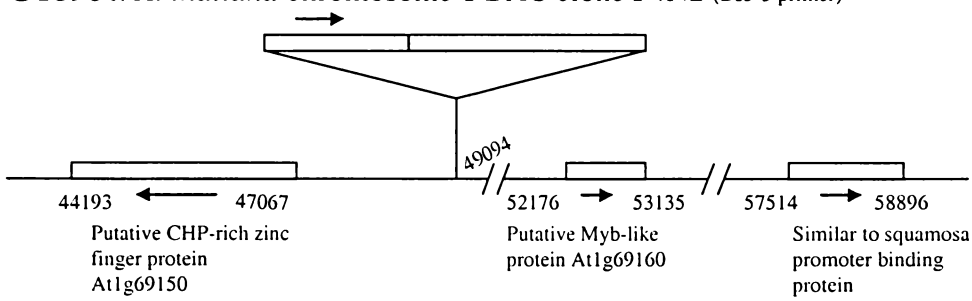
GT5914: *A. thaliana* chromosome 1 BAC T12M4 (Ds5-4 primer)



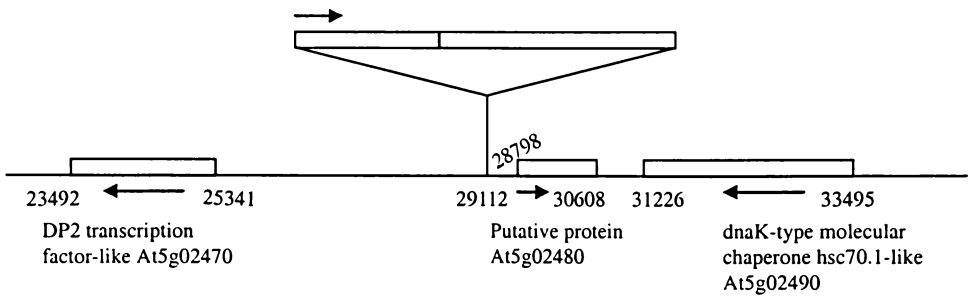
GT5957: *A. thaliana* chromosome 3 BAC F26K9 (3' flanking sequence of GT5957)



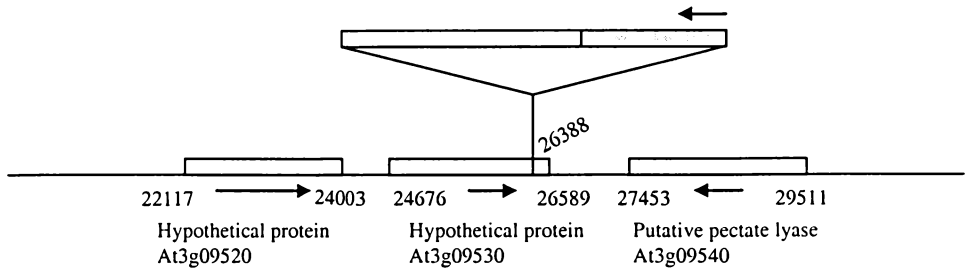
GT5964: *A. thaliana* chromosome 1 BAC clone F4N2 (Ds5-3 primer)



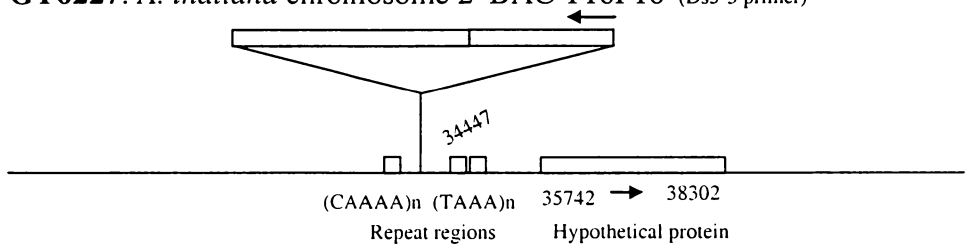
GT6021: *A. thaliana* chromosome 5 BAC clone T22P11 (Ds5-3 primer)



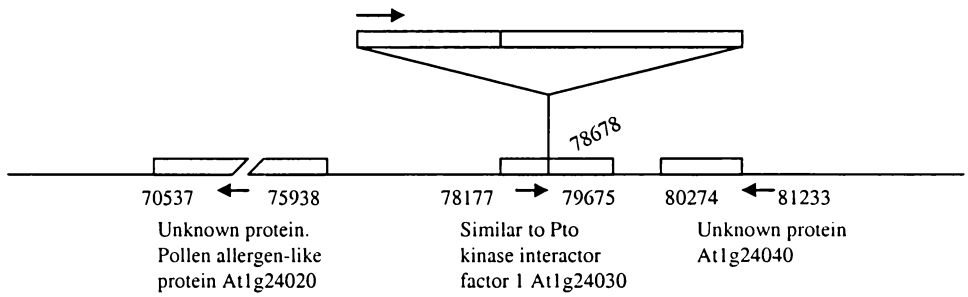
GT6039: *A. thaliana* chromosome 1 BAC f11F8 (Ds5-3 primer)



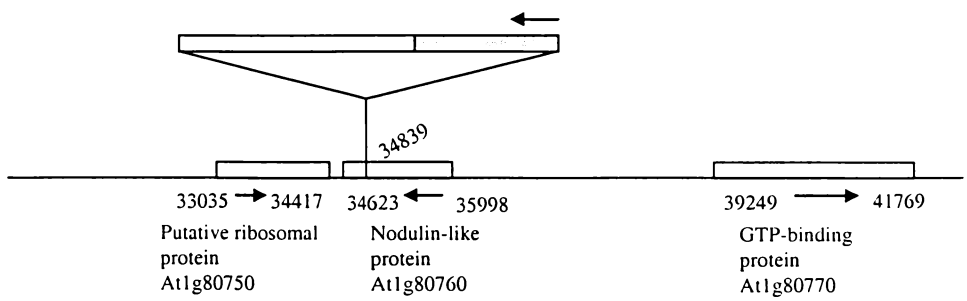
GT6227: *A. thaliana* chromosome 2 BAC T16F16 (Ds5-3 primer)



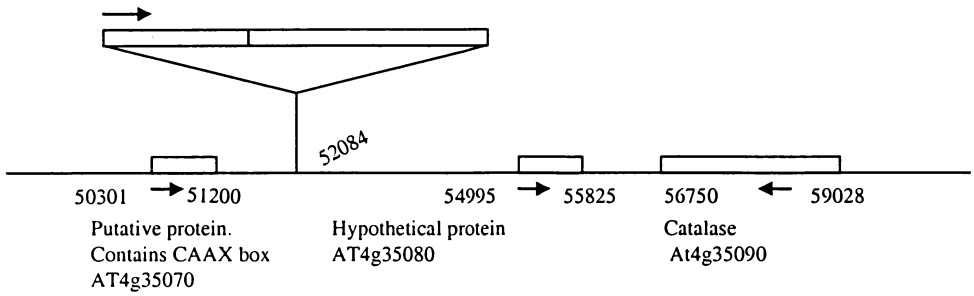
GT6228: *A. thaliana* chromosome 1 BAC T23E23 (Ds5-3 primer)



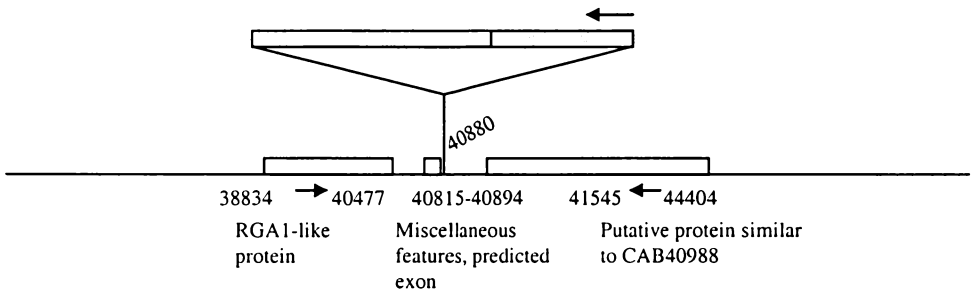
GT6241: *A. thaliana* chromosome 1 clones F23A5 (3' flanking sequence of gt6241)



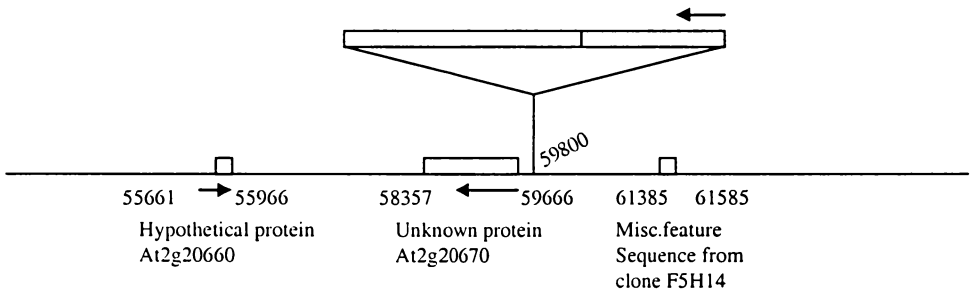
GT6281: *A. thaliana* chromosome 4 BAC M4E13 (Ds5-3 primer)



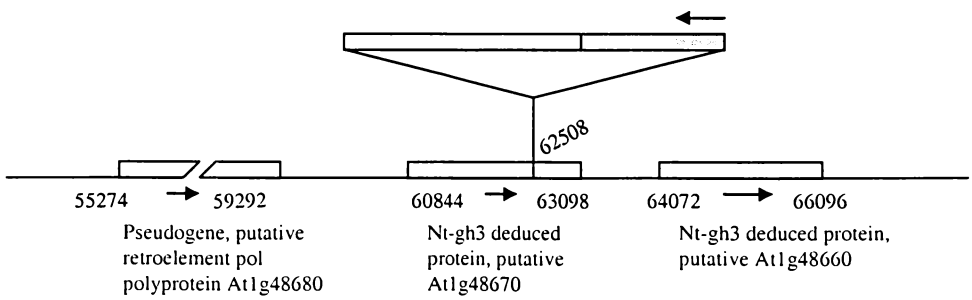
GT6338: *A. Thaliana* Chromosome 3 BAC T21P5 (Ds5-3 primer)



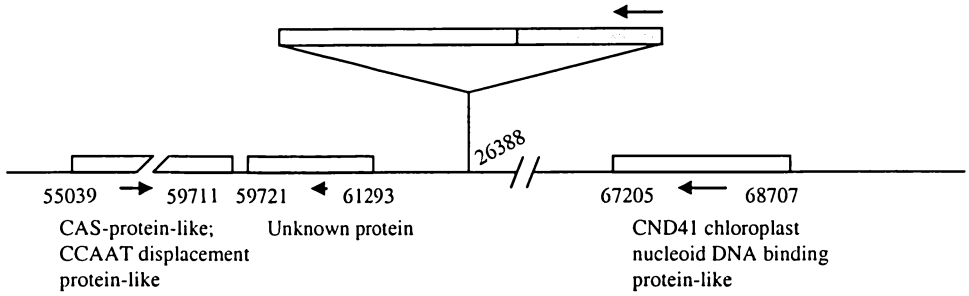
GT6341: *A. thaliana* chromosome 2 clones F23N11, F5H14 (Ds5-3 primer)



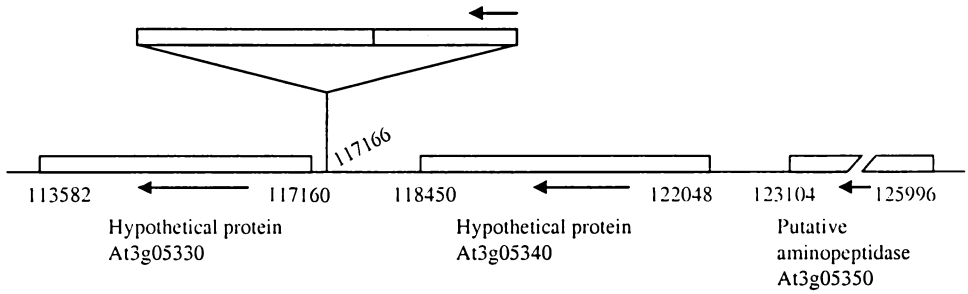
GT6372: *A. thaliana* chromosome 1 BAC F11I4 (Ds5-4 primer)



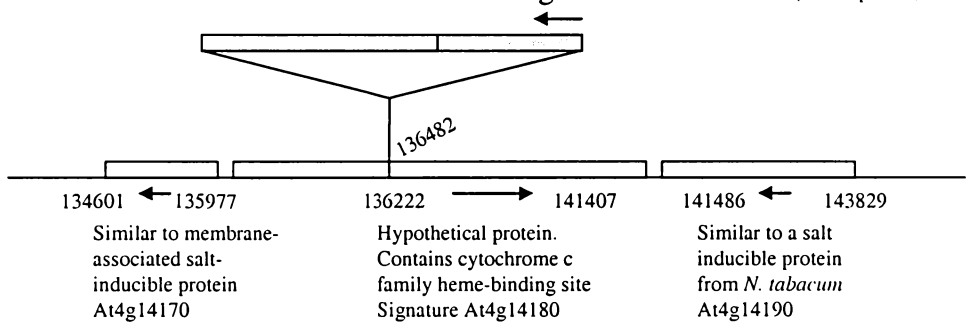
GT6407: *A. thaliana* chromosome 3 P1 clone MYF24 (Ds5-3 primer)



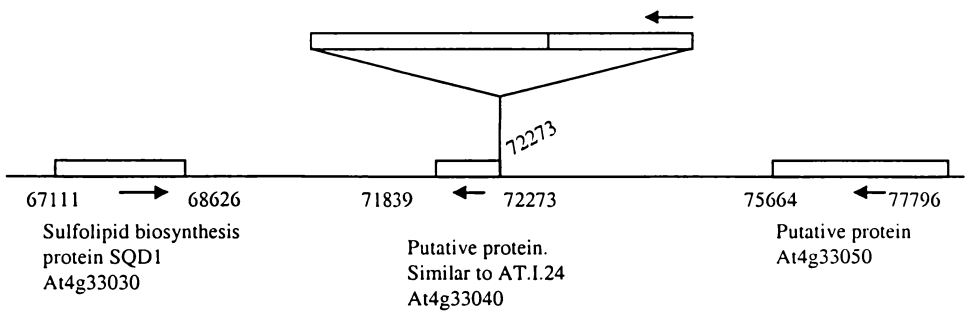
GT6634: *A. thaliana* chromosome 3 BAC T12H1 (Ds5-3 primer)



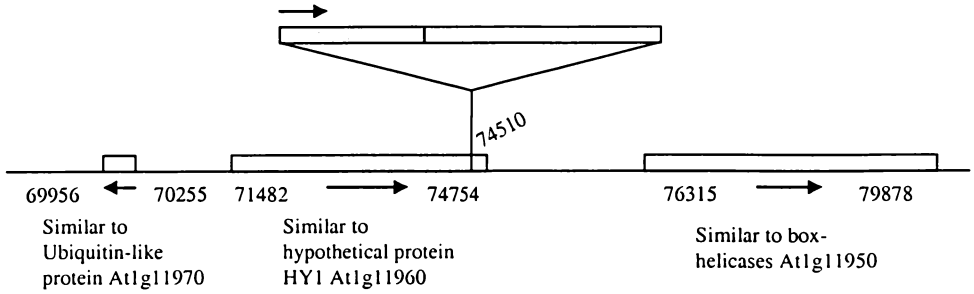
GT6647: *A. thaliana* chromosome 4 contig 0 clone ATFCA0 (Ds5-4 primer)



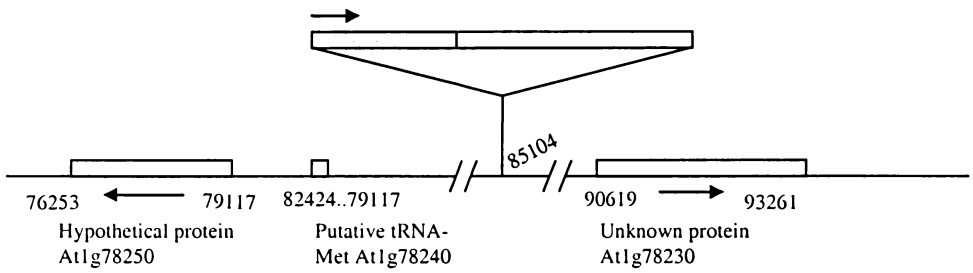
GT6671: A. thaliana chromosome 4 BAC F26P21 (Ds5-3 primer)



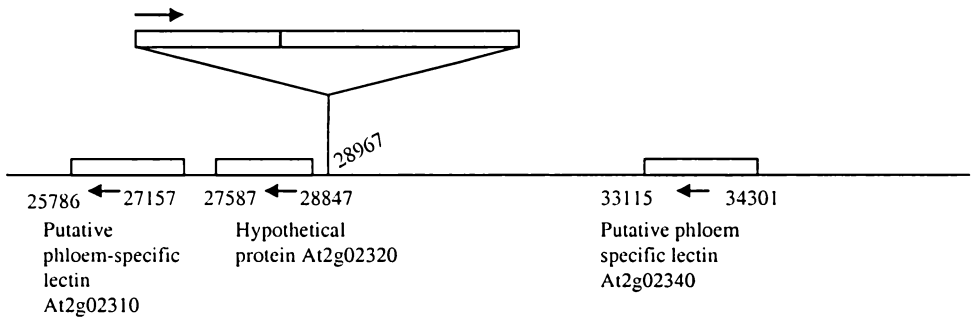
OK001.23:A. thaliana chromosome 1, BAC F12F1 (Ds5-3 primer)



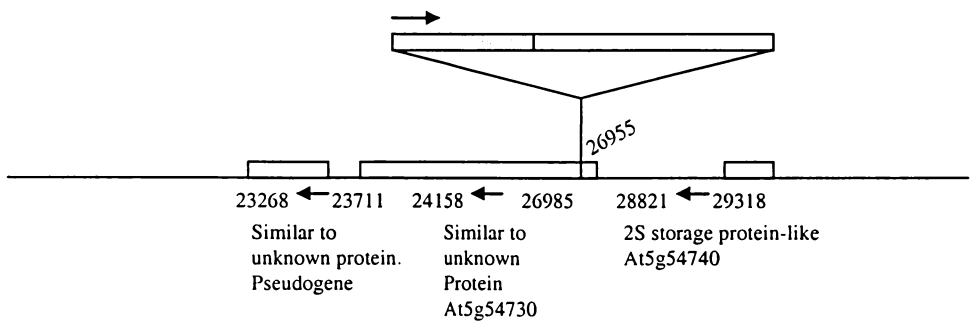
OK003.19:A. thaliana chromosome 1, BAC F3F9 (Ds5-3 primer)



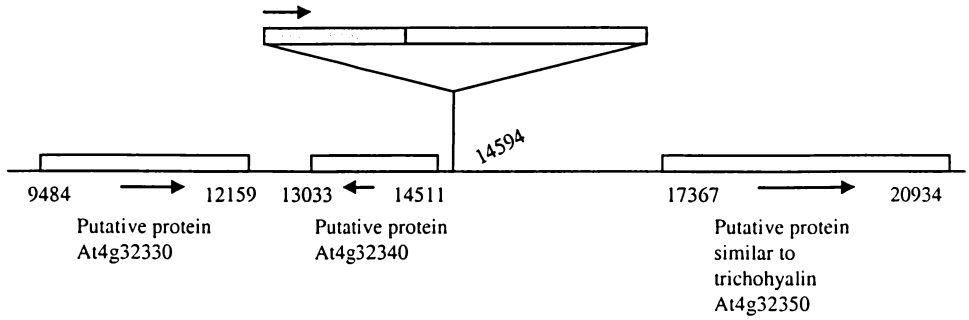
OK003.34: A. thaliana Chromosome 2 clones F504, T16F16 (Ds5-3 primer)



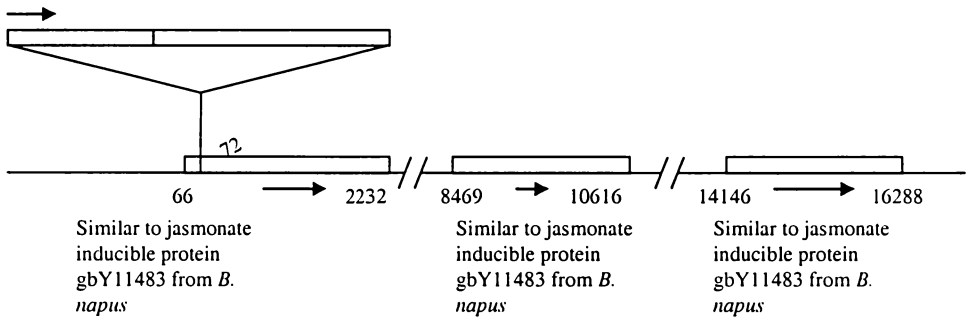
LH211.16: A. thaliana chromosome 5, TAC clone K5F14 (Ds5-3 primer)



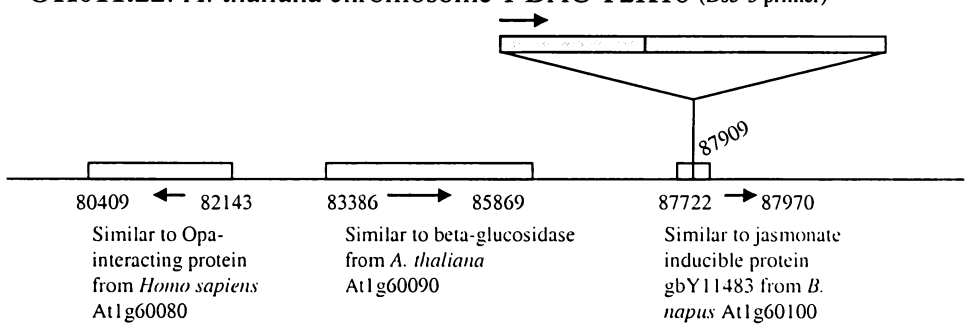
AJ146.55: A. thaliana chromosome 4 BAC F8B4 (Ds5-3 primer)



Ok011.22: A. thaliana chromosome 1 BAC T13D8 (Ds5-3 primer)



OK011.22: A. thaliana chromosome 1 BAC T2K10 (Ds5-3 primer)



Appendix 3

a. Classification of the differentially expressed genes in the cL-LD experiment. Out of 339 genes with differential expression greater than 2 or smaller than –2, 163 genes were annotated only as a clone number. These genes were not included in this table since it was impossible to determine the functional category

b. Classification of the differentially expressed genes in the cD-DL experiment. Out of 66 genes with differential expression greater than 2 or smaller than -2, 29 genes were annotated only as a clone number. These genes were not included in this table given the inability to determine the functional category.

Appendix 3A. Classification of the differentially expressed genes in the cL-LD microarray experiment

Gene product	Induction	Standard deviation	Functional category
NADPH:protochlorophyllide oxidoreductase	4		
Chloroplastic outer envelope protein	2.8		
Chloroplast 31KD ribonucleoprotein	2.6		**Chloroplast-general**
Chloroplast mRNA binding protein	2.4		
Chloroplast DNA binding protein	2.35	1.48	
Phosphoenolpyruvate carboxykinase	9.7		**Chloroplast- dark reactions**
Carbonate dehydratase	2.35	0.21	
Glyceraldehyde 3-phosphate DH	2.3	0.85	
Lhcb6 protein – A. thaliana	3.6		**Chloroplast – light reactions**
Chlorophyll a/b binding protein	2.7		
Cell wall protein	3.7		
Tubulin alpha chain – A. thaliana	3.5		
Cell division control	3.3	1.13	
Actin (*Brassica oleracea*)	3.3		
Mitogen activated protein kinase	3.2		
Cellulase	3.05	1.77	
Histone H3.3-line protein	3		**Cell division and expansion**
Cellular apoptosis susceptibility protein	2.6		
HSMU11 histone H1.1	2.5		
Similar to Picea histone H2A	2.5		
Putative pectin methylesterase	2.5		
Xyloglucan endotransglycosylase	2.2	1.13	
Beta-tubulin 4- A. thaliana	2.1	0.85	
3-phosphoshikimate 1-carboxyvinyltransferase	4.1		
T10M13.13	3.9		
ATP: Pyruvate phosphotransferase	3.7		
Acyl-CoA binding protein	3.4	1.27	**Metabolism-primary**
mRNA for Class III ADH	3.4		
Alanine transaminase Proso millet	3.3		
Aspartate aminotransferase	3.3		
Phosphoribosylformylglycinamidine	3.3		
HvB12D gene product – *H. vulgare*	3.25	1.2	
ADP-ribosylation factor	3.2		

Continuation Appendix 3A.

Gene product	Induction	Standard deviation	Functional category
Beta-glucosidase BGQ60 precursor	3.1		
Hypothetical protein - *Asparagus*	2.9		**Metabolism – primary**
Cytosolic phosphoglycerate kinase 1	2.8		
4-alpha-glucanotransferase – potato	2.8		
Amylogenin – *Oryza sativa*	2.8		
Nucleotide sugar epimerases	2.8		
Glucose 6-phosphate/ translocator precursor	2.7		
Similar to glycosyltransferase	2.7		
Catalase – *Vigna radiata*	2.7		
Aspartate aminotransferase – *A. thaliana*	2.7		
Putative transketolase precursor – *A. thaliana*	2.5		
Putative glucan synthase	2.5		
Invertase inhibitor homolog – *A. thaliana*	2.5		
Phosphopyruvate hydratase enolase	2.1	0.85	
S-Adenosylmethionine decarboxylase	2.05	0.35	
Flavonol synthase – *Solanum tuberosum*	4.4		
Mevalonate kinase	3.3		
Isochorismate synthase	2.95	1.63	
Aspartate kinase	2.8		
Acetyl-CoA carboxylase/biotin carboxylase	2.8		**Metabolism - secondary**
Vicia sativa CYP94A1	2.8		
AMP binding protein – *Brassica napus*	2.6		
Peroxisomal 2,4-dienoyl-CoA reductase	2.5		
Isoflavonoid reductase homolog	2.5		
Dimethylallyl diphosphate isomerase	2.5		
Ubiquinol-cytochrome-c-reductase	4.8		**Mitochondria/ oxidative Phosphorylation**
Ubiquinol-cytochrome c-reductase	3.2		
Ubiquinol cytochrome c-reductase	2.15	0.78	

Continuation Appendix 3A

Gene product	Induction	Standard deviation	Functional category
Hexameric polyubiquitin	4		
Similar to cysteine protease	3.2		
26S proteasome associated pad1 homolog	3		
Ubiquitin extension protein	2.9		
Ubiquitin-specific protease	2.9		
Peptidyl-prolyl cis trans isomerase	2.85	0.78	**Protein processing/ folding**
Ubiquitin conjugating enzyme	2.8		
Methionyl-tRNA synthetase	2.7		
Arginine/serine-rich splicing factor	2.7		
Arginine/serine-rich splicing factor	2.6		
Ubiquitin extension protein	2.55	0.49	
Chaperonin hsp60 precursor	2.05	0.78	
Hypothetical protein YEL031w	5.7		
ADP, ATP carrier protein	4.4		
Beta COP [*Rattus norvegicus*]	4.3		
AVP3 vacuolar proton pyrophosphatase	4.1		
c-subunit of V-ATPase [*Nicotiana tabacum*]	4.1		
Gamma tonoplast intrinsic protein 2	3.6	0.85	
14-3-3 like protein ATF1	3.4		
H+ transporting ATPase type 1	3.3		
NADH dehydrogenase	3.3		**Transport**
H+ transporting ATP synthase beta-1 chain	3.2		
Vacuolar type ATPase subunit A	2.8		
H+ translocating pyrophosphatase	2.8		
Mitochondrial ATPase beta subunit	2.8		
Brefeldin A sensitive Golgi protein-like	2.7	0.57	
V-ATPase subunit G	2.7		
P-glycoprotein 2	2.7		
Sec13-related protein	2.6		
Non-specific lipid transfer protein	2.15	0.49	
Non-specific lipid transfer protein	2	0.57	
Similar to human RNA polymerase complex	3.2		
RNA helicase	3		
RNA polymerase II 13.6 kDa chain	2.7		

Continuation Appendix 3A

Gene product	Induction	Standard deviation	Functional category
Similar to MADS box transcription factors	2.05	1.34	**Transcription**
Putative 40S ribosomal protein S25	4.7		
60S ribosomal protein L5	4.6		
60S ribosomal protein L26	4.4		
Putative ribosomal protein L10	3.9		
Ribosomal protein S21 – *Zea mays*	3.8		**Translation**
Ribosomal protein L38 isolog	3.6		
Ribosomal protein	3.6		
40S ribosome protein S7 - *Avicennia marina*	3.5		
30S ribosomal protein S9- *Synechocystis sp.*	3.5		
Translation elongation factor eEF-1	3.5		
Putative ribosomal protein L18A	3.45	0.21	
Cytoplasmic ribosomal protein L18	3.45	2.19	
40S ribosome protein S7 - *Avicennia marina*	3.4		
Translation elongation factor eEF-1	3.35	2.05	
60S ribosomal protein	3.3		
Similar to ribosomal protein	3.2	0.85	
60S ribosomal protein L37	3.1	1.41	
Cytoplasmic ribosomal protein S15	3	0.85	
Putative ribosomal protein S17	3		
40S ribosome protein S7-*Avicennia marina*	2.9		
Putative ribosomal protein L7	2.9		
40S ribosomal protein S29	2.9	1.13	
Similar to ribosomal protein L21	2.85	0.21	
Putative acidic ribosomal protein	2.85	0.49	
Putative ribosomal protein L7A	2.8	0.42	
Putative ribosomal protein L7A	2.75	0.78	
Ribosomal protein S11	2.75	1.06	
60S ribosomal protein L1	2.7		

Continuation Appendix 3A

Gene product	Induction	Standard deviation	Functional category
Ribosomal protein L16	2.6		
Elongation factor 2 – *Beta vulgaris*	2.6		
60S ribosomal protein L26	2.5	1.27	**Translation**
60S ribosomal protein L27a	2.5		
Ribosomal protein S15	2.5		
Elongation factor 2 – *Beta vulgaris*	2.5		
Ribosomal protein – *Rattus norvegicus*	2.25	0.92	
Translation elongation factor eEF-1	6.2		
Translation initiation factor eIF3 p40	2.9		
Eukaryotic translation initiation factor 3	2.9		
Putative serine carboxypeptidase	5.2		
Phi-1 – *Nicotiana tabacum*	4.6		
Profilin 2	3		
Auxin-induced protein	2.9		**Signal transduction**
Calcium-dependent protein kinase 6	2.7	0.57	
Putative kinase	2.7		
Lysine-sensitive aspartate kinase	2.5		
Nucleoside diphosphate kinase type 1	2.4	0.99	
Heat shock cognate protein 70-1	7.45	0.35	
Metallotionein 2a	5.4	0.42	
Omega 3-fatty acid desaturase CF3	4.9		
TMP-A transmembrane protein	4.1		
Germin-like protein	3.9		
Drought induced protein Di21	3.9		
Putative osr40 – *A. thaliana*	3.8		**Stress response**
Peroxidase – *A. thaliana*	3.7		
Pdr1 *A. thaliana*	3.7		
Nitrilase 1	3.5		
Monosaccharid transport protein STP4	3.2		
Probable glutathione transferase	3.1		
Transmembrane protein TMP-B	3		
Salt stress inducible small GTP binding	2.9		
Stress-induced protein OZI1 precursor	2.8		
Dehydration-induced protein RD22	2.7		

Continuation Appendix 3A.

Gene product	Induction	Standard deviation	Functional category
Jasmonate inducible protein isolog	2.6		
Transmembrane protein TMP-B	3.3		
Putative disease resistance protein	2.5		**Stress**
AtRab 18 – *A. thaliana*	2.4		**Response**
Copper homeostasis factor	2.25	1.06	
Glutathione S-transferase	2.15	0.78	
Nitrilase 1	2.05	0.21	
Chaperonin-60 beta subunit	2	0.14	
Metallothionein 2b	2.65	0.35	
Cyclase associate protein	3.2		
Cyclophilin	3		
Shoot forming PKSF1	2.9		
Hemolysin – *Aquifex aeolicus*	2.7		
Proline-rich protein APG	2.7		
Similar to mouse brain protein E46	2.7		**Other**
Bem46-like protein	2.6		
BIPOSTO	2.6		
Similar to Mycobacterium RlpF	2.5		

Appendix 3B. Classification of the differentially expressed genes in the cD-DL microarray experiment.

Upregulated genes		
Gene Product	**Induction**	**Functional category**
Tic22-like protein	11.1	**Chloroplast - general**
Photosystem I reaction center subunit III	4.4	
Monogalactosyldiacylglycerol synthase	4.5	
Photosystem II oxygen-evolving complex	3.2	
Photosystem II reaction center	3	
PSBY precursor – oxygen evolution	2.9	
PSI-H subunit – *Brassica rapa*	2.3	
Photosystem II - Oxygen evolving complex	2	
Carbonate dehydratase	3.9	**Chloroplast - dark reactions**
Ribulose bisphosphate carboxylase/oxygenase activase	2.8	
Chlorophyll a/b binding protein - like	3.6	**Chloroplast - light reactions**
Chlorophyll a/b binding protein - like	3.3	
Photosystem II type I chlorophyll a/b binding protein	3.1	
Lhcb2 protein	2.1	
PSI type III chlorophyll a/b binding protein	3.6	
Glyceraldehyde 3-phosphate dehydrogenase	2.2	
Succinate CoA ligase	3.6	**Metabolism – primary**
Thiamin biosynthesis protein thi4	3.3	
Aldehyde dehydrogenase like protein	2.5	
Ammonium transporter	2.5	
Plastidic aldolase – *Nicotiana paniculata*	2.3	
Phenylalanine ammonia-lyase	5	**Metabolism – secondary**
Putative malonyl-CoA:Acyl transacylase	2.3	
Acyl carrier-like protein	2.1	
Putative proteinase inhibitor II	2.1	**Protein processing-folding**
Selenium-binding protein	2.5	**Transport**
Putative DNA binding protein	3.5	**Transcription**
Homeotic protein Athb-6	2.9	
Plastid ribosomal protein	7.2	**Translation**

Continuation Appendix 3B.

Upregulated genes		
Gene Product	**Induction**	**Functional category**
Cytochrome P450	9.2	**Stress response**
Membrane associated salt-inducible protein	3	
Glutathione conjugate transporter	2.1	
Downregulated genes		
Aldehyde dehydrogenase (NAD+)-like protein	-2	**Metabolism - primary**
Alcohol dehydrogenase-like protein	-4.2	
Ribophorin I-like protein	-2.1	**Transport**
Arginine/serine-rich splicing factor	-2.1	**Transcription**

Appendix 4

See attached CD ROM (Plate 1) for appendix material.

The CD in the Appendix is the database of the GUS staining of lines analyzed for this work under cL, cD, cR, cFR, Rp, R-FRp, FRp, FR-R, LD and DL. In order to use these data, you need the ' Microsoft Access®' program for the pc. Simply insert the CD in the CD ROM drive. Open 'Access' and then click: 'File', 'Open Database'. In the window: 'enter database password', write: project. Click 'Open view reports' to do a search, or go to 'window', 'MBL database'. Click the tab 'Reports' and do your search. You can also use 'Insert', 'Query' and follow the instructions.

Vita

Olga Ruiz Kopp was born in Manizales (Caldas, Colombia) on September 11, 1959. She grew up in El Líbano (Tolima) and went to grade school to the 'Escuela Blanca Sáenz'. She graduated from High School in 1976 at the 'Colegio Nuestra Señora del Carmen' and then she went to the 'Universidad Nacional de Colombia' where she received a bachelor's in Biology in 1991. After graduation, she worked in 'Flores Colombianas' as Director of the Plant Tissue Culture Laboratory for two years, and then she came to the University of Tennessee to pursue higher education. She obtained her Master Degree at the Department of Ornamental Horticulture and Landscape Design in 1995 and her doctorate at the Department of Botany in Spring 2002.