Expression patterns of Arabidopsis PHYD and PHYE phytochrome genes antisense inhibition of the PHYB gene
by Lakshmi Ananthakrishnan Tirupathipanayam

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Biological Sciences
Montana State University

Abstract:
Phytochrome is the photoreceptor in plants that senses red and far-red light. In Arabidopsis, it is encoded by a small gene family with 5 members: PHYA, PHYB, PHYC, PHYD and PHYE. The regulation of the PHYA and PHYB genes has been characterized to an extent, but as yet little is known about the PHYD and PHYE genes. These two genes are more closely related to PHYB and form a sub-family of PHYB-like genes. So, the spatial, temporal and photoregulation of the PHYD and PHYE gene promoters was studied in comparison with that of the PHYB gene. Using the promoter-reporter fusion system with the gus reporter gene, I show here that, in spite of the amino acid sequence homology between them, there are distinct differences in the expression patterns of the three genes. The PHYB gene is expressed throughout the plant at medium to high intensity and shows no regulation by light. On the other hand, the PHYD gene shows light-dependent regulation in the roots. It is also expressed to a much lower intensity, the level of expression being strongly correlated to the transgene copy number. The PHYE gene also shows a very low intensity of expression. The difference in the pattern of expression of these 3 genes is very apparent in the male gametophyte: PHYB is expressed strongly in the filament, the anther and the pollen, while PHYD is expressed only in the pollen, and PHYE is expressed only in the anther.

In order to determine whether the activity of an individual phytochrome gene can be selectively suppressed with an antisense transgene, inhibition of the PHYB gene was attempted. Multiple copies of the antiB transgene succeeded in decreasing the amount of PHYB protein in the plant by about 50%. The phenotype observed in these plants was intermediate between the wild-type and that shown by the phyB null mutant, hy3, in characteristics of hypocotyl length and chlorophyll levels. However, certain variant phenotypes were also observed that deviated from both wild-type and mutant phenotypes. It is possible that these variant phenotypes might be caused by the antiB transgene interfering with the expression of other members of the phytochrome gene family.

EXPRESSION PATTERNS OF *ARABIDOPSIS PHYD* AND *PHYE* PHYTOCHROME GENES AND ANTISENSE INHIBITION OF THE *PHYB* GENE

by

Lakshmi Ananthakrishnan Tirupathipanayam

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Biological Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

April, 1994

APPROVAL

of a thesis submitted by

LAKSHMI ANANTHAKRISHNAN TIRUPATHIPANAYAM

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

5/3/94
Date

Robert A Shaw
Chairperson, Graduate Committee

Approved for the Major Department

3 May 1994
Date

Robert S Moore
Head, Major Department

Approved for the College of Graduate Studies

5/6/94
Date

R L Brown
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature Lakshmi. T. A.

Date May 3, 1994.

I dedicate this thesis to my beloved parents,
Mrs. Suseela Ananthakrishnan and Mr. T.V. Ananthakrishnan.

ACKNOWLEDGEMENTS

I wish to acknowledge with gratitude the members of my advisory committee, Dr. Robert A. Sharrock, Dr. Thomas J. McCoy and Dr. Ernest R. Vyse, for their help and advice. In particular, I would like to thank Dr. Bob Sharrock for lending me his help and support whenever I needed it. My sincere thanks also to Dr. Richard G. Stout for his friendship and moral support throughout the period of my stay at MSU.

I thank my colleague, Sarah Mathews, for her friendship and help. My thanks also to Ted Clack and Lynn Wester for their timely and expert technical assistance. I am grateful to Dr. Michael Ivie for letting us use his microscope to generate photographs. A special thanks to the user services assistants at the computer lab in Reid Hall for helping me with my computer work.

I wish to acknowledge with deep love and gratitude, the encouragement and support given to me by my parents, Mrs. Suseela Ananthakrishnan and Mr. T.V. Ananthakrishnan, and by my husband, Aiyappa Palecanda. My thanks also to my sister, Vidhya, and my grandmother, Mrs. Namagiri Sounderarajan, for their love and support.

Finally, I would like to thank my sponsor, the Rotary Foundation for sponsoring my graduate studies at Montana State University.

TABLE OF CONTENTS

TABLE OF CONTENTS-Continued

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Phytochrome is the photoreceptor in plants that senses red and far-red light. In *Arabidopsis*, it is encoded by a small gene family with 5 members: *PHYA, PHYB, PHYC, PHYD* and *PHYE*. The regulation of the *PHYA* and *PHYB* genes has been characterized to an extent, but as yet little is known about the *PHYD* and *PHYE* genes. These two genes are more closely related to *PHYB* and form a sub-family of *PHYB*-like genes. So, the spatial, temporal and photoregulation of the *PHYD* and *PHYE* gene promoters was studied in comparison with that of the *PHYB* gene. Using the promoter-reporter fusion system with the *gus* reporter gene, I show here that, in spite of the amino acid sequence homology between them, there are distinct differences in the expression patterns of the three genes. The *PHYB* gene is expressed throughout the plant at medium to high intensity and shows no regulation by light. On the other hand, the *PHYD* gene shows light-dependent regulation in the roots. It is also expressed to a much lower intensity, the level of expression being strongly correlated to the transgene copy number. The *PHYE* gene also shows a very low intensity of expression. The difference in the pattern of expression of these 3 genes is very apparent in the male gametophyte: *PHYB* is expressed strongly in the filament, the anther and the pollen, while *PHYD* is expressed only in the pollen, and *PHYE* is expressed only in the anther.

In order to determine whether the activity of an individual phytochrome gene can be selectively suppressed with an antisense transgene, inhibition of the *PHYB* gene was attempted. Multiple copies of the antiB transgene succeeded in decreasing the amount of PHYB protein in the plant by about 50%. The phenotype observed in these plants was intermediate between the wild-type and that shown by the *phyB* null mutant, *hy3*, in characteristics of hypocotyl length and chlorophyll levels. However, certain variant phenotypes were also observed that deviated from both wild-type and mutant phenotypes. It is possible that these variant phenotypes might be caused by the antiB transgene interfering with the expression of other members of the phytochrome gene family.

CHAPTER I

INTRODUCTION

Photomorphogenesis in Plants

Light is the primary source of energy for green plants. In addition to this, however, light also functions as a regulator of plant development. Plants are sedentary organisms that are unable to relocate under unfavourable conditions. Therefore, they have evolved to sense environmental cues and vary their responses accordingly. Light provides a number of cues by way of changes in light intensity, direction, duration and spectral quality. These are sensed by photoreceptors in the plants, and used to regulate the developmental program, a process commonly known as photomorphogenesis. Many diverse plant responses throughout the entire life of the plant have been found to be light-regulated including seed germination, tropic responses, chloroplast development, stem growth, and flowering, to mention just a few (1, 2, 3, 4).

Plant Regulatory Photoreceptors

Response to a light cue occurs via photodetection and subsequent signal transduction. Photoregulation of plant responses depends, therefore, on the ability of the plant photoreceptors to detect accurately the quality and quantity

of light received. Plant photoreceptor molecules are present in a wide variety of cells and not confined to a few specialized cells, as is the case with the animal light-sensory pigment rhodopsin (3). There are three major kinds of photoreceptors in plants: a) phytochrome, sensing red (600-700 nm) and far-red (700-780 nm) light, b) cryptochrome, sensing UV-A /blue light (320-520 nm), and c) a UV-B absorbing pigment (280-320 nm). Of these, phytochrome is the best-characterised (2, 3, 4, 5,) and the subject of this study. The first gene for a blue-light photoreceptor, has recently been isolated in *Arabidopsis thaliana* (6). The protein encoded by this gene shows considerable homology to microbial photolyases, which are flavoproteins that catalyse blue light-dependent reactions. The UV-B photoreceptor has not yet been isolated. The action spectra for a wide variety of plant photoresponses show that the chlorophylls do not play a major part in photomorphogenesis (3).

Biological Activity Of Phytochrome

Phytochrome was first identified in 1959 on the basis of its unique spectral properties (7) and is now the best characterised of the three photoreceptors. It is a regulatory photoreceptor that functions as a binary molecular switch controlling plant gene expression in response to light signals from the environment (2). The phytochrome molecule can assume two spectrally distinct

forms: Pr, which has maximum absorbance at 666 nm, and Pfr, which has maximum absorbance at 730 nm (4). These two forms are photointerconvertible: red light converts the Pr form to the Pfr form and far-red light reverses this reaction. Phytochrome is synthesised as Pr, which is inactive for most phytochrome-mediated responses. When converted to the Pfr form by photons of red light, phytochrome is activated and initiates a transduction process that culminates in altered expression of selected genes and ultimately in altered growth and development appropriate for the prevailing light environment (2). Studies of genes that are regulated by phytochrome show that the alteration of expression may be through transcriptional activation or repression of the genes. Many of the phytochrome-mediated responses can be cancelled by the subsequent reconversion of Pfr back to Pr (4). On repeated irradiation of plant tissue with alternating doses of R and FR, a simple response has been shown to be completely dependant upon the nature of the last stimulus given (3).

Recently, it has been shown that the red-absorbing (Pr) form of phytochrome B is required for normal hypocotyl gravitropism in *Arabidopsis* (9). This demonstrates that distinct biological activities can be associated with each of the two interconvertible forms of phytochrome and raises the possibility that the Pr form may not be completely inactive as has been the concept so far.

Since the interconversion of Pr and Pfr is rapid and reversible, the proportion of phytochrome in either form reflects a photoequilibrium based upon the proportion of red and far-red (R and FR) light entering the cell. This equilibrium is dependent on the ratio of the amount of R to FR light, and can change rapidly in response to changing light conditions. Thus, the phytochrome system is sensitive not merely to the presence or absence of light, but also to the spectral composition of the radiation (10).

The first phytochrome to be studied was the light-labile Type I phytochrome, now called PHYA phytochrome. Much more is known about this phytochrome than any of the other members of this receptor family. In the absence of any activating light signal, PHYA phytochrome accumulates in the red-absorbing Pr form. PHYA Pr is very stable with a half-life of about 100 hours, while the Pfr form is 50 to 100-fold less stable in the cell. Hence, within seconds of Pfr formation following illumination, the majority of this type of phytochrome becomes associated with discrete, amorphous subcellular bodies within the cytoplasm and is degraded (3). PHYA phytochrome photoreceptor is found in almost all plant tissues and organs examined, but varies in abundance in different cell types and in different regions of the plant. Cells that have been newly derived from shoot and root meristematic regions have larger amounts of this phytochrome compared to other cell types. Within

the cell, PHYA phytochrome in the Pr form shows no preferential association with any subcellular compartment (11).

Molecular Properties Of Phytochrome

Most of the biochemical characterization of phytochrome has been done on the Type I PHYA form and it is not yet clear whether other members of the phytochrome family conform to all of the properties listed here. Phytochrome is a soluble homodimeric chromoprotein. Each monomer is an approximately 120 kD protein (1100 amino acids) consisting of two discrete domains: an N-terminal domain of about 74 kD that contains a linear tetrapyrrole chromophore covalently linked via a thiol-ether linkage to a Cys residue in a hydrophobic pocket and a C-terminal domain of about 55 kD that possesses a site(s) responsible for dimerization. These two domains are linked together by a proteolytically vulnerable hinge region (4, 12). On dimerization, the molecule is thought to assume a Y-shaped structure (4). The chromophore is a bile pigment, phytochromobilin (13). Attachment of the chromophore occurs autocatalytically (in vitro), indicating that the protein has an intrinsic chromophore lyase activity. Other posttranslational modifications include the removal of the N-terminal Met, N-acetylation of the penultimate Ser and possible phosphorylation and glycosylation (4). The phototransformation between the Pr and the Pfr forms occurs

in the absence of any additional factors. The chromophore is the site of the R/FR light absorption and sets off a number of changes in the molecule. These include a cis- to trans-isomerization of one of the double bonds within the chromophore, a 31^0 reorientation of the chromophore relative to the polypeptide, and multiple conformational changes within the polypeptide, especially near the N terminus (3, 4).

Efforts made to delineate activities of various domains of the phytochrome through deletion analysis of the carboxy terminus of the oat *PHYA* gene have revealed the presence of separate domains required for spectral and biological activity (14).

The Phytochrome Gene Family

Data from early physiological and biochemical experiments suggested that there were at least two pools of phytochrome in plants, one that predominates in dark-grown, etiolated plants and the other which predominates in light-grown plants. These were referred to as Type I and Type II phytochrome respectively. The first phytochrome gene to be sequenced was the *PHYA* gene, which was found to encode the Type I or etiolated tissue or light-labile phytochrome (17). The Type II or green tissue or light-stable phytochrome was found to be heterogeneous and encoded by more than one gene. In *Arabidopsis*, it was shown that phytochrome is encoded by a small family of genes (8).

This gene family was found to include five members, *PHYA*, *PHYB*, *PHYC*, *PHYD* and *PHYE*, that are at least 50% identical with each other at the amino acid sequence level (15). These genes, though variable in sequence, are conserved in terms of their basic structure. They are all predicted to be soluble proteins located in the cytoplasm. No large non-homologous domains have been found in any one of them. The regions of highest sequence identity are around the portions of the phytochrome required for chromophore attachment and spectral integrity of the molecule. Analysis of amino acid sequence similarity among the five *PHY* genes shows *PHYA, PHYB* and *PHYC* to be equally divergent, and *PHYD* and *PHYE* to be more closely related to *PHYB* than to the other two (15; Table 1). In fact, the PHYD polypeptide shows a striking 80% sequence similarity to PHYB. The evolutionary pattern of the *PHY* genes derived from all these data is that *PHYA*, *PHYB* and *PHYC* presumably arose by duplications of an ancestral *PHY* gene either before or very early in angiosperm evolution and that the *Arabidopsis* PHYB, PHYD and PHYE phytochromes constitute a subfamily of proteins that are more recently derived and more structurally related (15). The high degree of sequence similarity among the five *PHY* genes suggests that they may work via similar signal transduction mechanisms. Characterisation of phytochrome mutants which are deficient in only one phytochrome protein (e.g. the *hy3* mutant that

	PHYA	PHYB	PHYC	PHYD	PHYE
PHYA	--	52	52	52	48
PHYB		--	52	80	56
PHYC			--	50	46
PHYD				--	55
PHYE					--

Table 1. Percent amino acid sequence identity among the *Arabidopsis* phytochrome polypeptides (from Clack et al. (15)).

lacks only PHYB protein (18, 19) and the *hy8* mutant that lacks only PHYA protein (20, 21)) has shown that these genes are not redundant. Nonetheless, it is likely that they interact in synergistic or complementary ways to control plant responses (15). Analysis of the five *PHY* mRNA levels under different light conditions, in different plant organs and in different stages of growth indicates that the phytochrome genes are relatively constitutive in expression, and do not exhibit diurnal cycling under a 12 hour photoperiod (15).

The *PHYA* gene is highly controlled by light and codes for the light-labile protein which is most abundant in dark-grown plants. Since *PHYA* transcription is rapidly repressed by red light and the Pfr form of PHYA is approximately 100-fold less metabolically stable than Pr, the level of the PHYA protein in green plants is as much as 100 times lower than in etiolated plants (4). The *hy8/fre1* mutants of *Arabidopsis* that are deficient in the PHYA protein show a lack of the far-red high irradiance response (FR-HIR) as dark-grown seedlings, but are nearly unaffected as light-grown plants. Thus, PHYA protein appears to be active mainly in etiolated seedlings and primarily responsible for enhancing the photosensitivity of seedlings before they emerge from the soil. In *Arabidopsis*, the *PHYA* gene is located on chromosome 1.

In contrast to *PHYA*, the *PHYB* and *PHYC*, and perhaps

PHYD and *PHYE*, genes encode proteins that are light-stable (15, 18) and can be classified under the Type II or light-stable forms. A mutant in the *PHYB* gene has been identified in *Arabidopsis* (*hy3*), that shows normal FR-HIR but shows alteration of red and white-light-controlled responses including cell elongation, chloroplast differentiation and timing of flowering as light-grown plants (16, 19). In *Arabidopsis*, this gene is located on chromosome 2. The physiological roles of the other phytochrome genes, *PHYC*, *PHYD* and *PHYE*, are yet to be determined.

Regulation Of Plant Genes By Phytochrome

Several plant genes have been found to be positively regulated by phytochrome. Among them are the cab genes that encode chlorophyll a/b binding proteins (23, 26), the *rbcS* genes that encode the small subunit of the ribulose-1,5-bisphosphate carboxylase-oxygenase enzyme (22, 24, 25), and the gene encoding the enzyme NADPH-protochlorophyllide reductase (27). Genes that are negatively regulated by phytochrome have also been identified and include the gene for asparagine synthetase (28, 29) and the NPR (Negatively phytochrome regulated) genes in *Lemna gibba* (30). Short DNA sequences in the promoter regions of these light-responsive genes have been shown to function as cis-acting elements involved in photocontrol (22, 23).

Though the light-responsive nature of phytochrome is

now well understood, little is known about the signal transduction mechanisms that follow the light perception and lead to gene regulation. Calcium/calmodulin and a heterotrimeric G protein have been implicated in several in vivo phytochrome responses (31). In addition, the phytochrome of the moss *Ceratodon* has been found to carry a protein kinase domain (32), and a C-terminal section of higher plant phytochromes has been found to share sequence homologies with the transmitter modules of bacterial sensor proteins (33). However, the actual mechanism of signal transduction by phytochrome molecules still remains to be understood.

Agrobacterium-mediated Plant Transformation

In order to understand the complexities of plant gene expression, plant gene cloning must be coupled with efficient methods of gene transfer into plants. The most important characteristic that differentiates plant gene transfer systems from animal systems is the totipotency that is observed in plant cells (34). This means that a typical plant cell can give rise to a complete, differentiated organism. Thus, gene transfer into plants does not have to use germ-line cells like in animals. Also, isolated plant cells do not necessarily retain a differentiated state determined by the tissue of origin, as is the case with animal cells. Thus, any plant cell in culture is capable of de-differentiating from the tissue of origin and

re-differentiating to give a whole, new plant. Consequently, transgenic plants can be produced in large numbers more easily than transgenic mice or *Drosophila* (35).

The production of morphologically normal plants that contain and express foreign genes is most easily accomplished through the use of the natural gene-transfer capacity of *Agrobacterium tumefaciens*. *Agrobacterium tumefaciens* is a soil bacterium that causes crown gall disease in plants (36). This bacterium is naturally capable of transferring a piece of DNA into the genomes of most dicotyledonous plants. The expression of genes located on this transferred DNA (T-DNA) inside the plants causes an imbalance of endogenous hormones leading to tumour formation. Foreign genes inserted on this T-DNA can be co-transferred, and can integrate into the plant genome (35, 36).

Furthermore, it is possible to 'disarm' the *Agrobacterium* Ti plasmids such that gene transfer can occur without disturbing the endogenous hormone balance in the plants. It has been shown that, for gene transfer to be accomplished, only the T-DNA borders and some flanking sequences need to be present in cis. Thus, the normal T-DNA-encoded genes have been deleted and easily selectable markers, such as those for bacterial antibiotic resistance, have been added to these T-DNA vectors. In this way, transgenic plants that differ from the wild-type only by the

presence of a transferred gene can be generated and studied (34).

Arabidopsis as the Model Plant System

Arabidopsis thaliana or 'Thale cress' as it is commonly called, belongs to the Cruciferaceae or the Mustard family. It is a small and inconspicuous weed, whose unique properties have rendered it a model plant system, ideal for use both in classical and molecular genetics. *Arabidopsis* is important to plant scientists because it has the smallest, simplest known genome of any angiosperm. In its haploid state, it has 5 small chromosomes. It has a haploid genome measuring close to 100,000 kilobases with a very low level of dispersed repetitive DNA (37). This genome is about 15 times that of *E. coli*, 5 times that of *Saccharomyces cerevisiae* and less than half that of *Drosophila melanogaster*, for which it is sometimes referred to as *Drosophila botanica* or a botanical *Drosophila* (38). The recombinational length of the genome is approximately 500 cM. RFLP maps of the genome are already available (37). Thus, this plant is ideal for molecular genetic work.

Other factors that make this plant favourable to plant scientists are its small stature, its ability to grow well indoors, and its short generation time of 5 to 6 weeks. Thus it is possible to grow a large number of these plants in a small area. The plant is self-pollinated and yields up to 100 seeds per seed pod or silique. Seeds are very small

and easy to store. The small size of the seed also makes it easy for mutagenesis on a large scale and ensures that the effects of mutations are carried to seed. This has resulted in a large number of mutants in various genes, both previously known and unknown. Finally, *Arabidopsis* is readily amenable to transformation with the *Agrobacterium tumefaciens*' Ti plasmid (37, 38).

Aims of Our Study

It is now well-known that the PHYA phytochrome is light-labile, whereas the PHYB and PHYC phytochromes are light-stable (18). Furthermore, the *PHYA* gene is transcriptionally light-regulated (40). The structure of the 5' flanking sequence of the oat and *Arabidopsis PHYA* genes has been studied and found to contain the promoter to the gene. A large intron (1.2 kbp) has been found within the 5'untranslated region of *PHYA* genes (39, 40). The photoregulation of the *PHYA* gene promoter has been extensively studied in monocots, notably the expression of the oat promoter in rice using the reporter gene chloramphenicol acetyl transferase (CAT) (40). However, little is known as yet about the photoregulation of the other *PHY* genes. One of the objectives of this study was, therefore, to elucidate the temporal, spatial and photoregulation of the *PHYD* and *PHYE* genes, and to compare them to the expression patterns of the *PHYB* gene.

The other objective was to observe the effects of

reduced PHYB protein levels in *Arabidopsis* plants. The null mutation in the *PHYB* gene in *hy3* strains of *Arabidopsis* causes alterations in hypocotyl elongation, chlorophyll content and flowering time (19). These alterations in the wild-type phenotype can be restored to normal by introduction of a *PHYB* transgene, while more than one copy of the transgene results in exaggeration of many of the same responses (16). *Arabidopsis* plants with reduced PHYB levels were generated using an antisense construct to that gene to look for deviations from both normal receptor functions and null mutant phenotypes. Furthermore, developing the antisense system to the *PHYB* gene would contribute to future work on making antisense to the *PHYD* and *PHYE* genes by determining if the activity of an individual phytochrome gene can be selectively suppressed.

In short, the objectives of our study were:

I. To see if the upstream flanking sequences of the Arabidopsis genes *PHYD* and *PHYE* contain promoter sequences using the *gus* (β-glucuronidase) reporter gene and if so, to characterise the expression of these promoters; also, to compare the expression patterns of these genes with that of the *PHYB* gene promoter.

II. To generate *Arabidopsis* plants that have reduced levels of the *PHYB* gene product using antisense to that gene and observe the effects.

CHAPTER II

PROMOTER ANALYSIS

Introduction to Promoter-reporter Gene Fusion System

One of the ways to elucidate the temporal and spatial regulation of genes is through a promoter-reporter gene fusion system. The concept for this technique is that one gene with a product that is easily detectable/assayable is used to infer the behaviour of another gene that is functionally fused to it (41). The functional fusion could be a transcriptional fusion, in which transcription of the reporter gene (with the easily detectable protein) starts from the +1 nucleotide of the gene whose regulation is to be studied, or a translational fusion in which the start codon for the reporter gene's protein is that of the regulated gene. This system can also be used in deletion analysis, in which truncated promoter sequences fused to the reporter gene can indicate positive and negative cis-acting sequence elements.

In most cases, the transferred genes are expressed under the control of the cotransferred cis-acting regulatory sequences and, at least qualitatively, retain their tissue- and developmental-specific expression patterns. However, occasionally in plants, the evolutionary distance between monocots and dicots appears to prevent such expression,

probably because the host plant does not contain appropriate trans-acting factors (35).

The most common reporter genes used in studying plant genes are:

1. the CAT or Chloramphenicol acetyl transferase gene (42)
2. the NPTII or the Neomycin phospho-transferase gene (43)
3. the LUC or the Luciferase gene (44), and
4. the GUS or the β-glucuronidase gene (45).

CAT and NPTII reporter gene systems have often been used to study light-regulation of plant genes (22, 23). Luciferase (from fireflies) is another reporter gene which is most useful in situations that require transient expression assays (44). Of these reporter genes, the β-glucuronidase or GUS reporter system is the most frequently used tool for the analysis of plant gene expression. The major advantages in this system include fast and non-radioactive analysis which is extremely sensitive, and also the potential to obtain both quantitative and qualitative data with the same reporter gene.

Vectors that Carry the *gus* Gene

The gene encoding GUS or β-glucuronidase is the *gusA* gene (formerly known as *uidA*), originally isolated from *Escherichia coli* (46).

For the purposes of promoter analysis, a set of vectors containing the *gus* gene and based on the binary plasmid pBIN19 (47) has been developed by Jefferson (45). The use

of these vectors is greatly facilitated by the ability of the β-glucuronidase enzyme to tolerate large amino-terminal additions for translational fusions. These allow both transcriptional fusions and translational fusions in all three reading frames. The vectors that were used in this study are pBI101.1 which contains a promoterless *gus* gene to which putative promoters can be fused, and pBI121, a derivative of pBI101 containing the 35S promoter of the cauliflower mosaic virus fused to *gus*.

The enzyme, β-glucuronidase, has a monomeric molecular weight of approximately 68,000, but exists in vivo as a tetrameric species. It has a wide specificity range for β-conjugated glucuronides, but will not cleave other glycosides, such as α- or β-glucoside subtrate types (48). The appropriate substrates all contain the sugar D-glucopyranosiduronic acid attached by glycosidic linkage to a hydroxyl group (usually a phenolic hydroxyl) of a chromogenic, fluorogenic, or other detectable molecule (48).

The *gus* Reporter System in Plants

One of the most attractive features of the *gus* reporter gene system in plant molecular biology is the virtual absence of background activity in a broad spectrum of higher plants (45, 46). There exists, however, a small amount of endogenous GUS activity. Some plants may lack GUS activity in vegetative organs but show activity in reproductive tissue (49). Several solanaceous plants show

activity of the enzyme in the male gametophyte (50). Endogenous GUS activity has also been found in a wide variety of species at pH 5 (51).

In *Arabidopsis thaliana*, native GUS activity was found in leaves, stems, roots and flowers of untransformed plants at pH values below 7. However, no activity was observed at pH 7 and 8. On the other hand, GUS-transformed plants showed *gusA* derived activity at all pH values, but the activity was higher at values equal and above pH 7. Thus, endogenous GUS activity is pH-dependent and can be suppressed by increasing the pH of the staining solution (52). Use of 20% methanol in the staining solution also serves this purpose (52).

Assaying for GUS Activity

The GUS reporter gene system has a number of advantages over other reporter gene systems. First among them is the nonradioactive nature of the analysis. The assay is fast and extremely sensitive, and it is possible to obtain both quantitative and qualitative data with the same reporter gene. Qualitative data, i.e., specificity of expression in tissues and organs, can be obtained using histochemical substrates such as X-gluc, whereas quantitative data, i.e., level of gene expression, can be derived by using fluorigenic substrates such as 4-MUG. It is even possible to use GUS as a selectable marker using certain chemical modifications of the substrate (52). The

most popular substrate for the histochemical assay is X-gluc which is 5-bromo-4-chloro-3-indolyl β-D-glucuronide. The insoluble blue precipitate, dichloro-dibromoindigo, is formed as a result of a two-step process involving production of a colorless intermediate by the GUS enzyme which then undergoes oxidative dimerization. The insoluble nature of the dye ensures its localization in vivo at the sites of enzyme activity. The GUS enzyme is very stable. It is in fact so stable that even very weak expression will eventually lead to significant staining (52).

Fresh tissue is usually fixed before it can be assayed. Low concentrations of wetting agents like Triton ensure that the substrate enters the plant organs. Oxidative catalysts like potassium ferricyanide and ferrocyanide are included in the substrate mix to accelerate the oxidative dimerization of the colorless cleavage intermediate into the colored final product, ClBr-indigo (53).

Quantification of enzyme activity can be done with maximum accuracy using fluorometry. The most widely used fluorigenic substrate for the GUS enzyme is 4-methylumbelliferyl β-D-glucuronide or 4-MUG (48).

GUS assays can be performed with either whole tissue or with homogenized or ground-up plant tissue, since there are protocols for both destructive and nondestructive assays.

Since the β-glucuronidase reporter gene system was first described in 1987, it has been used to study temporal and spatial patterns of gene expression as well as define cis-acting regulatory regions of several gene promoters (55, 56, 57, 58).

Disadvantages of the GUS Reporter System

One of the advantages of using the GUS enzyme is its stability. But this can lead to overinterpretation of data since even very weak expression will eventually lead to significant staining, also making it difficult to differentiate between early onset of transcription and strong promoter activity (52). Cell size and vacuolization can also affect the results of a GUS assay. Assuming that GUS is localized in the cytoplasm, small active cells, like meristematic cells,with few small vacuoles will appear to express the enzyme more than large, highly vacuolated cells. The histochemical staining of GUS transgenics can also show a preferential staining of vascular tissue, which can be attributed to accumulation of the substrate or product, and to the permeability of substrate in the tissue (52).

Another problem encountered is the variable results seen in some GUS transgenic lines. Sometimes, a decrease and even complete cessation of GUS expression is noticed in some lines. This may be due to an increased methylation of the introduced DNA, a phenomenon that is reported in expression of genes in transgenic plants in general (54).

Materials and Methods

Plant Material and Growth Conditions

The plant material used in all the transformations was *Arabidopsis thaliana* ecotype Nossen (No-0 WT).

For all seedling assays, seeds were surface sterilized for 30 min in 15% bleach/0.2% SDS, rinsed at least 5 times with sterile water, and plated in 150x25mm Petri dishes containing GM medium (60) (see Appendix for media content). Plates were treated at 4°C in the dark for 2-5 days and then transferred to continuous white light. For dark-grown seedlings, germination was induced by placing the plates in white light for 10 min prior to transferring them to darkness at 24°C. Experiments requiring continuous white light used white light from a bank of 40 watt fluorescent bulbs (50 μm m^{-1} sec^{-2}). For one experiment involving mature plants, seedlings grown on petri dishes were transferred to GM contained in magenta boxes. For mature plant assays and for generating seed of the next generation, seedlings were moved to pots containing potting soil overlaid with vermiculite, and grown at 24°C under a 16 hour-photoperiod.

Construction of the P_DGUS Translational Fusion

pBI 101.1 is the vector that was used in the translational fusion experiments. This is part of the pBI 101 series. pBI 101 is a 'promoter-less' GUS cassette in

the *Agrobacterium* binary plasmid vector pBIN19 (47). It consists of the GUS gene blunt-end ligated into a filled-in KpnI site of the pBIN19 polylinker upstream of a 260 bp SstI - EcoRI fragment containing the polyadenylation signal from the nopaline synthase gene (*nos*) of the *Agrobacterium tumefaciens* Ti plasmid. Upstream of the ATG of the GUS cassette is a polylinker region containing Hind III, Sph I, Pst I, Sal I, Xba I, Bam HI and Sma I sites. This allows plant promoters to be easily cloned upstream of GUS and transferred to plants with all the advantages of the binary vector systems. This vector has a lowcopy RK2 origin of replication, and confers kanamycin resistance, due to the presence of the Neomycin phosphotransferase II gene (*npt II*) with the *nos* gene promoter and termination sequences. The kanamycin resistance can be used both in bacteria and in plants. The pBI 101.1, 101.2 and 101.3 represent the 3 reading frames in which the GUS gene can be read meaningfully (45).

The *PHYD* upstream sequence, that contains the putative *PHYD* gene promoter, the 5' untranslated region and the first 65 amino acids of the *PHYD* open reading frame was taken from the genomic clone λD 6-1. This 3 kb fragment, flanked by BglII restriction enzyme ends had been cloned into the BamHI site of M13mp18 in the reverse orientation and designated mpTC304. The *PHYD* promoter fragment was excised out of mpTC304 at the flanking SphI and SmaI

restriction sites and ligated into the SphI and EcoRV sites of the pGEM5Z+ vector. The ligated DNA was used to transform competent cells of *E. coli* strain HB 101 through heat-shock method (59). Selection of transformant colonies was done by screening the transformation on LB plates containing ampicillin at 100 μg/ml. DNA from selected colonies was minipreped according to the protocol described by Maniatis et al. 1982 (59). One clone thus obtained proved to contain the *PHYD* promoter fragment in the right orientation and was designated pLT1.

Plasmid pLT1 was then digested with SalI and the *PHYD* promoter fragment was inserted into the Sal I site in the polylinker 5' to the GUS gene of pBI101.1. The ligated DNA was used to transform *E. coli* HB 101 competent cells and selected on LB-kan_{50} plates. Transformants carrying the translational fusion of the *PHYD* promoter and coding sequence to the GUS gene were identified and designated P_DGUS (Fig. 1).

Presence of the P_DGUS construct in the transformed colonies on LB-kan plates was determined by colony hybridization with a radioactive probe containing a 900 bp fragment of the *PHYD* promoter region (59). The hybridization revealed 2 transformant colonies with the P_DGUS construct. The two clones were designated P_DGUS3 and P_DGUS12, and used to transform *Agrobacterium tumefaciens* strain LBA4404 (60). *Agrobacterium tumefaciens* carries the

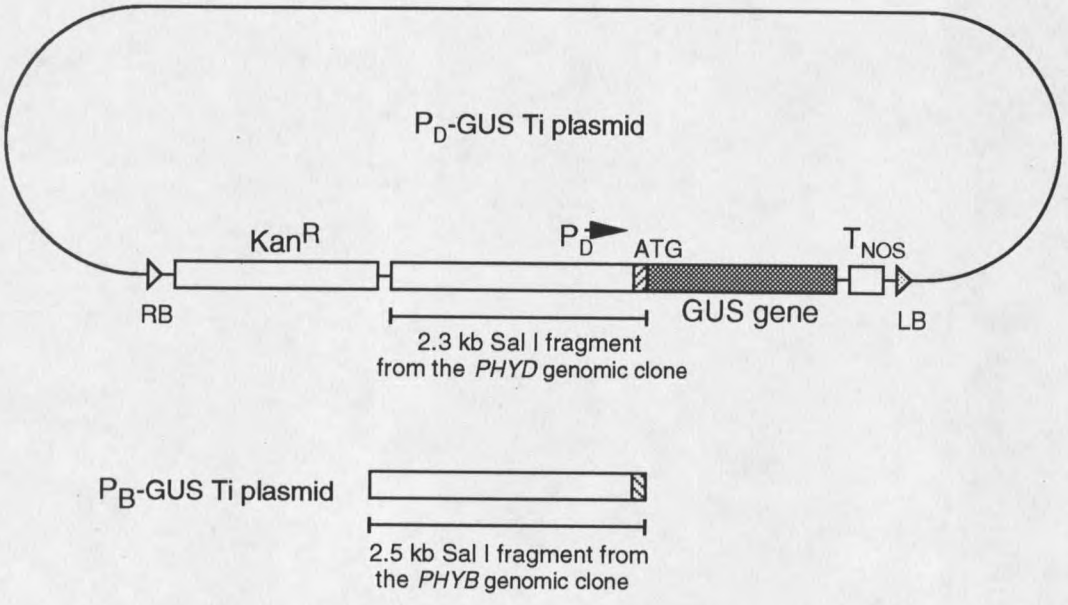

Fig. 1. Construction of the *PHYD*-GUS translational fusion. LB, left border; RB, right border; Kan^R, kanamycin resistance; P_D, *PHYD* gene promoter; T_{NOS}, nopaline synthase terminator. In the *PHYB*-GUS fusion, the 2.5 kb Sal I fragment from the *PHYB* gene was used to make a similar fusion with the GUS gene.

gene for resistance to streptomycin, so selection of transformed *Agrobacteria* was accomplished by plating out the transformation on LB $kan_{50}str_{25}$. The 2 transformed clones of *Agrobacteria*, P_DGUS3 and P_DGUS12, were then used to transform *Arabidopsis thaliana* (see below).

Construction of P_EGUS

The 2.6 kb sequence 5' to the *PHYE* gene (containing the putative promoter to the *PHYE* gene, P_E, and the first 21 amino acids) had originally been cloned into M13mp19 between the Hind III and Sal I sites. This sequence was cut out of M13mp19 and cloned into the pBI101.1 vector using the same restriction sites on the polylinker. This construct was designated P_EGUS (Fig. 2). The ligated DNA was used to transform competent cells of *E. coli* HB 101 by electroporation (59) and selected on LBkan_{50} plates. The insertion of the P_E sequence in the polylinker of the vector in the selected cells was confirmed by restriction digests using the enzyme BamHI.

The DNA from the transformants was used to transform *Agrobacterium tumefaciens* strain LBA 4404 cells by electroporation (59). Screening for transformants was on LB $kan_{50}str_{25}$ and yielded 2 positive clones, designated P_EGUS 1 and P_EGUS 2, which were then used to transform *Arabidopsis thaliana* (see below).

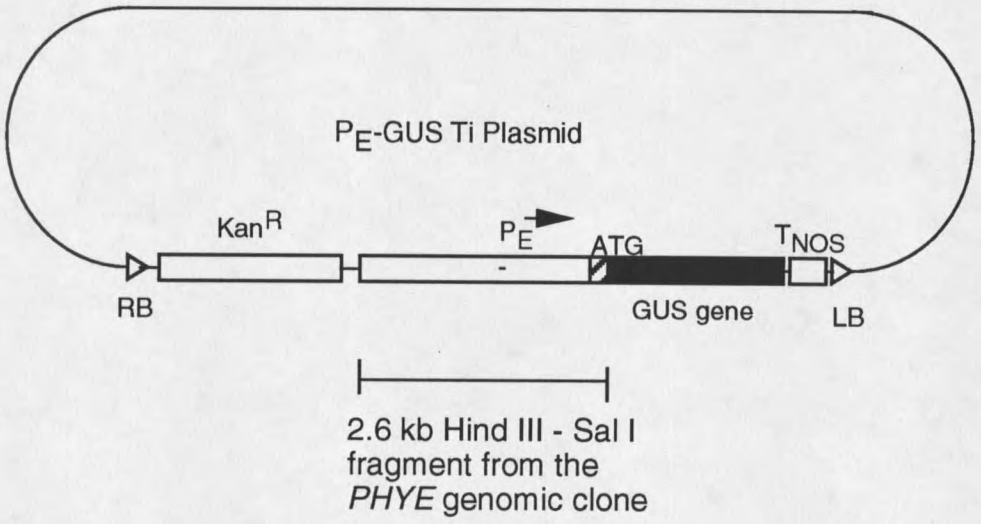

Fig. 2. Construction of the *PHYE*-GUS translational fusion. LB, left border; RB, right border; Kan^R, kanamycin resistance; P_E, *PHYE* gene promoter; T_{NOS}, nopaline synthase terminator.

Transformation of *Arabidopsis thaliana*

Transformation of Arabidopsis can be done using a variety of plant parts such as leaf discs, meristems, root bits, etc. (60). All transformations done for these projects used 2-week-old root bits as explants. Seeds of No-O WT were first subjected to a 70% ethanol wash for about 5 minutes and rinsed with sterile water. Then they were sterilized with 15% clorox/0.2% SDS solution for about 30 minutes with shaking. This was followed by 5 rinses with sterile water. The seeds were then transferred into 50 mls of liquid GM medium contained in 250 ml conical flasks and incubated under low light conditions at 100 rpm for 2 weeks. The medium was changed once a week. At the end of two weeks, the roots from the seedlings were harvested and placed individually on solid CIM plates. These plates were then sealed with air-permissive tape and incubated for 3 days. On the same day as roots were put on CIM (Callus-Inducing Medium) plates, the relevant strain of *Agrobacterium* containing the required construct was streaked out on LB $kan_{50}str_{25}$ plates and incubated at room temperature. After 2 days, a liquid culture of the streaked-out *Agrobacterium* was started in 2xYT $kan_{50}str_{25}$ medium and incubated at room temperature at 100 rpm. Thus, in 3 days' time, the explants and the *Agrobacterium* were both ready for the inoculation process. On the 4th day, the CIM-treated roots were taken out and cut into 0.5 cm

explants and stored in 20 mls of B5 medium. To this, 1 ml of the liquid *Agrobacterium* culture was added and the mixture shaken gently for 2 minutes. Following this, explants were blotted on sterile filter paper and transferred to new (solid) CIM plates in small bundles of about 5 rootbits each, in which they were incubated aerobically for 2 days.

After 2 days, the root segments bundles were lifted out carefully and washed with liquid B5 medium containing sterile carbenicillin (500 μg/ml). This wash was meant to kill all the bacteria adhering to the explants. After the antibiotic treatment, the root segments were once again blotted on sterile filter paper and transferred to solid SIM C_{500} (Shoot-Inducing Medium containing Carbenicillin at 500 μg/ml and kanamycin at 50μg/ml) plates. These plates were incubated aerobically in a plant tissue culture incubation chamber.

Plantlets that arise from transformed tissue began to appear after about a month of incubation. As they appeared, they were transferred to new solid SIM with a lower Carbeniclin content of 200 μg/ml. Healthy plantlets that formed rosettes were then transferred to solid GM media containing kanamycin at 50 μg/ml concentration (GM Kan_{50}) in tall magenta boxes and covered with two layers of Mira cloth held down by the lid of the magenta box. These plantlets were maintained for the rest of the life cycle in this

manner in the tissue culture incubator. In case of contamination appearing in the medium, the plantlets were transferred into rehydrated Jiffy 7 peat pellets (#703 Hummert) along with a plug of agar and watered regularly. Seeds were harvested from ripe siliques that these plantlets produced.

Analysis of Transformants

The T_2 seeds from the original transformants (the T_1 generation) (63) were sterilized with 15% chlorox/0.2% SDS with shaking for about 30 minutes and rinsed 5 times with sterile water. They were plated on solid GM plates containing kanamycin at 50 μg/ml concentration and placed at 4°C for 2 to 5 days. Then they were unwrapped and transferred to continuous light. The ratio of kanamycin-resistant to susceptible seedlings was counted on the 7th day from the date of transfer. After a couple more days, the resistant seedlings were transferred onto peat soil (Sunshine mixture) soaked with nutrient solution and covered with a layer of vermiculite. The same procedure was used for selection in all the transformations carried out and was repeated for the seeds of successive generations. Individual transgenic lines were selfed out to the T_3 or T_4 generation by which point they were no longer segregating kanamycin-sensitive progeny and were likely homozygous for the transgenes. Histochemical GUS assays were performed on the T_3 or T_4 progeny which are identified

in the text. For example, P_D12#1:1-1 is the T_4 progeny of a P_D12 original transformant.

Arabidopsis Transgenic Plants used as Control Lines

The plasmid pBI101.1 was used to transform *Arabidopsis thaliana* and the transformant was designated CTL 4. The plasmid pBI101.1 was the same as the one used to construct the *PHYD*-GUS and the *PHYE*-GUS fusions; in this case, the promoter-less GUS cassette was used as a negative control.

Three independent lines of *Arabidopsis* carrying a translational fusion between the upstream sequences of the *PHYB* gene (containing the first 63 amino acids along with the putative promoter sequences) and the β-glucuronidase gene (designated P_BGUS) had been developed earlier by Sharrock (unpublished) (Fig. 1). These were used to compare the expression of the *PHYB* gene promoter to *PHYD* and *PHYE* gene promoters.

Two independent lines of *Arabidopsis* carrying the plasmid pBI121 (which contains the *gus* gene under the control of the cauliflower mosaic virus 35S promoter) were also used to compare the expression of the *PHY* gene promoters to that of the constitutive CaMV 35S promoter.

Southern Analysis of Transformants

It has been observed in transgenic plants that sometimes more than one copy of the transgene gets inserted into the plant genome. The presence of multiple copies may

then affect the pattern or degree of expression of the transgene. Therefore, the transgenic lines obtained from transformation with the P_DGUS constructs were subjected to analysis by Southern blotting (62). Plant genomic DNA was extracted using the methods described by K. Edwards and co-workers (61). The P_DGUS genomic DNA was digested with EcoRI, fractionated on a gel and probed with a fragment from 5' flanking sequence of the *PHYD* gene (see Fig. 1). This analysis was done with the help of Ted Clack.

Histological GUS Staining

During the course of study, a wide variety of plant tissues from P_DGUS and P_EGUS transgenic plants were assayed for GUS activity. At different points of time, seedlings, rosette leaves, cauline leaves, portions of stem, inflorescence clusters, single flowers, etc. were stained either as whole tissue or as large sections as in the case of large rosette leaves. In all the assays performed, the same histological staining procedure was used, regardless of the type of transgenic (P_BGUS/P_DGUS/P_EGUS) and the plant tissue assayed.

Assay Protocol

After the excision of the plant tissue from the parent plant, it was first fixed in a solution containing 100mM NaH2PO4/Na2HPO4 buffer (pH 7.0), 0.1% formaldehyde, 0.1% Triton X-100 and 0.1% dithiotreitol for 15 minutes under

vacuum, and then rinsed for 5 minutes with shaking, first with 100mM NaH2PO4/Na2HPO4 buffer (pH 7.0), and then with the same buffer at 50mM concentration. The tissue was then infiltrated with the GUS substrate for 6 minutes under vacuum. The GUS substrate was made up with 50 mM NaH2PO4/Na2HPO4 buffer (pH 7.0), 0.5 mM potassium ferri-/ferrocyanide, 1.0 mM Na2EDTA and 1mg/ml X-gluc. The tissue in the X-gluc solution was incubated at 37° C for the required time period. At the end of the incubation, the substrate was removed and the tissue rinsed with the Na2HPO4/NaH2PO4 buffer, after which the tissue was fixed again, this time with 5% formaldehyde, 5% acetic acid and 20% ethanol for at least 2 hours (45). In order to observe the GUS staining patterns in tissues better, the tissue was clarified in increasing concentrations of ethanol, i.e., in successive steps with 50%, 70% and 95% ethanol. Observations were then made of the tissue with the aid of a dissecting microscope. The tissues were then stepped down through the decreasing concentrations of ethanol (95%, 70%, 50% and 25%) and transfered into 25% glycerol and finally into 50% glycerol. This was done to facilitate photography of the tissues.

Results

Construction of *PHYD*-GUS Transgenic Lines

The 5' flanking region of the *PHYD* gene that contains about 2.3 kb of upstream DNA and includes the start codon

was cloned from the λ D-6 genomic clone into the plasmid pBI101.1 in a translational fusion (Fig. 1). This construct was designated P_DGUS. The independent clones found to contain this construct were P_DGUS3 and P_DGUS12. These P_DGUS constructs were used to transform *Arabidopsis thaliana* root-bits, and transgenic plantlets were obtained.

The number of kan-positive original transformants were:

P_DGUS3 =20

P_DGUS12=10

Total =30.

Of these, some turned out to be chimaeric on the basis of their phenotype on kanamycin-containing media, and were neither carried through subsequent generations nor assayed. Thus, 6 P_DGUS3 lines and 3 P_DGUS12 lines were carried on. Of these, even in the T_3 and T_4 generations, some of the transgenic lines like P_DGUS3#12:1-1 were neither homozygous, nor showed the normal 3:1 segregation pattern on GMkan, expected when the transgene is inserted at a single locus (Table 2). Moreover, even the lines that appear to have only a single locus could be carrying more than one copy of the transgene at that locus because of tandem insertion. The copy numbers of all the transgenic lines of both P_BGUS and P_DGUS transgenics were therefore examined through Southern analysis.

Transgenic Line	Generation	Ratio (Kan^R/Kan^S)
P_BGUS#6	T3	35:0
P_DGUS3#1	T4	33:0
P_DGUS3#2	T3	30:17
P_DGUS3#5	T4	35:0
P_DGUS3#6	T4	29:0
P_DGUS3#10	T3	49:9
P_DGUS3#12	T4	30:1
P_DGUS12#1	T4	40:0
P_DGUS12#2	T4	35:1
P_DGUS12#8	T4	28:0
CTL4#2	T3	43:0

Table 2. Segregation ratios (kan^r/kan^s) of progeny of the 9 P_DGUS lines that were studied. It can be seen that while some of the lines appear to be homozygous, others show ratios that vary from the normal heterozygotic ratio of 3:1, indicating multicopy insertions in the genome. The ratios of the P_BGUS#6 line and the CTL4#2 line which were used as controls are also presented. The ratios give the actual number of seedlings evaluated.

Southern Analysis of the P_BGUS and P_DGUS Transgenic Lines

Southern blot analysis of genomic DNA from transgenic P_DGUS lines cut with EcoRI was performed. This enzyme cuts within the T-DNA so the copy number of the transgene can be estimated from the number of bands on the blot, using the endogenous *PHYD* promoter fragments as a single copy control (Fig. 3).

P_DGUS lines: Of the nine P_DGUS lines, P_DGUS3#5 and P_DGUS12#8 proved to have single copy insertions. P_DGUS3#1 contained 2 copies, and the lines P_DGUS3#2, P_DGUS3#6 and P_DGUS3#12 had 3 to 4 copies of the transgene. P_DGUS3#10 and P_DGUS12#1 had multiple copies. However, the line P_DGUS12#2 was found to lack detectable transgene sequences even though it was GUS positive. When the blot shown in Fig. 3 was stripped and rehybridized with a probe for the GUS gene, the single bands appeared for all lines as expected, but again no hybridization was observed for line P_DGUS12#2. The origin of the GUS activity in this strain is not known.

P_BGUS lines: The copy number of the P_BGUS transgene in the three lines used in these experiments were also determined by similar Southern analysis. Line P_BGUS#3:1 was found to carry a single copy of the transgene, while P_BGUS#6:1 and P_BGUS#7:5 had multiple copies (data not shown).