Plant Molecular Biology **20**: 589–600, 1992.

Isolation and characterization of a cDNA-clone coding for potato type B phytochrome

Arnd Heyer and Christiane Gatz*
*Institut für Genbiologische Forschung Berlin GmbH, Ihnestrasse 63, 1 Berlin 33, Germany (*author for correspondence)*

Received 5 February 1992; accepted in revised form 25 May 1992

Key words: phytochrome, *Solanum tuberosum*

Abstract

We have isolated and sequenced overlapping genomic and cDNA clones encoding the apoprotein of a potato phytochrome. Based on the deduced amino acid sequence, which shows 77% identity to the *Arabidopsis phyB* and 50% identity to the potato *phyA* open reading frame, we suggest that these clones encode *phyB* phytochrome. However, the size of the deduced open reading frame of 1133 amino acids is smaller than the size of the other two *phyB* open reading frames characterized so far in higher plants, which contain 1171 or 1187 amino acids. The intron/exon structure within the coding region is conserved in *phyA* and *phyB* genes of various species. Southern blot analysis indicates that potato *phyB* is a single-copy gene. *PhyB* mRNA levels do not differ among different organs or different light regimes. Transcription initiation starts from two different start points which are 63 bp apart.

Introduction

Phytochrome is the best characterized of the regulatory photoreceptors that control plant development in response to light [24, 26, 40]. The molecule is a chromoprotein, consisting of two subunits with a molecular mass of 120–127 kDa depending on the plant species [46]. Each subunit contains a linear tetrapyrrole chromophore covalently attached to a cysteine residue [25]. In dark-grown plants phytochrome is synthesized as the physiologically inactive Pr form, which is converted upon illumination with red light (λ_{max}:660–670 nm) to the active Pfr form. The Pfr form can be transformed back to the Pr form by far-red light (λ_{max}: 730 nm). Though a considerable amount of data is available on the phytochrome protein, very little is known on how it enables plants to adapt such diverse processes as the determination of hypocotyl, petiole and internode length, the formation of leaf primordia, the development of plastids and the induction of flowering to given light conditions [34].

Immunological [39, 45], spectroscopic [45], genetic [23] as well as molecular studies have established that plants contain more than one type of phytochrome. At least three different phytochrome polypeptides are encoded by a group of divergent *phy* genes [38]. The most abundant species of phytochrome in plants is the molecule

found in etiolated tissues, often referred to as type 1 or type A phytochrome, which is encoded by *phyA*. In retrospect it is now clear that the vast majority of molecular and biochemical phytochrome research conducted both *in vivo* and *in vitro* has involved characterization of phytochrome A [32]. A characteristic feature of type 1 phytochrome is that it is rapidly degraded in the light. In contrast, relatively little is known about the molecular properties of the other phytochrome types. Extracts of fully green oat tissue contain a different type of phytochrome (type 2 phytochrome) that has different spectral properties and is more stable *in vivo* in the presence of light than the etiolated-tissue form [39, 45].

The possibility that the different types of phytochrome possess unique roles in plant development is only beginning to be explored [1, 42, 47]. One promising approach involves the generation and analysis of mutants with altered photomorphogenic properties [23]. By taking advantage of *Arabidopsis* genetics [29], two independent laboratories were able to show that it is possible to obtain mutants that are deficient for a specific molecular species of phytochrome. The *hy3* mutant of *Arabidopsis* for instance, shows reduced levels of phyB product, whereas phyA protein levels appear to be normal [31, 41].

A useful addition to classical genetics involves the transfer of defined phytochrome genes into the genome of homologous or heterologous plants. Thus the amount of one specific type of phytochrome can be either enhanced by expressing the gene under the control of a strong promoter [6, 7, 20, 22], or conversely inhibited by expressing specific 'anti-sense' RNAs. We are using this strategy for the analysis of the phytochrome system of potato. The complexity of the genome of this important crop plant renders classical genetics rather complicated, but the availability of well established transformation and regeneration protocols [12] makes 'reverse genetics' feasible. The onset of tuberization [4] as well as the growth and the size distribution of tubers depend on the prevailing light conditions [43, 44]. Like the majority of plants grown as crops, potato falls into the shade-avoiding category. Shade, which can be perceived by phytochrome, results in increased shoot/root ratios and soluble sugar/starch ratios, thus leading to a decrease in deposition of photoassimilates into storage organs [21]. We are interested in defining the role of the different phytochromes in these processes. This might enable us to influence carbon partitioning and thus potato yield by manipulating one of the key regulatory photoreceptors.

As a first step, we have isolated different phytochrome genes from potato. Recently we have reported the isolation and characterization of a complete cDNA encoding type A phytochrome [17]. This cDNA has been used to isolate a related phytochrome gene, which shows a different expression pattern and exhibits similarities to the rice [9] and *Arabidopsis phyB* gene [38].

Materials and methods

Plant material

Commercial potato (*Solanum tuberosum* L.) cv. Désirée was obtained through Vereinigte Saatzuchten (3112 Ebsdorf, FRG). Potato line AM 80/5793 is a homozygous diploid derivative of the tetraploid *S. tuberosum* cv. Désirée and was kindly provided by the Max Planck Institut für Züchtungsforschung, Köln, FRG. Seeds were germinated for 10 days in complete darkness or for 10 days in cycles of 16 h light and 8 h dark on MS medium without sucrose. Tubers were allowed to sprout for 10 days either in the dark or under greenhouse conditions. Leaves were harvested from plants grown in the greenhouse. For re-etiolation experiments plants were kept in complete darkness for 4 days. Plant material that had been kept in the dark was harvested under dim green safelight [3]. Tissue for RNA isolation was frozen in liquid nitrogen and stored at −80 °C.

Recombinant DNA techniques

Standard procedures were used for recombinant DNA work [28].

Isolation of phyB *cDNA and genomic clones*

An EMBL3 genomic library from potato AM80/5793 was provided by Dr Xian Liu and Dr L. Willmitzer (IGF, Berlin, FRG). This library was screened with a radioactively labelled DNA fragment (Multiprime labelling system, Amersham Buchler, Braunschweig, FRG) spanning the region encoding amino acid (aa) 332 to aa 695 of the potato *phyA* gene [17]. Hybridization was carried out at 42 °C in hybridization buffer containing 30% formamide [37]. Washing was performed twice for 15 min in 4 × SSC at 37 °C. A λZAPII cDNA library from *S. tuberosum* cv. Désirée leaves was provided by Dr J. Kossmann and Dr L. Willmitzer (IGF). The library was screened with a radioactively labelled DNA fragment of the *phyB* genomic clone, which encodes aa 286 to 684. Hybridization was carried out at 42 °C in hybridization buffer containing 50% formamide. Washing was performed three times for 20 min in 2 × SSC, 0.5% SDS at 65 °C.

Sequencing

Suitable restriction fragments were subcloned into pUC18 and sequenced according to the dideoxy chain-termination method [5] using the T7 Polymerase Sequencing Kit from Pharmacia, Uppsala, Sweden. Templates were primed with M13 universal or reverse primers (Boehringer, Mannheim, FRG) or *phyB* -specific oligonucleotides synthesized using a DNA synthesizer. Computer analysis of DNA sequences was performed using programs made available by the University of Wisconsin Genetics Computer Group [10].

Northern blot analysis

Northern blot analysis was performed as described [17].

Southern blot analysis

High-molecular-weight DNA was prepared from cv. Désirée plants and potato AM80/5793 plants [33]. Genomic DNA was digested with various restriction enzymes and the resulting DNA fragments were separated electrophoretically. Subsequently the gel was incubated for 25 min in 250 mM HCl. DNA fragments were transferred in 10 × SSC to Hybond N membranes (Amersham/Buchler, Braunschweig, FRG). Hybridization with the radioactively labelled internal probe was carried out at 42 °C in hybridization buffer containing 30% formamide [37]. Washing was performed twice for 15 min in 4 × SSC at 37 °C. Hybridization with the radioactively labelled border probe was carried out at 42 °C in hybridization buffer containing 50% formamide [36]. Washing was performed three times for 20 min in 2 × SSC, 0.5% SDS at 65 °C.

Primer extension analysis

10 μg of poly(A)$^+$ RNA were annealed to 1 pmol of α-^{32}P-5′-end-labelled primer in 40 mM PIPES, pH 6.4, 400 mM NaCl, 1.25 mM EDTA, 30% formamide for 12 h at 30 °C. The sequence encoding aa 1 to 10 of the *phyB* open reading frame was used to design the oligonucleotide (GGAATGCTTTGTTCTACTTCCAGAAGCCAT). Reverse transcription reactions were done as described [2]. Primer extension products were analysed on a 6% sequencing gel, using a sequencing reaction performed with the primer as a size marker.

Results

Isolation and structure of potato phyB *clones*

A subclone of the potato *phyA* gene encoding aa 332 to 695 (λ6) was chosen to screen an EMBL3 library constructed from genomic potato DNA. As the corresponding region of the *Arabidopsis type A* phytochrome shows 62% amino acid identity to the *Arabidopsis* type B phytochrome it seemed likely that the potato *phyB* gene could be identified with the λ6 fragment under conditions of low stringency. Members of a class of weakly

hybridizing clones were analysed by digestion with *Eco* RI and *Hin*d III and subsequent Southern blot analysis. A 1200 bp *Hin*d III fragment of one of those clones hybridized to the λ6 fragment. Sequence analysis of this fragment revealed high sequence similarity to *phyB* from *Arabidopsis* [38]. Thus, it was used to screen a λZAPII cDNA library, constructed using poly(A)$^+$ RNA prepared from green cv. Désirée leaves. After screening of 200 000 plaques two hybridizing clones (J-1, J-22) were isolated and found to have identical sequence to each other and to the genomic clone. Both cDNA clones were fully sequenced and found to be incomplete. J-22 encodes 666 carboxy terminal amino acids of the *phyB* open reading frame and a 3′ untranslated region of 517 bp. To exclude that the long 3′ untranslated leader was due to a cloning artefact during the construction of the cDNA library, we verified by restriction analysis that it was also present in the genomic *phyB* clone (data not shown). This analysis also revealed that there was no intron in this region. J-1 is a partial cDNA clone which encodes aa 6 to 702 of the *phyB* open reading frame. The missing sequence from the 5′ end was obtained from the homologous *phyB* genomic clone. Figure 1 shows the restriction maps of the two cDNA clones and the genomic clone, as well as the location of the introns, which were identified by comparing the sequence of the genomic clone with the sequence of the cDNA clones. As the introns were not fully sequenced, their size was determined by restriction analysis.

Sequence of the phyB *gene*

Figure 2 shows the combined nucleotide sequence and full-length derived amino acid sequence of the *phyB* open reading frame, 997 bp of the genomic sequence preceding the ATG and 517 bp of the 3′ untranslated region. Alignment of the complete deduced potato *phyB* amino acid sequence with all available phytochrome sequences from oat [15, 16], zucchini [27], pea [35], rice [9, 19], maize [8], potato [17], *Arabidopsis* [38] and *Selaginella* [14] allows the construction of a phylogenetic tree as shown in Fig. 3. Like *Arabidopsis phyB*, potato *phyB* is in 74% of the amino acid positions identical to the rice *phyB*, confirming the observation from Dehesh and Quail [9] that monocot and dicot *phyB* sequences have a higher value of identity (74%) than monocot and dicot *phyA* sequences (65%). Potato *phyB* shows 78% amino acid identity to *Arabidopsis phyB*. *PhyB* sequences of higher plants reveal a 61–62% (rice, *Arabidopsis*) or 65% amino acid (potato) identity to the phytochrome of the spike moss *Selaginella*.

The nucleotide sequence encodes a predicted polypeptide of 1133 aa (125 kDa). The numbers of amino acids of phytochrome open reading

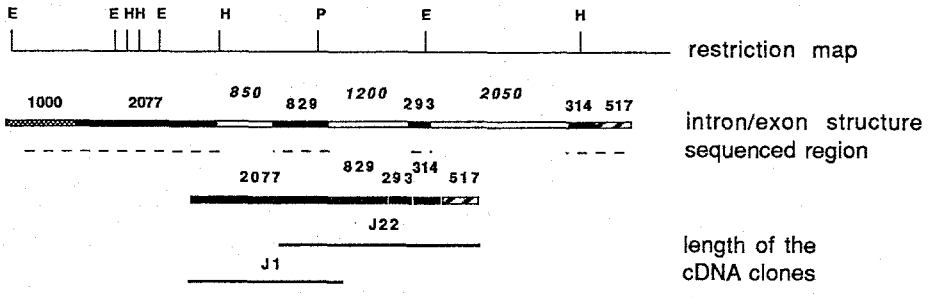
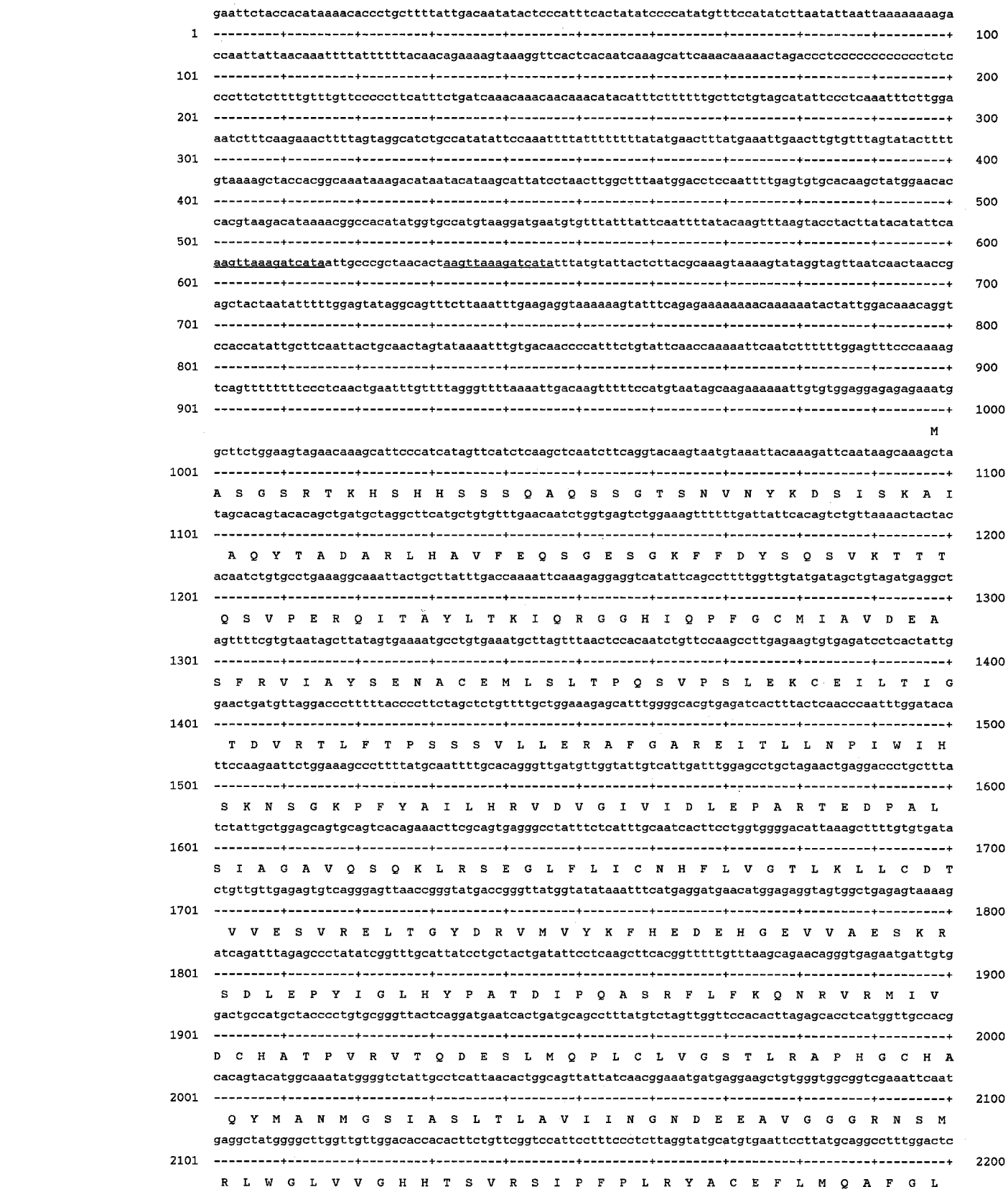


Fig. 1. Physical map of the potato *phyB* gene. The location of the recognition sites for *Eco* RI (E), *Hin*d III (H) and *Pst* I (P) on the genomic clone are indicated in the restriction map. To illustrate the gene structure, exons are shown in black and introns in white. The grey box at the 5′ end of the gene indicates that the intron/exon structure in this region is not known. The shaded box at the 3′ end of the gene indicates the 3′-untranslated region which contains no intron. Numbers indicate the number of bp. The dotted lines indicate the sequenced region of the gene. The length of the two partial cDNA clones J22 and J1 is indicated by thick lines.



```
caattgaacatggagttgcaattggcgtcacagttgtctgagaaacatgttttaaggacacaaacactgttatgtgacatgctccttcgagactctccac
2201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2300
Q  L  N  M  E  L  Q  L  A  S  Q  L  S  E  K  H  V  L  R  T  Q  T  L  L  C  D  M  L  L  R  D  S  P  P
cggggattgttacccaaagccccagtattatggaccttgtgaagtgcgatggtgctgctctatactaccaggggaagtactatccattaggcgtaacacc
2301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2400
  G  I  V  T  Q  S  P  S  I  M  D  L  V  K  C  D  G  A  A  L  Y  Y  Q  G  K  Y  Y  P  L  G  V  T  P
aactgaagctcagataaaagacattgtggagtggttattggcttaccatggagactcaacaggtttaagtactgacagtttgcctgatgctgggtatcct
2401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2500
 T  E  A  Q  I  K  D  I  V  E  W  L  L  A  Y  H  G  D  S  T  G  L  S  T  D  S  L  P  D  A  G  Y  P
ggagcagcttcacttggtgatgcagtttgtggtatggctgtcgcttatataacttctaaagatttcttgttttggtttcgctcccacacagcgaaagaaa
2501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2600
G  A  A  S  L  G  D  A  V  C  G  M  A  V  A  Y  I  T  S  K  D  F  L  F  W  F  R  S  H  T  A  K  E  I
taaagtggggtggtgcaaagcatcatcctgaagacaaggatgacggacagagaatgcatccacgttcttccttcaaggcatttctggaagttgttaaaag
2601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2700
  K  W  G  G  A  K  H  H  P  E  D  K  D  D  G  Q  R  M  H  P  R  S  S  F  K  A  F  L  E  V  V  K  S
tcggagctcaccatgggaaaatgccgaaatggatgcaatccactctttgcagctaattctgcgagattcatttaaggatgctgaggcaagtaattctaag
2701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2800
 R  S  S  P  W  E  N  A  E  M  D  A  I  H  S  L  Q  L  I  L  R  D  S  F  K  D  A  E  A  S  N  S  K
gctattgtgcatgctcatcttggggaaatggagttgcaagggatagatgaactgagttctgttgccagagaaatggttagattgatcgaaactgcaacag
2801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2900
A  I  V  H  A  H  L  G  E  M  E  L  Q  G  I  D  E  L  S  S  V  A  R  E  M  V  R  L  I  E  T  A  T  A
ctcccatatttgctgttgatgttgaaggtcgcataaatgggtggaatgcaaaggtcgctgaattgacaggtgtatcagttgaagaagcaatggggaagtc
2901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3000
  P  I  F  A  V  D  V  E  G  R  I  N  G  W  N  A  K  V  A  E  L  T  G  V  S  V  E  E  A  M  G  K  S
cttggttcatgatcttgtgtacaaagaatcacaggagactgctgagaagcttctgtacaatgctctaagaggcgaggaagataaaaatgtagaaataaag
3001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3100
 L  V  H  D  L  V  Y  K  E  S  Q  E  T  A  E  K  L  L  Y  N  A  L  R  G  E  E  D  K  N  V  E  I  K
ttgaggacatttggagctgaacaactggagaaagctgtttttgtggtggttaatgcttgcgctagagattacacaaacaacattgttggtgtttgctttg
3101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3200
L  R  T  F  G  A  E  Q  L  E  K  A  V  F  V  V  V  N  A  C  A  R  D  Y  T  N  N  I  V  G  V  C  F  V
ttgggcaggatgttactggggaaaaagttgttatggacaagtttattaacatccaaggtgattacaaggccattgtgcacagccccaatcctctgatccc
3201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3300
  G  Q  D  V  T  G  E  K  V  V  M  D  K  F  I  N  I  Q  G  D  Y  K  A  I  V  H  S  P  N  P  L  I  P
tccaatatttgcatcagatgagaacacttgttgctccgagtggaacactgccatggaaaaactcactggttggtctagaggggagattgttggaaaaatg
3301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3400
 P  I  F  A  S  D  E  N  T  C  C  S  E  W  N  T  A  M  E  K  L  T  G  W  S  R  G  E  I  V  G  K  M
ttagttggtgagatttttggaagttgttgtcgcctcaagggcccagatgccatgacaaagttcatgatcgtgttgcataatgcaattggaggccaggata
3401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3500
L  V  G  E  I  F  G  S  C  C  R  L  K  G  P  D  A  M  T  K  F  M  I  V  L  H  N  A  I  G  G  Q  D  T
cagacaagtttccattttccttttttgaccgaaatgggaaatatgtgcaagctcttttgactcgaaacaagagagttaatatggagggcgatactattgg
3501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3600
  D  K  F  P  F  S  F  F  D  R  N  G  K  Y  V  Q  A  L  L  T  R  N  K  R  V  N  M  E  G  D  T  I  G
ggctttctgtttcatacagatagccagtcccgaactgcagcaagctctaagagttcaaaggcaacaggaaaagaagtgctattctcagatgaaagagctg
3601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3700
 A  F  C  F  I  Q  I  A  S  P  E  L  Q  Q  A  L  R  V  Q  R  Q  Q  E  K  K  C  Y  S  Q  M  K  E  L
gcatacatttgtcaggaaataaaaagtcctcttaatggtatacgctttacaaattcattgttggaggccacaaatttgacagaaaatcagaagcagtatc
3701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3800
A  Y  I  C  Q  E  I  K  S  P  L  N  G  I  R  F  T  N  S  L  L  E  A  T  N  L  T  E  N  Q  K  Q  Y  L
tagagacaagtgctgcttgtgagaggcagatgtctaagatcattagggatattgatctggaaaacattgaggacgggtcactgacccttgagaaagaaga
3801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3900
  E  T  S  A  A  C  E  R  Q  M  S  K  I  I  R  D  I  D  L  E  N  I  E  D  G  S  L  T  L  E  K  E  D
ttttttttcttgggagtgtaatagatgctgttgttagccaagtgatgttattgctgagggaaaaaggcgtgcagttaatccgtgatataccagaggaaatt
3901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4000
 F  F  L  G  S  V  I  D  A  V  V  S  Q  V  M  L  L  L  R  E  K  G  V  Q  L  I  R  D  I  P  E  E  I
aagacattaacagtacatggtgatcaagtgagaattcaacaggtcttggcagatttcttgttgaacatggtacggtatgcaccatcacctgatgggtggg
4001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4100
K  T  L  T  V  H  G  D  Q  V  R  I  Q  Q  V  L  A  D  F  L  L  N  M  V  R  Y  A  P  S  P  D  G  W  V
```

```
     tagaaatccaacttcgaccaagtatgatgccaatatctgatggagtaactgttgtgcatattgaactaggattatatgcgccgggaaggcttcctcctga
4101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4200
       E  I  Q  L  R  P  S  M  M  P  I  S  D  G  V  T  V  V  H  I  E  L  G  L  Y  A  P  G  R  L  P  P  E
     attggttcaagacatgttccacagcagtcggtgggtaactcaggaaggcctaggactgagcatgtgcagaaaaatgttaaagcttatgaatggagaaatc
4201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4300
       L  V  Q  D  M  F  H  S  S  R  W  V  T  Q  E  G  L  G  L  S  M  C  R  K  M  L  K  L  M  N  G  E  I
     cagtatatcagagaatcagaaagatgctatttcctgattatccttgacctgccaatgacccgcaaaggtccaaagagtgttggctaggagcttacagcac
4301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4400
       Q  Y  I  R  E  S  E  R  C  Y  F  L  I  I  L  D  L  P  M  T  R  K  G  P  K  S  V  G  *
     acagtaagagaagaacctgtaatatctcttaaattgcatgctggtgaaattctggtaatgtgtctgatcctttaatggaacttgacacaggagagagcaa
4401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4500
     gtatgggggattgatcggcatcattctaacgctattgttgcgacttgtatttgctggcatcttttacaactctgaactagtattctctcagcatgtgcaa
4501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4600
     gttgagaaagctgttgctccttcgtgtatgtttttaccaaagattaattcttgacatagtaagtgctaatggccaaataattggaaaatggttagcatga
4601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4700
     ttccaaattacttgatcccagagctgttgtgtgtagcataacaggtgtattgttgtataatactagtatgtattaaaaggtgggaataaatgtcatcctt
4701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4800
     tccatggtctggatgagttggtgtctcccactctttgtaacattaggatttgtgtaacatgaacatactgagaatcttatatgactaattagctcctctt
4801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4900
     ggtg
4901 ----  4904
```

Fig. 2. Nucleotide and derived amino acid sequence of potato *phyB*. The amino acid sequence is presented below the nucleotide sequence in single-letter code. Black triangles mark the position of the introns. Two 15 bp direct repeats (AAGTTAAAGATCATA) in the potential promoter region are underlined, as well as the polyadenylation signal AATAAA in the 3′-untranslated region.

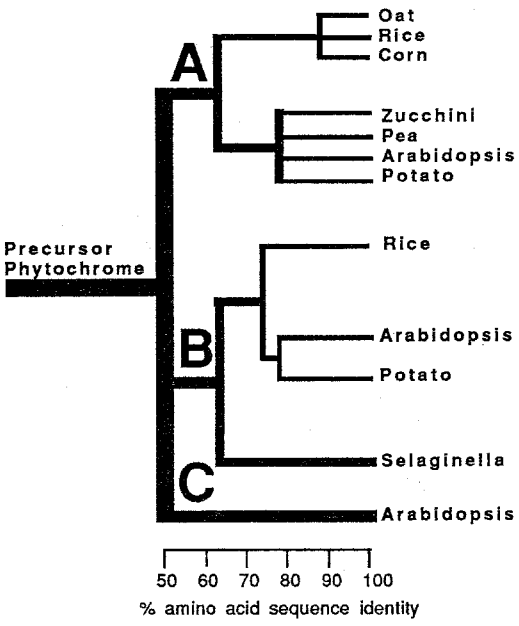


Fig. 3. Deduced phylogeny of phytochrome polypeptides. Amino acid sequence identity (%) between each pairwise comparison of the various phytochromes was used to group related sequences as a measure of phylogenetic distances in constructing the tree.

frames from various species are compiled in Table 1. The higher molecular weight of the two *phyB* polypeptides from rice and *Arabidopsis* is mainly due to a hydrophilic N-terminal extension which is missing in all other phytochromes.

Within the coding region the location of the introns in *phyA* and *phyB* genes of monocots, dicots and the spike moss *Selaginella* are strictly conserved. So far the genomic clones of *phyA* genes of the three monocots maize [8], oat [16] and rice [19] and the dicot pea [35] have been

Table 1. Number of amino acids of the different *phy* open reading frames.

	phyA	*phyB*	*phyC*
Oat	1128		
Rice	1126	1171	
Maize	1126		
Zucchini	1123		
Pea	1124		
Arabidopsis	1130	1187	1125
Potato	1123	1129	
Selaginella		1123	

characterized. All of these contain an intron in the 5′-untranslated region. As the cDNA clone J-1 does not cover the 5′-untranslated region, we do not know whether *phyB* from potato contains an intron 5′ of the start codon of the open reading frame. The intron found in the 3′-untranslated region of monocot *phyA* genes are not found in genomic clones from dicot *phyA* or *phyB* clones or in the *Selaginella* phytochrome gene.

Southern blot analyses are indicative of a single phyB *gene*

Potato genomic DNA from the tetraploid *S. tuberosum* cv. Désirée as well as from the homozygous diploid cv. AM80/5793 was digested with three different restriction enzymes and subjected to Southern blot analysis (Fig. 4). Using an internal restriction fragment as a probe the homologous fragments showed the strongest hybridization signals. As hybridization was carried out under less stringent conditions, *phyA* restriction fragments (780 bp *Eco* RI fragment and 895 bp *Hin*d III fragment) lighted up, as well as other bands, again indicating the presence of as yet unidentified phytochrome genes in the potato genome. The detection of only a single band with a border probe fragment under stringent conditions reveals, that the *phyB* gene characterized here is a single-copy gene.

PhyB *is constitutively expressed in different organs*

We used a probe derived from sequences 3′ of the *Pst* I site of the cDNA clone J-22 to determine the pattern of *phyB* expression in potato. This probe was hybridized to a northern blot of poly(A)$^+$ RNA isolated from different organs of potato, grown either in complete darkness or under a 16 h light and 8 h dark regime. As shown in Fig. 5, *phyB* is constitutively expressed in all organs, irrespective of the light conditions. This contrasts with the expression profile of *phyA*, which shows highest levels of expression in dark-grown seedling and sprouts. To detect *phyB* signals, we had

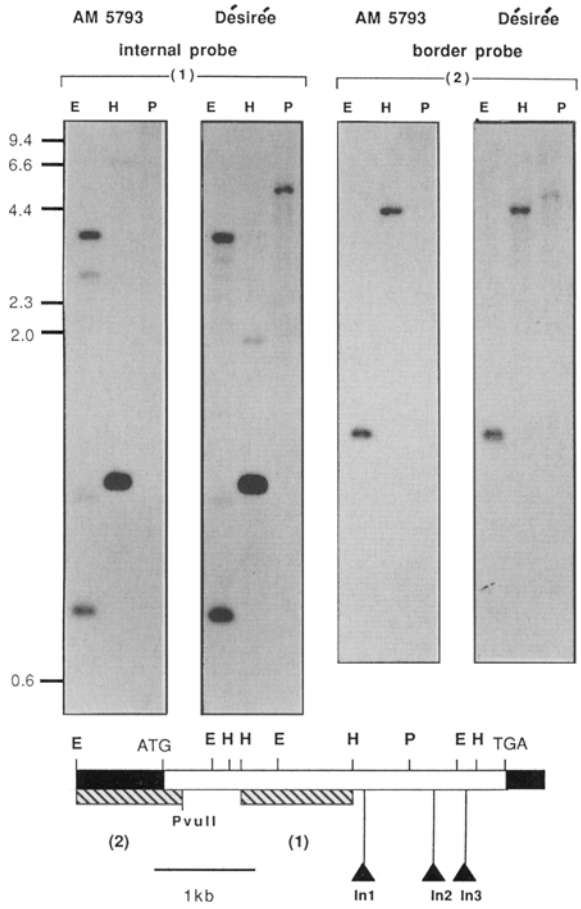


Fig. 4. Southern blot analysis. Total DNA from two different potato lines (Désirée and AM5793) was digested with *Eco* RI (E), *Hin*d III (H) and *Pst* I (P). Hybridization with an internal probe (1) was performed using a *Hin*d III fragment (1194 bp) which includes the chromophore binding site. Hybridizing genomic *Eco* RI and *Hin*d III fragments have the same size as the corresponding cDNA restriction fragments. Hybridization with a border probe (2) was performed using a 1040 bp *Eco* RI-*Pvu* II fragment. The positions of the probes within the combined cDNAs are shown in the diagram below the blots. Untranslated sequences are shown in black.

to expose the blot ten times longer to an X-ray film than we did for the detection of *phyA* signals, indicating that at least in etiolated seedlings and sprouts *phyA* transcripts are more abundant than *phyB* transcripts.

10 μg of the poly(A)$^+$-RNA from green leaves were used for a primer extension experiment in order to determine the length of the 5′-untranslated leader as well as the number of transcrip-

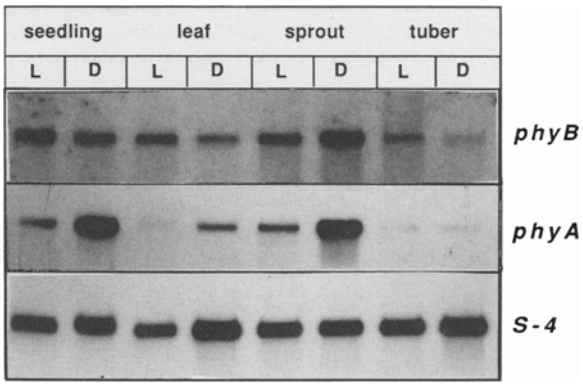


Fig. 5. Northern blot hybridization of the *phyB* mRNA present in poly(A)$^+$ RNA from different potato organs kept under different light regimes (see Materials and methods; D, dark; L, light). The expression pattern of *phyA* is shown below [17]. The blot was also hybridized with a restriction fragment encoding the potato ribosomal protein S-4 [11] in order to show that comparable amounts of RNA were loaded.

tional start sites. The primer was designed to be complementary to the nucleotide sequence encoding the first 10 amino acids of the *phyB* open reading frame. As shown in Fig. 6, two different primer extension products were detected, indicating the presence of two different 5′ termini in transcripts of the *phyB* gene. Comparison with the sequence ladder generated with the primer and the genomic clone revealed that these termini are located 73 and 126 nucleotides upstream from the initiation codon ATG. Thus the *phyB* mRNAs are either 3980 bp or 4043 bp long. This finding corresponds well with the lengths seen on northern blots, which show that the *phyB* mRNA runs slightly higher than the *phyA* mRNA (3614 bp).

Discussion

We have isolated and characterized two partial cDNA clones and a genomic clone for a potato phytochrome gene. Taking amino acid sequence identity as a criterion, phytochrome genes have been grouped into different classes (types A, B and C), which share 50% sequence identity [38]. Based on the deduced amino acid sequence of the new potato clones, which shows 77% amino acid identity to the *Arabidopsis phyB* open reading

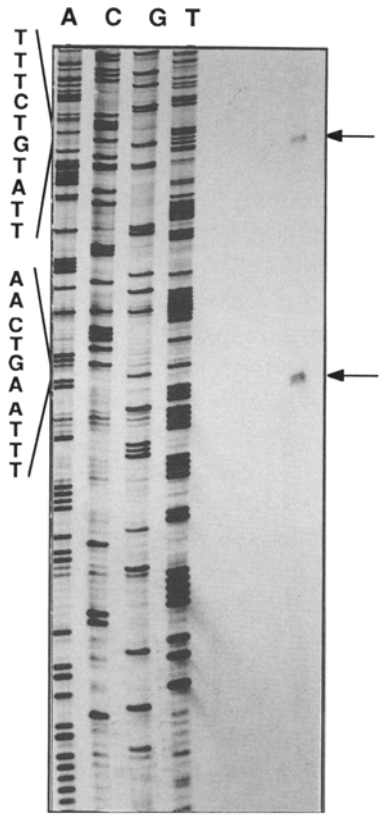


Fig. 6. Primer extension analysis of the 5′ termini of *phyB* mRNAs. The primer, which was designed to be complementary to the nucleotide sequence encoding the first 10 amino acids of the *phyB* open reading frame, was used for the sequencing reaction shown on the left. The sequence was used to determine the length of the untranslated region. The arrows point to the two resulting primer extension products.

frame [38], we have classified these clones as potato *phyB* phytochrome clones. Like the *Arabidopsis phyB* gene it shows 74% identity to the rice *phyB* [9] gene, which contrasts with the finding for *phyA* genes which are only 65% identical between monocots and dicots. Especially the carboxy terminal domains of the *phyB* genes (residues 800 to the very C-terminal amino acid) of monocots and dicots show a higher degree of identity (67%) than the respective *phyA* genes (47%). Thus *phyB* genes seem to diverge more slowly than *phyA* genes, which might be due to stricter structural constraints at the carboxy terminus of the molecule. On the other hand, the 77% amino acid identity between potato and

Arabidopsis phyB indicates that within dicots *phyB* genes diverge to the same extent as *phyA* genes. This high level of divergence seems unexpected in view of a slower evolution of *phyB*.

Another unexpected feature of the potato *phyB* sequence relates to the amino terminal part of the molecule. *Arabidopsis phyB* and rice *phyB* both contain a strongly hydrophilic amino-terminal extension, which is lacking in all *phyA* genes, and was therefore assumed to be characteristic of *phyB*. This extension is missing in the potato *phyB* gene. Hanelt *et al.* provided the phytochrome sequence of the spike moss *Selaginella* to the EMBL data bank [14]. *Selaginella* is a member of the Pteridophyta, which appeared earlier during evolution than Spermatophyta. *Selaginella* phytochrome is more closely related to *phyB* of higher plants than to *phyA* or to *phyC*. It shows 65% amino acid identity to potato *phyB* and 61% and 62% amino acid identity to the respective genes from rice and *Arabidopsis*. *Selaginella* phytochrome also does not contain the amino terminal extension, which is present in the monocot rice and the dicot *Arabidopsis*, but not in the dicot potato. It could well be that rice *phyB* and *Arabidopsis phyB* might have acquired the extension independently, because of no structural constraints for the very amino terminus of the phyB polypeptide. However, it cannot be excluded either that the *phyB* gene, after having already diverged from the *phyA* and *phyC* genes, underwent another gene duplication which gave rise to divergent *phyB* genes. The Southern blot (Fig. 4) does not contradict this interpretation, because under conditions of low stringency, bands that hybridize stronger than the 780 bp *Eco* RI restriction fragment and the 895 bp *Hin*d III restriction fragment, which are specific for *phyA*, can be identified. These bands might represent DNA sequences that are more related to *phyB* than to the other classes. However, sequencing of more *phyB* genes from both monocots and dicots, as well as the characterization of as yet unidentified phytochrome genes in *Arabidopsis*, rice and potato are necessary to support the assumption of divergent *phyB* genes being present in the genome of higher plants.

The introns within the coding region of the potato *phyB* gene are located in the same positions as within the *phyA* genes from higher plants [8, 15, 35] and *Selaginella phy* [14]. Generally, the introns of the potato gene are larger than those of the other species. In contrast, the intron located within the 3′ untranslated leader of monocot *phy* genes [15, 19] is neither found in the *Pisum phyA* gene [35], nor in potato or *Selaginella phyB* [14]. Thus it seems likely that this intron was acquired by monocots shortly after the divergence of monocots and dicots. As our cDNA clones were not complete, we do not know at the moment, whether the intron in the 5′ untranslated region, which is present in all *phyA* genes is also present in *phyB* genes.

At the RNA level, phyB shows a different expression pattern than *phyA* (Fig. 5). *phyA* mRNA levels vary in different organs, being highest in etiolated sprouts and seedlings, where the protein is involved in triggering the light-induced developmental switch from heterotrophic to phototrophic growth. *phyB* expression is not significantly altered in different organs or under different light regimes. As *phyB*-specific antibodies, generated using a recombinant *phyB* gene product as an antigene, recognize a light-stable phytochrome pool in *Arabidopsis* [31, 41], it can be assumed that *phyB* genes in general encode at least part of the *typeB* or type 2 phytochrome pool, being present in green tissues of higher plants. Whereas *phyA* senses the onset of light and is then degraded, a stable phytochrome can constantly perceive the given light conditions by adjustment of the photoequilibrium between the Pfr and the Pr form.

Using primer extension analysis we have defined two mRNA start points, which are located 73 and 126 nucleotides upstream from the initiation codon ATG. Taking into account that our cDNA clones were incomplete and that the genomic phytochrome clones characterized so far (*phyA* from rice [19], oat [16], maize [8] and pea [35]) all contain an intron in the 5′ untranslated region, we are unable to decide whether the sequences preceding the ATG contain the promoter. Assuming no intron in the potato *phyB*

gene there is only a TATA-box-like sequence (TATAAAA) 23 bp upstream of the 126 nt start point, but no similar sequence upstream of the 73 nt start point. Two 15 bp direct repeats (AAGTTAAAGATCATA) are located 184 and 214 nt upstream of the putative TATA box. If these sequences have any relevance to the expression of the gene remains to be determined.

Having isolated and characterized *phyA* and *phyB* phytochrome genes in potato, we are now in the position to determine the functions of these genes *in vivo*. We are approaching this question by expressing phytochrome 'sense' and 'anti-sense' RNAs in transgenic potato plants. We are using a tightly repressible promoter that can be activated with low levels of tetracycline [12]. If a strong inhibition of the expression of *phy* 'sense' or 'anti-sense' RNAs interferes with the regeneration process, this promoter might turn out to be advantageous. In addition, the combination of this regulated promoter with phytochrome sequences allows to define the role of the different phytochromes at different stages of the plant's life cycle.

Acknowledgements

We thank Regina Wendenburg for excellent technical assistance and Drs J Kossmann, Xinag Liu and L. Willmitzer for potato libraries. This work was supported by a grant from the Bundesministerium für Forschung und Technologie (BCT 0389).

References

1. Adamse P, Kendrick RE, Koornneef M: Photomorphogenetic mutants of higher plants. Photochem Photobiol 48: 833–841 (1988).
2. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (eds) Current protocols in molecular biology, Vol. 1, chapter 4.8.1, New York: John Wiley and Sons (1989).
3. Bolton GW, Quail PH: Cell-free synthesis of phytochrome apoprotein. Planta 155: 212–217 (1982).
4. Batutis EJ, Ewing EE: Far-red reversal of red light effect during long-night induction of potato (*Solanum tuberosum* L.) tuberization. Plant Physiol 69: 672–674 (1982).
5. Biggin MD, Gibson TJ, Hong GF: Buffer gradient gels and 35S label as an aid to rapid DNA sequence determination. Proc Natl Acad Sci USA 80: 3963–3965 (1983).
6. Boylan MT, Quail PH: Oat phytochrome is biologically active in transgenic tomatoes. Plant Cell 1: 765–773 (1989).
7. Boylan MT, Quail PH: Phytochrome A overexpression inhibits hypocotyl elongation in transgenic *Arabidopsis*. Proc Natl Acad Sci USA 88: 10806–10810 (1991).
8. Christensen AH, Quail PH: Structure and expression of a maize phytochrome-encoding gene. Gene 85: 381–3935 (1989).
9. Dehesh K, Tepperman J, Christensen AH, Quail PH: *PhyB* is evolutionarily conserved and constitutively expressed in rice seedling shoots. Mol Gen Genet 225: 305–313 (1991).
10. Devereux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programs for the VAX. Nucl Acids Res 12: 387–395 (1984).
11. Devi KRG, Chan Y-L, Wool IG: The primary structure of rat ribosomal protein S4. Biochim Biophys Acta 1008: 258–262 (1989).
12. Eckes P, Schell J, Willmitzer L: Isolation and characterization of a light inducible, organ-specific gene from potato and analysis of its expression after tagging and transfer into tobacco and potato shoots. Mol Gen Genet 205: 14–22 (1986).
13. Gatz C, Kaiser A, Wendenburg R: Regulation of a modified CaMV 35S promoter by the Tn*10*-encoded Tet repressor in transgenic tobacco. Mol Gen Genet 227: 229–237 (1991).
14. Hanelt S, Braun B, Marx S, Schneider-Pötsch H: Phytochrome evolution: a phylogenetic tree with the first complete sequence of phytochrome of a cryptogamic plant (*Selaginella martensii* Spring). Photochem Photobiol, in press (1992).
15. Hershey HP, Barker RF, Idler KB, Lissemore JL, Quail PH: Analysis of cloned cDNA and genomic sequences for phytochrome: Amino acid sequences for two gene products expressed in etiolated *Avena*. Nucl Acids Res 13: 8543–8559 (1985).
16. Hershey HP, Barker RF, Idler KB, Murray MG, Quail PH: Nucleotide sequence and characterization of a gene encoding the phytochrome polypeptide from *Avena*. Gene 61: 339–348 (1987).
17. Heyer A, Gatz C: Isolation and characterization of a cDNA clone coding for potato type A phytochrome. Plant Mol Biol, in press (1992).
18. Kay SA, Keith B, Shinozaki, Chye M-L, Chua N-H: The rice phytochrome gene: Structure, autoregulated expression, and binding of GT-1 to a conserved site in the 5′ upstream region. Plant Cell 1: 351–360 (1989).
19. Kay SA, Keith B, Shinozaki, Chua N-H: The sequence

of the rice phytochrome gene. Nucl Acids Res 17: 2865–2866 (1989).
20. Kay SA, Nagatani A, Keith B, Deak M, Furuya M, Chua N-H: Rice phytochrome is biologically active in transgenic tobacco. Plant Cell 1: 775–782 (1989).
21. Keiller D, Smith H: Control of carbon partitioning by light quality mediated by phytochrome. Plant Sci 63: 25–29 (1989).
22. Keller JM, Shanklin J, Vierstra RD, Hershey HP: Expression of a functional monocotyledonous phytochrome in transgenic tobacco. EMBO J 8: 1005–1012 (1989).
23. Kendrick RE, Nagatani A: Phytochrome mutants. Plant J 1: 133–139 (1991).
24. Kronenberg GHM, Kendrick RE: The physiology of action. In: Kronenberg GHM, Kendrick RE (eds) Photomorphogenesis in Plants, pp. 99–114. Martinus Nijhoff Publishers, Dordrecht (1986).
25. Lagarias JC, Rapoport H: Chromopeptides from phytochrome. The structure and linkage of the Pr form of the phytochrome chromophore. J Am Chem Soc 102: 4821–4828 (1980).
26. Lagarias JC: Progress in the molecular analysis of phytochrome. Photochem Photobiol 42: 811–820 (1985).
27. Lissemore JL, Colbert JT, Quail PH: Cloning of cDNA for phytochrome from etiolated *Cucurbita* and coordinate photoregulation of the abundance of two distinct phytochrome transcripts. Plant Mol Biol 8: 485–496 (1987).
28. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1982).
29. Meyerowitz EM: *Arabidopsis*, a useful weed. Cell 56: 263–268 (1989).
30. Mohr H, Drumm-Herrel H, Oelmüller R: Coaction of phytochrome and blue/UV light photoreceptors. In: Senger H (ed) Blue Light Effects in Biological Systems, pp. 6–19. Springer-Verlag, Heidelberg/Berlin (1984).
31. Nagatani A, Chory J, Furuya M: Phytochrome B is not detectable in the *hy3* mutant of *Arabidopsis*, which is deficient in responding to end-of-day far-red light treatments. Plant Cell Physiol 32: 1119–1122 (1991).
32. Quail PH: Phytochrome: a light-activated molecular switch that regulates plant gene expression. Annu Rev Genet 25: 389–409 (1991).
33. Rogers SO, Bendich AJ: Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5: 69–76 (1985).
34. Salisbury FB, Ross CW: Plant Physiology. Wadsworth, Belmont (1985).
35. Sato N: Nucleotide sequence and expression of the phytochrome gene in *Pisum sativum*: Differential regulation by light of multiple transcripts. Plant Mol Biol 11: 697–710 (1989).
36. Sharrock RA, Lissemore JL, Quail PH: Nucleotide and amino acid sequence of a *Cucurbita* phytochrome cDNA clone: Identification of conserved features by comparison with *Avena* phytochrome. Gene 47: 287–295 (1986).
37. Sharrock RA, Parks BM, Koornneef M, Quail PH: Molecular analysis of the phytochrome deficiency in an *aurea* mutant of tomato. Mol Gen Genet 213: 9–14 (1988).
38. Sharrock RA, Quail PH: Novel phytochrome sequences in *Arabidopsis thaliana*: structure, evolution and differential expression of a plant regulatory photoreceptor family. Gene Devel 3: 1745–1757 (1989).
39. Shimazaki Y, Pratt LH: Immunochemical detection with rabbit polyclonal and mouse nonoclonal antibodies of different pools of phytochrome from etiolated and green *Avena* shoots. Planta 164: 333–344 (1985).
40. Shropshire W Jr, Mohr H (eds) Photomorphogenesis, Encyclopedia of Plant Physiology, New Series 16A, pp. 119–151. Springer-Verlag, Heidelberg/Berlin (1983).
41. Somers DE, Sharrock RA, Tepperman JM, Quail PH: The *hy3* long hypocotyl mutant of *Arabidopsis* is deficient in phytochrome B. Plant Cell 3: 1263–1274 (1991).
42. Smith H, Whitelam GC: Phytochrome, a family of photoreceptors with multiple physiological roles. Plant Cell Environ 13: 695–707 (1990).
43. Struik PC; Effect of shading during different stages of growth and development, yield and tuber-size distribution of *Solanum tuberosum* L. Am Potato J 63: 457–458 (1986).
44. Struik PC: Effects on shading on the initiation, growth and size distribution of potato tubers. Potato Res 30: 136–137 (1987).
45. Tokuhisa JG, Quail PH: Phytochrome in green tissue: spectral and immunochemical evidence for two distinct molecular species of phytochrome in light-grown *Avena sativa* L. Planta 164: 521–528 (1985).
46. Vierstra RD, Cordonnier M-M, Pratt LH, Quail PH: Native phytochrome: immunoblot analysis of relative molecular mass and in vitro proteolytic degradation for several plant species. Planta 160: 521–528 (1984).
47. Whitelam GC, Smith H: Retention of phytochrome mediated shade avoidance responses in phytochrome-deficient mutants of *Arabidopsis*, cucumber and tomato. J Plant Physiol; in press.