Structure and Expression of Maize Phytochrome Family Homeologs

Moira J. Sheehan,*,† Phyllis R. Farmer* and Thomas P. Brutnell*†

*Boyce Thompson Institute and †Plant Biology Department, Cornell University, Ithaca, New York 14853

Manuscript received December 23, 2003
Accepted for publication April 8, 2004

ABSTRACT

To begin the study of phytochrome signaling in maize, we have cloned and characterized the phytochrome gene family from the inbred B73. Through DNA gel blot analysis of maize genomic DNA and BAC library screens, we show that the PhyA, PhyB, and PhyC genes are each duplicated once in the genome of maize. Each gene pair was positioned to homeologous regions of the genome using recombinant inbred mapping populations. These results strongly suggest that the duplication of the phytochrome gene family in maize arose as a consequence of an ancient tetraploidization in the maize ancestral lineage. Furthermore, sequencing of Phy genes directly from BAC clones indicates that there are six functional phytochrome genes in maize. Through Northern gel blot analysis and a semiquantitative reverse transcriptase polymerase chain reaction assay, we determined that all six phytochrome genes are transcribed in several seedling tissues. However, expression from PhyA1, PhyB1, and PhyC1 predominate in all seedling tissues examined. Dark-grown seedlings express higher levels of PhyA and PhyB than do light-grown plants but PhyC genes are expressed at similar levels under light and dark growth conditions. These results are discussed in relation to phytochrome gene regulation in model eudicots and monocots and in light of current genome sequencing efforts in maize.

As sessile organisms in an ever-changing environment, plants have evolved sophisticated signaling networks to perceive and respond rapidly to changes in the external environment. Higher plants have at least three photoreceptor systems that enable them to detect red (R)/far-red (FR), blue (B)/UV-A, and UV-B light (Gyula et al. 2003). The primary R/FR photoreceptors of plants, the phytochromes, are the best characterized of all light receptor families (Quail 2002). The phytochrome holoprotein consists of a light-sensing linear tetrapyrrole chromophore, phytochromobilin, covalently attached to the phytochrome apoprotein (Terry et al. 1993). In angiosperms the phytochrome apoproteins are encoded by a small nuclear gene family of at least three members (Mathews et al. 1995; Mathews and Sharrock 1996, 1997). Phytochrome-regulated responses include seed germination, seedling deetiolation, leaf expansion, shade avoidance, and time to flowering. As these responses contribute to ecologically and agronomically important traits, phytochrome has been the subject of intense study for several decades (for reviews see Smith 1995; Ballare and Casal 2000).

Over the past few years detailed genetic and biochemical studies of eudicots have greatly increased our understanding of how phytochrome signaling pathways regulate plant development (Casal et al. 2003). In Arabidopsis, detailed mutant and molecular genetic analyses have defined roles for all five phytochrome apoproteins present in the genome. phyA is the primary photoreceptor controlling high-irradiance FR light responses such as hypocotyl elongation, cotyledon expansion, and seed germination (Nagatani et al. 1993; Parks and Quail 1993; Whitelam et al. 1993). phyB is the primary red light photoreceptor mediating many of the shade-avoidance syndrome responses that result in petiole elongation, internode elongation, reduced leaf area, and early flowering (Reed et al. 1993). phyC mediates seedling responses to red light and represses flowering under short-day conditions (Franklin et al. 2003; Monte et al. 2003). Although phyD and phyE are structurally similar to phyB, they are not functionally redundant (Sharrock et al. 2003a). Responses mediated by phyD and phyE include petiole and internode elongation and the control of flowering time (Aukerman et al. 1997; Devlin et al. 1998, 1999).

Despite our detailed understanding of phytochrome signaling in eudicots, little is known of monocot light signal transduction. The monocots and, in particular, the grasses are by far the most agronomically important group of plants. Rice, maize, wheat, and sorghum are staple food crops throughout the world and have been under intense human selection for the last 5000–10,000 years. Furthermore, traits such as yield, flowering time, and stature are all influenced by red/far-red light (Smith 1995). Early flowering varieties of barley (Hanumappa et al. 1999), sorghum (Childs et al. 1997), rice (Izawa...
et al. 2000), and maize (Sawers et al. 2002) have all recently been shown to contain lesions in the phytochrome signal transduction pathway. In maize, light response is a highly variable trait and a reduced light response is associated with early flowering North American and European inbred lines (Markelz et al. 2003). Together, these studies indicate that phytochrome response pathways contribute to the agronomic performance of these important crop species.

The phytochrome gene family structure in Arabidopsis contains five members (PHYA–E); however, characterization of the gene family in other eudicots has revealed variation in gene family composition (Howe et al. 1998; ALBA et al. 2000; LI and CHINNAPPA 2003). Unlike eudicots, the phytochrome apoprotein gene family in all monocots examined contains only three lineages: PhyA, PhyB/D, and PhyC (Matthews and Sharrock 1996, 1997). In maize, a recent polyploidization event has resulted in large segmental duplications (Gaut and Doebley 1997). Recent mapping experiments suggest that two PhyB sequences (PhyB1 and PhyB2) and two PhyA gene sequences (PhyA1 and PhyA2) arose from this genome duplication and are located in homeologous regions of the maize genome (Christensen and Quail 1989; Childs et al. 1997; Basu et al. 2000). Preliminary studies also suggest that as many as three copies of PhyC-like genes may be present in the maize genome (Basu et al. 2000). However, only the PhyA2 gene of maize has been characterized in any detail (Christensen and Quail 1989).

To examine evolutionary and functional relationships between phytochrome gene family members in maize, a detailed analysis of phytochrome gene family structure and expression has been performed. We report on the cloning and sequence analysis of full-length PhyA, PhyB, and PhyC genes in maize. Screens of bacterial artificial chromosome (BAC) clones generated from the standard maize inbred B73 indicate that the entire phytochrome gene family is duplicated in maize, resulting in closely related PhyA, PhyB, and PhyC gene pairs. These sequences have been mapped to homeologous regions of the maize genome, strongly suggesting that duplication events are a consequence of an ancestral tetraploidization event (Gaut and Doebley 1997). To examine the expression of phytochrome gene family members, we developed a semiquantitative reverse transcriptase polymerase chain reaction (RT-PCR) assay to differentiate between each Phy homeolog. In conjunction with Northern gel blot analysis we show that all phytochrome genes are expressed in several maize tissues. The structure and function of the maize phytochrome gene family is discussed in relation to the evolutionary history of maize and current genome sequencing efforts.

MATERIALS AND METHODS

Plant materials and growth conditions: Seeds from the maize (Zea mays L. ssp. mays) inbred line B73 were surface sterilized with a 10% (v/v) bleach solution for 15 min, rinsed thoroughly, and imibed overnight in sterile water. Imibed seeds were planted in 12 × 6-cell Rootrainer trays (Hummert, St. Louis) filled with vermiculite and grown either in continuous darkness (D) or under continuous white light (W) at 100 μmol m⁻² sec⁻¹ at 28°C for 10 days in a Conviron (Winnipeg, Manitoba, Canada) incubator. Fluence rates were measured with an IL1400A radiometer (International Light, Newburyport, MA) equipped with a SEL033 silicon probe (detection range: 200–1100 nm). The white light source was a combination of incandescent and cool-white fluorescent lamps. Tissue was harvested under green safe lights for D-grown tissues or ambient light for W-grown tissues. The second and third leaf blades (blade), second and third leaf sheaths (sheath), and mesocotyls from 20 seedlings were harvested and tissue pooled prior to RNA extractions.

Maize (Z. mays) phytochrome chromophore attachment region and 3'-untranslated region probes: The 456-bp PhyA chromophore attachment region (CAR) probe was generated from B73 genomic DNA using primers TBp23 (5'-GCTTGGTTCGCC AAGCCAACCTCAGA-3') and TBp24 (5'-GCTCGGGTTGAT CACTGGTCATGA-3') designed to the published PhyA sequence (GenBank accession no. AF260865; Christensen and Quail 1989). The 1055-bp PhyB CAR probe was generated from B73 genomic DNA with primers 19356f (5'-AACCGATATGGATTGCA-3') and 12417r (5'-AGAGTGACACCCCATGTAGATTTT-3') designed to a partial maize PhyB sequence (GenBank accession no. AF137332). A 293-bp PhyC CAR probe was generated from B73 genomic DNA using primers TBp29 (5'-TTTCTGTTTATGAGACAAATHTGT-3') and TBp30 (5'-TTTCTGTTTATGAGACAAATHTGT-3') designed to a partial maize phyC sequence (GenBank accession no. U61220). PCR was conducted with 2.5 units rTaq (Panvera, Madison, WI), 0.04% DMSO, 500 μM each dNTP, 1.0 μM primers, 1 X PCR buffer, 0.25 μM each dNTP. PCR reactions were performed on B73 genomic DNA under the following cycling conditions: 30 sec at 94°C, 30 sec at 62°C, 4 min at 72°C, 35 cycles; 10-min final extension at 72°C. PCR products were cloned into pCR4-TOPO vector and between phytochrome gene family members in maize, we transformed into TOP10 Escherichia coli cells (Invitrogen, Carlsbad, CA) according to the manufacturer’s protocol to generate the clones pPhyA-CAR (PhyA), pPhyB-CAR (PhyB), and pPhyC-CAR (PhyC). The identity of each clone was confirmed by sequence analysis using T3 or T7 sequencing primers.

Recombinant inbred mapping: DNA from 94 individuals derived from the Inbred B73 × Mo17 (IBM) 94 (Lee et al. 2002) population were digested with EcoRI (PhyA and PhyB) or HindIII (PhyC) and hybridized to gene-specific probes. Initial hybridizations were performed with CAR probes, which resulted in two segregating polymorphisms. These polymorphisms were scored and submitted to CIMDE (http://www.maizemap.org/CIMDE/cimde.html) to determine the map positions of Phy genes. DNA probes derived from the 3'-untranslated regions (3'-UTRs) of PhyA, PhyB, and PhyC were then used to distinguish homeologs, on the basis of differential hybridization and restriction site polymorphisms. The 3'-UTRs were amplified from B73 genomic DNA for each gene and the products were cloned into pCR4-TOPO according to the manufacturer’s specification. The 407-bp PhyA3'-UTR product was generated with ZmPAI_3'-UTR (5'-CCACTTGCGAAGAATGAGA-3') and ZmPAI_3'-UTR-R (5'-CCCATATTAAAAAAAACTCCA CACTACTAAC-3') and then cloned into pCR4-TOPO to generate the clone pPhyA3'-UTR. The 356-bp PhyA2 3'-UTR product was generated with ZmPA2_3'-UTR-F (5'-CTCAAGAG TTCATTGGGAACCGCAAGT-3') and ZmPA2_3'-UTR-R (5'-GAAGTGAAATAGAAGAAGGCT-3') and then cloned into pCR4-TOPO to generate the clone pPhyA2-3'-UTR. The 390-bp PhyB1 3'-UTR product was generated with ZmPB1_3'-UTR-F (5'-AGAGGGGCTGCTAACCAAGAGA-3') and ZmPB1_3'-UTR-R (5'-CAACCGGTGAAGTGGAAACA-3').

Plant materials and growth conditions: Seeds from the maize (Zea mays L. ssp. mays) inbred line B73 were surface sterilized with a 10% (v/v) bleach solution for 15 min, rinsed thoroughly, and imibed overnight in sterile water. Imibed seeds were planted in 12 × 6-cell Rootrainer trays (Hummert, St. Louis) filled with vermiculite and grown either in continuous darkness (D) or under continuous white light (W) at 100 μmol m⁻² sec⁻¹ at 28°C for 10 days in a Conviron (Winnipeg, Manitoba, Canada) incubator. Fluence rates were measured with an IL1400A radiometer (International Light, Newburyport, MA) equipped with a SEL033 silicon probe (detection range: 200–1100 nm). The white light source was a combination of incandescent and cool-white fluorescent lamps. Tissue was harvested under green safe lights for D-grown tissues or ambient light for W-grown tissues. The second and third leaf blades (blade), second and third leaf sheaths (sheath), and mesocotyls from 20 seedlings were harvested and tissue pooled prior to RNA extractions.
Phytochrome Family Homeologs of Maize

CCTC-3') and then cloned into pCR4-TOPO to generate the clone pPhyB1-3'-UTR. The 313-bp PhyB2 3'-UTR product was generated with ZmPB2-3'-UTR-F (5'-GGTCTGATTAGTACGTTCTACGAGTTCAAC-3') and ZmPB2-3'-UTR-R (5'-GGCCGCTCTTGACATACTACTCTG-3') and then cloned into pCR4-TOPO to generate the clone pPhyB2-3'-UTR. The 410-bp PhyCl 3'-UTR product was generated with ZmPC1-3'-UTR-F (5'-GAAGAGTGTGGTGTGAT-3') and ZmPC1-3'-UTR-R (5'-GAGAGATTGTCAGTTGCACGACGCG-3') and then cloned into pCR4-TOPO to generate the clone pPhyCl-3'-UTR. The 3353-bp PhyC2 3'-UTR product was generated with ZmPC2-3'-UTR-F (5'-GCCGCTTGGGTGCTG-3') and ZmPC2-3'-UTR-R (5'-GGAGCTDIIIATCAGGCGAAGAAC-3') and then cloned into pCR4-TOPO to generate the clone pPhyC2-3'-UTR.

**BAC library screens:** A maize HindIII BAC library, ZmMBBb, was screened to identify all potential Phy genes of maize (Tomkins et al. 2000, 2002). Radiolabeled CAR probes were generated by Klenow polymerase (Promega, Madison, WI) incorporation of [α-32P]dCTP (New England Nuclear, Boston) using random hexamers (Sambrook et al. 1989). Hybridsizations were performed as described (Church and Gilbert 1984) at 65° in a bottle incubator (Thermo Hybaid, Needham Heights, MA) using extra-large bottles (300 × 70 mm, Continental Lab Products, San Diego). Probes were purified using Micro BioSpin P-30 Tris chromatography columns (Bio-Rad, Hercules, CA) and incorporation efficiency was estimated using a Beckman LS 5801 liquid scintillation counter (Beckman Coulter, Fullerton, CA). Filters were exposed to PhosphoImage screens (Amersham Biosciences, Sunnyvale, CA) overnight and scanned at 50-μm resolution using the Storm 840 scanner (Molecular Dynamics, Sunnyvale, CA). Clone scoring was performed manually by ImageQuant (IQMac v1.2) software (Molecular Dynamics). BAC clones were ordered from Clemson University Genomics Institute (Clemson, SC; http://www.genome.clemson.edu/orders/), and single colonies were isolated for overnight culture. Minipreps of putative positive BAC clones were grown in 4.0 mluria-Bertani medium (LB) with 12.5 μg ml⁻¹ chloramphenicol at 37° with 250-rpm continuous agitation overnight. Cells were pelleted and the BACs were isolated using a standard alkaline lysis protocol (Sambrook et al. 1989). DNA isolated from putative Phy, PhyB, and PhyC clones was digested with restriction enzymes (New England Biolabs, Boston) EcoRI (PhyB) or HinIII (PhyB and PhyC) overnight at 37° and fractionated on 0.8% (w/v) agarose gels. DNA was transferred to Hybond-N⁺ nylon membranes (Amersham Biosciences). DNA gel blot analysis was performed as described above.

**BAC sequencing:** Sequence analysis was performed directly on BAC clones to minimize cloning and Taq-generated errors. A single representative BAC clone served as a template for sequencing each of the six Phy gene family members. Phytochrome gene sequence was generated from the following BAC clones: PhyA1, ZMBBb00242P1; PhyA2, ZMBBb0020P08; PhyB1, ZMBBb0022H07; PhyB2, ZMBBb000124; PhyCl, ZMBBb009204; and PhyC2, ZMBBb0026K01. Midiprep cultures were generated for the desired BAC clone by direct inoculation from plates to 250 ml LB with 12.5 μg ml⁻¹ chloramphenicol. Cultures were grown overnight at 37° at 250 rpm. An alkaline lysis midiprep kit (Qiagen, Valencia, CA) was used following the protocol for very low copy plasmids and following the modified protocol of Kirschner and Stratakis (1999). BigDye terminator mix (Perkin-Elmer, Wellesley, MA) and subsequent purification were performed according to the manufacturer’s protocol for BAC templates. Samples were fractionated on an ABI-3100 (Perkin-Elmer), and sequence analysis was performed using Sequencher 4.1.2 (Gene Codes, Ann Arbor, MI) and LaserGene 1.66 software (DNASTAR, Madison, WI). The majority of Phy gene sequences were http://wuyi.clemson.edu/phycensseq/ initiated at the central CAR regions of the genes. Most of the Phy gene sequence was obtained with the standard BigDye protocols; however, several regions were refractory to sequence analysis and modifications were made to obtain sequence from CG- and AT-rich regions of the genes. Namely, 0.04% (v/v) DMSO and 500 mM Betaine were substituted for water in the linear amplification stage and eliminated the problem in many cases. Overnight digestion of the template DNA with frequent four-base cutting restriction enzymes (HindIII, MspI, and SstII; New England Biolabs) prior to linear amplification was also utilized to obtain promoter sequence in half of the BAC clones. The accession numbers of the maize phytochrome genes as follows: PhyA1, AY234822; PhyA2, AY260865 (Christensen and Quail 1989); PhyB1, AY234827; PhyB2, AY234828; PhyCl, AY234829; and PhyC2, AY234830.

**Predicted peptide alignments:** Protein sequence alignments were generated from full-length predicted sequences from *Sorghum bicolor*, *Orzya sativa*, and both Z. mays homeologs using MegAlign in LaserGene 1.66 (DNASTAR). MegAlign Clustal W parameters were as follows: Gap Penalty = 10, Gap Extension penalty = 0.2, protein weight matrix was the Gonnet Series. Accession numbers of sequences used are as follows: ShbPHYA, AAB41397; ShbPHYB, AAB41398; ShbPHYC, U56731; OsPHYA, X141172; OsPHYB, X57563; and OsPHYC, AF141942. Domain architecture was defined using the Conserved Domain Architecture Retrieval Tool (CDART; http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi/cmd=cds) with the default parameters.

**RNA gel blot analysis:** Total RNA was extracted from ~1 g of tissue as previously described (Van Tunen et al. 1988) and quantified using a Beckman DU 530 spectrophotometer. Approximately 20 μg total RNA per lane was fractionated on 6.8% (v/v) formaldehyde, 1.5% (w/v) agarose gel. Hybridizations and image analyses were performed as described for DNA gel blot analysis (see above).

**Quantitative RT-PCR and DNA gel blot analysis:** A semi-quantitative PCR assay was developed to monitor low abundance Phy gene transcripts and to identify transcripts encoded by homeologous genes. Fifty micrograms total RNA from light- and dark-grown sheath, blade, and mesocotyl tissues was first treated with DNaseI (Invitrogen) at 25° for 1 hr. RNA was purified using phenol:chloroform:isoamyl alcohol (25:24:1, v/v) and precipitated with 95% (v/v) ethanol and 1/10 volume 3 M sodium acetate. The RNA pellet was washed overnight at 4° in 70% (v/v) ethanol to remove excess salts and RNA samples were resuspended in pyrocarbonic acid diethyl ester (DEPC)-treated water at a concentration of 1.0 μg ml⁻¹. Five micrograms of total RNA was added to first strand cDNA synthesis reactions performed with SuperScript II kit (Invitrogen) according to the manufacturer’s protocol along with “no-RT” controls. After a 2-hr incubation, all samples were treated with RNaseH (Invitrogen) according to the manufacturer’s recommended protocol. The cDNA mixture from each sample was diluted 10-fold in DEPC-treated water and 10 μl was used in PCR amplifications of phytochrome genes (only 5 μl was used for *ubi* control). To obtain a semiquantitative estimate of relative transcript levels, PCR reactions were performed for a limited number of cycles. HotStarTaq (Qiagen) was used according to the suggested PCR protocol: 15-min initial denaturation at 95°; denaturation at 94° for 30 sec, annealing at 55° for 30 sec, 1-min extension at 72° for either 20 or 25 cycles; 10-min final extension at 72°; and 4° soak. A region of the maize *ubi* gene was also amplified to control for equal addition of RNA templates. Amplification of *ubi* was as described above except that 15 cycles of PCR were performed. The primers are as follows: PhyA-F (5'-GAGGAGATCCATGAAAGAAAAGTGTTTAC-3'), PhyA-R (5'-GAAGGTGGCATCGCCGAGCTTG-3'); PhyB1-F (5'-GCTGCGTTACTGCTG-3'), PhyB-R (5'-GGCCGCTCTTGACATACTCTG-3'), PhyCl-F (5'-GCCGCTCTTGACATACTCTG-3'), PhyCl-R (5'-GCTGCGTTACTGCTG-3'), PhyC-F (5'-GCTGCGTTACTGCTG-3'), PhyC-R (5'-GCTGCGTTACTGCTG-3').
The *chrome gene was first assessed through DNA gel blot analysis of *PhyA*, *PhyB*, and *PhyC* loci. The most conserved regions were identified through the use of several restriction enzymes and hybridized to DNA digested with EcoRI (E), HindIII (H), or SacI (S) from two inbred lines, B73 and W22.

A restriction enzyme digest of PCR products was used to distinguish homeologs. The enzyme *MdfI* was used to fractionate *PhyA* from *PhyA2* products, *AvaII* was used to distinguish *PhyA1* and *PhyB1*, and *AvaII* was used to distinguish *PhyC1* from *PhyC2*. Ten microliters of the PCR products was digested overnight according to manufacturer’s recommendations (New England Biolabs). Products were fractionated on 2% (w/v) agarose gel and transferred by capillary action in 20× SSC to nylon membranes (Amersham). Hybridization and image analysis were as described above for DNA blot analysis. Bands were visualized with [α-32P]dCTP-labeled RT-PCR products derived from *PhyA*, *PhyB*, *PhyC*, or Ubi-specific amplification products.

**RESULTS**

**Characterization of *Phy* gene family structure:** Although partial expressed sequence tag (EST) and partial genomic sequence for phytochrome genes of maize have been described, only one family member, *PhyA2*, has been fully sequenced (Christensen and Quail 1989). Thus, it was first necessary to identify and sequence all phytochrome gene family members. The copy number of each phytochrome gene was first assessed through DNA gel blot analysis. DNA from the inbred lines B73 and W22 was digested with three restriction enzymes and hybridized to [α-32P]-labeled DNA derived from the conserved CAR of *PhyA*, *PhyB*, or *PhyC*. The CAR regions (see MATERIALS AND METHODS) represent the most conserved regions of phytochrome genes (Sharrock et al. 1986) and are thus ideally suited for use as an ortholog-specific gene probe. The use of several restriction enzymes revealed polymorphic sites between homeologs in two standard inbreds, confirming that all loci are in fact present in more than one copy in the maize genome (Figure 1). On the basis of the number of restriction fragments detected there appear to be two copies of *PhyA*, two copies of *PhyB*, and two copies of *PhyC* in at least two standard maize inbred lines.

**Isolation of genomic BAC clones containing phytochrome loci:** A BAC library was screened to identify genomic clones containing all phytochrome genes in maize (see MATERIALS AND METHODS). The library, ZmMBBb, is estimated to represent 13.5 genome equivalents on 10 filters (Tomkins et al. 2000, 2002). Thus, a single gene is expected to be present 1.3 times on each filter. For duplicated genes, the expected number of clones per filter is 2.6.

To estimate phytochrome gene copy number, tallies of total *Phy*-containing clones were made. Six BAC filters were sequentially hybridized with *PhyA*- and *PhyC*-CAR probes with the expectation that ∼16 clones (2.6 clones × 6 filters) would be recovered for each homeologous gene pair. DNA gel blot analysis of putative *phy*-containing BAC clones (see MATERIALS AND METHODS) confirmed the identities of 10 *PhyA*, 7 *PhyB*, and 10 *PhyC*-clones. Restriction fragment length polymorphism (RFLP) and DNA blot analyses of BAC clones indicated that there were two classes of *PhyA*, *PhyB*, and *PhyC*-hybridizing clones. These results are consistent with the DNA gel blot analysis and indicate that duplicated copies of *PhyA* (*PhyA1* and *PhyA2*), *PhyB* (*PhyB1* and *PhyB2*), and *PhyC* (*PhyC1* and *PhyC2*) are present in the B73 genome.

**Sequence analysis of BAC clones containing *Phy* loci:** Full-length genomic sequences for *PhyA1*, *PhyB1*, *PhyB2*, *PhyC1*, and *PhyC2* were obtained by direct sequencing of BAC clones as detailed in MATERIALS AND METHODS. Annotated *Phy* gene sequences were deposited in GenBank. The full-length *PhyA2* sequence was previously reported (Christensen and Quail 1989). Putative promoter regions and intron/exon junctions were identified through comparisons of previously published *Phy* gene family members of sorghum (Childs et al. 1997) and existing maize EST sequences. Comparison of the predicted genomic structure of *Phy* homeologs reveals a striking conservation of gene organization. Figure 2 shows a schematic representation of the phytochrome gene family in maize. Despite 94% (*PHYB* and *PHYC*) to 96% (*PHYA*) shared amino acid identity between homeologs, several indels and nucleotide substitutions clearly differentiate each *Phy* gene pair.
share 94% nucleotide identity across the coding region of the genes and extensive sequence similarity throughout putative 5' and 3' UTRs. However, the intron in the 5' UTR of *PhyA1* is approximately double the size of the intron in *PhyA2*. *PhyB* homeologs also share 94% sequence identity in coding regions but show greater heterogeneity in intron size. For instance, *PhyB1* contains a 1.8-kb retroelement insertion (nucleotides 5876–7713) in intron II with similarity to *Ty3-gypsy*-like retrotransposon. This sequence is not present in intron II of *PhyB2* or in *PhyB* from sorghum (Childs et al. 1997), suggesting a relatively recent origin or that *PhyB2* is the homolog of sorghum *PhyB*. Intron III of *PhyB1* has also expanded relative to intron III of *PhyB2* and *PhyB* in sorghum (Childs et al. 1997), although there is no obvious insertion element to account for this increase in size. *PhyC* homeologs have similar intron/exon structure and show little similarity in 5'-noncoding regions but not in the 3' UTRs.

**Maize phytochrome homeologs map to syntenic regions of the genome:** Gene-specific probes were generated from the coding region or 3'-UTR sequences for each of the six *Phy* gene family members and positioned on the IBM recombinant inbred population (Lee et al. 2002; Figure 3). Although the probes for *phyA* and *phyB* could not distinguish between homeologs due to the high sequence similarity in the 3'-UTR, restriction site differences enabled the assignment of RFLPs to one or the other homeolog. The *PhyB1* locus placed on chromosome 1S (1.03) and *PhyB2* placed on chromosome 9L (9.05/9.06). *PhyC1* was positioned to 1L (1.12) and *PhyC2* was positioned to 5S (5.02/5.03). We have found no evidence of a third *PhyC* copy in the B73 genome as suggested by Basu et al. (2000) in their comparative mapping study. During the course of this work, we discovered that the published *PhyA* sequence is located on chromosome 5 and not chromosome 1 as originally described in Christensen and Quail (1989). To more accurately reflect the likely ancestry of this sequence, the published *PhyA1* sequence (GenBank accession no. AY260865) was renamed *PhyA2* (P. H. Quail, personal communication). That is, phytochrome genes on chromosome 1 (*PhyA1*, *PhyB1*, and *PhyC1*) were most likely derived from one ancestral genome and the phytochrome genes located on chromosomes 9 (*PhyB2*) and 5 (*PhyA2* and *PhyC2*) were likely derived from another ancestral genome following an ancient allopolyploidization event (Gaut and Doebley 1997; Wilson et al. 1999; Gaut 2001).

**Predicted protein structure of phytochrome gene family in maize:** Alignments of *Phy* gene pairs with phytochrome transposon. This sequence is not present in intron II of *PhyB2* or in *PhyB* from sorghum (Childs et al. 1997), suggesting a relatively recent origin or that *PhyB2* is the homolog of sorghum *PhyB*. Intron III of *PhyB1* has also expanded relative to intron III of *PhyB2* and *PhyB* in sorghum (Childs et al. 1997), although there is no obvious insertion element to account for this increase in size. *PhyC* homeologs have similar intron/exon structure and show little similarity in 5'-noncoding regions but not in the 3' UTRs.

**Figure 2:** Schematic of *Phy* gene structure. Alignments of genomic sequences for each phytochrome gene are shown with gaps removed. Exons are represented by boxes; solid boxes denote the 5'- and 3'-UTRs where known. Stippled box represents a 1.8-kb retroelement insertion with similarity to *Ty3-gypsy*-like elements. Black lines represent noncoding regions.

**Figure 3:** Comparative maps of maize and related grass species, sorghum, and rice showing synteny relationships based on shared RFLP marker placement.
vast array of signal transduction proteins (ARAVIND and PONTING 1997; HO et al. 2000). Flanking the GAF domain is a conserved Phytochrome domain found in all phytochrome genes and two Per-ARnt-Sim domains found in many sensory transduction proteins (PONTING and ARAVIND 1997; PELLEQUER et al. 1998; WHITMORE et al. 1998). The conserved histidine kinase-related domain that functions as a serine/threonine kinase (YEH and LAGARIAS 1998) and a phytochrome-like ATPase (HATPase) domain are also present in each of the maize phytochrome gene family members. Furthermore, critical amino acids required for phyA (DEHESH et al. 1993; REED et al. 1994; QUAIL et al. 1995; XU et al. 1995) and phyB activity in Arabidopsis (WAGNER and QUAIL 1995; KRAAL and REED 2000) are also conserved in maize PHYA1/PHYA2 and PHYB1/PHYB2 proteins (asterisks in Figure S1, A and B, at http://www.genetics.org/supplemental/). Taken together, these data strongly suggest that all six maize Phy genes have the potential for encoding fully functional PHY proteins.

Despite the striking similarity among all Phy gene homologs, a number of features differentiate the PhyB gene family homologs (see Figure S1B at http://www.genetics.org/supplemental/). Several H repeats and a series of G repeats near the amino terminal region of PHYB define a hypervariable region among monocot PhyB sequences. This likely reflects the variable expansion of several microsatellite sequences near the 5’ end of the PhyB genes. For instance, a variable number of CAC repeats account for the presence of three to seven H residues near the N terminus of ZmPHYB1, ZmPHYB2, SbPHYB, and OsPHYB. Similarly, slipped mispairing or polymerase stuttering accompanying DNA replication through long stretches of GC residues (FARABAUGH and MILLER 1978) may account for the variable number of G residues near the N terminus of monocot PHYB. The predicted PHYB1 protein is five amino acids shorter than PHYB2 due to additional G repeats at the N-terminal region of the protein and a +1-bp frameshift 9 bases before the predicted stop codon.

RNA gel blot analysis of the phytochrome gene family:
RNA gel blot analysis was used to examine the light regulation of Phy gene transcript accumulation (Figure 4). Due to the strong sequence similarity among homeologous genes, we were unable to develop gene-specific probes that could unambiguously discriminate each homeolog. Thus, the transcript pools examined by Northern analysis reflect the accumulation profiles of homeologous gene pairs and do not address the contributions by single gene family members. Nevertheless, hybridizations with gene-specific probes (see MATERIALS AND METHODS) detected transcripts of ~4 kb for PhyA, PhyB, and PhyC, indicating that mature transcripts for PhyA, PhyB, and PhyC accumulate in seedling tissues. As observed previously (CHRISTENSEN and QUAIL 1989), transcripts encoded by PhyA were more abundant in D-grown tissues than in W-grown tissues. PhyB transcripts also accumulated to higher levels in D than in W. Interestingly, PhyC transcript levels were maintained at similar levels under light and dark growth conditions, indicating that the mechanisms regulating PhyC transcript accumulation may differ from those regulating PhyA and PhyB. In addition to light-induced changes in expression, tissue-specific differences in expression were also observed. Most notably, transcripts for PhyB and PhyC accumulated to similar levels in light- and dark-grown mesocotyl tissues and PhyA transcripts were much less abundant in dark-grown mesocotyl tissues relative to dark-grown blade and sheath tissues.

Semiquantitative RT-PCR of phytochrome gene expression: To examine the transcriptional profiles of each Phy gene, we developed a RT-PCR assay that differentiates between each homeolog (see MATERIALS AND METHODS). In brief, primers were designed to regions of perfect identity that are shared between homeologs and that span the last intron to control for genomic DNA contamination. Product lengths (~300–450 bp) for each homeolog were also identical to prevent selective amplification of one or the other homeolog. A limited number of PCR cycles were then performed to enable semiquantitative analysis of gene expression profiles. cDNA products were digested with restriction enzymes that could discriminate between cDNA derived from homeologs due to the presence of single nucleotide polymorphisms (see schematic in Figure 5). Products were fractionated on agarose gels and transferred to nylon membranes and hybridized to gene-specific probes. The relative abundance was estimated from densitometric scans of radiolabeled filters.
As shown in Figure 5, all six maize phytochrome genes are expressed in several seedling tissues including leaf blade, leaf sheath, and mesocotyl, but the light- and tissue-specific patterns of expression differ significantly among homeologous genes. Densitometric analysis of transcript profiles indicated that PhyA1 transcripts accumulated to higher levels than PhyA2 transcripts in all tissues examined (data not shown). However, the tissue-specific patterns of PhyA1 and PhyA2 expression were very similar to one another, and transcripts encoded by both genes accumulated to higher levels in all dark-grown tissues examined relative to light-grown tissues. The similarity between PhyA1 and PhyA2 accumulation profiles may be attributable to small regions of conserved sequence in promoter and 5′-UTR (see Figure S2A at http://www.genetics.org/supplemental/) or the 3′-UTR (see Figure S2B at http://www.genetics.org/supplemental/). For example, putative phytochrome responsive PE3 and RE1 elements (Bruce et al. 1991) were identified in both PhyA1 and PhyA2 promoters, indicating a similar mode of PhyA transcriptional control (see Figure S2C at http://www.genetics.org/supplemental/).

Gene expression profiles for the PhyB genes were similar to those observed for the PhyA genes. Expression from only one of the PhyB homeologs predominated in all seedling tissues examined, and transcripts for PhyB1 and PhyB2 were more abundant in D-grown relative to W-grown blade and sheath tissues. However, both PhyB1 and PhyB2 were expressed at similar or slightly higher transcript levels in D-grown mesocotyl tissue relative to W-grown mesocotyl tissue. As observed for PhyA and PhyB homeologs, transcripts from the phyC homeolog located on chromosome 1 predominated in all seedling tissues examined. PhyC transcripts accumulated to slightly higher levels in D-relative to W-grown blade tissue but accumulated to similar levels in D- and W-grown sheath tissues. Together, these studies indicate that phytochrome gene family members are under both tissue- and light-dependent regulation during maize seedling development.

**DISCUSSION**

Phytochromes are the primary red/far-red photoreceptors in higher plants and play an essential role in the regulation of plant development in response to envi-
rnonmental change (Neff et al. 2000). In higher plants, phytochromes are encoded by a small nuclear gene family that varies in size and complexity throughout the angiosperm lineage. In many diploid eudicots (e.g., Arabidopsis, poplar, tomato, and Stellaria longipes) the phytochrome gene family is represented by three to five gene family members (Mathews and Sharrock 1997; Howe et al. 1998; Alba et al. 2000; Li and Chinnappa 2003). However, in the monocots rice and sorghum the family is represented by three lineages, PhyA, PhyB/D, and PhyC (Mathews and Sharrock 1996, 1997; Childs et al. 1997). Here we show that in maize the phytochrome gene family is composed of six members, PhyA1, PhyA2, PhyB1, PhyB2, PhyC1, and PhyC2. Each gene family member is predicted to encode a functional phytochrome apoprotein and is expressed in several seedling tissues. The position of Phy paralogs (e.g., PhyA1 and PhyA2) in homeologous regions of the maize genome strongly suggests that these gene duplications were generated as a consequence of an ancient tetraploid event (Gaut and Doebley 1997; Wilson et al. 1999; Gaut 2001). Thus, the PhyA1, PhyB1, and PhyC1 genes on chromosome 1 are most accurately defined as the homeologs of PhyA2, PhyB2, and PhyC2, respectively, present on chromosomes 5S (PhyA2 and PhyC2) and 9L (PhyB2). The striking conservation of sequence and overlapping expression patterns of phytochrome homeologs suggest that functional redundancy of phytochrome gene action has been maintained during the evolution and domestication of modern maize.

Although the function of any Phy gene has yet to be determined in maize, it is tempting to speculate on some of the consequences of maintaining such similar homeologs. Perhaps one of the most intriguing possibilities is that the homeologs could form active heterodimers that mediate responses different from those regulated by homodimers. Early studies of oat phytochrome proteins indicated that the biologically active forms of phytochrome function as homodimers (Jones and Quail 1986). Evidence strongly suggests that these dimerized forms interact with nuclear localized transcription factors to mediate phytochrome responses (Sakamoto and Nagatani 1996; Kircher et al. 1999, 2002; Yamaguchi et al. 1999; Huq and Quail 2002; Huq et al. 2003). In the case of Arabidopsis phyB, dimerized and nuclear localized N-terminal sequences (amino acids 1–651) are sufficient to mediate a number of red light responses in a phyB mutant background (Matsushita et al. 2003). Although many regions of the mature protein assist or strengthen dimerization, the amino acids that are critical for dimerization of Arabidopsis phyB lie between 1116 and 1161 (Quail 1997). Due to the high degree of sequence identity between homeologous gene pairs, especially within this dimerization domain, it is likely that the maize PHY homeologs will also form heterodimers. Clearly, the consequences of heterodimerization are dependent on the, as yet, untested functional divergence of each protein. Nevertheless, the overlapping expression domains of maize Phy homeologs are consistent with the possibility of direct interaction of Phy homeolog proteins.

The high sequence conservation between homeologous phy gene pairs also highlights some of the challenges that face maize genome sequencing projects. Searches of The Institute for Genome Research (TIGR) Maize Gene Index (ZmGI) revealed tentative consensus (TC) sequences for PhyA, PhyB, and PhyC genes of maize. This database curates assemblies of maize ESTs (TIGR contigs or TCs) and identified two potential PhyA genes. For example, TC185264 incorporates several EST sequences derived from multiple cDNA libraries and includes the previously published PhyA2 sequence (Christensen and Quail 1989). Comparisons of the individual EST clone sequences that constitute TC185264 with the PhyA1 and PhyA2 sequences reported here reveal that both PhyA1 and PhyA2 sequences are represented in the assembly. A tentative consensus sequence (TC171984) was also derived from ESTs that were all most similar to PhyA1 sequence. Thus, TC185264 is most similar to PhyA2 whereas TC171984 is most similar to PhyA1 sequence, but neither assembly is identical to the predicted PhyA1 gene reported here or to the previously published PhyA2 sequence (Christensen and Quail 1989).

Attempts to distinguish highly similar gene homeologs such as the phytochrome genes are further complicated by the presence of repetitive elements within genes. As shown in Figure 2, a nearly 2-kb retroelement insertion is present in the second intron of PhyB1. Although gene islands are anticipated to be scattered throughout a sea of retrotransposons in maize (SanMiguel et al. 1996), the presence of large retroposon sequences within genes could prove problematic in genomic assemblies (Rabinowicz et al. 2003). To enrich for gene-rich regions of the genome, current maize genome sequencing projects are utilizing methyl filtration technologies and high C,t sequences to reduce representation of repetitive LTR retroelements (Rabinowicz et al. 2003). Since we did not examine the methylation status of this element, it is difficult to predict whether or not PhyB1 intron sequences would be represented in methylated genomic libraries (Rabinowicz et al. 1999). Regardless, the repetitive nature of the element and its large size would likely lead to its exclusion from genomic assemblies. Recent methods developed to link gene-rich regions should prove particularly useful in assembling genes such as PhyB1 in maize (Yuan et al. 2002).

In addition to confounding bioinformatics-based approaches to genome annotation, the duplications of the maize Phy gene family also highlight challenges faced in defining syntenic relationships among grass genomes. The relatively recent polyploidization in the maize lineage has likely contributed to the duplication of >80% of the maize genome as colinear regions of homeologous
maize chromosomes (Gaut 2001). These highly similar sequences have the potential to greatly complicate the analysis of orthologous relationships in the grasses. For instance, in a recent study, a PhyA-containing BAC clone of maize was characterized as a syntenic region with a PhyA gene sequence from sorghum (Morishige et al. 2002). The authors characterized a single maize BAC clone and assigned this gene and all predicted gene sequences residing on this BAC to chromosome 1L, corresponding to the published map position of PhyA2 (Christensen and Quail 1989). However, as we report here, the original assignment of this homeolog to chromosome 1 was incorrect and is instead located on chromosome 5S. Furthermore, because only one of the two PhyA orthologs was characterized, the previously reported synteny comparisons made between the sorghum and maize PhyA regions are incomplete (Morishige et al. 2002). It is tempting to speculate that because all Phy genes located on chromosome 1 are more highly expressed in seedling tissues than are their homeologs located on chromosomes 5 and 9, genes tightly linked to these Phy genes may show similar homeolog-specific biases in gene expression.

An analysis of Phy transcript profiles was conducted to examine potential light- and tissue-specific regulation of Phy gene activity in maize. Transcripts for both PhyA1 and PhyA2 accumulated to much higher levels in dark relative to white light in blade, sheath, and mesocotyl tissues. This light-dependent decrease in phyA transcripts has been observed in several grass species (Colbert et al. 1983; Hershey et al. 1984; Christensen and Quail 1989; Kay et al. 1989; Dehesh et al. 1991) and likely reflects a common regulatory mechanism that acts to maintain high levels of phyA transcript in the dark (Bruce et al. 1989). Furthermore, several elements have been defined in monocot PhyA promoters that likely mediate a high transcription rate of phyA in the dark (Bruce and Quail 1990; Bruce et al. 1991). The presence of PE3 and RE1 elements in the promoter region of maize PhyA homeologs suggests a similar mechanism of transcriptional control.

Despite the common features shared by monocot phyA promoters, the kinetics of light-mediated reduction of phyA transcript levels differs among grass species (Colbert et al. 1983; Lissemore and Quail 1988; Christensen and Quail 1989; Kay et al. 1989). In oat (Lissemore and Quail 1988) and rice (Kay et al. 1989) there is a rapid decrease in phyA transcript levels that can largely be attributed to a decreased rate of phyA transcription. In maize the reduction in phyA transcript accumulation is less rapid (Christensen and Quail 1989) and indicates that post-transcriptional regulation may also contribute to phyA transcript accumulation. Here we have shown that this light-dependent reduction in phyA transcript accumulation is intrinsic to phyA1 and phyA2 regulation and operates in several seedling tissues.

Although little is known of PhyB or PhyC gene regulation in monocots, detailed characterizations of organ- and tissue-specific regulation of phytochrome gene expression have been performed in several model eudicots. In Arabidopsis, PHYA, PHYB, PHYC, PHYD, and PHYE transcripts accumulate in several plant organs (Clack et al. 1994), and promoter-Gus fusions indicate that the promoters determine tissue specificity, light inducibility, and magnitude of gene expression (Somers and Quail 1995a,b; Goosey et al. 1997). However, translational and/or post-translational mechanisms largely determine the accumulation of phytochrome proteins in Arabidopsis (Clack et al. 1994; Sharrock and Clack 2002; Sharrock et al. 2003b). In potato, a uniform pattern of PhyB expression was observed in several mature plant organs (Heyer and Gatz 1992). In tomato, however, temporal tissue- and organ-specific differences were observed for all phytochrome genes (Häuser et al. 1997, 1998). Also, in general, PhyA transcripts are most abundant in several seedling and mature plant tissues, with the exception of ripening fruit, where PhyB2 and PhyF transcripts predominate (Häuser et al. 1997). In tobacco, PHYB1-Gus promoter fusions also suggest tissue-specific regulation of PHYB1 transcription in developing seedlings and mature plants. In particular, PHYB1 expression appeared to be restricted to phloem cells of the stem, petiole, and midrib tissues (Adam et al. 1996). The organ- and tissue-specific expression profiles of PHYB genes in tomato (Häuser et al. 1997) and tobacco (Adam et al. 1996) and maize (this study) contrast with the relatively uniform pattern of PHYB expression in Arabidopsis (Somers and Quail 1995b) and potato (Heyer and Gatz 1992) and suggest that there may be little conservation in the transcriptional regulation of PHYB and PHYC accumulation between monocots and dicots.

In rice (Dehesh et al. 1991; Basu et al. 2000) and sorghum (Childs et al. 1997) PhyB and PhyC transcripts are maintained at relatively constant levels under dark and white light conditions, but nothing is known of tissue- or organ-specific expression profiles. In maize, we have shown that tissue-specific accumulation of PhyB is light dependent. That is, PhyB transcripts are most abundant in blade and sheath tissues of dark-grown plants but accumulate to similar levels in mesocotyl tissues of light- and dark-grown plants (see Figure 4). Thus, light appears to play an important role in regulating the organ-specific patterns of PhyB gene expression in maize. PhyC transcripts were the least abundant of all Phy genes in dark-grown tissues and were maintained at low levels in light-grown tissues as well.

The identification and sequence analysis of all phytochrome gene family members reported here will now permit the utilization of association methods to examine the role of phytochrome gene function in this important crop plant. In addition, the elucidation of phytochrome gene family organization in maize will guide reverse genetic screens to disrupt all phytochrome gene
family members in maize. The single, double, and higher order mutants will provide valuable insights into the functional divergence of the phytochrome gene family in maize and the role of phytochrome in the regulation of agronomically important traits.

The authors thank Denise Costich, Ruairidh Sowers, Shanna Moore, Jeffery Doyle, and Martha Hamblin for their helpful suggestions during the course of this work and for their careful review of this manuscript. M.J.S. also thanks Jeffery Doyle, Stephen Kresovich, and Gregory Martin for their support and advice on this project. This work was funded by National Science Foundation grant IBN-0110297. M.J.S. was supported by a Plant Cell and Molecular Biology training grant (Cornell University) and a Graduate Assistance in Areas of National Need fellowship (Cornell University).

LITERATURE CITED


Kirschen, R. S., and C. A. STRATFORD, 1999 Large-scale preparation


Communicating editor: J. A. Birchler