Orobanchaceae were traditionally circumscribed to contain only holoparasitic genera such as Orobanchaceae, as redefined by Young et al. (1999), are a morphologically diverse family of predominantly herbaceous, parasitic plants. The majority of species are facultative or obligate root parasites, which may be photosynthetic (hemiparasites) or totally dependent on the host plant (holoparasites). Several genera, such as Striga, Alectra, and Orobanche, are serious agricultural pests of cereal and legume crops, particularly in the dry tropics and subtropics, where they may reduce crop yields substantially (Riches and Parker, 1995). The family has a worldwide distribution, including species from extreme northern latitudes (e.g., Pedicularis dasyantha Hadac, endemic to the European-Russian arctic; Odasz and Savolainen, 1996), but the main centers of distribution are the Mediterranean, southern Africa, the Himalayas, and western North America. Some genera are distributed over several continents, such as Euphrasia (Europe, North and South America, Oceania), Bartsia (Africa, Europe, North and South America), and Buchnera and Melasma (Africa, Asia, North and South America). Currently 89 genera, containing ca. 2061 species, are recognized in Orobanchaceae (Nickrent, 2006). Thus, Orobanchaceae are the most species-rich of parasitic angiosperm families—sizes of other families range from two to 905 species—and are of a size comparable to the order Santalales (Nickrent, 2006). A robust phylogenetic hypothesis for the family will enable a greater understanding of the factors, such as parasitism, that may underlie the success of Orobanchaceae and also will serve as a context in which to better understand the implications of plastid genome evolution (e.g., dePamphilis, 1995; Wolfe and dePamphilis, 1998; Bungard, 2004; Young and dePamphilis, 2000, 2005).

Orobanchaceae were traditionally included in Scrophulariaceae (e.g., Bentham, 1846, 1876; von Wettstein, 1891). However, these classifications did not deal adequately with holoparasitic taxa such as Lathraeae and Hybanche, which were placed in Scrophulariaceae by some authors on account of their bilocular gynoecia and in Orobanchaceae by others on account of their holoparasitic nature (Boeshore, 1920; Kuijt, 1969). Some authors included Orobanchaceae entirely within Scrophulariaceae (Hallier, 1903; Bellini, 1907). More recently, these relationships were investigated in a series of phylogenetic analyses of plastid sequence data (dePamphilis et al., 1997; Nickrent et al., 1998; Young et al., 1999; Young and dePamphilis, 2000, 2001) and of nuclear 18S rDNA (Nickrent et al., 1998) and ITS (Wolfe et al., 2005) sequence data. Hemiparasitic species, formerly placed in Scrophulariaceae subfamily Rhinantheoideae, and Orobanchaceae were shown to comprise a monophyletic group. Parasitism is believed to have evolved once in the group, followed by multiple independent origins of holoparasitism from hemiparasitic ancestors (Young et al., 1999). Lindenbergia philippensis, from a genus of 12 nonparasitic species from northeast Africa and Asia (Hjertson, 1995) previously classified in Scrophulariaceae subfamily Antirrhinoideae (von Wettstein, 1891), was resolved as sister to the parasitic species (Nickrent et al., 1998; Young et al., 1999; Young and dePamphilis, 2000; Olmstead et al., 2001) and included in a clade definition of Orobanchaceae (Young et al., 2001; Bennett et al., 2002). Partial sequences of the nuclear gene encoding the photoreceptor phytochrome A (PHYA) are used to reconstruct relationships within Orobanchaceae, the largest of the parasitic angiosperm families. The monophyly of Orobanchaceae, including nonphotosynthetic holoparasites, hemiparasites, and nonparasitic Lindenbergia is strongly supported. Phytochrome A data resolve six well-supported lineages that contain all of the sampled genera except Brandisia, which is sister to the major radiation of hemiparasites. In contrast to previous plastid and ITS trees, relationships among these major clades also are generally well supported. Thus, the robust phylogenetic hypothesis inferred from the PHYA data provides a much better context in which to evaluate the evolution of parasitism within the group. Ninety-eight species of Orobanchaceae, representing 43 genera, are included and Brandisia, Bungea, Cymbaria, Esterhazyza, Nesogenes, Phtheirospermum, Radamaea, Siphonostegia, and Xylocalyx are confirmed as members of Orobanchaceae. The earliest diverging lineage of hemiparasites is identified for the first time; it contains Bungea, Cymbaria, Monochasma, Siphonostegia, and the monotypic Schwalbea, which is federally endangered. This basal clade is marked by the presence of two novel introns. A second, apparently independent gain of one of these introns marks a clade of largely European taxa. There is significant rate heterogeneity among PHYA sequences, and the presence of multiple PHYA in some taxa is consistent with observed ploidy levels.

Key words: introns; Orobanchaceae; parasitism; phylogeny; phytochrome A; Scrophulariaceae.
al., 1999). These molecular analyses confirmed the view that holoparasitic and hemiparasitic species are closely related, but the multiple independent losses of photosynthetic ability in the group did not support the conclusion of Boeshore (1920) that there was a single progressive degradation from autotrophs through hemiparasites to holoparasites.

Additional evidence for a close relationship between hemiparasites and holoparasites comes from studies of pollen morphology (Inceoglu, 1982; Minkin and Eshbaugh, 1989; Abu Shaib et al., 1994). The pollen of both holoparasites and hemiparasites is tricolpate with reticulate exine sculpturing, whereas that of nonparasitic species of Scrophulariaceae sensu von Wettstein (1891) is tricolporate with a reticulate exine (Minkin and Eshbaugh, 1989). The pollen of Linderbergia is also tricolporate with a reticulate exine (Hjertson, 1995), but their pattern of corolla aestivation is the same as that of the parasitic species, with the abaxial lobes folded over the adaxial lobes (Brühl, 1920; Armstrong and Douglas, 1989).

Although analyses of plastid sequence data have considerably advanced our understanding of relationships and the evolution of parasitism in Orobanchaceae, the taxonomic sampling from within Orobanchaceae for these studies was limited, and there was a lack of resolution within the family. Moreover, the loss of many plastid genes from some parasitic taxa (dePamphilis and Palmer, 1990; Morden et al., 1991; Wimpee et al., 1991, 1992; Wolfe et al., 1992a–d) limits their utility as phylogenetic markers. Internal transcribed spacer data have provided insight, but they also failed to provide robust resolution of major lineages within the family (Wolfe et al., 2005). On average, protein-coding nuclear genes evolve faster than plastid and mitochondrial genes (Wolfe et al., 1987) and thus may contain variation suitable to address unresolved relationships within Orobanchaceae. In this study we utilize partial sequences from a nuclear gene, phytochrome A, to identify the composition of the major lineages within the family and address relationships among them. Phytochromes regulate growth and developmental responses to light signals in the red and far-red regions of the visible spectrum (Smith, 2000), some of which are significantly altered in parasites. Thus, while our focus is phylogenetic, the accumulation of phytochrome data sets will provide an important resource not only for investigation of molecular evolution at nuclear loci in parasites to complement studies of plastid genes (dePamphilis, 1995; dePamphilis et al., 1997; Nickrent et al., 1998; Young et al., 1999; Young and dePamphilis, 2000; Olmstead et al., 2001; Wolfe et al., 2005). Forty-three of the 89 genera currently recognized in Orobanchaceae (Nickrent, 2006) were included.

**MATERIALS AND METHODS**

**Taxon sampling and outgroup selection**—Partial PHYA sequences were obtained from 98 species of Orobanchaceae (Appendix), including multiple accessions of some genera (e.g., *Pedicularis, Euphrasia, Orobanchus*). The data set comprised representatives of all the major clades inferred from previous analyses of Orobanchaceae (dePamphilis, 1995; dePamphilis et al., 1997; Nickrent et al., 1998; Young et al., 1999; Young and dePamphilis, 2000; Olmstead et al., 2001; Wolfe et al., 2005). Forty-three of the 89 genera currently recognized in Orobanchaceae (Nickrent, 2006) were included. *Bungea, Cymbaria, Escobedia, Esterhazya, Monoschisma, Nesogonea, Odontites, Radamaea, Siphonostegia*, and *Xylocalyx* are here included for the first time in a molecular phylogenetic study.

Several species of Lamiales were also included in the maximum parsimony (MP) analyses (families are given according to Olmstead et al. [2001] and Oxelman et al. [2005]): *Ajuga reptans* (Lamiaceae), *Alonsoa sp.* (Scrophulariaceae), *Asparagus pumilus* (Scrophulariaceae), *Chelone obliqua* (Plantaginaceae), *Glehoma hederaeae* (Lamiaceae), *Paulownia tomentosa* (Paulowniaceae), *Penstemon cobaea* (Plantaginaceae), *Schropharia arguta* (Scrophulariaceae), and *Strobilanthes attenuata* (Acanthaceae), and trees were rooted with *Solanum tuberosum* (Solanaceae). To reduce computation time only *Paulownia* and *Solanum* were included as outgroups in the maximum likelihood (ML) analyses.

**DNA extraction, cloning and sequencing**—Total genomic DNA was extracted using the DNAeasy Plant Mini Kit protocol (QIAGEN, Valencia, California, USA) from fresh or dried tissue. Phytochrome fragments were amplified and cloned using the protocols detailed in Mathews et al. (2000). In some cases, PHYA products were doublets or triplets that were cloned separately, with two clones being sequenced from each fragment. From a subset of species, up to 12 clones were screened. From most accessions, a 2.0–2.5-kb region including intron 1 was successfully amplified using the forward primer 212f and reverse primer a832r.2 (Table 1). For eight sequences, two smaller fragments were amplified using different combinations of the primers listed in Table 1 and assembled in Sequencher (GeneCodes, Ann Arbor, Michigan, USA). In four cases, there was substantial overlap (120–555 base pairs [bp]) and the fragments had >99% identity. In the other four cases the separate fragments did not overlap, but they were sister sequences in preliminary analyses. They were combined into a single terminal for subsequent analyses.
**Phylogenetic analyses**—Exon boundaries were inferred based on the conserved gene structure of land plant phytochromes, and the introns were removed before the coding sequences were aligned manually. A small variable region near to the chromophore attachment site was difficult to align unambiguously and was deleted. Autopomorphic gaps also were deleted, leaving a final matrix of 1824 nucleotide sites, of which 1141 (ca. 62%) were phylogenetically informative in the parsimony analysis. The 121 sequences analyzed in the parsimony analysis represented 98 species of Orobanchaceae and 11 outgroup species.

Parsimony analyses using PAUP* version 4.0b10 (Swofford, 2002) comprised 1000 random-addition replicates with tree-bisection-reconnection (TBR) branch swapping and MULPARS on so that multiple most parsimonious trees are saved during branch swapping. Gaps were scored as missing data. One thousand bootstrap replicates were performed using 100 random-addition replicates and MULPARS off following the strategy of DeBry and Olmstead (2000).

Maximum likelihood analyses were conducted using PAUP* 4.0b10 (Swofford, 2002). The general time-reversible model (Rodriguez et al., 1990) with rate variation among nucleotides following a discrete gamma distribution and assuming a proportion of invariant sites was chosen by a hierarchical likelihood ratio test as implemented in Modeltest version 3.7 (Posada and Crandall, 1998). Model parameters were set to those estimated by Modeltest 3.7. Tree-bisection-reconnection branch-swapping was employed on a starting tree obtained using random stepwise addition. One hundred bootstrap replicates were analyzed using nearest-neighbor interchange (NNI) branch swapping and Asls stepwise addition.

**Templeton and Shimodaira–Hasegawa tests**—To explore the alternative placements of the basalmost clades, the nonparametric Templeton test (Larsen, 1994) and the likelihood-based Shimodaira–Hasegawa test (SH test; Shimodaira and Hasegawa, 1999) were implemented in PAUP* 4.0b10. We compared the trees obtained from the parsimony analysis (clade I sister to the rest of Orobanchaceae; Fig. 1) with trees obtained from constrained analyses in which (1) clades I and II are sister to the rest of Orobanchaceae (the topology obtained in the ML analysis, Fig. 2) and (2) clade II is sister to the rest of Orobanchaceae. For the SH test, the test distribution was obtained using the reestimated log likelihoods (RELL) approximation with 1000 nonparametric bootstrap replicates.

### RESULTS

**Phytochrome A sequence characteristics**—A single PHYA sequence was amplified and cloned from most taxa, but two or more PHYA sequences were cloned from Castilleja sulphurea, Escobedia grandiflora, Euphrasia stricta, Lathraea clandestina, L. squamaria, Melasma scabrum, Striga linearifolia, Xylocalyx asper, and X. carterae.

Most PHYA fragments spanned a region that included intron I, located at amino acid position 745 of the alignment of Mathews et al. (1995). Intron I was present in all of these sequences and was of variable length (Appendix). Additionally, PHYA coding sequences from some species were interrupted by up to two additional putative introns. The first, intron A, is located at amino acid position 482, downstream of the chromophore-binding site (amino acid position 374) and was found in all species of clades II and V (Figs. 1–3). The second novel intron, intron B at position 575, was found only in species of clade II.

There was large length variation between the same introns from different species (Appendix), and it was impossible to unambiguously align the sequences across the whole family. Alignment of introns from species across major clades was also difficult. However, it was possible to align intron sequences from species of the same genus (e.g., within Castilleja or Rhinanthus). Three intron sequences, intron B from Schwalbea americana and intron I from Odontites himalayicus and Glechoma purpurea, did not have the canonical 5′ GT and 3′ AG that are recognized by the spliceosome. Schwalbea and Odontites had a single base substitution at the 5′ end to GA and GC, respectively, and Glechoma had a substitution at the 3′ end to AA. If these introns were not spliced correctly a frame shift or stop codon would result in an altered protein product.

**Phylogenetic analyses**—Parsimony analysis yielded four equally parsimonious trees of length 9754 that differed only in the relationships among Castilleja rubicundula, C. tenuifolia, and Triphysaria pusilla, and among the outgroups. A schematic tree showing the relationships between the major clades from the parsimony analysis is shown in Fig. 1. The topology of the ML tree (Figs. 2–3) was nearly identical to that of the parsimony tree, except that the ML tree unites clades I and II, while the parsimony tree places them as successive sister taxa to the rest of the family. Relationships within clade VI also differ between the two analyses, although these do not receive bootstrap support in either analysis. Bootstrap support for most nodes is similar between the parsimony and ML trees, and only 12 nodes differed by more than 10 percentage points;

### Table 1. Primer sequences used in this study for the amplification of PHYA.

<table>
<thead>
<tr>
<th>Forward Primers (5′–3′)</th>
<th>Reverse Primers (5′–3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>212f</td>
<td>TCWGnaARCCNTTTfAYGC</td>
</tr>
<tr>
<td>a230f.oro</td>
<td>GAYTYYGACCYGNNADCCYYAYG</td>
</tr>
<tr>
<td>a236f.oro</td>
<td>CCYAYAGKTBCCATGASYG</td>
</tr>
<tr>
<td>377f</td>
<td>CASTHYTGCCNAMAYATG</td>
</tr>
<tr>
<td>444f</td>
<td>CARCTTWTTGCTAHPACAYG</td>
</tr>
<tr>
<td>a624f.oro</td>
<td>GAYTWYGARATGGAYGCRAT</td>
</tr>
</tbody>
</table>

Fig. 1. Phylogeny of Orobanchaceae inferred from a parsimony analysis of PHYA. Species relationships within the major clades are not shown. Parsimony bootstrap values are indicated.
we report both MP and ML bootstrap values for the nodes discussed in the text below.

The monophyly of Orobanchaceae sensu Young et al. (1999) is well supported in both MP and ML analyses with bootstrap (BS) percentages of 100%. Six major clades plus *Brandisia hancei* are consistently resolved within Orobanchaceae in all analyses (Figs. 1–3). Clades I and II are well supported as monophyletic in both the MP and ML analyses (100% BS), and species relationships within these clades also receive high support. However, there is only weak support for the position of clade I as sister to all other Orobanchaceae in the MP analysis (50% BS; Fig. 1). In the ML tree (Fig. 2), clade I is sister to clade II (67% BS), which together are sister to the rest of the family with high BS (100%). We conducted a Bayesian analysis to determine whether an ML-based search using a codon model would yield a tree consistent with the ML tree inferred using a nucleotide model. The Bayesian tree also unites clades I and II, with a posterior probability of 0.92 (data not shown). The alternative placements of clades I and II from the constrained analyses cannot be rejected at a level of significance ≤ 0.2 by the Templeton and SH tests (data not shown).
The next diverging clade, clade III, contains most of the holoparasitic species and is supported in both MP and ML analyses (85% MPBS, 91% MLBS). Within this clade Orobanche is monophyletic (100% MPBS and MLBS) and is sister to a clade of Epifagus, Conopholis, and Boschniakia (87% MPBS, 72% MLBS). Brandisia hancei is the next lineage sister to the rest of the family (84% MPBS, 72% MLBS).

The remaining species fall into a large clade (100% MPBS and MLBS) containing clades IV, V, and VI. Clades IV and VI receive bootstrap support of 100% in both MP and ML analyses. Clade V receives only 66% BS support in the MP analysis, but 100% BS support in the ML analysis. Clades V and VI are supported as sister taxa (99% MPBS, 82% MLBS). Clade IV is well supported as sister to clades V and VI in both the MP and ML analyses (100% BS).

Within clade IV, Pedicularis is resolved as monophyletic with 100% BS support in both MP and ML analyses and is sister to the remaining species. Within Pedicularis, two monophyletic groups are resolved, one containing the Asian species P. densispica, P. gyrorhyncha, and P. kansuensis (100% MPBS and MLBS) and the other containing four European species, P. foliosa, P. tuberosa, P. julica, and P. kernerii (96% MPBS, 74% MLBS). Phtheirospermum japoni-
cum is resolved as sister to the remaining species. Castilleja and Orthocarpus PHYA sequences occur in two clades, with Cordyanthus and Triphysaria on the one hand and Lamourouxia, Seymeria, Agalinis, Aureolaria, and Esterhazyza on the other. Each of the two Castilleja clades contains one of the two PHYA sequences from C. sulphurea, suggesting that the polyphyly of Castilleja in the PHYA tree results from differential sampling of paralogs. There is 100% BS support for a clade of Agalinis, Aureolaria, Esterhazyza, and Seymeria in both the MP and ML analyses.

From some clade V taxa (Larthraea clandestina, L. squamaria, Euphrasia stricta, and Bartsia alpina), multiple copies of PHYA were detected. The two sequences from E. stricta are not sister in the Euphrasia clade, suggesting that other Euphrasia species may also contain additional PHYA copies that were not detected in our study. Three sequences from L. squamaria, L. clandestina, and B. alpina form a well-supported monophyletic group (clade Vb) in both the MP and ML analyses (100% BS) that is separated from clade Vb by a long branch (Fig. 3). Bootstrap support for clade Vb is high (100% MPBS, 99% MLBS), suggesting that the sequences of clade Vb represent a divergent copy that was not detected in the other species that were sampled. Support for clade V as a whole, including both Vb and Vb*, is only 66% in the MP analysis but 100% in the ML analysis.

Within clade Vb, species relationships are generally well resolved. The five species of Melampyrum form a monophyletic group (87% MPBS, 95% MLBS) that is sister to the rest of the clade (100% MPBS, 99% MLBS). Rhinanthus is monophyletic (100% MPBS and MLBS) and is sister to four sequences from Larthraea clandestina and L. squamaria. Bartsia alpina is sister to a group of Rhinanthus, Larthraea, and Rhyynchocorys (in clade Vb only), but this relationship does not receive much support (52% MPBS, 60% MLBS). Euphrasia also is monophyletic (100% MPBS and MLBS). Bartsia trixago from Europe and B. crenata from Bolivia are paraphyletic to Parentucellia and are separate from B. alpina.

Relationships within clade VI receive less support than relationships within other clades, and they differ in the MP and ML trees. Multiple PHYA sequences were cloned from Melasma scabrum, Escobedia grandiflora, Xylocalyx asper, X. cartereis, and Striga linearifolia. The monophyly of Sophiba and Striga is well supported (99–100% MPBS and MLBS), and Radamaea montana is supported as sister to the two species of Nesogenes (99% MPBS, 95% MLBS).

**DISCUSSION**

This is the most extensive sampling of Orobanchaceae to date, representing a broad geographic and taxonomic range of taxa and including several genera that have not been included in previous phylogenetic studies. PHYA sequence data provide greater evidence of phylogenetic structure within the family than have analyses of plastid (dePamphilis, 1995; dePamphilis et al., 1997; Nickrent et al., 1998; Young et al., 1999; Young and dePamphilis, 2000; Olmstead et al., 2001). ITS (Wolfe et al., 2005), and nuclear 18S rDNA sequence data (Nickrent et al., 1998). Sixty-seven percent of the nodes in the parsimony bootstrap analysis and 71% in the maximum likelihood bootstrap analysis receive bootstrap support of over 90%. Support is greatest towards the terminal portions of the tree, although many of the nodes along the spine are also well supported.

The monophyly of Orobanchaceae, containing holoparasites, hemiparasites, and nonparasitic species, is well supported (100% MPBS and MLBS). This corroborates previous analyses of plastid and nuclear ribosomal DNA sequences and provides further evidence that the traditional classifications of Orobanchaceae and Scrophulariaceae (e.g., Bentham, 1846, 1876) do not reflect the phylogenetic relationships of these families.

The PHYA data support six major clades (66–100% MPBS, 91–100% MLBS). Relationships among these clades also are supported by the PHYA data. Clades V and VI are sister taxa, and these are sister to clade IV. The position of Brandisia as the sister taxon to these three clades receives less support, but the position of the completely holoparasitic clade III as sister to the majority of the hemiparasitic and holoparasitic taxa is well supported. Thus, the major exception to the almost universal support for broad phylogenetic structure within Orobanchaceae in the PHYA tree is at the base of the tree, where the PHYA data are ambiguous as to the branching order of clades I (the nonparasitic genus, Lindenbergia) and II (the earliest diverging clade of hemiparasites) (cf. Figs. 1 and 2).

These results differ from those obtained in previous studies. An analysis of combined plastid data supported a clade of mostly European genera that is consistent with PHYA clade V and an African clade that is consistent with PHYA clade VI (the tropical and subtropical genera), but provided low support for a clade corresponding to PHYA clade III (Orobanchaceae and related holoparasites); Castilleja and Pedicularis were not united in the plastid tree as they are in PHYA clade IV, and bootstrap support for relationships among major clades was below 50% (Young et al., 1999). However, the one species of the nonparasitic Lindenbergia that was included was well supported as the sister taxon of the remaining species, which included the monotypic Schwalbea (Young et al., 1999). An analysis of ITS data also supported clades that are consistent with PHYA clades V and VI, but provided only low support for a clade corresponding to PHYA clade IV. Orobanchaceae and holoparasitic relatives were resolved with bootstrap support below 50% as paraphyletic to the ITS African–SW Asian clade, and the position of Lindenbergia as sister to all parasites received low support (Wolfe et al., 2005). As with the plastid data, relationships among the major clades received bootstrap support below 50%, except for a sister group relationship between clades that correspond to PHYA clades IV (Castilleja, Pedicularis, and relatives) and V (mostly European genera), which received 71% MP bootstrap support (Wolfe et al., 2005).

The basal branches of Orobanchaceae—Recent molecular studies have placed Lindenbergia philippensis as sister to the rest of Orobanchaceae but they included no other species of Lindenbergia (Nickrent et al., 1998; Young et al., 1999; Young and dePamphilis, 2000; Olmstead et al., 2001; Wolfe et al., 2005). This would suggest that parasitism evolved only once in the family and was retained by all members. Lindenbergia and other Orobanchaceae share the same patterns of corolla derivation, although the pollen of Lindenbergia differs from that of the parasitic species and is more similar to nonparasitic members of Lamiales (Armstrong and Douglas, 1989; Minkin and Eshbaugh, 1989).

 Parsimony analysis of PHYA data resolves four species of Lindenbergia as the sister group of the rest of the family,
followed by clade II (Fig. 1), but this topology is not supported (50% MPBS). The alternative topology from the ML analysis (Fig. 2), in which Lindenbergia and clade II are sister taxa and form the first branch of the family, also receives low bootstrap support (67% MLBS). Topology tests indicate that neither topology can be rejected by the data. The exact relationship between clades I and II therefore remains unresolved in this study, but PHYA data provide the first strong evidence that clade II is one of the earliest diverging branches of the family. Moreover, the PHYA data provide new insight into the relatives of the monotypic Schwalbea americana. A close relationship among Bungea, Cymbaria, Monochasma, Schwalbea, and Siphonostegia had been suggested by earlier taxonomists (e.g., Bentham, 1876; Franchet and Savatier, 1879; Hedge, 1979a, b; Greuter, 1987), but until the PHYA study, they had not been included in any of the molecular phylogenetic studies. All five genera of clade II are species poor and have disjunct distributions. Schwalbea is monotypic and is restricted to the southeastern United States (Musselman and Mann, 1977), while Bungea, with two species, occurs in Iran, Turkey, and central Asia (Hedge, 1979b). The three other genera have Asian or disjunct distributions—Cymbaria with four species in the Ukraine, Russia, Mongolia, and China (Golubkova, 1955); Monochasma, with two species from Japan and central China (Yamazaki, 1993); and Siphonostegia, with two species in China, Korea, Taiwan, and Japan (Kusano, 1908; Yamazaki, 1993), and one in the Mediterranean (Hedge, 1979a) the last sometimes treated as a separate genus Lesquerueuxia. Schwalbea americana is federally listed as endangered. Its position at the base of the Orobanchaceae tree, in a small clade with Asian and Mediterranean species, highlights its importance as a resource of unique genetic diversity in the family.

The distribution of holoparasitism in Orobanchaceae—

The PHYA tree indicates that the split of clade III (Orobanche and its holoparasitic relatives) from the rest the family predates the major diversification of hemiparasites. Clade III contains ca. 195 species and is by far the most species-rich group of holoparasites in Orobanchaceae. Orobanche, the largest genus, is monophyletic and is sister to a clade that includes Boscniakia, Conopholis, and Epibagus. This conflicts with both the plastid (Young et al., 1999) and ITS (Wolfe et al., 2005) trees, in which Orobanche is not monophyletic. The PHYA sequences are more informative about relationships in clade III than are the ITS data (Wolfe et al., 2005), suggesting that expanded sampling of PHYA data from this clade may prove useful in placing genera such as Cistanche Hoffmans, & Link and Diphylopaea Nicolson, which are unresolved in the plastid (Young et al., 1999) and the ITS (Schneeweiss et al., 2004a; Wolfe et al., 2005) trees.

Other transitions to holoparasitism in the family have occurred in clades V and VI. The partial or complete loss of photosynthetic ability has occurred several times within clade VI. Hyobanche and Aeginetia are entirely heterotrophic, whereas genera such as Striga and Alectra have reduced photosynthetic capacity (de la Harpe et al., 1981; Graves et al., 1992). In contrast, Lathraea, in clade V, is the only European genus that is entirely holoparasitic. Tozzia alpina may represent a transitional state between hemi- and holoparasitism. It has a prolonged holoparasitic subterranean phase of up to 10 years, but when the plant emerges it becomes photosynthetic (Weber, 1973).
The 

\textit{Lamourouxia} from a study of its morphology, and the rps2 analysis of Young et al. (1999) did not resolve its closest relative from among the Castillejinae. \textit{Esterhazya}, a small genus from Brazil, Bolivia, and Paraguay, is sister to \textit{Agalinis}, which is consistent with the high degree of morphological similarity between the two genera (Barringer, 1985).

The \textit{PHYA} tree contains most of the genera that likely comprise clade V; \textit{Hedbergia Molau, Nothobartsia Bolliger} & Molau and the four segregate genera of \textit{Odontites} (\textit{Macro-syringion} Rothm., \textit{Oddontiella} Rothm., \textit{Bormuellera} Rothm., and \textit{Bartsiella} Bolliger) (Rothmaler, 1943; Bolliger, 1992), are putative members that remain unsampled. The phylogeny of the multiple \textit{PHYA} copies found in some species of clade V is consistent with a history of gene or genome duplication. The split of \textit{V_a} and \textit{V_b} indicates a gene/genome duplication early in the history of clade V, with later duplications/doubleings occurring in \textit{Lathraea} and \textit{Euphrasia}. The hexaploid number of \textit{Lathraea} (Gates and Latter, 1927) is consistent with such a scenario. It is possible that the clade \textit{V_b} genera have a \textit{PHYA} that is closely related to \textit{Lathraea PHYA \#3} and \textit{Bartsia alpina \#2}, but it is also possible that this copy has been silenced, a well-documented process in polyploids (e.g., Adams and Wendel, 2005) and the expected fate of duplicate genes (Walsh 1995; Lynch and Conner, 2000).

Within clade \textit{V_b}, species relationships are well supported and are consistent with results from other analyses. \textit{Melampyrum} is monophyletic and, consistent with the ITS and plastid trees (Young et al., 1999; Olmstead et al., 2001; Wolfe et al., 2005), is sister to the remaining species. However, in contrast to the ITS tree (Wolfe et al., 2005), \textit{Lathraea} and \textit{Rhinanthis} are unambiguously resolved as sister taxa, with the Eurasian genus \textit{Rhynochocorys} as their sister taxon. \textit{Bartsia alpina}, an enigmatic species that was placed in its own section of \textit{Bartsia} by Molau (1990), is sister to these three genera. A second major clade of mostly European genera that is well supported by the \textit{PHYA} data includes \textit{Bartsia} (except \textit{B. alpina}), \textit{Odontites}, \textit{Parentucellia}, \textit{Tozzia}, and \textit{Euphrasia}. Both \textit{Bartsia} (~49 species; Molau, 1990) and \textit{Euphrasia} (~350 species; Fischer, 2004) are broadly distributed, occurring in both the New and Old World and in both the northern and southern hemispheres. Phytochrome data support the monophyly of \textit{Euphrasia}, but \textit{Parentucellia} is nested within \textit{Bartsia}. Denser sampling will be required to resolve relationships of \textit{B. alpina} with the other European and Andean \textit{Bartsia} species and with allied genera such as \textit{Hedbergia} from Africa, regarded as a primitive member of the group by Molau (1988), and the European \textit{Nothobartsia} (Bolliger and Molau, 1992). \textit{Pseudobartsia} D.Y. Hong is probably not related to \textit{Bartsia}, but may be related to \textit{Phalteiraspernum} (Deyuan et al., 1998).

The \textit{PHYA} tree unites clade V with clade VI (99% MPBS, 82% MLBS). Several of the genera belonging to clade VI, of mostly tropical to subtropical distribution, have not been sampled. The phylogeny of \textit{Bartsia} is nested within the family. \textit{Nesogenes} from Madagascar is sister to \textit{Nesogenes}, a genus of one species in Tanzania and seven species confined to islands in the South Pacific and Indian Ocean (Marais, 1981), which accords with Hemsley (1913), who noted a similarity in habit in the genera. \textit{Cyclocheliton}, another African genus of uncertain affinity (Marais, 1981), has also recently been identified as a member of Orobanchaceae (Oxelman et al., 2005), but its placement within the family was not determined and we did not have the material necessary to include it in the \textit{PHYA} study. It has three species from Somalia and Ethiopia (Marais, 1981) and may therefore also be a member of clade VI. Our results also confirm that the previously unsampled South American genus \textit{Escobedia} belongs in this clade. Clade VI thus represents a morphologically diverse group of tropical and subtropical lineages from both the New and Old World, with the bulk of the species in the southern hemisphere. Several of the relationships suggested by the topology of this clade are consistent with those suggested by the ITS tree (Wolfe et al., 2005). \textit{Sophitia} is the sister group of the remaining genera, and \textit{Alectra} and \textit{Melasma} are sister genera in both the \textit{PHYA} and ITS trees; the \textit{PHYA} data additionally suggest that \textit{Escobedia} belongs in a clade with these two genera. The placement of the Asian \textit{Aeginetia} in a clade with \textit{Hyobanche} is well supported by the \textit{PHYA} and ITS data, as is the position of \textit{Buchnera} in a clade with \textit{Striga}. However, the position of \textit{Cycnium} in the two trees differs in a way that cannot be accounted for by differences in taxonomic sampling, and this needs to be explored with additional data.

Few of the genera in clade VI are rich in species: \textit{Buchnera} (100 species), \textit{Sophitia} (40 species), and \textit{Striga} (40 species) are the largest (Philcox, 1965, 1990; Hansen, 1975). Branch lengths in the clade are among the longest in the \textit{PHYA} tree, and there is evidence of gene duplication in the taxa that we sampled. It is possible that the combination of small, divergent genera, gene duplication and loss, codon bias, and/or elevated rates of nucleotide evolution (discussed later) have contributed to the lack of strong support for relationships among genera in clade VI. Sampling of additional taxa that might belong to this clade should be a priority for future studies of Orobanchaceae.

**Molecular evolution of phytochrome A in Orobanchaceae**—We detected no stop codons in any of the \textit{PHYA} clones that we sequenced, most of which were approximately 1.8 kb of coding sequence (half of the total coding sequence length). We also have found that Orobanchaceae \textit{PHYA} sequences appear to be evolving under constraints similar to \textit{PHYA} sequences in autotrophs, although selection appears to be slightly relaxed in holoparasites (J. R. Bennett and S. Mathews, unpublished manuscript). Finally, a full-length cDNA has been cloned from \textit{Orobanche} \textit{minor} that, when expressed in \textit{phyA} null mutants of \textit{Arabidopsis}, restores some \textit{phyA} functions (Okaawara et al., 2005; Trakulnaleamsai et al., 2005). These observations suggest that while some phytochrome-mediated responses in holoparasites are lost or altered, retention of a functional \textit{phyA} is important for their survival. They also suggest that the transition to parasitism involves modifications in light-signaling pathways and that further studies of \textit{Orobanche} phytochromes will bear on our understanding of the mechanisms that underlie physiological changes that take place during this transition.

The discovery of novel introns in \textit{PHYA} sequences in some species of Orobanchaceae was unexpected. Phytochrome gene structure is generally well conserved across land plants and coding sequences are usually interrupted by three introns. Exceptions include \textit{PHYB} of the green alga \textit{Mesotaenium caldariorum} (Laghr.) Hansg. with 11 introns (Lagarias et al., 1995) and \textit{PHYC} of \textit{Arabidopsis thaliana}, which lacks intron III (Cowl et al., 1994). One of the Orobanchaceae introns, intron A, was found in clades II and V, a distribution
that, along with their low sequence similarity, suggests independent gain in these two clades, despite their identical position of insertion into the coding sequence.

As noted, we found evidence of gene or genome duplication in clades IV, V, and VI. This is not unexpected given the documentation of polyplody in the family, including in Aeginetia (Schneeweiss and Weiss, 2003), Castilleja (Heckard and Chuang, 1977), Euphrasia (Barker et al., 1988), Lathraea (Gates and Latter, 1927), Odontites (Bolliger et al., 1990), Orobanche (Schneeweiss et al., 2004b), Orthocarpus (Chuang and Heckard, 1982), and Striga (Iwo et al., 1993). Nonetheless, the species phylogenies suggested by the topologies of the PHYA clades are consistent with other lines of evidence. While the topology of the gene tree could result from differential gene silencing (Adams et al., 2003; Lai et al., 2004; Adams and Wendel, 2005) following a whole-genome duplication event that predated the origin of Orobancheae (Aagaard et al., 2005), the pattern also is consistent with more recent single gene duplication, segmental, or genome-doubling events.

Significant rate heterogeneity was observed among lineages in maximum likelihood analyses using PAML (Yang, 1997) (J. R. Bennett and S. Mathews, unpublished manuscript). Increases in substitution rates are found in the branch leading to clade VI, in several branches within this clade, and in the branch leading to the small clade of duplicate PHYA sequences from Lathraea clandestina, L. squamaria, and Bartisia alpina. While clade VI is not species-rich compared with other clades of Orobancheae, it encompasses the greatest degree of diversity in habit and morphology. It thus would be interesting to determine whether this diversity is correlated with a more general, genome-wide elevation in rates of genetic change (e.g., Barralough and Savolainen, 2001; Webster et al., 2003). Additionally, clade VI is characterized by a significant increase in codon bias compared with the rest of the family (J. R. Bennett and S. Mathews, unpublished manuscript). Codon bias can have a major effect on the molecular evolution of proteins (Singer and Hickey, 2000) and may also confound the reconstruction of phylogeny (Lockhart et al., 1992; Foster and Hickey, 1999; Mosers and Holmes, 2000; Inagaki et al., 2004; Christianson, 2005) and is often correlated with reduced rates of synonymous substitution. The unusual combination of elevated rates of nucleotide evolution and codon bias in clade VI may have implications for functional divergence in phyA of members of this clade.

Conclusions—PHYA sequences have provided a robust hypothesis of phylogenetic structure within Orobancheae, based on the most extensive sampling of genera to date. The tree places all genera within well-supported major clades and identifies relationships among these clades. In several cases, including Pedicularis and Orobanche, it appears that PHYA data will be useful for resolving relationships within genera. The extended taxon sampling of this study reveals the composition of the earliest diverging lineage of hemiparasites and highlights the instability in the position of the root. This has important implications for the evolution of parasitism and represents a result that needs to be explored with additional sampling of characters and taxa. Furthermore, despite the high levels of both MP and ML bootstrap support for relationships in the PHYA tree, which are consistent with results from previous molecular analyses, it is a single gene tree that needs to be tested with additional data.

LITERATURE CITED


VON WETTSTEIN, R. 1891. Scrophulariaceae. In A. Engler and K. Prantl
APPENDIX. Species sampled, voucher specimens and intron lengths for taxa included in this study. — cDNA sequence; n/s = intron not sequenced because fragment was amplified with primers that do not span the intron region; * sequence obtained from GenBank.
AM233970; Bennett 57 (FHO); 547, absent, n/s. L. clandestina L. #2—AM233971; Bennett 57 (FHO); 411, absent, 318. L. clandestina L. #3—AM233972; AM234040; Bennett 57 (FHO); 283, absent, 137. L. squamaria L. #1—AM233973; Frajan s.n. (LJU); 298, absent, 325. L. squamaria L. #2—AM233974; Frajan s.n. (LJU); 515, absent, 317. L. squamaria L. #3—AM233975; Frajan s.n. (LJU); n/s, n/s, 137. Lindenberga indica (L.) Vatke—AM233976, AM234041; Dickason 5437 (GH); absent, 291. L. muraria (Roxb.) Brühl—AM233977; Grierson & Long 2312 (E); absent, 290. L. philippensis (Cham. & Schltdl.) Benth.—AM233978; Armstrong 1163 (ISU); absent, 696. L. sokotrana Vierh.—AM233918; Miller 1127 (E); absent, n/s.

Melampyrum arvense L.—AM233919; Schwein. 1317/02 (WU); 257, absent, 103. M. carstiense Fritsch—AM233979; Krajesk s.n. (LJU); 257, absent, 243. M. nemorum L.—AM233980; Braysatraker 96-M17-1 (RNG); 97, absent, 72. M. pratense L.—AM233981; Nyffler s.n. (GH); 253, absent, 306. M. sylvaticum L.—AM233982; Krajesk s.n. (LJU); 88, absent, 72. Melasoma scabrum P.J. Bergius #1—AM233983; Steiner 3496 (CAS); absent, n/s. M. scabrum P.J. Bergius #2—AM233984; Steiner 3496 (CAS); absent, 323. Monochauma shearei Franch. & Sav.—AM233985, AM234042; Ye 3921 (MO); 227, 108, 175.

Nesogenes africanus G. Taylor—AM233986; Bidgood 2063 (MO); absent, n/s. N. dupontii Hems.—AM233987; Jongkind 968 (MO); absent, n/s. Odontites himalayicus Pennell—AM233988; Webster 6290 (GH); 507, absent, 307. Orobanche californica Cham. & Schltdl.—AM233989; Colwell AC-97-CA-BLO (WTU); absent, 503. O. corymbosa (Ryd.) Ferris—AM233990; Leihold s.n. (WTU); absent, 559. O. densiflora Reut.—AM233991; Bennett 65 (FHO); absent, n/s. O. fasciculata Nutt.—AM233992; Beardsley s.n. (WTU); absent, 494. O. gracilis Sm.—AM233993; Bennett 82 (GH); absent, n/s. O. pinorum Geyer—AM233994; Colwell AC96-WA-1C (WTU); absent, n/s. O. raddeana Beck—AM233995; Schwein. 342 (WU); absent, 613. O. ramosa L.—AM233996; Hughes 2191 (FHO); absent, 574. Orthocarpus bracteosus Benth.—AM233997; Wolfe 571 (OS); absent, 947. O. tenuifolius A. Gray—AM233998, AM234043; Lavin 72892 (MONT); absent, n/s. Parentucellia latifolia (L.) Caruel—AM233999; Jury 18765 (RNG); 484, absent, 312. P. viscosa (L.) Caruel—AM234000; Bennett 66 (FHO); 490, absent, 312. Paulownia tomentosa (Thunb.) Steud.—AM234001; Olimstead 88-008 (WTU); absent, 172. Pedicularis densisepica Maxim.—AM234002; Boufford et al. 28416 (GH); absent, 899. P. foliosa L.—AM234003; dePamphilis s.n.; absent, 971. P. gryrophylcha Maxim.—AM234004; Boufford et al. 28435 (GH); absent, 896. P. julca E. Mayer—AM234005; Krajesk s.n. (LJU); absent, 896. P. kansuensis Maxim.—AM234006; Boufford et al. 27695 (GH); absent, 889. P. kernerii Dalla Torre—AM234007; Nyffler s.n. (GH); absent, 819. P. tuberosa L.—AM234008; Nyffler s.n. (GH); absent, 817. Penstemon cobaea Pennell—AM234009; Bennett 77 (GH); absent, 460. Phtheirospermum japonicum (Thunb.) Kantz.—AM234010; Tsugaru & Sawaiha 28464 (GH); absent, 857. Radamaea montana Bentham.—AM234011, AM234044; Rakutamaala 1304 (MO); absent, n/s. Rhinanthus alectorophorus (Scop.) Pollich—AM234012; Nyffler s.n. (GH); 264, absent, 290. R. freynii Fiori.—AM234013; Bennett 88 (GH); 305, absent, 302. R. glacialis Personn.—AM234014; Nyffler s.n. (GH); 298, absent, 252. R. minor L.—AM234015; Bennett 84 (GH); 359, absent, 290. Rhynchosyris orientalis Bentham.—AM234016; Albach & Schwein. s.n. (WU); 237, absent, 264.

Schwarzia americana L.—AM234017; Kirkman s.n. (PAC); 619, 421. 714. Scrophularia arguta Sol.—AM234018; Miller 17017A (E); absent, 92. Seymouria lacinata (M. Martens & Galeotti) Standl.—AM234019; Egger 1201 (UOM); absent, n/s. Siphonostegia chinensis Bentham.—AM234020; Bartholomew et al. 1435 (GH); 652, 410, n/s. Solanum tuberosum L. *—SA4872; —,—. —,—. Sophonia cana Harv.—AM234021; Randle 135 (OS); absent, 229. S. lanata Engl.—AM234022; Bidgood 3601 (MO); absent, n/s. S. ramosa Hochst.—AM234023; Kayumbo 2752 (MO); absent, 234. Striga asiatica (L.) Kunze—AM234024; Parker 1213 (RNG); absent, n/s. Stylochaeta bifiabata Kunze—AM234025; Steiner 3273 (NBG); absent, n/s. S. gesneroides (Willd.) Vatke—AM234026; Jury 19619 (RNG); absent, n/s. S. linearia (Schumach. & Thonn.) Hepper #1—AM234027; Bybeeiess B2334 (RNG); absent, 392. S. linearifolia (Schumach. & Thonn.) Hepper #2—AM234028; Bybeiers B2334 (RNG); absent, n/s. S. orobanchoides (R. Br.) Bentham.—AM234029; Steiner s.n. (NBG); absent, n/s. S. strobilanthus attenuata Nees.—AM234030; Chilterns Seeds, UK; absent, 83. Tocchia alpina L.—AM234031; Bennett 87 (GH); 389, absent, 309. Triphysaria pusilla (Benth.) T.I. Chuang & Heckard—AM234032; Colwell 92-CA-TYC (MO); absent, 822. Xylocalyx asper Balf. f. #1—AM234033; Miller 17076 (E); n/s, n/s. X. asper Balf. f. #2—AM234034; Miller 17076 (E); n/s, n/s. X. carterae Thulin #1—AM234035; Wieland 4387 (MO); n/s, n/s. X. carterae Thulin #2—AM234036; Wieland 4387 (MO); n/s, n/s.