

The Utility of Nuclear *gapCp* in Resolving Polyploid Fern Origins

Eric Schuettpelz,¹ Amanda L. Grusz, Michael D. Windham, and Kathleen M. Pryer

Department of Biology, Duke University, Durham, North Carolina 27708 U.S.A.

¹Author for correspondence (ejs7@duke.edu)

Communicating Editor: James F. Smith

Abstract—Although polyploidy is rampant in ferns and plays a major role in shaping their diversity, the evolutionary history of many polyploid species remains poorly understood. Nuclear DNA sequences can provide valuable information for identifying polyploid origins; however, remarkably few nuclear markers have been developed specifically for ferns, and previously published primer sets do not work well in many fern lineages. In this study, we present new primer sequences for the amplification of a portion of the nuclear *gapCp* gene (encoding a glyceraldehyde-3-phosphate dehydrogenase). Through a broad survey across ferns, we demonstrate that these primers are nearly universal for this clade. With a case study in cheilanthoids, we show that this rapidly evolving marker is a powerful tool for discriminating between autopolyploids and allopolyploids. Our results indicate that *gapCp* holds considerable potential for addressing species-level questions across the fern tree of life.

Keywords—allopolyploidy, autopolyploidy, ferns, *gapCp*, GAPDH, molecular systematics, nuclear marker.

Polyploidy—the multiplication of entire chromosome sets—has been documented in nearly all major eukaryotic lineages (Otto and Whitton 2000; Gregory and Mable 2005; Tate et al. 2005). The process is especially prevalent in ferns, which exhibit both the highest known gametic chromosome numbers (Abraham and Ninan 1954) and some of the highest incidences of polyploidy (Walker 1966; Manton and Vida 1968; Löve et al. 1977; Walker 1984). Nearly 50% of fern species that have been studied cytogenetically are polyploids of recent origin (neopolyploids, following Ramsey and Schemske 2002), exhibiting chromosome numbers that are multiples of those documented in closely related species (Vida 1976; Walker 1984). In addition, it is estimated that at least 95% of fern species have undergone polyploidization at some point in their evolutionary history (Grant 1981; Haufler 1987). Polyploidy has left an indelible mark on fern evolution and continues to serve as a dynamic source of genetic variability, ecological innovation, and species diversity (e.g. Klekowski and Baker 1966; Klekowski 1972; Walker 1984; Werth et al. 1985; Soltis and Soltis 1987; Werth and Windham 1991; Haufler et al. 1995).

Evolutionary biologists typically recognize two major categories of neopolyploids (see Soltis et al. 2007 and references therein). Organisms containing multiple genomes from a single diploid species are generally called autopolyploids; those that incorporate genomes derived from two or more diploid species are called allopolyploids. Although these categories are widely used, the distinction can be difficult to operationalize because it requires detailed knowledge of diploid progenitor populations (which may be unknown or extinct) and is dependent on the species concept applied (Soltis et al. 2007). Nonetheless, it is important to be able to assign individuals or taxa to these categories, because autopolyploids and allopolyploids exhibit fundamental differences in their genetics, ecology, and evolutionary potential (Levin 1983; Thompson and Lumaret 1992; Soltis and Soltis 2000; Wendel 2000; Ramsey and Schemske 2002; Osborn et al. 2003).

Through the years, evolutionary biologists have used a variety of techniques to discriminate between autopolyploids and allopolyploids. Morphology, chromosome pairing behavior, and patterns of genetic segregation (based primarily on allozyme data) all provide important clues regarding polyploid origins (Grant 1981; Jackson 1982; Soltis and

Rieseberg 1986). However, these approaches are inherently phenetic, relying exclusively on the genetic similarity of the genomes involved. As we move toward more integrated species concepts, it becomes increasingly important to place the genomes found in polyploid organisms in a phylogenetic context. Thus, DNA sequencing—specifically of nuclear and organellar markers in combination—is emerging as a powerful tool for revealing polyploid origins (Ge et al. 1999; Sang and Zhang 1999; Hoot and Taylor 2001; Popp and Oxelman 2001; Popp and Oxelman 2007). Biparentally-inherited nuclear markers provide sequences unique to individual diploid species, which can be isolated from polyploids through cloning and analyzed in a phylogenetic context. This allows for a more objective assessment of whether constituent genomes came from a single diploid species (autopolyploid) or more than one diploid species (allopolyploid). Maternally-inherited (Sears 1980; Gastony and Yatskievych 1992) organellar markers, in turn, can distinguish the maternal from paternal species.

We now have many primer sets at our disposal for the amplification of organellar (primarily plastid) DNA from ferns (Small et al. 2005), but there are currently few options available for obtaining nuclear sequences in this group of plants. To date, most nuclear sequencing in ferns has focused on ribosomal DNA (rDNA) markers. Although the small ribosomal subunit gene (18S) is broadly amplifiable using published primers (Bult et al. 1992; Wolf 1995), it is of limited utility even at the deepest phylogenetic levels (Schuettpelz et al. 2006), and therefore certainly not useful within a species complex. The large subunit gene (26S) and internal transcribed spacers (ITS) may provide increased phylogenetic signal, but researchers have had little success in sequencing these regions in ferns (Gastony and Rollo 1998; Van den heede et al. 2003; Reid et al. 2006). In any case, rDNA sequences are not an ideal choice for assessing polyploid origins in ferns (or any other group), as the effects of concerted evolution may obscure evidence of reticulation (Wendel et al. 1995; Álvarez and Wendel 2003; Kovarik et al. 2004, 2005; Lihová et al. 2006; Lim et al. 2007; Volkov et al. 2007). Single- or low-copy nuclear genes are not nearly as susceptible to this process (Wendel 2000; Ma and Gustafson 2005), and therefore offer the greatest promise for revealing polyploid origins (Small et al. 2004).

Two single-copy nuclear markers have been developed for

ferns, with primers published for the amplification of portions of the *pgiC* (Ishikawa et al. 2002) and *gapCp* (Ebihara et al. 2005) genes. Unfortunately, our efforts to amplify these markers using the published primers were unsuccessful for cheilanthoid ferns—a large clade within the Pteridaceae (Smith et al. 2006; Schuettelpelz et al. 2007) wherein polyploidy is especially rampant (Windham and Yatskievych 2003). So too were our attempts to amplify other low-copy nuclear markers (e.g. portions of the *adh*, *gapC*, *leafy*, and *waxy* genes) from cheilanthoids with primer sets not specifically designed for ferns (Frohlich and Meyerowitz 1997; Mason-Gamer et al. 1998; Small et al. 1998; Wall 2002).

In an effort to obtain a marker for use in identifying polyploid origins in cheilanthoid ferns, we developed new primers for the amplification and sequencing of a second region of the nuclear *gapCp* gene, adjacent to the segment utilized previously (Ebihara et al. 2005). Here, we first assess the universality of our primers through a broad survey across ferns. Then, with a case study focused on a cheilanthoid species complex (see Discussion for details on this complex), we demonstrate the utility of this newly-developed marker for discriminating between autopolyploid and allopolyploid species.

MATERIALS AND METHODS

Primer Design—To identify conserved priming sites that would potentially allow for amplification of *gapCp* across ferns, we compared a published *gapCp* cDNA sequence from the fern *Marsilea* L. (GenBank accession AJ003783; Meyer-Gauen et al. 1998) to published *gapCp* gene sequences from the seed plant *Pinus sylvestris* L. (GenBank accession AJ001706; Meyer-Gauen et al. 1998) and several other embryophyte species (Petersen et al. 2003). Suitable (i.e. conserved) priming sites were located within exons 8 and 11 (of the *Pinus gapCp* gene), and primers (ESGAPCP8F1 and ESGAPCP11R1) were designed to amplify portions of these exons and the elements of the *gapCp* gene between them (for a map and primer sequences, see Fig. 1).

Taxonomic Sampling—For our broad survey, to assess the universality of our primers across ferns and to determine whether multiple copies of the *gapCp* gene were present, we sampled a single individual from each of 11 fern species—representing a variety of leptosporangiate lineages (Appendix 1). For our case study, to demonstrate the utility of this marker in assessing polyploid origins, we focused on a small cheilanthoid species complex—*Cheilanthes* subgenus *Physapteris* section *Fendlera* (Reeves 1979). From this complex we included three of the four known diploids (*C. covillei* Maxon, *C. fendleri* Hook., and *C. lindheimeri* Hook.) and two apomictic triploids of uncertain parentage (*C. lindheimeri* Hook. and *C. wootonii* Maxon); the fourth known diploid (*C. clevelandii* D. C. Eaton) is closely related to *C. covillei* (A. L. Grusz et al. unpubl. results) and its exclusion here does not impact our findings. From outside of this complex, we included one additional diploid (*C. parryi* (D. C. Eaton) Domin). Two individuals were sampled from each taxon in the case study (Appendix 2).

DNA Isolation, Amplification, Cloning, and Sequencing—For each sampled individual, genomic DNA was extracted from silica-dried material using the DNeasy Plant Mini Kit (Qiagen, Valencia, California), following the protocol described in Schuettelpelz and Pryer (2007). Amplifications from these extractions were carried out using the polymerase chain reaction (PCR), with 1 × PCR buffer IV containing MgCl₂ (ABgene, Epsom, United Kingdom), 200 μM each dNTP, 100 μg/ml BSA, 50 U/ml Taq polymerase, 0.5 μM primer ESGAPCP8F1, 0.5 μM primer ESGAPCP11R1, and 1 μl template DNA eluate in a 25 μl reaction. Ther-

mocycling programs entailed an initial denaturation step (94°C for 5 min) followed by 35 denaturation, annealing, and elongation cycles (94°C for 1 min, 55°C for 1 min, 72°C for 2 min) and a final elongation step (72°C for 10 min). The results of these reactions were visualized on an agarose gel.

PCR products were cloned using the TOPO TA Cloning kit (Invitrogen, Carlsbad, California). Cloning reactions included 0.33 μl salt solution, 0.33 μl water, 0.33 μl TOPO vector, and 0.66 μl PCR product, and were incubated for 30 minutes at room temperature. Transformations included only 0.66 μl of the cloning reaction and 16.7 μl of One Shot TOP10 competent cells, but otherwise followed the manufacturer's protocol. The transformed cells were spread on selective (ampicillin plus X-gal) LB plates and incubated overnight at 37°C. A minimum of 16 white colonies from each plate were individually picked and resuspended in 25 μl PCR reactions prepared as above, but with the M13 Forward (–20) and M13 Reverse primers supplied by Invitrogen. Thermocycling was conducted as described above, but with elongation cycles of 3 min. The results of these reactions were again visualized on an agarose gel. For the broad survey, multiple reactions from each observed size class were selected for sequencing. For the case study, multiple reactions from the appropriate size class only (see Results section below) were selected for sequencing.

Selected PCR products were purified using Montage PCR Centrifugal Filter Devices (Millipore, Billerica, Massachusetts), following the manufacturer's protocol. Sequencing of the cleaned PCR products employed the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, California). Each 10 μl reaction incorporated 0.375× BigDye Terminator Ready Reaction Mix (Applied Biosystems), 0.625× BigDye Terminator Sequencing Buffer (Applied Biosystems), 1 μM primer (for inserts up to 700 bp, only the ESGAPCP8F1 and ESGAPCP11R1 primers were used; for longer inserts, the M13 Forward (–20) and M13 Reverse primers were also employed), and 2 μl purified PCR product. Thermocycling and reaction purification followed the manufacturer's protocol. Sample electrophoresis and analysis were performed using an ABI 3730xl DNA Analyzer (Applied Biosystems). The multiple sequencing reads obtained as chromatograms from each individual purified PCR product were assembled and edited separately using Sequencher 4.5 (Gene Codes Corporation, Ann Arbor, Michigan). All consensus sequences (219 newly obtained) were subsequently deposited in GenBank (Appendices 1, 2).

Sequence Alignment and Phylogenetic Analysis—The consensus sequences obtained for the broad survey and the case study were separately aligned using MacClade 4.08 (Maddison and Maddison 2005). For the broad survey, *gapC* (GenBank accession L07501) and *gapCp* (GenBank accession AJ001706) exon sequences from *Pinus* were also included to aid in alignment and analysis. Ambiguously aligned regions (all three introns in the case of the broad survey; a few regions within the introns in the case study) were excluded from subsequent analyses. The included portions of the broad survey and case study data sets had 0.0% and 0.3% missing data, respectively.

Both data sets (TreeBASE study number S2015) were phylogenetically analyzed using GARLI version 0.951 (Genetic Algorithm for Rapid Likelihood Inference; Zwickl 2006), employing the GTR + I + G model of DNA sequence evolution. The analyses comprised eight independent runs, each from a random starting topology and using the default settings recommended by the author. To assess branch support, non-parametric bootstrap analyses (with 100 replicates) were conducted, also in GARLI (with the model and settings as above).

To further assess branch support, both data sets were also analyzed using a Bayesian Markov chain Monte Carlo (B/MCMC) approach in MrBayes 3.1.1 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003), again employing the GTR + I + G model of DNA sequence evolution. These analyses comprised four independent runs, each with four chains (one cold and three heated), and used the default (i.e. flat) priors. Chains were run for 10 million generations and trees were sampled from the cold chain every 1000 generations. To identify when analyses had reached stationarity, the standard deviation of the split frequencies among the independent runs (as calculated by MrBayes) was examined, and the output parameter estimates were plotted using Tracer 1.4 (Rambaut and Drummond 2007). Based on these convergence diag-



FIG. 1. Portion of the *gapCp* gene utilized in this study, with priming sites indicated. Exon and intron numbers follow those given for the *Pinus gapCp* gene (GenBank accession AJ001706; Meyer-Gauen et al. 1998). Note that although the exon lengths are to scale, the intron lengths are not (Table 1). ESGAPCP8F1 (5' to 3') = ATYCCAAGYTCAACTGGTGTCTGC; ESGAPCP11R1 (5' to 3') = GTATCCCCAYTCRTTGTCTACC.

nostics, the first 2.5 million generations were (very conservatively) excluded from each analysis before obtaining clade posterior probabilities with the "sumt" command.

RESULTS

Broad Survey—Initial amplifications with primers ESGAPCP8F1 and ESGAPCP11R1 (Fig. 1) yielded products of one to three distinct lengths for any given taxon, suggesting that either the nuclear *gapCp* gene underwent duplication at some point in the evolutionary history of ferns, or that some other gene (perhaps coding for another glyceraldehyde-3-phosphate dehydrogenase) was being amplified simultaneously in some instances. Through cloning we were able to isolate and sequence individual fragments, which ranged in size from about 500–1700 bp (Table 1; Appendix 1). All sequences obtained were readily alignable to the *gapC* and *gapCp* exon sequences from *Pinus*, without the need for any insertions or deletions in these protein-coding regions, indicating that they encoded glyceraldehyde-3-phosphate dehydrogenases. The presence of three amino acid sequence signatures in the (translated) exon sequences (proline at position 235, asparagine at position 236, and cysteine at position 244; Petersen et al. 2003) further revealed that nearly all were *gapCp* genes. Introns varied considerably in length (Table 1) and were not alignable across ferns, but all conformed to the GT–AG rule (i.e. all intron sequences began with GT and ended with AG; Breathnach et al. 1978).

Phylogenetic analysis of the exon data (a total of 296 characters) resulted in a tree with mixed levels of branch support (Fig. 2A). Sequences of a given size class, from a particular taxon, were essentially identical (generally differing only by a few substitutions) and thus were always well supported as monophyletic. The relationships of these sequence clusters to one another, however, were often without support, which was expected given the limited number of characters. The *gapC/gapCp* duplication was well supported, and it is clear that some sequences obtained from *Cheilanthes* and *Adiantum* L. were of *gapC* rather than *gapCp*. Furthermore, our data suggest that a *gapCp* duplication—perhaps yielding "long"

and "short" versions of the gene—occurred relatively early in the evolutionary history of ferns, although its precise phylogenetic position cannot yet be determined (Fig. 2B; see Discussion section below).

Case Study—As in the broad survey (Table 1), amplifications in the *Cheilanthes* case study yielded products of more than one length. With a desire to focus our efforts on a single paralog, and recognizing that *gapC* and the "long" version of *gapCp* in *Cheilanthes* were not readily discernable on a gel (with ranges of 929–930 and 917–967 bp, respectively; Table 1), we targeted only the "short" version of *gapCp* (594–602 bp; Appendix 2). Both the coding and non-coding regions of these sequences were readily alignable by eye, and it was necessary to exclude only 33 of 609 characters from the phylogenetic analyses due to questionable alignment. This was highly suggestive of homology, as the *gapCp* "short" introns from *Cheilanthes* are not even alignable to those from its closest sampled relative (*Adiantum*; Fig. 2); homology of these sequences was confirmed through phylogenetic analysis (results not shown).

Phylogenetic analysis of the *Cheilanthes* data set resulted in a tree with four distinct clades (three of which were strongly supported as monophyletic), corresponding to the four sampled diploid species (*C. covillei*, *C. fendleri*, *C. lindheimeri*, and *C. parryi*; Fig. 3); all sequences obtained from any given diploid fell within a single clade. All sequences obtained from our two accessions of the triploid *C. lindheimeri* (15 plus 19 sequences) fell within the diploid *C. lindheimeri* clade (intermixed with diploid *C. lindheimeri* sequences), supporting an autopolyploid ancestry for this taxon. By contrast, sequences obtained from two different accessions of the triploid *C. wootonii* (13 plus 15 sequences) grouped with both the *C. fendleri* and *C. lindheimeri* clades (again, intermixed with diploid sequences), indicating allopolyploid ancestry.

DISCUSSION

Glyceraldehyde-3-Phosphate Dehydrogenase Gene Duplications—Extant land plants possess four distinct glyceraldehyde-3-phosphate dehydrogenase (GAPDH) genes (*gapA*,

TABLE 1. Summary of sequences obtained in the broad survey using primers ESGAPCP8F1 and ESGAPCP11R1. Assignments to *gapC* or *gapCp* are based on the presence of amino acid sequence signatures (see text; Petersen et al. 2003). The "long" and "short" designations for *gapCp* are based solely on product length; if only one version of *gapCp* was retrieved, no length designation is given (see text). Exon and intron numbers follow those given for the *Pinus gapCp* gene (GenBank accession AJ001706; Meyer-Gauen et al. 1998). For exons 8 and 11, lengths provided are for the amplified and sequenced portions only.

Genus	Product	Clones	Exon 8 (bp)	Intron 8 (bp)	Exon 9 (bp)	Intron 9 (bp)	Exon 10 (bp)	Intron 10 (bp)	Exon 11 (bp)	All exons (bp)	All introns (bp)	Total (bp)
<i>Adiantum</i>	<i>gapC</i>	7	4	225	143	86–88	84	123–148	65	296	434–461	730–757
<i>Adiantum</i>	<i>gapCp</i> "long"	4	4	78	143	204	84	76	65	296	358	654
<i>Adiantum</i>	<i>gapCp</i> "short"	6	4	87	143	68	84	84–85	65	296	239–240	535–536
<i>Cheilanthes</i>	<i>gapC</i>	8	4	437–438	143	112	84	84	65	296	633–634	929–930
<i>Cheilanthes</i>	<i>gapCp</i> "long"	2	4	388–438	143	152	84	81	65	296	621–671	917–967
<i>Cheilanthes</i>	<i>gapCp</i> "short"	16	4	121–122	143	90–93	84	89	65	296	300–304	596–600
<i>Dennstaedtia</i>	<i>gapCp</i>	10	4	104–108	143	79–81	84	322–343	65	296	507–526	803–822
<i>Diplazium</i>	<i>gapCp</i>	7	4	245–246	143	416	84	100	65	296	761–762	1057–1058
<i>Dryopteris</i>	<i>gapCp</i> "long"	3	4	90	143	1175–1256	84	89	65	296	1354–1435	1650–1731
<i>Dryopteris</i>	<i>gapCp</i> "short"	8	4	115	143	158–170	84	78	65	296	351–363	647–659
<i>Hymenasplenium</i>	<i>gapCp</i>	5	4	80	143	107	84	86	65	296	273	569
<i>Hymenophyllum</i>	<i>gapCp</i>	5	4	314	143	337	84	91	65	296	742	1038
<i>Lindsaea</i>	<i>gapCp</i>	5	4	95	143	393	84	84	65	296	572	868
<i>Lygodium</i>	<i>gapCp</i>	8	4	203–205	143	231	84	82	65	296	516–518	812–814
<i>Pteris</i>	<i>gapCp</i> "long"	5	4	239–240	143	275	84	75	65	296	589–590	885–886
<i>Pteris</i>	<i>gapCp</i> "short"	5	4	79–90	143	88	84	79–83	65	296	250–257	546–553
<i>Sticherus</i>	<i>gapCp</i>	4	4	299	143	424	84	90	65	296	813	1109

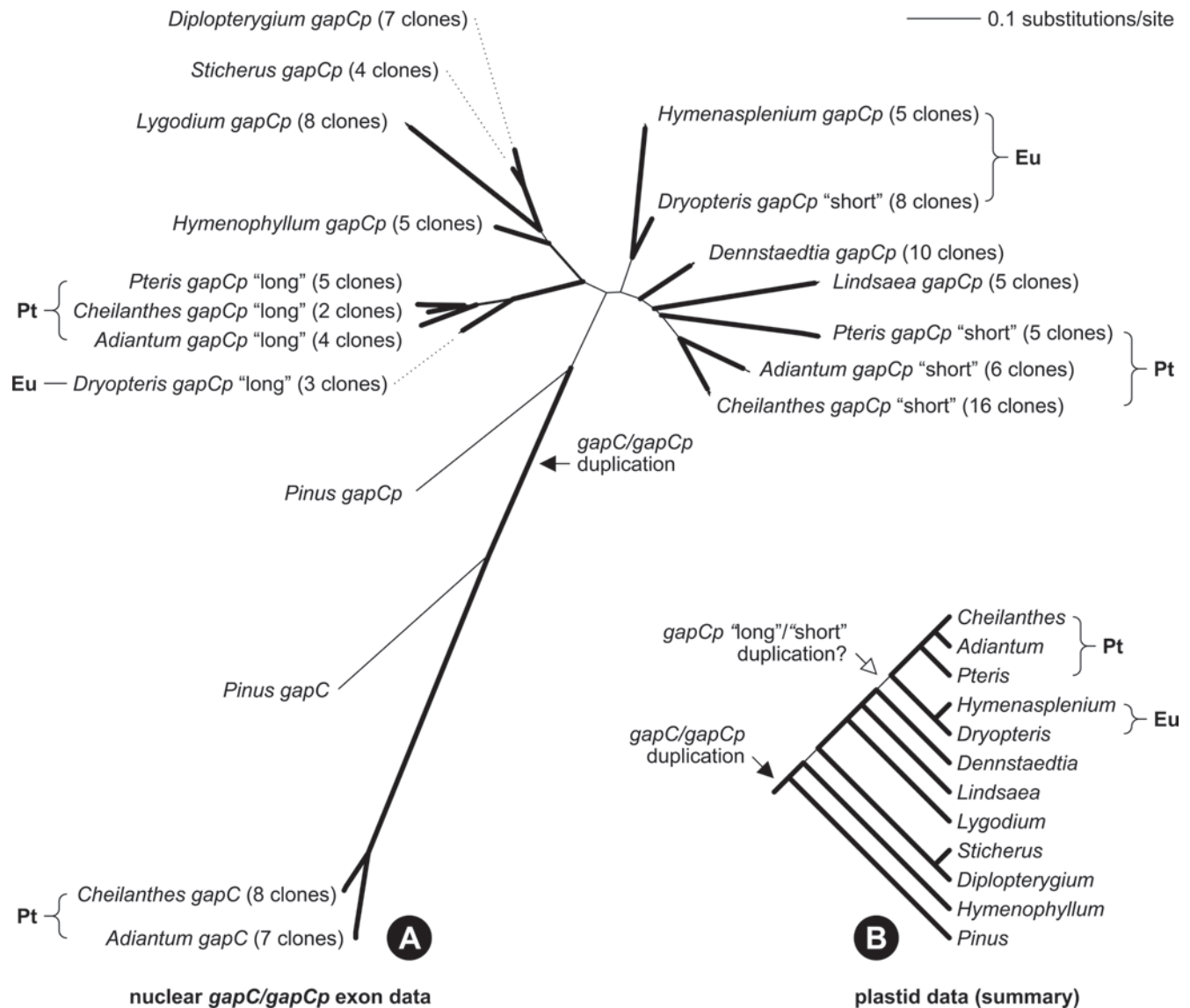


FIG. 2. Genealogy (A) resulting from maximum likelihood analysis of the broad survey *gapC/gapCp* data set (exons only; 296 bp), and summary cladogram (B) resulting from previous analyses of plastid data (Pryer et al. 2004; Schuettpelz et al. 2006; Schuettpelz and Pryer 2007). Genealogy resulting from Bayesian analysis of the broad survey data set yielded an essentially congruent topology (not shown). In both the genealogy (A) and the summary cladogram (B), heavily thickened branches (most subterminal branches in *gapCp* genealogy and nearly all branches in summary cladogram) received both a maximum likelihood bootstrap score $\geq 70\%$ and a Bayesian posterior probability ≥ 0.95 ; slightly thickened branches received either a maximum likelihood bootstrap score $\geq 70\%$ or a Bayesian posterior probability ≥ 0.95 . In the *gapC/gapCp* genealogy (A), the number of clones in each terminal clade is indicated in parentheses. The *gapC/gapCp* gene duplication (filled arrow) occurred well before the divergence of seed plants (*Pinus*) from ferns (all other genera), and represents the root of the presented genealogy. A *gapCp* duplication (open arrow, B) occurred sometime before the divergence of pteroid ferns (Pt) from eupolypod ferns (Eu), as "long" and "short" versions of the gene were recovered from both of these sister clades; however, its precise phylogenetic position remains unclear (see text).

gapB, *gapC*, and *gapCp*). The *gapA* gene is believed to be of plastid origin and the *gapC* gene is thought to be of mitochondrial origin (Martin and Cerff 1986; Brinkmann et al. 1987; Martin and Schnarrenberger 1997), but both were transferred to the host cell nucleus following primary endosymbiosis and both were subsequently duplicated (to yield *gapB* and *gapCp*, respectively) well before the emergence of embryophytes (Brinkmann et al. 1989; Meyer-Gauen et al. 1994; Meyer-Gauen et al. 1998; Petersen et al. 2003; Petersen et al. 2006). The early *gapC/gapCp* duplication can be visualized in the results of our broad survey across ferns (Fig. 2A). Although our primers were generally *gapCp*-specific, we also amplified *gapC* from two pteroid genera (*Adiantum* and

Cheilanthes). As expected, these sequences are more closely related to *gapC* from *Pinus* than they are to *gapCp* sequences from ferns (Fig. 2A).

More recent GAPDH gene duplications have been uncovered in several groups of plants (see Petersen et al. 2003 for a summary), but none have yet been characterized within ferns. Our broad survey, however, indicates that a *gapCp* duplication occurred relatively early in the evolutionary history of this clade. We recovered both "long" and "short" *gapCp* sequences from all sampled pteroid genera (*Adiantum*, *Cheilanthes*, and *Pteris* L.), as well as from one of two sampled eupolypod genera (*Dryopteris* Adans.; Table 1). Because pteroids and eupolypods are understood to be sister clades

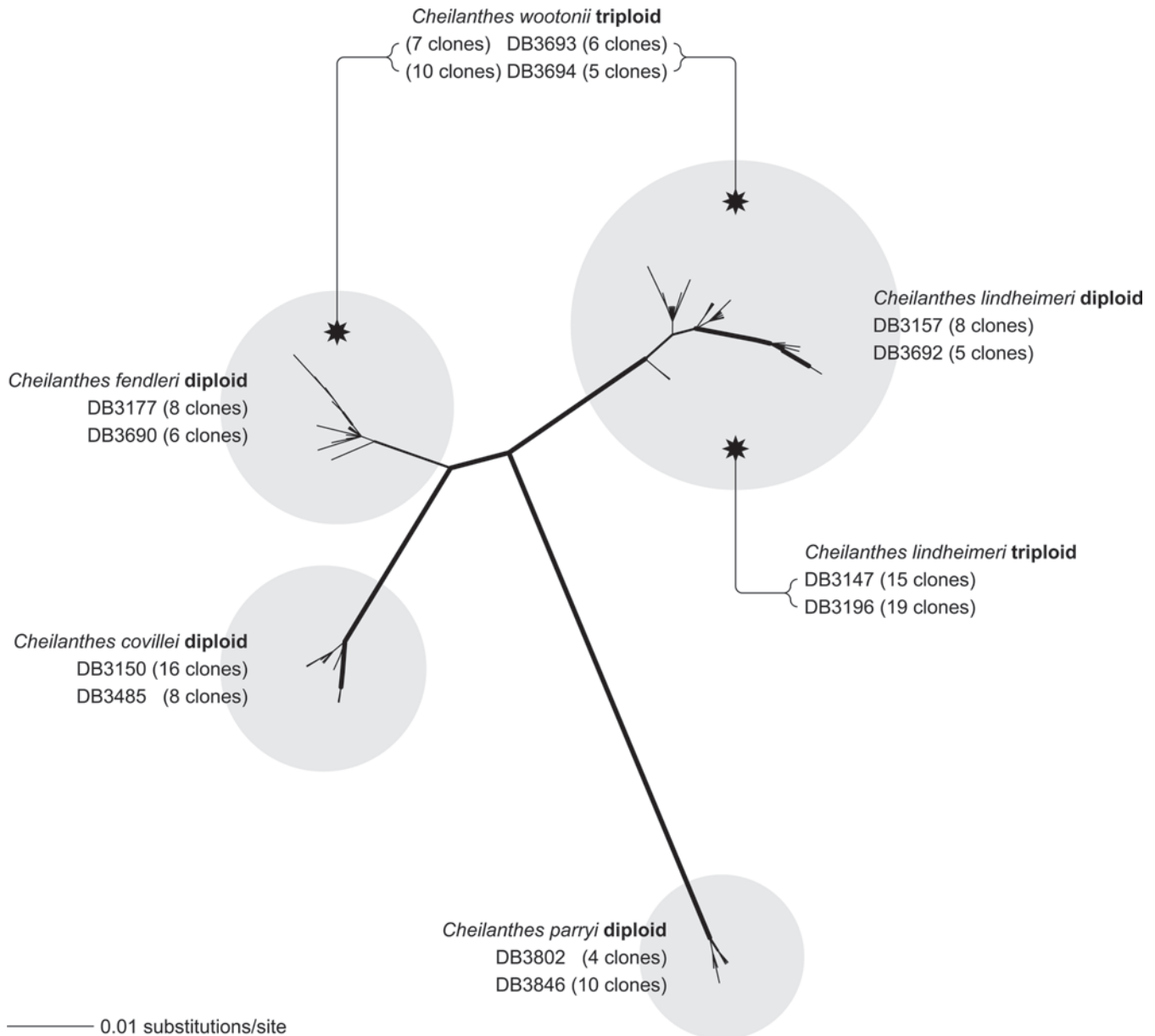


FIG. 3. Genealogy resulting from maximum likelihood analysis of the case study data set. Genealogy resulting from Bayesian analysis of the case study data set yielded an essentially congruent topology (not shown). Branch thickening follows Fig. 2. Numbers prefaced by "DB" refer to sampled individuals (see Appendix 2). The number of clones, from each individual, in each terminal clade is indicated in parentheses. The triploid *C. wootonii* is inferred to be of allopolyploid origin, as clones from this species are resolved in both the *C. fendleri* and *C. lindheimeri* clades. The triploid *C. lindheimeri* is inferred to be of autopolyploid origin, as all clones from this species are resolved within the diploid *C. lindheimeri* clade.

(Schuettpelz and Pryer 2007), one might hypothesize that the *gapCp* duplication (yielding "long" and "short" versions of the gene) occurred immediately prior to their divergence from one another (Fig. 2B). However, the *gapC/gapCp* genealogy is not entirely consistent with this scenario, because the "long" and "short" versions do not form reciprocally monophyletic sister clades (Fig. 2A). While this result is likely the product of too few data (only 296 bp were analyzed and there is little support across the backbone of the genealogy), it does cast doubt on the working hypothesis. It is entirely possible that the *gapCp* duplication occurred much earlier in the evolutionary history of ferns, followed by the loss of one copy in most of the sampled lineages. Alternatively, the observed pattern may simply reflect our inability to recover all copies from all lineages or a much more complex evolu-

tionary history (with many underlying duplications and losses).

Identifying Polyploid Origins in Cheilanthoid Ferns—Cheilanthoid ferns compose a large, mostly xeric-adapted clade within the Pteridaceae (Smith et al. 2006; Schuettpelz et al. 2007). Both polyploidy and apomixis are rampant, and evolutionary histories can be very complex (Windham and Rabe 1993; A. L. Grusz et al. unpubl. results). Despite considerable morphological, cytogenetic, and allozyme data, the origins of many polyploid species remain unclear. Such was the case for the two apomictic triploids included in our case study: *C. lindheimeri* and *C. wootonii*.

The apomictic triploid *Cheilanthes lindheimeri* has a broad range extending from southern Arizona and central Texas to Hidalgo, Mexico. Because its rather extreme morphology

(Reeves 1979) cannot be readily explained by hybridization among any other species of *Cheilanthes*, *C. lindheimeri* was generally assumed to be an autopolyploid whose diploid progenitor was undiscovered (due to its morphological similarity to the widespread triploid). Interestingly, the existence of this diploid "*C. lindheimeri*" was also implicit in the proposed origin of another apomictic triploid. Based on morphological comparisons, Reeves (1979) hypothesized that *C. wootonii* (subsequently shown to be an apomictic triploid; Windham and Yatskievych 2003) was produced via hybridization between *C. lindheimeri* and *C. fendleri*. However, there is no known mechanism by which the apomictic triploid *C. lindheimeri* could contribute to a hybrid with the same ploidy level (Gastony and Windham 1989). Thus, if *C. lindheimeri* was indeed one of the parents of *C. wootonii*, the former had to include more than one ploidy level. A recent survey of herbarium specimens (A. L. Grusz et al. unpubl. results) resulted in the discovery of three diploid populations of *C. lindheimeri*. Two individuals from these populations were included here, as hypothetical parents of both focal triploids.

All of the morphological and cytogenetic hypotheses outlined above were corroborated by our *gapCp* analysis. The sequences isolated from diploid and triploid individuals of *C. lindheimeri* were intermixed with one another and composed a single well-supported clade (Fig. 3), supporting an autopolyploid origin for the latter. Sequences obtained from accessions of *C. wootonii* revealed contributions from two phylogenetically divergent diploid species, indicating that *C. wootonii* is indeed an allopolyploid and that *C. lindheimeri* and *C. fendleri* were involved in its formation (Fig. 3). Additional work is underway to more fully characterize the origins of triploid *C. lindheimeri*, *C. wootonii*, and other polyploid species of *Cheilanthes* (A. L. Grusz et al. unpubl. results).

The Utility of the Nuclear *gapCp* Gene—It is clear from the lack of branch support in our broad survey genealogy (Fig. 2A) that the small portion of the nuclear *gapCp* gene employed herein will be of little use for the deepest phylogenetic questions (only 296 bp of exon data are available for analysis; introns are unalignable at this level). However, our findings in cheilanths (Fig. 3), combined with the apparent universality of our primers (Table 1), indicate that this marker holds considerable potential for addressing species-level problems across the fern tree of life. Whether or not *gapCp* will be useful for addressing relationships at the generic level (i.e. within families) remains to be explored.

When working with *gapCp*, as with any nuclear marker, it will be important for researchers to appreciate the possibly confounding nature of multiple copies (Small et al. 2004). We have not yet determined the precise phylogenetic position of the *gapCp* duplication (Fig. 2); however, it is obvious that many ferns will have both "long" and "short" versions (Table 1), and there is certainly the potential for other more recent duplications. In addition to amplifying multiple copies of *gapCp*, our primers may even (in certain circumstances) amplify the nuclear *gapCp* gene. This mixture of PCR products in a single reaction could be viewed as problematic, but may also be seen as beneficial. More markers (i.e. more data) can be obtained from a single amplification. Cloning is required regardless, and can effectively isolate individual gene copies. Most *gapCp* paralogs are easily separated on an agarose gel (Table 1). If a specific copy of *gapCp* is desired, it can be sequenced and other products ignored. In any case, the risk of constructing a data set that combines different *gapCp* para-

logs is alleviated by the fact that they are highly divergent, with introns that are not at all alignable. Of course, in some instances (e.g. if the desired product is not preferentially amplified) it would be advisable to develop copy-specific primers.

In our study, more than two unique but highly similar sequences corresponding to a particular paralog, and forming a clade, were occasionally recovered from a single diploid individual. Similarly, more than three sequences were sometimes recovered from a single triploid individual. As only two (or three) allelic sequences should have been found, these results suggest that either errors had been introduced by PCR amplification or that a species-specific duplication had occurred. Because relatively few clones were sequenced from any given individual, we made no attempt here to discriminate among allelic diversity, substitutional artifacts, and recent duplication events. However, future studies will need to consider these possibilities. Furthermore, those choosing to work with *gapCp* or any other nuclear marker must also be wary of the possible formation of chimeric sequences, especially when working with polyploid species (Bradley and Hillis 1997; Cronn et al. 2002).

ACKNOWLEDGMENTS. We are grateful to Donovan Bailey, Ann Kelsey, Jordan Metzgar, Nathalie Nagalingum, and Harald Schneider, as well as the New Mexico State University Herbarium, United States Department of Agriculture—Forest Service, United States Department of Defense—Fort Huachuca, United States Department of the Interior—Bureau of Land Management, and University of Utah Herbarium for their assistance in the acquisition of material newly utilized in this study. James Beck, Gordon Burleigh, Layne Huiet, and Carl Rothfels provided valuable comments and criticism. We thank David Hearn and Channa Pickett for their help in evaluating the utility of previously published primer sets. Funding for this study was provided in part by an NSF award to K.M.P. and M.D.W. (DEB-0717398), an NSF DDIG award to K.M.P. and E.S. (DEB-0408077), and an NSF CAREER award to K.M.P. (DEB-0347840).

LITERATURE CITED

- Abraham, A. and C. A. Ninan. 1954. Chromosomes of *Ophioglossum reticulatum* L. *Current Science (Bangalore)* 23: 213–214.
- Álvarez, I. and J. F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417–434.
- Bradley, R. D. and D. M. Hillis. 1997. Recombinant DNA sequences generated by PCR amplification. *Molecular Biology and Evolution* 14: 592–593.
- Breathnach, R., C. Benoist, K. O'Hare, F. Gannon, and P. Chambon. 1978. Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proceedings of the National Academy of Sciences USA* 75: 4853–4857.
- Brinkmann, H., R. Cerff, M. Salomon, and J. Soll. 1989. Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits *GapA* and *GapB* of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach. *Plant Molecular Biology* 13: 81–94.
- Brinkmann, H., P. Martinez, F. Quigley, W. Martin, and R. Cerff. 1987. Endosymbiotic origin and codon bias of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. *Journal of Molecular Evolution* 26: 320–328.
- Bult, C., M. Källersjö, and Y. Suh. 1992. Amplification and sequencing of 16/18S rDNA from gel-purified total plant DNA. *Plant Molecular Biology Reporter* 10: 273–284.
- Cronn, R., M. Cedroni, T. Haselkorn, C. Grover, and J. F. Wendel. 2002. PCR-mediated recombination in amplification products derived from polyploid cotton. *Theoretical and Applied Genetics* 104: 482–489.
- Ebihara, A., H. Ishikawa, S. Matsumoto, S.-J. Lin, K. Iwatsuki, M. Takamiya, Y. Watano, and M. Ito. 2005. Nuclear DNA, chloroplast DNA, and ploidy analysis clarified biological complexity of the *Vandenboschia radicans* complex (Hymenophyllaceae) in Japan and adjacent areas. *American Journal of Botany* 92: 1535–1547.

- Frohlich, M. W. and E. M. Meyerowitz. 1997. The search for flower homeotic gene homologs in basal angiosperms and Gnetales: a potential new source of data on the evolutionary origin of flowers. *International Journal of Plant Sciences* 158: S131–S142.
- Gastony, G. J. and D. R. Rollo. 1998. Cheilantheid ferns (Pteridaceae: Cheilantheoideae) in the southwestern United States and adjacent Mexico: a molecular phylogenetic reassessment of generic lines. *Aliso* 17: 131–144.
- Gastony, G. J. and M. D. Windham. 1989. Species concepts in pteridophytes: the treatment and definition of agamosporous species. *American Fern Journal* 79: 65–77.
- Gastony, G. J. and G. Yatskiyevych. 1992. Maternal inheritance of the chloroplast and mitochondrial genomes in cheilantheid ferns. *American Journal of Botany* 79: 716–722.
- Ge, S., T. Sang, B.-R. Lu, and D.-Y. Hong. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proceedings of the National Academy of Sciences USA* 96: 14400–14405.
- Grant, V. 1981. *Plant speciation*. New York: Columbia University Press.
- Gregory, T. R. and B. K. Mable. 2005. Polyploidy in animals. Pp. 428–501 in *The Evolution of the Genome*, ed. T. R. Gregory. San Diego: Elsevier Academic Press.
- Haufler, C. H. 1987. Electrophoresis is modifying our concepts of evolution in homosporous pteridophytes. *American Journal of Botany* 74: 953–966.
- Haufler, C. H., M. D. Windham, and E. W. Rabe. 1995. Reticulate evolution in the *Polypodium vulgare* complex. *Systematic Botany* 20: 89–109.
- Hoot, S. B. and W. C. Taylor. 2001. The utility of nuclear ITS, a LEAFY homolog intron, and chloroplast *atpB-rbcL* spacer region data in phylogenetic analyses and species delimitation in *Isoetes*. *American Fern Journal* 91: 166–177.
- Huelsenbeck, J. P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics (Oxford, England)* 17: 754–755.
- Ishikawa, H., Y. Watano, K. Kano, M. Ito, and S. Kurita. 2002. Development of primer sets for PCR amplification of the *pgiC* gene in ferns. *Journal of Plant Research* 115: 65–70.
- Jackson, R. C. 1982. Polyploidy and diploidy: new perspectives on chromosome pairing and its evolutionary implications. *American Journal of Botany* 69: 1512–1523.
- Klekowski, E. J. 1972. Genetical features of ferns as contrasted to seed plants. *Annals of the Missouri Botanical Garden* 59: 138–151.
- Klekowski, E. J. and H. G. Baker. 1966. Evolutionary significance of polyploidy in the Pteridophyta. *Science* 153: 305–307.
- Kovarik, A., R. Matyasek, K. Y. Lim, K. Skalická, B. Koukalová, S. Knapp, M. Chase, and A. R. Leitch. 2004. Concerted evolution of 18–5.8–26S rDNA repeats in *Nicotiana* allotetraploids. *Biological Journal of the Linnean Society* 82: 615–625.
- Kovarik, A., J. C. Pires, A. R. Leitch, K. Y. Lim, A. M. Sherwood, R. Matyasek, J. Rocca, D. E. Soltis, and P. S. Soltis. 2005. Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. *Genetics* 169: 931–944.
- Levin, D. A. 1983. Polyploidy and novelty in flowering plants. *American Naturalist* 122: 1–25.
- Lihová, J., K. K. Shimizu, and K. Marhold. 2006. Allopolyploid origin of *Cardamine asarifolia* (Brassicaceae): incongruence between plastid and nuclear ribosomal DNA sequences solved by a single-copy nuclear gene. *Molecular Phylogenetics and Evolution* 39: 759–786.
- Lim, K. Y., R. Matyasek, A. Kovarik, and A. Leitch. 2007. Parental origin and genome evolution in the allopolyploid *Iris versicolor*. *Annals of Botany* 100: 219–224.
- Löve, Á., D. Löve, and R. E. G. Pichi Sermolli. 1977. *Cytotaxonomical atlas of the Pteridophyta*. Vaduz: Cramer.
- Ma, X.-F. and J. P. Gustafson. 2005. Genome evolution of allopolyploids: a process of cytological and genetic diploidization. *Cytogenetic and Genome Research* 109: 236–249.
- Maddison, D. R. and W. P. Maddison. 2005. *MacClade: Analysis of Phylogeny and Character Evolution*, version 4.08. Sunderland: Sinauer Associates.
- Manton, I. and G. D. Vida. 1968. Cytology of the fern flora of Tristan da Cunha. *Proceedings of the Royal Society of London. Series B. Biological Sciences* 170: 361–379.
- Martin, W. and R. Cerff. 1986. Prokaryotic features of a nucleus-encoded enzyme. cDNA sequences for chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenases from mustard (*Sinapis alba*). *European Journal of Biochemistry* 159: 323–331.
- Martin, W. and C. Schnarrenberger. 1997. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. *Current Genetics* 32: 1–18.
- Mason-Gamer, R. J., C. F. Weil, and E. A. Kellogg. 1998. Granule-bound starch synthase: structure, function, and phylogenetic utility. *Molecular Biology and Evolution* 15: 1658–1673.
- Meyer-Gauen, G., H. Herbrand, J. Pahnke, R. Cerff, and W. Martin. 1998. Gene structure, expression in *Escherichia coli* and biochemical properties of the NAD⁺-dependent glyceraldehyde-3-phosphate dehydrogenase from *Pinus sylvestris* chloroplasts. *Gene* 209: 167–174.
- Meyer-Gauen, G., C. Schnarrenberger, R. Cerff, and W. Martin. 1994. Molecular characterization of a novel, nuclear-encoded, NAD⁺-dependent glyceraldehyde-3-phosphate dehydrogenase in plastids of the gymnosperm *Pinus sylvestris* L. *Plant Molecular Biology* 26: 1155–1166.
- Osborn, T. C., J. C. Pires, J. A. Birchler, D. L. Auger, Z. J. Chen, H.-S. Lee, L. Comai, A. Madlung, R. W. Doerge, V. Colot, and R. A. Martienssen. 2003. Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics* 19: 141–147.
- Otto, S. P. and J. Whitton. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401–437.
- Petersen, J., H. Brinkmann, and R. Cerff. 2003. Origin, evolution, and metabolic role of a novel glycolytic GAPDH enzyme recruited by land plant plastids. *Journal of Molecular Evolution* 57: 16–26.
- Petersen, J., R. Teich, B. Becker, R. Cerff, and H. Brinkmann. 2006. The *gapA/B* gene duplication marks the origin of Streptophyta (charophytes and land plants). *Molecular Biology and Evolution* 23: 1109–1118.
- Popp, M. and B. Oxelman. 2001. Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoeologous nuclear DNA sequences. *Molecular Phylogenetics and Evolution* 20: 474–481.
- Popp, M. and B. Oxelman. 2007. Origin and evolution of North American polyploid *Silene* (Caryophyllaceae). *American Journal of Botany* 94: 330–349.
- Pryer, K. M., E. Schuettpelz, P. G. Wolf, H. Schneider, A. R. Smith, and R. Cranfill. 2004. Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. *American Journal of Botany* 91: 1582–1598.
- Rambaut, A. and A. Drummond. 2007. Tracer, version 1.4. Edinburgh: University of Edinburgh (<http://tree.bio.ed.ac.uk/software/tracer/>). Downloaded October, 2007.
- Ramsey, J. and D. W. Schemske. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* 33: 589–639.
- Reeves, T. 1979. *A monograph of the fern genus Cheilanthes subgenus Physapteris (Adiantaceae)*. Ph.D dissertation, Arizona State University.
- Reid, J. D., G. M. Plunkett, and G. A. Peters. 2006. Phylogenetic relationships in the heterosporous fern genus *Azolla* (Azollaceae) based on DNA sequence data from three noncoding regions. *International Journal of Plant Sciences* 167: 529–538.
- Ronquist, F. and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics (Oxford, England)* 19: 1572–1574.
- Sang, T. and D. Zhang. 1999. Reconstructing hybrid speciation using sequences of low copy nuclear genes: hybrid origins of five *Paeonia* species based on *Adh* gene phylogenies. *Systematic Botany* 24: 148–163.
- Schuettpelz, E. and K. M. Pryer. 2007. Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. *Taxon* 56: 1037–1050.
- Schuettpelz, E., P. Korall, and K. M. Pryer. 2006. Plastid *atpA* data provide improved support for deep relationships among ferns. *Taxon* 55: 897–906.
- Schuettpelz, E., H. Schneider, L. Huiet, M. D. Windham, and K. M. Pryer. 2007. A molecular phylogeny of the fern family Pteridaceae: assessing overall relationships and the affinities of previously unsampled genera. *Molecular Phylogenetics and Evolution* 44: 1172–1185.
- Sears, B. B. 1980. Elimination of plastids during spermatogenesis and fertilization in the plant kingdom. *Plasmid* 4: 233–255.
- Small, R. L., R. C. Cronn, and J. F. Wendel. 2004. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17: 145–170.
- Small, R. L., E. B. Lickey, J. Shaw, and W. D. Hauk. 2005. Amplification of noncoding chloroplast DNA for phylogenetic studies in lycophytes and monilophytes with a comparative example of relative phyloge-

- netic utility from Ophioglossaceae. *Molecular Phylogenetics and Evolution* 36: 509–522.
- Small, R. L., J. A. Ryburn, R. C. Cronn, T. Seelanan, and J. F. Wendel. 1998. The tortoise and the hare: Choosing between noncoding plastome and nuclear ADH sequences for phylogeny reconstruction in a recently diverged plant group. *American Journal of Botany* 85: 1301–1315.
- Smith, A. R., K. M. Pryer, E. Schuettpelz, P. Korall, H. Schneider, and P. G. Wolf. 2006. A classification for extant ferns. *Taxon* 55: 705–731.
- Soltis, D. E. and L. H. Rieseberg. 1986. Autopolyploidy in *Tolmiea menziesii* (Saxifragaceae): genetic insights from enzyme electrophoresis. *American Journal of Botany* 73: 310–318.
- Soltis, D. E. and P. S. Soltis. 1987. Polyploidy and breeding systems in homosporous Pteridophyta: a reevaluation. *American Naturalist* 130: 219–232.
- Soltis, D. E., P. S. Soltis, D. W. Schemske, J. F. Hancock, J. N. Thompson, B. C. Husband, and W. S. Judd. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* 56: 13–30.
- Soltis, P. S. and D. E. Soltis. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences USA* 97: 7051–7057.
- Tate, J. A., D. E. Soltis, and P. S. Soltis. 2005. Polyploidy in plants. Pp. 371–426 in *The Evolution of the Genome*, ed. T. R. Gregory. San Diego: Elsevier Academic Press.
- Thompson, J. D. and R. Lumaret. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. *Trends in Ecology & Evolution* 9: 302–307.
- Van den heede, C. J., R. L. L. Viane, and M. W. Chase. 2003. Phylogenetic analysis of *Asplenium* subgenus *Ceterach* (Pteridophyta: Aspleniaceae) based on plastid and nuclear ribosomal ITS DNA sequences. *American Journal of Botany* 90: 481–495.
- Vida, G. 1976. The role of polyploidy in evolution. Pp. 267–294 in *Evolutionary Biology*, eds. V. J. A. Novák and B. Pacltová. Prague: Czechoslovak Academy of Science.
- Volkov, R. A., N. Y. Komarova, and V. Hemleben. 2007. Ribosomal DNA in plant hybrids: inheritance, rearrangement, expression. *Systematics and Biodiversity* 5: 261–276.
- Walker, T. G. 1966. A cytotaxonomic survey of the Pteridophytes of Jamaica. *Transactions of the Royal Society of Edinburgh* 66: 169–237.
- Walker, T. G. 1984. Chromosomes and evolution in pteridophytes. Pp. 101–141 in *Chromosomes in Evolution of Eukaryotic Groups Volume 2*, eds. A. K. Sharma and A. Sharma. Boca Raton: CRC Press.
- Wall, D. P. 2002. Use of the nuclear gene glyceraldehyde 3-phosphate dehydrogenase for phylogeny reconstruction of recently diverged lineages in *Mitthyridium* (Musci: Calymperaceae). *Molecular Phylogenetics and Evolution* 25: 10–26.
- Wendel, J. F. 2000. Genome evolution in polyploids. *Plant Molecular Biology* 42: 225–249.
- Wendel, J. F., A. Schnabel, and T. Seelanan. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proceedings of the National Academy of Sciences USA* 92: 280–284.
- Werth, C. R. and M. D. Windham. 1991. A model for divergent allopatric speciation of polyploid pteridophytes resulting from silencing of duplicate-gene expression. *American Naturalist* 137: 515–526.
- Werth, C. R., S. I. Guttman, and W. H. Eshbaugh. 1985. Electrophoretic evidence of reticulate evolution in the Appalachian *Asplenium* complex. *Systematic Botany* 10: 184–192.
- Windham, M. D. and E. W. Rabe. 1993. *Cheilanthes*. Pp. 152–169 in *Flora of North America North of Mexico, Volume 2*, ed. Flora of North America Editorial Committee. New York: Oxford University Press.
- Windham, M. D. and G. Yatskievych. 2003. Chromosome studies of cheilanthoid ferns (Pteridaceae: Cheilanthoideae) from the western United States and Mexico. *American Journal of Botany* 90: 1788–1800.
- Wolf, P. G. 1995. Phylogenetic analyses of *rbcl* and nuclear ribosomal RNA gene sequences in Dennstaedtiaceae. *American Fern Journal* 85: 306–327.
- Zwickl, D. J. 2006. *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion*. Ph.D. dissertation, The University of Texas at Austin.
- ucts obtained and clones sequenced (with length and GenBank accession information provided for each).
- Adiantum malesianum* J. Ghatak, DB2506, *Huieit* 111 (UC), *gapC*, clone 2, 730 bp, EU551457, clone 3, 757 bp, EU551458, clone 6, 730 bp, EU551459, clone 11, 730 bp, EU551460, clone 12, 730 bp, EU551461, clone 15, 757 bp, EU551462, clone 16, 730 bp, EU551463, *gapCp* “long”, clone 4, 654 bp, EU551254, clone 10, 654 bp, EU551259, clone 14, 654 bp, EU551261, clone 17, 654 bp, EU551262, *gapCp* “short”, clone 1, 535 bp, EU551253, clone 5, 536 bp, EU551255, clone 7, 535 bp, EU551256, clone 8, 536 bp, EU551257, clone 9, 536 bp, EU551258, clone 13, 535 bp, EU551260; *Cheilanthes covillei* Maxon, DB3150, *Schuettpelz* 443 (DUKE), *gapC*, clone 2, 929 bp, EU551464, clone 9, 929 bp, EU551465, clone 10, 929 bp, EU551466, clone 11, 929 bp, EU551467, clone 21, 930 bp, EU551468, clone 24, 930 bp, EU551469, clone 25, 930 bp, EU551470, clone 26, 929 bp, EU551471, *gapCp* “long”, clone 7, 967 bp, EU551268, clone 27, 917 bp, EU551280, *gapCp* “short”, clone 1, 600 bp, EU551263, clone 3, 600 bp, EU551264, clone 4, 596 bp, EU551265, clone 5, 600 bp, EU551266, clone 6, 600 bp, EU551267, clone 12, 596 bp, EU551269, clone 13, 597 bp, EU551270, clone 14, 600 bp, EU551271, clone 15, 596 bp, EU551272, clone 16, 600 bp, EU551273, clone 17, 600 bp, EU551274, clone 18, 596 bp, EU551275, clone 19, 600 bp, EU551276, clone 20, 600 bp, EU551277, clone 22, 600 bp, EU551278, clone 23, 596 bp, EU551279; *Dennstaedtia dissecta* (Sw.) T. Moore, DB2465, *Schuettpelz* 9 (DUKE), *gapCp*, clone 1, 805 bp, EU551392, clone 2, 808 bp, EU551393, clone 4, 807 bp, EU551394, clone 5, 807 bp, EU551395, clone 6, 807 bp, EU551396, clone 7, 807 bp, EU551397, clone 8, 803 bp, EU551398, clone 9, 805 bp, EU551399, clone 10, 808 bp, EU551400, clone 11, 822 bp, EU551401; *Diplopterygium bancroftii* (Hook.) A. R. Sm., DB172, *Smith* 2569 (UC), *gapCp*, clone 2, 1057 bp, EU551402, clone 4, 1058 bp, EU551403, clone 5, 1058 bp, EU551404, clone 6, 1058 bp, EU551405, clone 7, 1058 bp, EU551406, clone 8, 1058 bp, EU551407, clone 9, 1058 bp, EU551408; *Dryopteris marginalis* (L.) A. Gray, DB2979, *Schuettpelz* 334 (DUKE), *gapCp* “long”, clone 1, 1650 bp, EU551409, clone 2, 1731 bp, EU551410, clone 11, 1650 bp, EU551419, *gapCp* “short”, clone 3, 647 bp, EU551411, clone 4, 659 bp, EU551412, clone 5, 659 bp, EU551413, clone 6, 659 bp, EU551414, clone 7, 659 bp, EU551415, clone 8, 659 bp, EU551416, clone 9, 659 bp, EU551417, clone 10, 659 bp, EU551418; *Hymenasplenium unilaterale* (Lam.) Hayata, DB3470, *Ranker* 2072 (COLO), *gapCp*, clone 1, 569 bp, EU551420, clone 4, 569 bp, EU551421, clone 5, 569 bp, EU551422, clone 6, 569 bp, EU551423, clone 7, 569 bp, EU551424; *Hymenophyllum nephrophyllum* (G. Forst.) Ebihaara & K. Iwats., DB935, *Smith* 2606 (UC), *gapCp*, clone 1, 1038 bp, EU551425, clone 2, 1038 bp, EU551426, clone 3, 1038 bp, EU551427, clone 4, 1038 bp, EU551428, clone 5, 1038 bp, EU551429; *Lindsaea quadrangularis* Raddi, DB3304, *Christenhusz* 4018 (TUR), *gapCp*, clone 1, 868 bp, EU551430, clone 5, 868 bp, EU551431, clone 6, 868 bp, EU551432, clone 7, 868 bp, EU551433, clone 8, 868 bp, EU551434; *Lygodium japonicum* (Thunb.) Sw., DB2545, *Metzgar* 92 (DUKE), *gapCp*, clone 2, 812 bp, EU551435, clone 3, 814 bp, EU551436, clone 4, 814 bp, EU551437, clone 5, 814 bp, EU551438, clone 6, 814 bp, EU551439, clone 9, 812 bp, EU551440, clone 10, 814 bp, EU551441, clone 11, 814 bp, EU551442; *Pteris vittata* L., DB3400, *Christenhusz* 4195 (TUR), *gapCp* “long”, clone 6, 885 bp, EU551448, clone 7, 885 bp, EU551449, clone 8, 885 bp, EU551450, clone 9, 886 bp, EU551451, clone 10, 886 bp, EU551452, *gapCp* “short”, clone 1, 546 bp, EU551443, clone 2, 546 bp, EU551444, clone 3, 546 bp, EU551445, clone 4, 553 bp, EU551446, clone 5, 553 bp, EU551447; *Sticherus bifidus* (Willd.) Ching, DB176, *Smith* 2565 (UC), *gapCp*, clone 1, 1109 bp, EU551453, clone 2, 1109 bp, EU551454, clone 3, 1109 bp, EU551455, clone 4, 1109 bp, EU551456.

APPENDIX 2. Taxonomic sampling, voucher information, and GenBank accession numbers for our case study data set. For each species, sampled individuals are identified by their fern DNA database (www.pryerlab.net) record numbers (prefixed by “DB”). For each sampled individual, voucher information is provided, along with a list of *gapCp* “short” version clones sequenced (with length and GenBank accession information provided for each).

Cheilanthes covillei Maxon, **diploid**, DB3150, *Schuettpelz* 443 (DUKE), clone 1, 600 bp, EU551263, clone 3, 600 bp, EU551264, clone 4, 596 bp, EU551265, clone 5, 600 bp, EU551266, clone 6, 600 bp, EU551267, clone 12, 596 bp, EU551269, clone 13, 597 bp, EU551270, clone 14, 600 bp, EU551271, clone 15, 596 bp, EU551272, clone 16, 600 bp, EU551273, clone 17, 600 bp, EU551274, clone 18, 596 bp, EU551275, clone 19, 600 bp, EU551276, clone 20, 600 bp, EU551277, clone 22, 600 bp, EU551278, clone 23, 596 bp, EU551279, DB3485, *Windham* 2945 (UT), clone 1, 600 bp, EU551281, clone 2, 600 bp, EU551282, clone 3, 600 bp, EU551283, clone 4, 600 bp, EU551284, clone 5, 600 bp, EU551285, clone 6, 600 bp, EU551286,

APPENDIX 1. Taxonomic sampling, voucher information, and GenBank accession numbers for our broad survey data set. For each sampled species, voucher information is provided, along with a fern DNA database (www.pryerlab.net) record number (prefixed by “DB”) and a list of prod-

clone 7, 600 bp, EU551287, clone 8, 600 bp, EU551288; *Cheilanthes fendleri* Hook., **diploid**, DB3177, *Schuettpelz 470* (DUKE), clone 7, 602 bp, EU551289, clone 8, 602 bp, EU551290, clone 13, 602 bp, EU551291, clone 14, 602 bp, EU551292, clone 15, 602 bp, EU551293, clone 16, 602 bp, EU551294, clone 17, 602 bp, EU551295, clone 18, 602 bp, EU551296, DB3690, *Windham 3408* (UT), clone 1, 602 bp, EU551297, clone 2, 602 bp, EU551298, clone 3, 602 bp, EU551299, clone 4, 602 bp, EU551300, clone 5, 602 bp, EU551301, clone 6, 602 bp, EU551302; *Cheilanthes lindheimeri* Hook., **diploid**, DB3157, *Schuettpelz 450* (DUKE), clone 1, 601 bp, EU551318, clone 2, 601 bp, EU551319, clone 3, 601 bp, EU551320, clone 4, 601 bp, EU551321, clone 5, 601 bp, EU551322, clone 6, 601 bp, EU551323, clone 7, 601 bp, EU551324, clone 8, 601 bp, EU551325, DB3692, *Spellenberg 5065* (NMC), clone 4, 600 bp, EU551345, clone 5, 601 bp, EU551346, clone 6, 601 bp, EU551347, clone 7, 600 bp, EU551348, clone 8, 601 bp, EU551349, **triploid**, DB3147, *Schuettpelz 440* (DUKE), clone 2, 601 bp, EU551303, clone 3, 601 bp, EU551304, clone 5, 601 bp, EU551305, clone 6, 601 bp, EU551306, clone 7, 601 bp, EU551307, clone 8, 601 bp, EU551308, clone 9, 601 bp, EU551309, clone 10, 601 bp, EU551310, clone 11, 601 bp, EU551311, clone 12, 601 bp, EU551312, clone 13, 601 bp, EU551313, clone 14, 600 bp, EU551314, clone 15, 601 bp, EU551315, clone 16, 601 bp, EU551316, clone 17, 601 bp, EU551317, DB3196, *Schuettpelz 489* (DUKE), clone 1, 601 bp, EU551326, clone 2, 600 bp, EU551327, clone 3, 601 bp, EU551328, clone 4, 601 bp, EU551329, clone 5, 600 bp, EU551330, clone 6, 601 bp, EU551331, clone 7, 601 bp, EU551332, clone 9, 601 bp, EU551333,

clone 10, 601 bp, EU551334, clone 11, 601 bp, EU551335, clone 12, 601 bp, EU551336, clone 13, 601 bp, EU551337, clone 14, 601 bp, EU551338, clone 15, 601 bp, EU551339, clone 16, 601 bp, EU551340, clone 17, 601 bp, EU551341, clone 19, 601 bp, EU551342, clone 20, 601 bp, EU551343, clone 21, 601 bp, EU551344; *Cheilanthes parryi* (D. C. Eaton) Domin, **diploid**, DB3802, *Metzgar 149* (DUKE), clone 1, 596 bp, EU551350, clone 2, 596 bp, EU551351, clone 3, 596 bp, EU551352, clone 4, 595 bp, EU551353, DB3846, *Windham 3440* (UT), clone 1, 595 bp, EU551354, clone 2, 595 bp, EU551355, clone 3, 595 bp, EU551356, clone 4, 595 bp, EU551357, clone 5, 595 bp, EU551358, clone 6, 595 bp, EU551359, clone 7, 594 bp, EU551360, clone 8, 595 bp, EU551361, clone 9, 595 bp, EU551362, clone 10, 595 bp, EU551363; *Cheilanthes wootonii* Maxon, **triploid**, DB3693, *Windham 3409* (DUKE), clone 9, 600 bp, EU551379, clone 10, 602 bp, EU551380, clone 14, 600 bp, EU551381, clone 15, 600 bp, EU551382, clone 16, 602 bp, EU551383, clone 17, 599 bp, EU551384, clone 18, 602 bp, EU551385, clone 19, 602 bp, EU551386, clone 20, 600 bp, EU551387, clone 21, 600 bp, EU551388, clone 22, 602 bp, EU551389, clone 23, 602 bp, EU551390, clone 24, 602 bp, EU551391, DB3694, *Spellenberg 10407* (NMC), clone 1, 602 bp, EU551364, clone 3, 600 bp, EU551365, clone 5, 602 bp, EU551366, clone 6, 600 bp, EU551367, clone 7, 602 bp, EU551368, clone 8, 602 bp, EU551369, clone 9, 602 bp, EU551370, clone 10, 601 bp, EU551371, clone 11, 600 bp, EU551372, clone 12, 602 bp, EU551373, clone 13, 600 bp, EU551374, clone 14, 602 bp, EU551375, clone 16, 602 bp, EU551376, clone 17, 602 bp, EU551377, clone 18, 602 bp, EU551378.