American Fern Journal 97(2):47–65 (2007)

Characterization of Four Members of the Alpha-Tubulin Gene Family in *Ceratopteris richardii*

RODNEY J. SCOTT
Department of Biology, Wheaton College, Wheaton, IL 60187
GERALD J. GASTONY
Department of Biology, Indiana University, Bloomington, IN 47405-7005
JEREMY W. WEATHERFORD
Department of Computer Science and Engineering, Washington University, St. Louis, MO 63108
TAKUYA NAKAZATO
Department of Biology, Indiana University, Bloomington, IN 47405-7005

ABSTRACT.—Four members of the alpha-tubulin gene family were examined in *Ceratopteris richardii*. Genetic linkage mapping based on a population of nearly 500 Doubled Haploid Lines was able to position three or four members of this gene family on linkage groups 17, 24, and 28, respectively (two of the four observed polymorphic restriction fragments containing alpha-tubulin genes are either identical or map too close to each other on linkage group 17 to be distinguishable in map distance). Non-mapable monomorphic bands observed on probed Southern blots suggest that the alpha-tubulin gene family in this species is large. Four alpha-tubulin genes from *C. richardii* were sequenced and found to be fairly similar to each other in terms of their amino acid sequences, with their greatest diversity at the carboxy-terminal ends. BLAST comparisons found each of these four amino acid sequences more similar to an alpha-tubulin from a dicot, gymnosperm, or alga species than it was to any other alpha-tubulin sequence presently known from *Ceratopteris* or from the fern *Anemia phyllitidis* or the moss *Physcomitrella patens*. Bayesian phylogenetic analysis of nucleotide sequences placed three of the four *Ceratopteris* alpha-tubulin gene copies in a clade with copies from *Pseudotsuga* and *Anemia*, consistent with a history of two gene duplication events, one following and one preceding the divergence of ferns and seed plants. The fourth copy is robustly separated from the preceding three and placed in a clade of algal alpha-tubulin genes, suggesting its divergence from the ancestor of the other three before the divergence of algae and land plants. As characterized thus far, the alpha-tubulin gene family of *C. richardii* is relatively large as compared to the six copies known from fully sequenced *Arabidopsis thaliana*, a condition that may be correlated with the large genome size and diverse life history constraints of this homosporous fern species. These findings suggest several new opportunities for research into the evolution, function, and regulation of the alpha-tubulin gene family in *Ceratopteris*.

This report describes the application of DNA sequencing and genetic linkage mapping to the alpha-tubulin genes of *Ceratopteris richardii* and shows how such studies can further enhance the utility of this model system. Initial successes with *Ceratopteris* were in the areas of cytogenetics (e.g., Hickok, 1976, 1977a, 1977b, 1978, 1979a, 1979b; Hickok and Klekowski, 1973, 1974), physiology and development, and Mendelian genetics (reviewed in Hickok, 1987; Hickok *et al.*, 1987, 1995). More recent studies have focused on the molecular genetics of *Ceratopteris* (Munster *et al.*, 1997; Hasebe *et al.*, 1998; Aso *et al.*, 1999; Stout *et al.*, 2003; Rutherford *et al.*, 2004; Salmi *et al.*, 2005). These recent studies that employ robust molecular methods are especially encouraging, since the lack of a technique to induce stable genetic trans-

formation in *Ceratopteris* has undoubtedly impeded its function as a unifying model system. In this report, DNA sequences from four unique alpha-tubulin genes of *C. richardii* are described, and the phylogenetic relationships of these genes to other plant alpha-tubulin genes are inferred. Molecular linkage data for three polymorphic alpha-tubulin loci of *C. richardii* are used to position these loci on the new genetic linkage map of this species, and several potential strategies for using this new information in molecular studies are outlined. These perspectives should provide further encouragement for those seeking to utilize *Ceratopteris* as a model system.

The alpha-tubulin gene family was selected for these studies because of the biological significance and tractability of tubulin in the gametophyte generation of ferns and because of the new insights that may be gained for genomics by studying them. Inferring evolutionary relationships among members of gene families is generally problematic because distinguishing paralogues and orthologues is difficult without appropriate known outgroups to polarize the gene phylogeny. Tubulin genes are ideal for studying gene family evolution because 1) alpha-, beta-, and gamma-tubulin genes are known to have diverged well before the divergence of plant lineages, and 2) the highly conserved nature of the genes allows their sequences to be aligned. These features allow inference of gene phylogenies within a given tubulin gene group, using sequences from other tubulin groups as outgroups.

Tubulins and the microtubules they form are obviously essential components of all eukaryotic cells. In fern gametophytes, however, their roles are directly observable in various stages of development that can be easily studied. For example, one of the first events that signals preparation for spore germination is the migration of the nucleus within the cytoplasm (Banks, 1999). This event, which is critical to the continued development of the gametophyte, is inhibited in *Onoclea sensibilis* by several microtubule inhibitors, including colchicine (Vogelmann *et al.*, 1981). An important role for microtubules at a later stage of development has been suggested by the studies of Murata *et al.* (1997) on blue light-induced inhibition of cell growth in dark-grown *Ceratopteris* gametophytes. These investigators found that cortical microtubules reorient in response to blue light at the same time that inhibition of cell elongation occurs. Microtubules also serve a key role in the organization and function of fern sperm (Raghavan, 1989), and *Ceratopteris* sperm has been used extensively in studies characterizing these roles (Hoffman and Vaughn, 1995a, 1995b, 1996; Hoffman *et al.*, 1994; Renzaglia *et al.*, 2004).

Because homosporous ferns compose most of the sister group to seed plants (including flowering plants; Pryer *et al.*, 2001), increased knowledge of genome structure and organization in the genomes of homosporous ferns will significantly broaden our knowledge of genome structure and evolution in vascular plants. Furthermore, characterization of the alpha-tubulin gene family in *Ceratopteris* will ultimately help to answer questions specifically related to the origin and function of gene families in organisms with large genomes. At 11,294 Mb (Jo Ann Banks, personal communication), the genome

of *C. richardii* is ca. 110 times the genome size of *Arabidopsis thaliana*, 24 times that of rice, 12 times that of tomato, 4 times that of maize, and 2 times that of barley. Some relevant questions include the following. Do organisms with larger genomes have larger gene families? Do such organisms (with their apparent excess amounts of DNA) have more pseudogenes among the members of their gene families? How are the various members of gene families regulated during development in organisms with large genomes? The latter question is an especially interesting one as it relates to the ferns, whose alternation of generations features fully independent sporophyte and gametophyte generations.

This study employed two distinct strategies for characterizing alpha-tubulin genes from *C. richardii*. The first approach involved identifying, isolating, and sequencing alpha-tubulin genes that were expressed in the gametophyte generation. This was accomplished by using an antibody-based screening method with a cDNA library derived from the gametophyte generation. Bayesian Inference analyses determined the phylogenetic relationships of these newly obtained sequences to other plant alpha-tubulin genes. The second approach utilized a mapping strategy developed for the recently completed project of Nakazato *et al.* (2006) to generate a high-resolution linkage map for *C. richardii*. This mapping project used genetic polymorphisms present in a population of nearly 500 Doubled Haploid Lines (DHLs) generated from an initial cross between diploid inbred lines of highly diverged geographic races of *C. richardii* (Hickok *et al.*, 1995): ΦN8 (derived from a Nicaraguan collection) and Hα-PQ45, a mutant of Hn-n (derived from a Cuban collection). Together these two approaches provide detailed insights regarding a previously uncharacterized gene family in *C. richardii*. These new insights suggest novel strategies for studying the alpha-tubulin gene family at the level of development and gene regulation and also at the level of the genome.

Materials and Methods

cDNA library screening.—The cDNA library used in this study was a gift from Jo Ann Banks (Purdue University). It was made from 12-day-old cultures of *C. richardii* gametophytes of the Hn-n strain containing both males and hermaphrodites. The cDNA was cloned into the *Eco*RI site of the lambda ZipLox bacteriophage vector (Life Technologies). The aliquot used to screen for alpha-tubulin genes was from a sample that had been amplified from the original library. The library was screened using standard methods for detecting specific proteins via antibody labeling (Young and Davis, 1991). In brief, the bacteriophage vector was first grown on a lawn of bacteria and, after plaques were produced, gene expression was induced with IPTG. Proteins from the library were adhered to nylon membranes and detected by hybridization to an anti alpha-tubulin monoclonal antibody (Sigma, catalogue number T5168). The presence of this antibody was detected with a labeled secondary antibody. After initial detection and isolation of plaques that tested positive for expression of alpha-tubulin, several rounds of re-screening were conducted

to eliminate contaminating vectors that were not positive for alpha-tubulin. Bacteriophage vectors containing alpha-tubulin cDNAs were converted to plasmids via an *in vivo* excision process as described in technical reference materials supplied by Life Technologies.

DNA sequencing.—Plasmids with the cDNA inserts were used as the sequencing templates. For each plasmid, at least two sequencing reactions were conducted using the two primer sites that flank the *Eco*RI site where the cDNA is inserted. When the results from two sequencing reactions indicated that the insert was longer than 1220 base pairs, two additional pairs of primers were constructed based on the new sequence data. These new internal sequencing primers provided more reliable results for the central regions of these longer inserts. The sequence reads from either two or four reactions were assembled using the software program ContigExpress (NTI Vector). Ambiguities that resulted from discrepancies between overlapping sequences were resolved by relying on the sequence(s) with the clearest chromatogram.

BLAST comparisons.—The sequences were compared using the standard BLAST program for proteins (protein-protein BLAST) provided on the NCBI web site to compare individual sequences with other sequences contained in the GenBank database as described in the Results section. The Composition-based statistic option and the filter options were not enabled so that the results of the BLAST comparisons would be based solely on similarities and/or differences of individual amino acids. Since the GenBank database is constantly being updated, it should be noted that the reported BLAST comparisons were last confirmed on July 29, 2005.

Phylogenetic analysis of alpha-tubulin gene relationships.—A Bayesian Inference analysis of nucleotide sequences was used to infer the phylogenetic relationships of the four *C. richardii* alpha-tubulin genes with regard to other plant alpha-tubulin gene nucleotide sequences available in GenBank for Viridiplantae. Sequences used were from a broad selection of plants, including those corresponding to the results of the amino acid BLAST search above, except that nucleotide sequences precisely corresponding to the amino acid sequences of *Gossypium hirsutum* Q6VAG0 and *Chlamydomonas reinhardtii* P09204 could not be located in GenBank. Prior to analysis, nucleotide sequences were aligned visually. The first three nucleotides, which were invariable, and the last 45 nucleotides, which caused ambiguous alignments, were removed. MrModeltest 2.2, a modified version of Modeltest 3.6 (Posada and Crandall, 1998) determined that the suitable nucleotide substitution model for the analyzed sequences is GTR+I+G. Bayesian analysis was conducted using MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) with three million generations and a sample frequency of 1000. The first 300 ''burn-in'' trees were discarded after the analysis. Five independent runs using the same setting converged to identical trees, except that one node was resolved in only one run. Two beta-tubulin genes were used as an outgroup to root the resulting tree. The species used and their GenBank accession numbers are identified in the Results section, as are the Bayesian posterior probability confidence values for the tree's clades.

Linkage mapping of the alpha-tubulin genes.—A genetic linkage map of the *Ceratopteris richardii* genome was developed independently of this alpha-tubulin study (Nakazato *et al.*, 2006). The mapping population of ~500 Doubled Haploid Lines (DHLs) was generated by intragametophytic selfing of gametophytes derived from spores of an initial cross between diploid inbred lines of highly diverged geographic races of *C. richardii* (Hickok *et al.*, 1995): *Φ*N8 (derived from a Nicaraguan collection, *Nichols 1719*, GH) and Hα-PQ45, a mutant of Hn-n (derived from a Cuban collection, *Killip 44595*, GH). The map is based on analysis of Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), and allozyme markers. RFLPs used in the general mapping project were generated from genomic DNA of parental and DHL sporophytes digested with *Eco*RI and *Hind*III, separated on 0.8% agarose gels in 1X TAE, and Southern blotted to nylon membranes. Further details of the methods used in development of mapping materials and map construction are in Nakazato and Gastony (2006) and Nakazato *et al.* (2006). Alpha-tubulin gene copies were located on the linkage map in the following way. RFLPs containing the alpha-tubulin genes were detected by probing the Southern blots with an alpha-tubulin probe made cheiluminescent by digoxigenin (DIG)-labeling according to a protocol optimized for the *C. richardii* genome mapping project. The alpha-tubulin probe used is *C. richardii* cDNA clone Cri_10_E18_SP6 (GenBank sequence number BQ086953), from the cDNA library of *C. richardii* gametophytic tissue provided by Jo Ann Banks (Purdue University). This clone sequence corresponds to GenBank sequence AY231146 of TuaCR1 in Table 1, according to an NCBI BLAST search using blastn, which found sequence identity at 651/652 (99.85%) of the bases compared. Probing of the mapping project's Southern blots with the alpha-tubulin probe was carried out toward the end of the mapping project when the Southern blots were beginning to wear out. Thus the quality of the autorads presented here are suboptimal but nevertheless scorable. Parental alpha-tubulin RFLPs segregating in the DHL mapping population were scored and placed on the linkage maps by MAPMAKER/EXP 3.0 (Lander *et al.*, 1987) at a UNIX workstation at Indiana University, Department of Biology, at settings used for the general mapping project.

Results

Sequencing results.—A cursory comparison of the amino acid sequences for the four *Ceratopteris* alpha-tubulin genes (Fig. 1) suggests immediately that they are relatively similar to one another. Although only one of the four sequences is complete, it is possible to note at least two trends directly from the comparison presented (Fig. 1). First, for the portion where sequences are available for all four genes (i.e., residues 195 to 332 of TuaCR1), TuaCR3 and TuaCR4 are the most distinct. In this region, TuaCR3 contains five unique amino acid residues and TuaCR4 contains four, while the other two sequences each possess only a single unique amino acid. The second observation that can be made directly from these data relates only to the three sequences that

TABLE 1. A comparison of the ten *Ceratopteris* alpha-tubulin sequences identified in this study to other alpha-tubulin sequences recorded in GenBank.

Names for new gene	GenBank Accession number	Number of bases sequenced in DNA[1]	Size of Predicted ORF	Accession number, description, and identity for amino acid sequence(s)[2] most similar to new sequences[3]
TuaCR1	AY231146	1,676	450	1) Q6VAG0, Tubulin alpha-2 chain
	AY862561	1,159	306	*Gossypium hirsutum*, 436/450 (96.2%)
	AY862563	1,564	450	
TuaCR2	AY862565	1,015	266	1) AAK81858.1, alpha tubulin subunit
	AY862566	1,016	267	*Rosa* Hybrid cultivar, 260/267 (97.4%)
	AY862567	1,018	267	2) AAV92379.1, alpha tubulin 1
	AY862569	1,016	267	*Pseudotsuga menziesii*, 260/267 (97.4%)
TuaCR3	AY862562	938	257	1) BAA03955, alpha-tubulin 1
	AY862564	1044	257	*Chlorella vulgaris* 245/257 (95.3%)
				2) P09204, Tubulin alpha-1 chain
				Chlamydomonas reinhardtii, 245/257 (95.3%)
TuaCR4	AY862568	975	325	1) AAV92379.1, alpha tubulin 1
				Pseudotsuga menziesii, 316/325 (97.2%)

[1]This includes reliable DNA sequences before and after the open reading frame if present, but not the predicted polyA tail.

[2]If two or more versions of the same gene occurred in GenBank, only the most recent one is listed here. If several identical sequences were submitted by the same author(s) and are listed with the same submission date, only the one with the last accession number in the numerical sequence is listed here. The descriptions are given in the form in which they were submitted to GenBank.

[3]Sequence similarities were determined using the standard BLAST program for protein sequences to search the GenBank database. The Composition-based statistic option and the filter options were not enabled. The BLAST searches were performed on July 29, 2005.

include complete carboxy-terminal ends (i.e., TuaCR1, 2, and 3). In this region, these three sequences are highly diverse, with TuaCR3 once again exhibiting the greatest number of unique residues.

An additional sequence identified as an alpha-tubulin sequence from *Ceratopteris* has been deposited in GenBank by a separate research group (Salmi *et al.*, 2005). This 825 base-long sequence (GenBank accession BE642799) is a single pass sequence from a collection of expressed sequence tags. The bl2seq tool on the NCBI web site was used to compare this sequence with the four alpha-tubulin sequences described here (data not shown). A large region of this sequence (ranging from 574 to 640 bases) is 77 or 78% identical to the beginning portions of TuaCR1 and TuaCR4 respectively. It has no significant similarity to TuaCR3. The highest identity (88% for a region that is 195 bases long) occurs between this sequence and TuaCR2. This region of similarity occurs at the beginning of TuaCR2 and at the end of the sequence described by Salmi *et al.* (2005). Because this sequence was derived from a single sequencing experiment, it may be expected to contain a relatively large number of incorrect bases. Furthermore, the authors did not provide a deduced

```
TuaCR1   MRECISIHIGQAGIQVGNACWELYCLEHGIQPDGQMPSDKTVGGGDDAFNT
TuaCR2   --------------------------------------------------
TuaCR3   --------------------------------------------------
TuaCR4   -------**********************H******************

TuaCR1   FFSETGAGKHVPRAIFVDLEPTVIDEVRTGTYRQLFHPEQLISGKEDAANN
TuaCR2   --------------------------------------------------
TuaCR3   --------------------------------------------------
TuaCR4   **************V***************F**********N********

TuaCR1   FARGHYTIGKEIVDLCLDRIRKLADNCTGLQGFLVFNAVGGGTGSGLGSLL
TuaCR2   --------------------------------------------------
TuaCR3   --------------------------------------------------
TuaCR4   ************************************H**********A***

TuaCR1   LERLSVDYGKKSKLGFTVYPSPQVSTSVVEPYNSVLSTHSLLEHTDVAVLL
TuaCR2   --------------------------------**********S***
TuaCR3   ------------------------------------------*******M*
TuaCR4   **************************************************

TuaCR1   DNEAIYDICRRSLDIDRPTYTNLNRLVSQVISSLTASLRFDGALNVDVTEF
TuaCR2   ***************E******************************
TuaCR3   ***************E*****************************I***
TuaCR4   ***************E****N*****I**********************

TuaCR1   QTNLVPYPRIHFMLSSYAPVISAEKAYHEQLSVAEITNSAFEPSSMMAKCD
TuaCR2   **************************************************
TuaCR3   **********************************************A******
TuaCR4   **********************A*************S*************

TuaCR1   PRHGKYMACCLMYRGDVVPKDVNAAVATIKTKRTIQFVDWCPTGFKCGINY
TuaCR2   **************************************************
TuaCR3   **********M*************S************************
TuaCR4   ************************----------------------

TuaCR1   QPPTVVPGGDLAKVQRAICMISNSTSVAEVFSRIDFKFDLMYCKRAFVHWY
TuaCR2   *****************V*****************Y******A********
TuaCR3   *****************V********I******L*H******A********
TuaCR4   --------------------------------------------------

TuaCR1   VGEGMEEGEFSEAREDLAALEKDYEEVGAEGQDDDEPGD DEY
TuaCR2   ****************************DEG**GE*G***
TuaCR3   ***********************F*****DSTEG*GEDEGE**
TuaCR4   -----------------------------------------
```

FIG. 1. Alignment of the predicted amino acid sequences of four alpha-tubulin proteins from *Ceratopteris richardii*; the standard symbols for amino acids are used. Only the sequence for TuaCR1 is complete, and all other proteins are compared to it. Dashes represent unknown amino acids, asterisks represent known amino acids that are identical to those shown for TuaCR1, bold letters represent known amino acids that differ between TuaCR1 and at least one other TuaCR copy, and a blank space represents an apparent gap in the alignment.

protein sequence for this gene. For these reasons, and because it may represent an otherwise uncharacterized portion of one of the genes described here, it was not included in any of the comparisons described below.

A BLAST comparison of the four protein sequences for the *Ceratopteris* genes with other alpha-tubulin sequences recorded in GenBank reveals a somewhat diverse pattern (see the last column of Table 1). First, each of the four sequences was more similar to an alpha-tubulin from a different plant species than it was to any other *Ceratopteris* alpha-tubulin sequence. Second, the sequences from other plants that are most similar to the *Ceratopteris* sequences are quite diverse. The first two *Ceratopteris* amino acid sequences (those of TuaCR1 and TuaCR2) are most similar to sequences derived from dicotyledonous plants and/or a gymnosperm, the third is most similar to two different algal sequences, and the fourth is most similar to a sequence from a gymnosperm. These observations are perhaps more noteworthy in light of the fact that two distinct sequences for alpha-tubulin genes from the fern *Anemia phyllitidis* (one is a partial sequence) and two from the moss *Physcomitrella patens* are deposited in GenBank. When compared to the *Ceratopteris* genes (data not shown), some of the four lower plant genes have nearly as many identical amino acid sequences as do the sequences from the higher plants and algae shown in Table 1. However, when these high levels of identity exist, the lower plant sequences also have at least one additional amino acid or at least one less amino acid than the *Ceratopteris* sequences, resulting in gaps in the sequence homologies. The homologies indicated in Table 1 do not have such gaps and therefore the sequences they refer to are considered to be more similar to the *Ceratopteris* sequences.

For comparative purposes, BLAST searches were conducted for each of the known alpha-tubulin genes of *Arabidopsis* to determine what sequences from *Arabidopsis* or other organisms would show the highest similarities to each of these genes (data not shown). The entire genome of *Arabidopsis* has been sequenced, and six separate alpha-tubulin genes (TUA1–TUA6) have been located in its genome. The patterns of similarity for these genes are dramatically different from those of *Ceratopteris*. In general, the six *Arabidopsis* alpha-tubulin genes show much more similarity to one another than do the four *Ceratopteris* genes. Among these six genes are two pairs whose protein sequences are identical. The proteins of TUA2 and TUA4 are identical, and those of TUA3 and TUA5 are identical. Furthermore, TUA6 is almost identical to TUA2/4, sharing 448 out of 450 amino acids (99.6% identity). The amino acid sequences that are the next most similar to these five *Arabidopsis* genes (after comparing them to other *Arabidopsis* genes) are all sequences from angiosperms. The sequence of TUA2/4 is 98.7% identical to a sequence from *Brassica napus*, the sequence of TUA6 is 99.1% identical to the same sequence from *B. napus*, and the sequence of TUA3/5 is 97.1% identical to a sequence from *Oryza sativa*. Only the sequence of the remaining *Arabidopsis* gene product, that of TUA1, is more similar to the sequence of a different plant than it is to another *Arabidopsis* gene. The TUA1 protein shares 414 amino acids out of 450 (92% identity) with the grass *Miscanthus*.

Carboxy termini of the four protein sequences of alpha-tubulins from *Arabidopsis thaliana*:

TUA1	AREDLAALEKDYEEVG**G**EG**AE**D**DDE**EG**DE**Y
TUA2/4	AREDLAALEKDYEEVG**A**EG**GD**D**EDD**EG**EE**Y
TUA3/5	AREDLAALEKDYEEVG**A**EG**GD**D**EED**EG**ED**Y
TUA6	AREDLAALEKDYEEVG**A**EG**GD**D**EDD**EG**EE**Y

Carboxy termini of three inferred protein sequences of alpha-tubulins from *Ceratopteris richardii*:

TuaCR1	AREDLAALEKD**Y**EEVGA**EGQDD**D**EPGD D**EY
TuaCR2	AREDLAALEKD**Y**EEVGA**EGDEG**D**EGEDGD**EY
TuaCR3	AREDLAALEKD**F**EEVGA**DSTEG**D**GEDEGE**EY

FIG. 2. Comparison of the carboxy terminus regions of *Arabidopsis* and *Ceratopteris* alpha-tubulin genes. Bold letters represent regions where the sequences within each set are not identical; a blank space represents an apparent gap in the alignment.

To further compare the relative diversity of the *Ceratopteris* and *Arabidopsis* alpha-tubulin proteins, the sequences from the carboxy terminus region, which is considered to be highly variable in alpha-tubulin proteins (e.g., see Sullivan, 1988; Fosket and Morejohn, 1992), were aligned with one another (Fig. 2). Since only three of the four *Ceratopteris* sequences contain the coding region for the carboxy terminus portion of the protein, TuaCR4 was omitted from this analysis. For both species the amino acid sequence is relatively conserved until the last 14 residues at the carboxy terminus end, or in the case of TuaCR1, which appears to have a single amino acid deletion in this region, the last 13 amino acids. In this region the *Arabidopsis* sequences share 6 identical amino acids out of 14, while the *Ceratopteris* sequences share only 3 identical amino acids.

Gene phylogeny results.—The Bayesian tree (Fig. 3) infers phylogenetic relationships of nucleotide sequences of the four sequenced *C. richardii* alpha-tubulin genes in relation to alpha-tubulin genes from other plant species. This figure depicts the tree from the single run noted in Materials and Methods as resolving one of the nodes, the resolved subnodes being toward the top of the tree with probabilities of 0.80, 0.50, and 0.81. The central clade from *Zea mays* 22149 to *Oryza sativa japonica* cultivargroup 1136121 was not resolved in any of the Bayesian runs. Posterior probability values show that many clades of this gene tree are strongly to moderately supported, although some deeper branches had weak support. The unresolved and weakly supported clades, however, are inconsequential to this paper, which focuses instead on the four copies of *C. richardii* genes and their placement in strongly supported clades. All nodes separating the four *Ceratopteris* genes are strongly supported. Three *Ceratopteris* copies (TuaCR1, TuaCR2, and TuaCR4) toward the top of Fig. 3 are relatively closely related to each other in a strongly supported subclade that is part of the major clade stretching from *C. richardii* TuaCR1 29423812 to

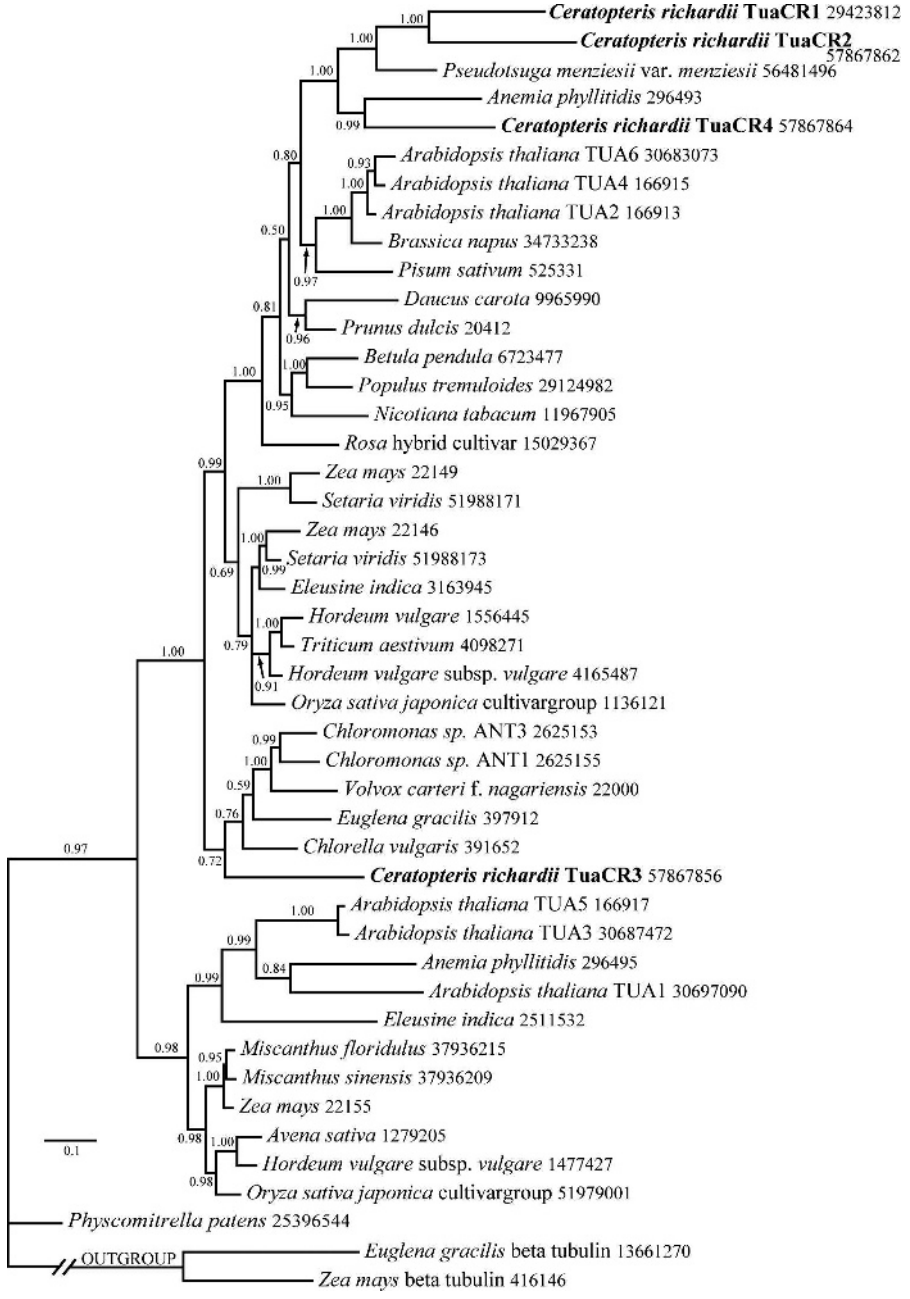

FIG. 3. Bayesian inference gene tree based on aligned nucleotide sequences of alpha-tubulin genes of Viridiplantae in GenBank (excluding the invariant first three nucleotides and the non-alignable last 45 nucleotides), showing relationships of the four sequenced *C. richardii* alpha-tubulin gene copies discussed. Taxon sources and GenBank gi numbers are given for each gene copy. Bayesian posterior probability values indicate support for clades.

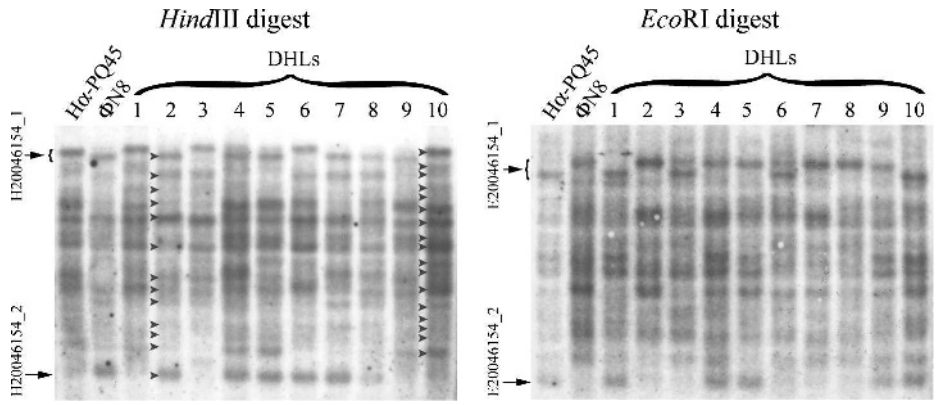

FIG. 4. Example autoradiograms of total genomic DNAs from the two parental *C. richardii* races and 10 DHLs derived from their cross, all probed with a DIG-labeled alpha-tubulin Cri_10_E18_SP6 cDNA clone after respective digestions of the genomic DNAs with *Hind*III or *Eco*RI. Arrowheads in DHL lanes 2 and 10 indicate at least 13 and 17 probable bands respectively. Segregating co-dominant and dominant RFLP bands containing alpha-tubulin genes are identified by arrows to the left of each autoradiogram and are discussed in the text and mapped in Fig. 5.

Oryza sativa japonica cultivargroup 1136121. This major clade is separated from its sister clade of miscellaneous algae and *C. richardii* TuaCR3 57867856 with a posterior probability of 1.00, although the precise positioning of *C. richardii* TuaCR3 in the algal clade is less strongly supported.

Mapping results.—Probings of parental and DHL DNAs with the DIG-labeled alpha-tubulin Cri_10_E18_SP6 cDNA clone yielded at least ca. 13–17 restriction fragment bands per DHL. The total number of bands cannot be determined with precision because of overlap and faintness of some of the bands. As an example, Fig. 4 shows the probing results for parental sporophytes Hα-PQ45 and ΦN8 and for ten DHLs whose genomic DNAs were cut with restriction enzymes *Hind*III and *Eco*RI, respectively. The DHL in lane 2 of the *Hind*III digest, for example, probably exhibits at least 13 bands (arrowheads) containing sequence to which the alpha-tubulin probe anneals and the DHL in lane 10 shows at least 17 bands. In some cases multiple bands may represent a single alpha-tubulin gene sequence that has been cut by the *Hind*III or *Eco*RI restriction enzyme. Although the sequence of the cDNA probe used in this alpha-tubulin study contains no *Hind*III or *Eco*RI restriction site, the coding region of the target DNA might contain such sites, and introns within the genes of the target DNA may contain *Hind*III or *Eco*RI sites. No data are available to address these possibilities.

Mapping can be performed only for those parental alpha-tubulin gene copies contained in restriction fragments that are polymorphic between the two parents and that therefore segregate in the DHLs. Four such unequivocal segregating sets of restriction fragments were observed in the genomic DNAs digested with *Hind*III and *Eco*RI in our mapping population. These are identified (Fig. 4) as H20046154_1 and H20046154_2 scored from the *Hind*III

digest and E20046154_1 E20046154_2 scored from the *Eco*RI digest. Polymorphic restriction fragment markers are usually expressed as co-dominant bands, meaning that the gene's presence is visualized in respective bands from both parents. In Fig. 4, this is exemplified in the *Hind*III digest by locus H20046154_1 where the larger fragment from parent Hα-PQ45 (also seen in DHL lanes 1, 3, 6, 10) is ca. 1 mm closer to the top of the figure than is the smaller fragment from parent ΦN8 (also seen in DHL lanes 2, 4, 5, 7, 8, 9). Locus E20046154_1 in the *Eco*RI digest in Fig. 4 shows a similar pattern. In this case the smaller fragment in parent Hα-PQ45 (also seen in DHL lanes 1, 3, 6, 10) is ca. 1.5 mm farther from the top of the autorad than is the larger fragment from parent ΦN8 (also seen in DHL lanes 2, 4, 5, 7, 8, 9). The two bands from the *Hind*III digest co-segregate with the two bands from the *Eco*RI digest in all of the mapping population's DHLs, indicating that they either mark the same identical locus visualized in the two different digests or mark two loci so closely linked that they show no crossover distance between them (i.e., they map to the same location). Locus H20046154_2 on the *Hind*III digest, on the other hand, is visualized as a dominant band in parent *Φ*N8 (also seen in DHL lanes 2, 4, 5, 6, 7, 8) meaning that no alternative band expression is visualized in the Hα-PQ45 parent (and in DHL lanes 1, 3, 9, 10). The reason for this dominant pattern is presently unknown but may be because the H20046154_2 gene copy in the Hα-PQ45 parent has been lost or has been moved to a different part of the genome where it cannot be scored because it overlaps with a different band, etc. In the *Eco*RI digest, locus E20046154_2 is also expressed as a dominant marker, present in parent Hα-PQ45 (and in DHL lanes 1, 4, 5, 9, 10), but with an alternative band lacking in parent *Φ*N8 (and in DHL lanes 2, 3, 6, 7, 8). Clearly the bands marking H20046154_2 and E20046154_2 do not co-segregate in the DHLs of the two digests, indicating that they mark different loci. The mapping program places H20046154_2 on linkage group (LG) 28 and E20046154_2 on LG 24 (Fig. 5 at arrows), whereas H20046154_1 and E20046154_1 map to the same position on LG 17 (Fig. 5 arrow).

Discussion

Map positions of alpha-tubulin genes.—The genetic linkage mapping project for *C. richardii*, fully described by Nakazato *et al.* (2006), has identified 41 linkage groups that partly correspond to the 39 chromosomes per haploid set in this species. Alpha-tubulin gene copy H20046154_2 is located on LG 28, H&E20046154_1 is on LG 17, and E20046154_2 is on LG 24 (Fig. 5). Also seen on these three linkage groups are a large number of AFLP markers, each identified by a lowercase ''a'' followed by the primer pair used to generate it, and additional RFLP markers whose code numbers are the GenBank gi numbers of the cDNA library clones used as probes, prefaced with an ''H'' or an ''E'' to indicate whether that locus was visualized on the *Hind*III or the *Eco*RI digests. On LG 17 is also found the isocitrate dehydrogenase (IDH) isozyme locus, which maps to virtually the same position as H9960276_2. In

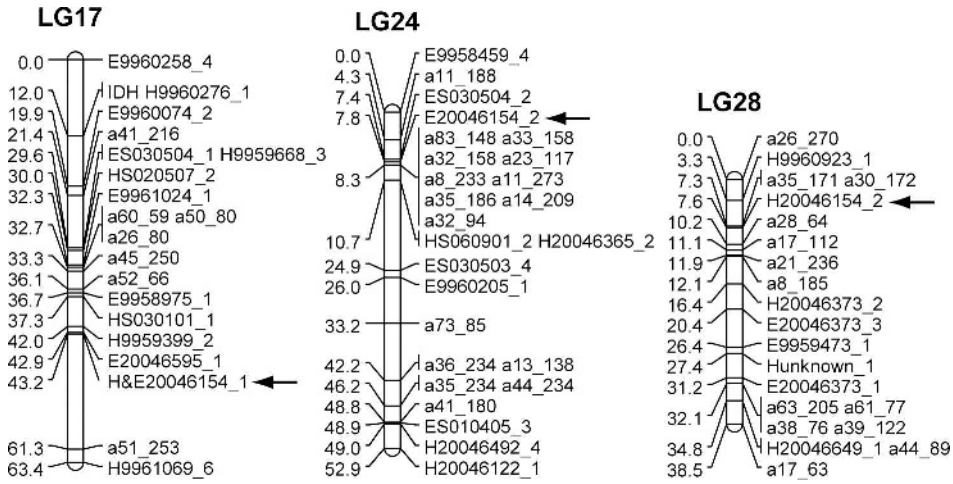

FIG. 5. Three of the 41 linkage groups presently identified in the *C. richardii* genome, showing at arrows the positions of the four segregating RFLP bands that contain alpha-tubulin genes identified in Fig. 4.

all cases, the cumulative distance in centiMorgans from the topmost maker is given to the left of each linkage group.

Because the alpha-tubulin clone used as a DIG-labeled probe here anneals to every alpha-tubulin gene copy with sufficient sequence identity, given the stringency conditions of our general mapping project, we cannot determine which of the genes in Table 1 are represented by respective markers H&E20046154_1, H20046154_2, and E20046154_2. The large number of bands (at least 13 to 17 in the indicated lanes in Fig. 4) suggests a large alpha-tubulin gene family in *C. richardii*, even if some genes are represented by more than one band. We were unable to map more than three or four alpha-tubulin loci because only four loci were contained in restriction fragments polymorphic in the two parents in *Hind*III and *Eco*RI digests.

Ceratopteris *alpha-tubulin gene phylogeny*.—The phylogenetic relationships of the four *C. richardii* genes in this paper are illustrated in Fig. 3. Bayesian posterior probability values show that copies TuaCR1 and TuaCR2 at the top of the tree are very strongly grouped as sister to each other and that both are in turn robustly sister to a copy from the conifer *Pseudotsuga menziesii*. Also with very strong support, the preceding are separated from TuaCR4 which is sister to an alpha-tubulin gene copy from the fern *Anemia phyllitidis*. The most parsimonious hypothesis to explain these present data is that TuaCR1 and TuaCR2 result from a recent duplication, perhaps after the divergence of ferns and seed plants, and that the gene duplication leading to the TuaCR4 lineage and the TuaCR1+TuaCR2+*Pseudotsuga* lineage preceded the divergence of ferns and seed plants. *Ceratopteris richardii* TuaCR3 (Fig. 3), on the other hand, is grouped in a clade of algal alpha-tubulin gene copies below the center of the tree. Although the precise relationships within this

clade of algal genes+TuaCR3 are not strongly supported, a 1.0 posterior probability value strongly separates this clade from the other three *Ceratopteris* genes in the major sister clade of miscellaneous seed plant alpha-tubulin gene copies above it. This indicates that TuaCR3 had already diverged from the common ancestor of the other three *Ceratopteris* alpha-tubulin genes by the time of the common ancestor of the ferns and algae, preceding the divergence of algae and land plants. The sampling of alpha-tubulin genes from plants in general, however, is currently very incomplete. The four *C. richardii* copies discussed here are simply those expressed in a cDNA library derived from 12-day-old gametophytes. Surely they do not represent the full range of alpha-tubulin gene copies expressed in the life cycle of *C. richardii*, just as the two alpha-tubulin gene copies sequenced thus far from the fern *Anemia phyllitidis* in Fig. 3 must represent only a fraction of the copies from that species. Inferred timings of duplications will likely change when more alpha-tubulin gene sequences become available. It may be noteworthy that *C. richardii* and the algae have motile gametes whereas all of the seed plants in Fig. 3 (including *Pseudotsuga*) lack motile sperm. This suggests TuaCR3 as a candidate for the gene copy functioning in sperm motility microtubules.

The occurrence of distantly related beta- and gamma-tubulin genes resulting from ancient duplications and the large number of alpha-tubulin gene copies seen in the *C. richardii* genomes in Fig. 4 indicate that there has been extensive and continuous duplication of tubulin genes throughout the evolution of organisms. It is also likely that deletion/silencing of tubulin gene copies is frequent, perhaps in lineage-specific ways. For example, despite intensive sequencing of alpha-tubulin genes in seed plants, gene copies from seed plants are absent from the lineage containing the TuaCR3 and algal gene copies in Fig. 3. This suggests that silencing/deletion of this copy may have occurred specifically in the seed plant lineage.

Diversity of the C. richardii *alpha-tubulin genes*.—In terms of amino acid sequences, particularly the carboxy termini, the four *Ceratopteris* alpha-tubulin sequences appear to be relatively diverse both when compared to one another (Figs. 1 and 2) and when compared to sequences from other plants (see Table 1, Fig. 2, and text of the Results section). The structural similarity of the *Ceratopteris* alpha-tubulin protein sequences to those of other diverse species is not surprising. This kind of similarity is often seen for alpha-tubulin proteins when sequences are compared using the BLAST algorithm (data not shown). The phylogenetic relationships of these four *C. richardii* sequences determined by Bayesian analysis of their nucleotide sequences excluding the invariant first codon and the non alignable 45 carboxy terminus nucleotides, however, indicates less diversity (Fig. 3). Phylogenetic analysis shows that TuaCR1 is closest to TuaCR2 and that, except for the alpha-tubulin copies from *Pseudotsuga* and *Anemia*, TuaCR4 is most closely related to TuaCR1 and TuaCR2. TuaCR3, on the other hand, is quite distant from the other three *C. richardii* copies and is most closely related to algal alpha-tubulin genes.

Potential insights from further studies of the Ceratopteris *alpha-tubulin gene family*.—Two main types of insights may be gained by studying the

diverse genes of organisms like *Ceratopteris*, and these are discussed in turn below. First, the diversity maintained in gene families like that of alpha-tubulin may provide new insights into the evolutionary history of specific genes as well as the mechanisms by which they evolve. The conventional wisdom regarding gene families is that the multiple versions of related genes develop by gene duplication and random mutation that produces variants subject to varying degrees of selection. In some cases, the different forms of these related genes assume different roles over the course of evolutionary history (e.g., see Zhang, 2003; Irish and Litt, 2005).

Consider the comparison presented earlier between the alpha-tubulin genes of *Arabidopsis* and those of *Ceratopteris*. The relative similarity in amino acid sequence of the six alpha-tubulin genes of *Arabidopsis thaliana* noted above suggests at least two likely interpretations. First, selection pressure may be very high for alpha-tubulin genes in this species, such that very little deviation is tolerated. Second, there may be relatively fewer potential roles for alpha-tubulin in the development and life history of *Arabidopsis* (for example, sperm are non-motile, unlike the situation in ferns) and therefore evolutionary factors have resulted in less diversity. In this respect, it is noteworthy that the *Arabidopsis* alpha-tubulin gene that differs most from the other five, TUA1, appears to be preferentially expressed primarily in pollen grains (Carpenter *et al.* 1992), while the others are apparently expressed in various tissues throughout the plant (Kopczak *et al.*, 1992).

By comparison with the alpha-tubulin genes of *Arabidopsis*, amino acid sequences of the four known alpha-tubulin genes of *Ceratopteris* appear somewhat diverse, suggesting selective pressures favoring the origin or maintenance of alpha-tubulin variation in this fern. This may relate in part to its fully independent gametophyte generation with perhaps more potential roles for alpha-tubulin (for example, as a component of the flagella in swimming sperm) than are found in its angiosperm counterpart, *Arabidopsis*. If these considerations prove correct and generally applicable to other genes, new insights regarding evolutionary mechanisms could be gained from studies focusing on *C. richardii* as a model homosporous vascular plant that could never be gained by studying plants like *Arabidopsis* alone.

A second type of information to be gained by studying diverse members of gene families, such as those of the *C. richardii* alpha-tubulin family, relates to gene expression. For example, what are the specific roles of the various members of the gene family? When during development and under which environmental conditions are these genes expressed? How is the expression of each gene controlled? To provide some insight into how *C. richardii* can be utilized to answer such questions, two potential research strategies utilizing the new sequence data presented in this report are described below.

A large number of morphological mutants exist from classical mutagenesis and selection screens using *Ceratopteris* (reviewed in Hickok, 1987; Hickok *et al.*, 1987, 1995). Some affect cell shape, like the *dwarf* or *bubbles* mutant (Hickok, personal communication), others affect intracellular organization, like the *polka-dot* mutant (Vaughn *et al.*, 1990), while yet others, like *sleepy*

sperm (Renzaglia *et al.*, 2004) affect the mobility of the sperm. Based on the phenotypes of the above mutants, it is likely that the defect associated with some of them may involve either an alpha-tubulin gene itself, a gene that controls expression of an alpha-tubulin gene, or a gene for another protein that interacts with alpha-tubulin. With the four sequences presented in this report, it will be relatively simple to assess the first two possibilities, using PCR to obtain sequence data and techniques like reverse transcription-PCR (RT-PCR) and/or real-time PCR to study relative levels of gene expression.

One may also study the roles of the various alpha-tubulin genes by utilizing recently developed methods for inducing gene silencing via RNA interference (RNAi) in spores and gametophytes of *Ceratopteris* (Stout *et al.*, 2003; Rutherford *et al.,* 2004). RNAi is a technique whereby specific genes may be silenced by the introduction of short single- or double-stranded RNA sequences into an organism's cells. These short RNA sequences, which must be highly similar to or identical to the complementary coding sequences within the targeted gene, trigger a cellular response that causes the degradation of the mRNA transcribed from the targeted gene.

In the two studies published so far using RNAi in *C. richardii*, the short RNAs have been introduced either by simply soaking spores directly in a solution containing the RNAs (Stout *et al.*, 2003) or by using the biolistic method to deliver the RNA into gametophytic cells (Rutherford *et al.,* 2004). The effects of gene silencing in *C. richardii* appear to be transmissible to cells descended from the cell into which the inhibiting RNA was initially introduced. In experiments where the inhibitory RNA was introduced into cells of young gametophytes by biolistic delivery, an observable phenotype was sometimes even transmitted to developing sporophytes (Rutherford *et al.*, 2004).

Each of the four alpha-tubulin genes described here has at least several portions of its sequence that are not shared by the other known members of this gene family (this is especially evident in the 3' regions of TuaCR1, 2, and 3). These sequences can be used to design short inhibitory RNAs that should be able to silence specifically the expression of their respective genes. Alternatively, it should be possible to silence two or more of these genes at the same time by using short RNA sequences that are shared by two or more genes. In either approach, such silencing can be confirmed by demonstrating an absence of (or reduction in) the targeted mRNA by using RT-PCR (Stout *et al.*, 2003) or real time PCR (Rutherford *et al.*, 2004), and phenotypes resulting from the silencing may also be observed in some cases (Rutherford *et al.*, 2004).

Several potential phenotypes could be easily observable as a result of silencing alpha-tubulin expression in *Ceratopteris*. Some may include such developmental defects as the inhibition of spore germination, or the generation of abnormal rhizoids, prothalial cells, or specialized cells such as those of antheridia or archegonia. Cytoskeletal defects generated by this approach may also lead to observable phenotypes—some may be associated with cell size or shape, while others may be as striking as the phenotype of the *polka-dot* mutant. Other observable abnormalities may affect the function of spermata-

zoids, generating phenotypes similar to the *sleepy sperm* mutant. The obvious advantage to generating such phenotypes through RNAi is that in each case, the phenotype will be directly associated with the silencing of one or more specific alpha-tubulin genes. This in turn may lead to the assignment of specific roles for each member of this gene family.

The linkage positions of the three mapped loci together with the specific sequencing data provided here should facilitate future cloning of genomic sequences containing alpha-tubulin genes in *C. richardii*. This will enable characterization of the organization and controlling elements of these genes, enhancing our understanding of the evolution, function, and regulation of the alpha-tubulin gene family in vascular plants in general. The research possibilities discussed here illustrate how robust molecular techniques can be applied to the *C. richardii* system, furthering its usefulness as a model system.

Acknowledgments

RJS and JWW thank Nalco Chemical Company for funds used by Wheaton College to purchase a DNA sequencer, and the Wheaton College Administration and the Wheaton College Alumni Association for funds used to conduct the sequencing work. GJG and TN thank the National Science Foundation for support of their mapping project through grant DEB-0128926.

Literature Cited

Aso, K., M. Kato, J. A. Banks and M. Hasebe. 1999. Characterization of homeodomain-leucine zipper genes in the fern *Ceratopteris richardii* and the evolution of the homeodomain-leucine zipper gene family in vascular plants. Molec. Biol. Evol. 16:544–552.

Banks, J. A. 1999. Gametophyte development in ferns. Annual Rev. Pl. Physiol. Pl. Molec. Biol. 50:163–186.

Carpenter, J. L., S. E. Ploense, D. P. Snustad and D. Silflow. 1992. Preferential expression of an α-tubulin gene of *Arabidopsis* in pollen. Pl. Cell 4:557–571.

Fosket, D. E. and L. C. Morejohn. 1992. Structural and functional organization of tubulin. Annual Rev. Pl. Physiol. Pl. Molec. Biol. 43:201–240.

Hasebe, M., C-K. Wen and J. A. Banks. 1998. Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*. Proc. Natl. Acad. Sci. U.S.A. 95:6222–6227.

Hickok, L. G. 1976. Cytological relationships between three diploid species of the fern genus *Ceratopteris*. Canad. J. Bot. 55:1660–1667.

Hickok, L. G. 1977a. An apomictic mutant for sticky chromosomes in the fern *Ceratopteris*. Canad. J. Bot. 55:2186–2195.

Hickok, L. G. 1977b. The cytology and derivation of a temperature-sensitive meiotic mutant in the fern *Ceratopteris*. Amer. J. Bot. 64:552–563.

Hickok, L. G. 1978. Homoeologous chromosome pairing and restricted segregation in the fern *Ceratopteris*. Amer. J. Bot. 65:516–521.

Hickok, L. G. 1979a. Apogamy and somatic restitution in the fern *Ceratopteris*. Amer. J. Bot. 66:1074–1078.

Hickok, L. G. 1979b. A cytological study of intraspecific variation in *Ceratopteris thalictroides*. Canad. J. Bot. 57:1694–1700.

Hickok, L. G. 1987. Applications of in vitro selection systems: whole plant selection using the haploid phase of the fern *Ceratopteris*, pp. 53–56. In H. M. LeBaron, R. O. Muma, R. C. Honeycutt and J. H. Duesing, eds. Biotechnology in agricultural chemistry. American Chemical Society, Washington, DC.

HICKOK, L. G. and E. J. KLEKOWSKI. 1973. Abnormal reductional and nonreductional meiosis in *Ceratopteris*: Alternatives to homozygosity and hybrid sterility in homosporous ferns. Amer. J. Bot. 60:1010–1022.

HICKOK, L. G. and E. J. KLEKOWSKI. 1974. Inchoate speciation in *Ceratopteris*: An analysis of the synthesized hybrid C. *richardii* × C. *pteridoides*. Evolution 28:439–446.

HICKOK, L. G., T. R. WARNE and R. S. FRIBOURG. 1995. The biology of the fern *Ceratopteris* and its use as a model system. Int. J. Pl. Sci. 156:332–345.

HICKOK, L. G., T. R. WARNE and M. K. SLOCUM. 1987. *Ceratopteris richardii*: Applications for experimental plant biology. Amer. J. Bot. 74:1304–1316.

HOFFMAN, J. C. and K. C. VAUGHN. 1995a. Using the developing spermatogenous cells of *Ceratopteris* to unlock the mysteries of the plant cytoskeleton. Int. J. Pl. Sci. 156:346–358.

HOFFMAN, J. C. and K. C. VAUGHN. 1995b. Post-translational tubulin modifications in spermatogeneous cells of the pteridophyte *Ceratopteris richardii*. Protoplasma 186:169–182.

HOFFMAN, J. C. and K. C. VAUGHN. 1996. Spline and flagellar microtubules are resistant to mitotic disrupter herbicides. Protoplasma 192:57–69.

HOFFMAN, J. C., K. C. VAUGHN and H. C. JOSHI. 1994. Structural and immunocytochemical characterization of microtubule organizing centers in pteridophyte spermatogenous cells. Protoplasma 179:46–60.

HUELSENBECK, J. P. and F. RONQUIST. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

IRISH, V. F. and A. LITT. 2005. Flower development and evolution: gene duplication, diversification and redeployment. Current Opinion in Genetics & Development 15:454–460.

KOPCZAK, S. D., N. A. HAAS, P. J. HUSSEY, C. D. SILFLOW and D. P. SNUSTAD. 1992. The small genome of *Arabidopsis* contains at least six expressed α-tubulin genes. Pl. Cell 4:539–547.

LANDER, E., J. ABRAHAMSON, A. BARLOW, M. DALY, S. LINCOLN, L. NEWBURG and P. GREEN. 1987. MAPMAKER: A computer package for constructing genetic linkage maps. Cytogenet. & Cell Genet. 46:642.

MUNSTER, T., J. PAHNKE, A. DI ROSA, J. T. KIM, W. MARTIN, H. SAEDLER and G. THEISSEN. 1997. Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants. Proc. Natl. Acad. Sci. U.S.A. 94:2415–2420.

MURATA, T., A. KADOTA and M. WADA. 1997. Effects of blue light on cell elongation and microtuble orientation in dark-grown gametophytes of *Ceratopteris richardii*. Pl. Cell Physiol. 38:201–209.

NAKAZATO, T. and G. J. GASTONY. 2006. High-throughput RFLP genotyping method for large genomes based on a chemiluminescent detection system. Plant Molec. Biol. Rep. 24:245a–245f.

NAKAZATO, T., M.-K. JUNG, E. A. HOUSWORTH, L. H. RIESEBERG and G. J. GASTONY. 2006. Genetic map-based analysis of genome structure in the homosporous fern *Ceratopteris richardii*. Genetics 173:1585–1597.

POSADA, D. and K. A. CRANDALL. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818.

PRYER, K. M., H. SCHNEIDER, A. R. SMITH, R. CRANFILL, P. G. WOLF, J. S. HUNT and S. D. SIPES. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409:618–622.

RAGHAVAN, V. 1989. Developmental Biology of Fern Gametophytes. Cambridge University Press, New York, NY.

RENZAGLIA, K. S., K. D. WOOD, G. RUPP and L. G. HICKOK. 2004. Characterization of the sleepy sperm mutant in the fern *Ceratopteris richardii*: A new model for the study of axonemal function. Canad. J. Bot. 82:1602–1617.

RONQUIST, F. and J. P. HUELSENBECK. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

RUTHERFORD, G., M. TANURDZIC, M. HASEBE and J. A. BANKS. 2004. A systemic gene silencing method suitable for high throughput, reverse genetic analyses of gene function in fern gametophytes. BMC Pl. Biol. 4:6. (online journal, see http://www.biomedcentral.com).

SALMI, M. L., T. J. BUSHART, S. C. STOUT and S. J. ROUX. 2005. Profile and analysis of gene expression changes during early development in germinating spores of *Ceratopteris richardii*. Pl. Physiol. 138:1734–1745.

STOUT, S. C., G. B. CLARK, S. ARCHER-EVANS and S. J. ROUX. 2003. Rapid and efficient suppression of gene expression in a single-cell model system, *Ceratopteris richardii*. Pl. Physiol. 131:1165–1168.

SULLIVAN, K. F. 1988. Structure and utilization of tubulin isotypes. Annual Rev. Cell. Biol. 4:687–716.

VAUGHAN, K. C., L. G. HICKOK, T. R. WARNE and A. C. FARROW. 1990. Structural analysis and inheritance of a clumped-chloroplast mutant in the fern *Ceratopteris*. J. Heredity 81:146–151.

VOGELMANN, T., A. BASSEL and J. MILLER. 1981. Effects of microtubule-inhibitors on nuclear migration and rhizoid differentiation in germinating fern spores (*Onoclea sensibilis*). Protoplasma 109:295–316.

YOUNG, R. A. and R. W. DAVIS. 1991. Gene isolation with λ gt11 system. Methods Enzymol. 194:230–238.

ZHANG, J. 2003. Evolution by gene duplication: an update. Trends Ecol. Evol. 18(6):292–298.