Assessment of gene copy number in the homosporous ferns *Ceratopteris thalictroides* and *C. richardii* (*Parkeriaceae*) by restriction fragment length polymorphisms

J. Mitchell McGrath, Leslie G. Hickok, and Eran Pichersky

Abstract: Homosporous ferns are generally considered polyploid due to high chromosome numbers, but genetically diploid since the expression of isozymes is generally controlled by a single locus. Gene silencing over evolutionary time is one means by which this apparent contradiction can be explained. A prediction of this hypothesis is that silenced gene sequences still reside in the genomes of homosporous ferns. We examined the genomes of *Ceratopteris richardii* and *C. thalictroides* for sequences which are similar to expressed gene sequences. Genomic DNA blots hybridized with *C. richardii* cDNA clones showed that the majority of these clones detected multiple fragments, suggesting that most gene-like sequences are duplicated in *Ceratopteris*. Hybridization signal intensity often varied between fragments of the same size between accessions, sometimes dramatically, which indicates that not all sequences are equivalent, and may represent the products of silenced genes. Observed reciprocal differences in intensity could be due to reciprocally silenced genes. In addition, an unusual segregation pattern for one locus followed by one probe may indicate homeologous chromosome pairing and segregation.

Since diploid homosporous ferns have high chromosome numbers, they are generally thought to be anciently polyploid. However, the expression of a large number of isozymes are each controlled by a single locus, thus there is no evidence for extensive gene duplication (Haufler & Soltis 1986 and references therein). To reconcile these contradictory observations, it has been hypothesized that genome-wide silencing of duplicate genes has occurred over evolutionary time (Haufler 1987). If gene silencing were an important feature of homosporous fern evolution, relic DNA sequences with similarity to expressed genes may be expected to still reside in their genomes. Evidence in favor of this hypothesis has been obtained by Pichersky & al. (1990) who showed that the majority of sequences similar to chlorophyll *a/b*-binding (CAB) protein genes isolated from *Polystichum munitum* (n = 41) were interrupted by deletions and premature stop codons. In contrast, many CAB genes have been characterized in numerous angiosperm species, but

only one defective CAB gene has so far been isolated (reviewed in GREEN & al. 1991).

To determine whether multiple copies of possibly defective genes other than CAB sequences exist in fern genomes, we have begun examining the level of duplicated nucleotide sequences among accessions of homosporous ferns within the genus *Ceratopteris*. *Ceratopteris richardii* (n = 39) is an attractive model system among homosporous ferns due to its ease of culture, relatively short generation time and the ability to recover homozygous mutants in a single generation (HICKOK & al. 1987). *Ceratopteris thalictroides* shares many of these attributes (KLEKOWSKI 1970) but is generally considered tetraploid (n = 77, 78, HICKOK 1979), and is thus an example of a neopolyploid. In a companion paper (MCGRATH & al. 1993 a, this volume), we describe the use of RFLPs to discriminate among *C. richardii* and *C. thalictroides*. In the present report, we present the data as it relates to genome organization of *Ceratopteris*.

Material and methods

The material and methods are those described in MCGRATH & al. (1994, this volume), with the following addition. For construction of a genetically segregating population, randomly-selected gametophytes from an intraspecific *C. richardii* hybrid (strain *Hn-n* as female parent by strain *D 176* as male; SCOTT & HICKOK 1987) were individually and separately transferred to Petri plates prior to sexual maturity, and self-fertilization (i.e., intra-gametophytic selfing) was accomplished. Each resulting sporophyte was expected to be homozygous at all loci, and in effect represented the doubled haploid genotype of each hybrid-derived gametophyte. This population was referred to as the "F_2" generation. Other accessions examined are listed in Table 1.

Results

Most sequences similar to expressed genes appear duplicated. The number of bands observed on a blot is a function of both the number of target sequences and the

Table 1. *Ceratopteris* accessions utilized in the present study. Chromosome numbers taken from HICKOK (1977, 1979) and HICKOK & KLEKOWSKI (1974)

Key	Accession	Location	Species	n =
1	*IN*	India	*C. thalictroides* (L.) BRONGN.	77
2	*19K1*	Hawaii	*C. thalictroides*	
3	*185 H*	Surinam	*C. thalictroides*	
4	*231-4*	New Guinea	*C. thalictroides*	77
5	*JDA 1*	Australia	*C. thalictroides*	
6	*KN 5*	Japan	*C. thalictroides*	
7	*NPA 6*	Puerto Rico	*C. richardii* BRONGN.	39
8	*PhiN 8*	Nicaragua	putative *C. richardii* × *C. pteridoides* hybrid	39
9	*RLMN 8*	Hawaii	*C. thalictroides*	78
10	*Hn-n*	Cuba	*C. richardii*	39
11	*D 176*	Guyana	*C. richardii*	39

distribution of restriction endonuclease sites within that sequence. Thus, the number of fragments detected is not a reliable indicator of the number loci (genes) carried in the genome. However, most plant genes, including homosporous fern CAB genes (PICHERSKY & al. 1990), are contained on genomic DNA fragments of 3 to 5 kb or less (including exons, introns and both upstream and downstream regulatory sequences). It is unlikely, therefore, that three or more fragments > 3 kb detected on genomic blots are derived from the same gene. The number of fragments we observed among accessions of *Ceratopteris* (Table 2) suggested that many sequences are duplicated. We consider the number of fragments given in Table 2 to be a conservative estimate since only the more strongly hybridizing bands observed after an exposure of 4 days were scored. Additional weak signals were often detected after longer exposures.

All *Ceratopteris* accessions showed four or more fragments > 3 kb when hybridized with 4 of the 11 cDNA probes (3-31-1, 3-31-13, 3-31-17, and 4-1-7; Table 2), and were considered as putative gene families. Two probes (1-20-2 and 1-20-4) consistently hybridized with fewer than four fragments, and were considered as single or low-copy number loci in all accessions. The remaining probes showed mixed results. With the exceptions of accessions *185 H* and *RLMN 8*, low-copy number sequences were detected with probe 1-20-16. Probe 4-1-14, and perhaps 1-20-11, was low-copy in the diploid species *C. richardii* (and the diploid accession *PhiN 8*) but appeared duplicated in the tetraploid species *C. thalictroides*. In contrast, probe 1-20-19 was generally scored as low-copy in *C. thalictroides*, but genetic evidence for its duplication in *C. richardii* was obtained (see below). These observations suggest that the majority of clones derived from expressed genes hybridize with more than one sequence in the *Ceratopteris* genome.

Table 2. Number of restriction fragments observed after 4 day exposure on Southern blots of *Ceratopteris* accessions. [1] see Table 1 for Key

Probe	Accession[1]										
	1	2	3	4	5	6	7	8	9	10	11
1-20-2	2	2	2	1	2	1	2	2	3	2	2
1-20-4	1	1	1	1	1	1	1	1	1	1	1
1-20-11	5	4	4	3	4	4	3	3	2	1	nd
1-20-16	3	3	4	3	2	2	2	2	9	3	3
1-20-19	2	5	3	2	2	2	3	1	6	4	4
3-31-1	4	4	4	6	7	7	5	4	7	6	6
3-31-2	6	3	4	3	9	7	6	4	11	7	7
3-31-13	11	11	12	12	12	12	10	8	10	5	nd
3-31-17	11	11	4	6	7	4	5	5	12	5	5
4-1-7	7	6	8	7	9	8	7	9	8	7	6
4-1-14	12	7	5	5	5	5	3	2	6	nd	3
n =	11	11	11	11	11	11	11	11	11	10	9
Mean	5.8	5.2	4.6	4.5	5.5	4.8	4.3	3.7	6.8	4.1	4.1
Std. dev.	3.97	3.34	3.01	3.24	3.62	3.43	2.65	2.69	3.66	2.28	2.03

A Student's t comparison between diploid and tetraploid species showed a slight but significant difference in the number of fragments detected over all probes. The mean number of fragments scored among diploids was 4.05 (n = 41, std. dev. = 2.37) and 5.31 among tetraploids (n = 77, std. dev. = 3.42; $t_{calculated}$ = 2.10 versus the critical $t_{0.05}$ = 1.65). If accessions showing the most extreme values (*RLMN8* and *PhiN8*, Table 2) were excluded from the analysis, the difference in mean fragment number (4.17 in diploids, 5.06 in tetraploids) was marginally significant at the 0.05% level ($t_{calculated}$ = 1.68 versus $t_{0.05}$ = 1.66).

Differences in intensity among fragments of the same size. The multiple fragments in each accession hybridizing to a given probe did not always show the same signal intensity. In some cases, signal intensity varied dramatically between fragments of the same size from different accessions. One explanation for this result may have been that since the cDNA library was constructed from *C. richardii* strain *Hn-n*, probes might have been expected to hybridize more strongly to DNA from *C. richardii* than *C. thalictroides*. This was observed only with probe 3-31-1.

With three probes (1-20-4, 3-31-13, and 4-1-7), the same-sized fragments showed differences in signal intensity among accessions of *C. thalictroides*. With probe 1-20-4, both Hawaiian *C. thalictroides* accessions displayed an intense signal, while all other accessions, including *C. richardii*, showed barely detectable signals. This observation could be explained by the presence of tandemly-duplicated sequences in the Hawaiian types. With the other two probes, the difference between band intensity among shared fragments of identical size was not correlated with geographical location. For example, strains *IN* and *19K1* both carry fragments of 9 kb and 8 kb detected with probe 4-1-7 (see McGrath & al. 1994, this volume: fig. 1 B), but differ in intensity such that the 9 kb fragment is more intense than the 8 kb fragment in *IN*, and vice versa in *19K1*. A more dramatic example is illustrated in Fig. 1. In this case a band of 4.5 kb detected with probe 3-31-13 is present in all *C. thalictroides*, but its intensity varies from high in accessions *IN* and *JDA 1* to barely detectable in strains *185 H* and *231-4*. Since signal strengths for shared fragments of 2.5, 3.8, and 9.0 kb are equivalent between these four accessions (indicating equal DNA loadings per lane), differential evolution of the 4.5 kb fragment must have occurred among *C. thalictroides* types.

Evidence for duplicated RFLP loci in *C. richardii*. Results reported previously (McGrath & al. 1994, this volume) indicated that the level of polymorphism among these *C. richardii* accessions was low. For eleven probes tested, only two showed an RFLP between strains *Hn-n* and *D 176*. The inheritance of one of these RFLPs was examined among the progeny from an *Hn-n* × *D 176* hybrid which had been derived by self-fertilizing single isolated gametophytes. Theoretically, these should have yielded sporophytes whose genetic constitution was identical to their progenitor gametophytes. This process was necessary to obtain sufficient material for DNA isolation and manipulation. Two important genetic consequences result from this mating scheme. First, the gametic output from the F_1 hybrid was sampled, thus the expected genetic segregation ratio for RFLP loci was 1 : 1. Second, since a single gametophyte was selfed in isolation through intra-gametophytic selfing, all loci were expected to be homozygous. The presence of two RFLPs in a single plant, therefore, could not be due to heterozygosity but must arise from two distinct loci.
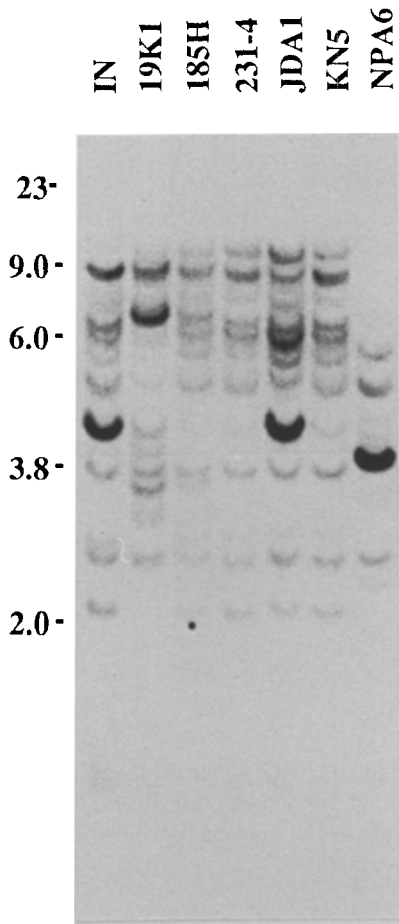
Fig. 1. Differences in signal intensity between the 4.5 kb fragments in different *Ceratopteris* accessions shown with probe 3-31-13. Note the high intensity in accessions *IN* and *JDA 1* and low intensity in *185 H* and *231-4*

We examined the segregation of two RFLP fragments (4.5 kb and 7 kb) detected with the cDNA insert from clone 1-20-19 in a population of 22 "F_2" progeny (Fig. 2). A total of 28 "F_2" plants were tested, but RFLP patterns of six plants were weak and could not be reliably scored (e.g., see lane marked with an open arrow in Fig. 2). The 4.5 kb fragment appeared as an RFLP between the parents, and its observed segregation ratio of this 4.5 kb fragment (10 present and derived from *D 176* to 12 absent) was not significantly different from the expected 1 : 1 ratio ($\chi^2 = 0.19$; p = 0.67). A second fragment of 7 kb was also segregating in this population. This observation is remarkable since this fragment was present in each parent and the F_1 hybrid, and was not expected to segregate in this material. Inheritance of this 7 kb fragment did not fit an expected 1 : 1 ratio, but rather its segregation was distorted (17 present: 5 absent; χ^2 for 1 : 1 ratio = 6.55; p = 0.01). Another RFLP fragment at 3.5 kb in strain *Hn-n* was observed, but its intensity was weak in most "F_2" plants and was not scored. Thus, at least two RFLP fragments were segregating in this population, and they likely represent separate loci since five "F_2" progeny co-inherited both fragments.

Discussion

Two important questions of homosporous fern genome organization are: (1) Is the level of sequence duplication in diploid homosporous ferns greater than that of
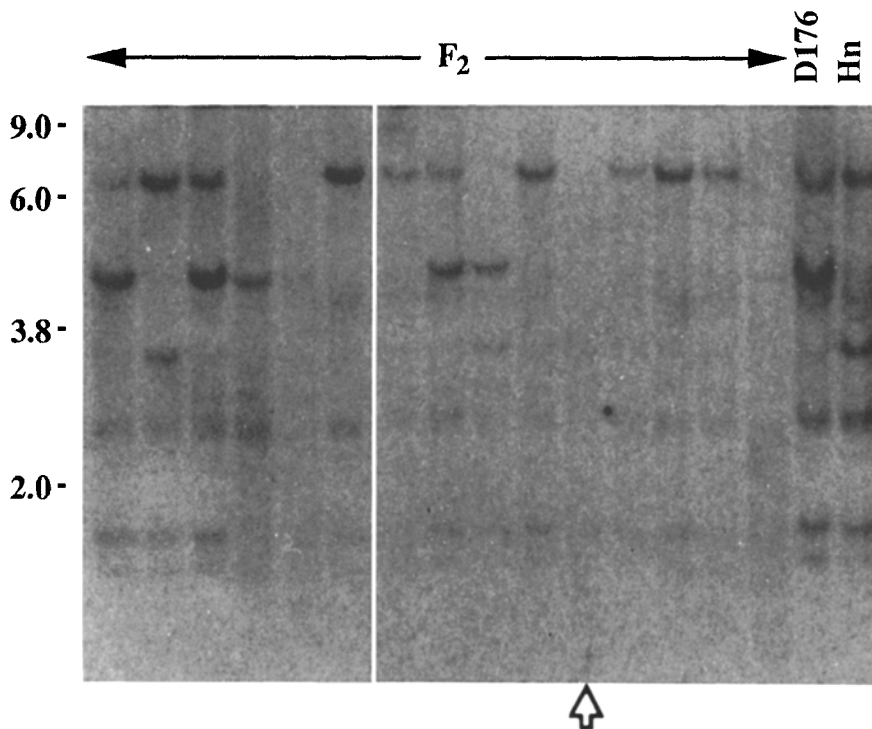
Fig. 2. Segregation at two RFLP loci (4.5 and 7 kb fragments) in *C. richardii* using probe 1-20-19. This image was obtained after a 20 day exposure, which revealed additional bands not observed in shorter exposures (Table 2) and also resulted in the high background signal. For arrow explanation see text

other diploid species, such as those within the flowering plants? (2) Is the level of sequence duplication in a neo-tetraploid species double that of a related diploid species?

Ceratopteris richardii appears to carry more duplicated sequences than diploid angiosperms. For example, a genetic analysis of sequences similar to cDNA clones in *Arabidopsis thaliana*, a small-genome sized crucifer, indicates that c. 15% of its genes are duplicated (MCGRATH & al. 1993), a level similar to other angiosperms as well [e.g., tomato (BERNATZKY & TANKSLEY 1986) and lettuce (LANDRY & al. 1987)]. In *C. richardii*, our evidence suggests that > 50% of the sequences recognized by cDNA clones are duplicated. This high level suggests that polyploidization was the process that created the gene duplications in *C. richardii*.

Firmer estimates of the level and distribution of sequences similar to expressed genes in homosporous ferns will require application of more probes and a further investigation of the inheritance of these sequences. Here we have demonstrated that some unusual results can be expected by examining the inheritance of RFLP fragments in *C. richardii*. For instance, one locus showed an unusual pattern of inheritance in that it was present in both parents and the hybrid, but segregated in the "F_2" progeny. The segregation ratio of this locus is consistent with models of tetrasomic inheritance involving pairing between homeologous chromosomes. In the simplest case, one pair of chromosomes would carry the observed fragment

and a homeologous chromosome pair would not. Unrestricted pairing and seg-regation between these four chromosomes would result in a gametic ratio of 5:1 (present vs. absent, respectively). This ratio is not significantly different from the 17:5 ratio we observed. Homeologous pairing and segregation has been previously reported in *Ceratopteris* (HICKOK 1978).

Our results also indicate that the level of sequence duplication is 30% higher in the neo-tetraploid species *C. thalictroides* relative to *Ceratopteris* diploids. One interpretation of this result is that *C. thalictroides* is an autotetraploid, perhaps resulting from intra- or inter-gametophytic fusion of 2n gametes. Although each locus would in fact be duplicated if this were the case, their detection would not be possible because fragments from the same locus would overlap. The 30% increase in fragment number over *C. richardii* may be due to divergence of duplicate se-quences after polyploidization in *C. thalictroides*. Alternatively, *C. thalictroides* may be an amphidiploid (allotetraploid) between *C. richardii* and another species whose genetic complement shares little sequence similarity with *C. richardii*. Since synthetic hybrids can be made between *C. richardii* and the two other described diploid *Ceratopteris* species (HICKOK & KLEKOWSKI 1974, HICKOK 1977), and *Ceratopteris* shares little affinity with other genera, this latter alternative seems unlikely.

Differences in signal intensity between accessions which share fragments of the same size were frequently encountered in comparing among accessions of *C. thal-ictroides*. These results indicate that these presumed homologous fragments have diverged in sequence, and may indicate that the less intense fragments are silenced copies. This interpretation is supported by pairs of shared fragments which show reciprocal intensity differences, and possibly are indicative of reciprocal gene si-lencing, as suggested by WERTH & WINDHAM (1987). This phenomenon was not seen in *C. richardii* perhaps because their accessions were drawn from geographical regions of close proximity. However, it may also be due to the diploid nature of *C. richardii*, such that the majority of duplicate genes have been anciently silenced and contemporary evidence of this process may only be detected in neopolyploids.

In sum, our data suggest that the genome of *Ceratopteris* contains multiple sequences similar to each cDNA tested. This interpretation is compatible with the hypothesis that homosporous ferns have a polyploid origin. The variation in hy-bridization signal seen with some probes suggests that some sequences have diverged since the original genome duplication event. Whether all observed fragments contain active coding loci is as yet unknown, but results from isozyme studies (HAUFLER & SOLTIS 1986) and the examination of CAB genes (PICHERSKY & al. 1990) suggest that many duplicate sequences are inactive in diploid species. Assuming that hy-bridization signals with the highest intensity represent the active gene, reciprocal differences in intensity between taxa which share same-sized bands may indicate that some duplicate copies have been reciprocally silenced. This process may proceed more rapidly than previously considered because the neo-tetraploids, but not the diploids, show these reciprocal differences. Finally, RFLP analysis may provide additional evidence for homeologous pairing among homosporous fern chromo-somes, and its potential to generate genetic diversity.

References

BERNATZKY, R., TANKSLEY, S. D., 1986: Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. — Genetics **112**: 887–898.

GREEN, B. R., PICHERSKY, E., KLOPPSTECH, K., 1991: Chlorophyll *a/b*-binding proteins: an extended family. − Trends Biochem. Sci. **16**: 181–186.

HAUFLER, C. H., 1987: Electrophoresis is modifying our concepts of evolution in homosporous pteridophytes. − Amer. J. Bot. **74**: 953–966.

− SOLTIS, D. E., 1986: Genetic evidence that homosporous ferns with high chromosome numbers are diploid. − Proc. Natl. Acad. Sci. USA **83**: 4389–4393.

HICKOK, L. G., 1977: Cytological relationships between three diploid species of the fern genus *Ceratopteris*. − Canad. J. Bot. **55**: 1660–1667.

− 1978: Homeologous chromosome pairing and restricted segregation in the fern *Ceratopteris*. − Amer. J. Bot. **65**: 516–521.

− 1979: A cytological study of intraspecific variation in *Ceratopteris thalictroides*. − Canad. J. Bot. **57**: 1694–1700.

− KLEKOWSKI, E. J., 1974: Inchoate speciation in *Ceratopteris*: an analysis of the synthesizes hybrid *C. richardii* × *C. pteridoides*. − Evolution **28**: 439–446.

− WARNE, T. R., SLOCUM, M. K., 1987: *Ceratopteris richardii*: applications for experimental plant biology. − Amer. J. Bot. **74**: 1304–1316.

KLEKOWSKI, E. J., 1970: Reproductive biology of the *Pteridophyta*. IV. An experimental study of mating systems in *Ceratopteris thalictroides* (L.) BRONGN. − Bot. J. Linn. Soc. **63**: 153–169.

LANDRY, B. S., KESSELI, R. V., FARRARA, B., MICHELMORE, R. W., 1987: A genetic map of lettuce (*Lactuca sativa* L.) with restriction length polymorphism, isozyme, disease resistance and morphological markers. − Genetics **116**: 331–337.

MCGRATH, J. M., HICKOK, L. G., PICHERSKY, E., 1994: Restriction fragment length polymorphisms distinguish among accessions of *Ceratopteris thalictroides* and *C. richardii* (*Parkeriaceae*). − Pl. Syst. Evol. **189**: 193–202.

− JANCSO, M. M., PICHERSKY, E., 1993: Duplicate sequences with similarity to expressed genes in the genome of *Arabidopsis thaliana*. − Theor. Appl. Genet. **86**: 880–888.

PICHERSKY, E., SOLTIS, D., SOLTIS, P., 1990: Defective chlorophyll *a/b*-binding protein genes in the genome of a homosporous fern. − Proc. Natl. Acad. Sci. USA **87**: 195–199.

SCOTT, R. J., HICKOK, L. G., 1987: Genetic analysis of antheridiogen sensitivity in *Ceratopteris richardii*. − Amer. J. Bot. **74**: 1872–1877.

WERTH, C. R., WINDHAM, M. D., 1987: A new model for speciation in polyploid pteridophytes resulting from reciprocal silencing of homeologous genes. − Amer. J. Bot. **74**: 713.

Addresses of the authors: J. MITCHELL MCGRATH, Department of Plant Pathology, 1630 Linden Dr., University of Wisconsin, Madison, WI 53706, USA. − ERAN PICHERSKY, Department of Biology, University of Michigan, Ann Arbor, MI 48109-1048, USA. − LESLIE G. HICKOK, Department of Botany, University of Tennessee, Knoxville, TN 37996, USA.