The evolution of chloroplast genome structure in ferns

Paul G. Wolf, Jessie M. Roper, and Aaron M. Duffy

Abstract: The plastid genome (plastome) is a rich source of phylogenetic and other comparative data in plants. Most land plants possess a plastome of similar structure. However, in a major group of plants, the ferns, a unique plastome structure has evolved. The gene order in ferns has been explained by a series of genomic inversions relative to the plastome organization of seed plants. Here, we examine for the first time the structure of the plastome across fern phylogeny. We used a PCR-based strategy to map and partially sequence plastomes. We found that a pair of partially overlapping inversions in the region of the inverted repeat occurred in the common ancestor of most ferns. However, the ancestral (seed plant) structure is still found in early diverging branches leading to the osmundoid and filmy fern lineages. We found that a second pair of overlapping inversions occurred on a branch leading to the core leptosporangiates. We also found that the unique placement of the gene *matK* in ferns (lacking a flanking intron) is not a result of a large-scale inversion, as previously thought. This is because the intron loss maps to an earlier point on the phylogeny than the nearby inversion. We speculate on why inversions may occur in pairs and what this may mean for the dynamics of plastome evolution.

Key words: Osmunda, Gleichenia, Lygodium, Vandenboschia, Dicksonia, Marsilea, Adiantum, genome evolution, inversion.

Résumé : Le génome plastidique (plastome) est une riche source de données phylogénétiques et autres chez les plantes. La plupart des plantes terrestres possèdent un plastome de structure semblable. Cependant, au sein d'un groupe majeur de plantes, les fougères, a évolué un plastome à la structure unique. L'ordre des gènes chez les fougères a été expliqué par une série d'inversions génomiques par rapport à l'organisation du plastome rencontré chez les plantes à graines. Dans ce travail, les auteurs examinent pour la première fois la structure du plastome à travers tout le groupe des fougères. Les auteurs emploient une stratégie PCR pour cartographier et partiellement séquencer les plastomes. Les auteurs ont observé qu'une paire d'inversions se chevauchant partiellement dans la région inversée répétée était présente chez l'ancêtre commun à la majorité des fougères. Cependant, la structure ancestrale (des plantes à graines) est encore observée chez les premières branches à diverger du tronc, lesquelles mènent aux osmundacées et aux hymenophyllacées. Une seconde paire d'inversions chevauchantes est présente dans l'embranchement qui mène aux principales fougères leptosporangiées. Les auteurs ont également trouvé que la position unique du gène *matK* chez les fougères (dépourvu d'un intron flanquant) n'est pas le résultat d'une inversion de grande taille comme cela a été suggéré antérieurement. Elle découle plutôt du fait que la perte de l'intron serait survenue plus tôt dans la phylogénie que l'inversion voisine. Les auteurs spéculent sur les causes possibles de l'occurrence des inversions en paires et ce que cela signifierait pour la dynamique de l'évolution des plastomes.

Mots-clés : Osmunda, Gleichenia, Lygodium, Vandenboschia, Dicksonia, Marsilea, Adiantum, évolution des génomes, inversion.

[Traduit par la Rédaction]

Introduction

Comparative biology, including phylogenetics, relies on differences among evolutionarily homologous structures. Generally, the deeper in time one delves, the harder these homologies are to recognize. One of the few suites of characters that transcend the entire green plant clade is the plastid (chloroplast) and its constituent genome (the plastome), which has been known for some time to be an ideal source of evolutionarily informative characters (Palmer 1985). In contrast to the nuclear genome, the plastid genome is small and contains a high proportion of DNA that codes for proteins and RNAs of known function. Over the last 20 years plastomes have provided a wealth of phylogenetic information. The result has been a greatly improved understanding of the green plant tree of life (Palmer et al. 2004). Early research focused on the structure of the plastid genome by mapping gene order with restriction mapping techniques (Raubeson and Jansen 1992). More recently, improved sequencing methods have shifted the focus away from plastome structure to the comparison of sequence data for selected plastid genes. The combination of sequence data with phylogenetic estimation methods is a well-established tool for inferring plant phylogeny. However, there are limits to the power of sequence data. For example, in attempts to resolve very ancient lineages, the power of sequence data is

P.G. Wolf,[1] J.M. Roper, and A.M. Duffy. Department of Biology, and Ecology Center, Utah State University, 5305 Old Main Hill, Logan, UT 84322, USA.

[1]Corresponding author (e-mail: wolf@biology.usu.edu).

diminished for several reasons. One problem is that there are only four nucleotides, so the probability of parallel and convergent substitutions increases rapidly as one goes back in time. Thus, inference of deeper branches can be enhanced by different types of markers, possibly those that do not evolve in a clock-like manner (Olmstead and Palmer 1994; Helfenbein and Boore 2004). It is proposed that complex genomic rearrangements resulting in a new gene order could represent such markers (Helfenbein and Boore 2004). However, genome rearrangements may not be free of homoplasy. Physical hotspots are known to occur on the plastid genome (Morton and Clegg 1993), and temporal destabilization might also be a possibility. It is encouraging that single inversion events have been used as reliable molecular markers. One classic example is a 30 kilobase (kb) inversion detected in the large single copy (LSC) of the bryophytes and lycophytes, relative to other land plants, indicating that lycophytes are a sister to all other extant vascular plants (Raubeson and Jansen 1992). With this example in mind it seems reasonable to assume that other, more complex changes in gene order are also candidates for use as phylogenetic markers. However, the apparent rarity of complex gene rearrangements makes it difficult to study the nature and utility of such events. Conveniently, the inverted repeat (IR) region of fern plastomes displays a reorganized structure that is best explained by several overlapping inversions (Hasebe and Iwatsuki 1990, 1992). Previous work based on morphology and DNA sequence data (Pryer et al. 2004; Schuettpelz et al. 2006) has resulted in what is believed to be a robust phylogeny on which these inversion events can be mapped (Fig. 1). In addition, at the time we designed this study, two complete fern plastomes (*Adiantum capillus-veneris* and *Angiopteris evecta*) were available for comparison. Since then, a third (*Alsophila spinulosa*) has been published (Gao et al. 2009).

The plastome of the marattioid fern *Angiopteris evecta* has similar size, gene order, and gene content to that of tobacco and other seed plants. In comparison, the plastid genome sequence of the polypod fern *Adiantum capillus-veneris* is highly rearranged within the inverted repeat (Hasebe and Iwatsuki 1990, 1992; Wolf et al. 2003). This rearranged gene order is believed to be the result of two main overlapping inversions (Hasebe and Iwatsuki 1992; Roper et al. 2007). Because these inversions span the IR, LSC, and SSC (small single copy) regions, it is possible to infer the end points and the sizes of the inversions as well as the order in which they occurred. The first of these was approximately 18 kb and the second was approximately 21 kb (Wolf et al. 2009). Mapping studies of *Polystichum acrostichoides* and *Cyathea furfuracea*, representative of two major branches within the core leptosporangiate ferns (sensu Pryer et al. 2004), indicate they share the rearranged gene order found in *A. capillus-veneris* (Stein et al. 1992; Raubeson and Stein 1995). This suggests the core leptosporangiates all have approximately the same gene order as *A. capillus-veneris*. However, the authors of these earlier studies designed their projects before a well-resolved fern phylogeny was available and, as a result, their taxon sampling was inadequate for several of the early-branching clades.

Our objective in this study was to map the IR region of a representative species of each major branch within the lepto-

sporangiate ferns (Fig. 1). We also examined a separate smaller series of inversions that cover an approximately 3.3 kb region in the LSC of fern plastomes. Plastomes contain regions that are commonly used for phylogenetic analyses (Graham and Olmstead 2000) as well as barcoding studies at a wide range of phylogenetic levels (Chase et al. 2007). Understanding gene order is necessary for designing functional primers, successful amplification, and DNA sequencing. For example, the inversions in the region of the IR in ferns appear to have disrupted the gene order in the region around the gene *matK* (Duffy et al. 2009). This gene has been targeted for use in bar coding (Chase et al. 2007), but the standard primers fail in ferns because of their unique plastome structure. Here our goal is to characterize plastome structure and its effect on gene placement in all major lineages of ferns.

By examining variation in genome structure in a phylogenetic context, we explore the phylogenetic utility and test the levels of homoplasy of complex genome inversion events. We ask whether the putative inversions in the fern plastid genome occurred on different branches of the fern phylogeny, in which case they are phylogenetically useful and interpretable. Alternatively, if these inversion events were clustered on the same branch, it would suggest that plastid genomes go through periods of high stability and destabilization, indicating that some structures and gene orders may be more stable than others. Such findings would have implications for both phylogenetic studies and our understanding of plastid, and ultimately nuclear, genome evolution.

Materials and methods

We used a PCR-based approach to map the overall structure of the plastome for each sample. Two sets of primers are needed to detect differences in gene orders that are hypothesized to differ by an inversion. The first primer set consists of the "anchored" primers, which are designed within genes close to, but not moved by, a putative inversion. The second primer set consists of "inversion" primers, designed within regions that are potentially moved by the inversion, and as close to the inversion boundaries as possible (Fig. 2). To test for gene order, each anchored primer is combined with each inversion primer. The combination of working and failing PCR amplifications will indicate which gene order is present. PCR results are confirmed by end sequencing of the PCR products. This approach has been used successfully to examine the distribution of a large inversion in moss plastomes (Sugiura et al. 2003). The availability of the complete plastome sequences of *A. evecta* and *A. capillus-veneris* enabled us to design PCR primers, some of which are universal for leptosporangiate ferns. A list of the main primers designed and used in this study is provided in Table 1.

Taxon sampling

We included representatives from each major lineage of the leptosporangiate ferns. The gene order was determined for a region within the LSC, within the IR, and at the LSC–IR and SSC–IR boundaries of *Osmunda* (= *Osmundastrum*) *cinnamomea*, *Vandenboschia radicans*, *Gleichenia japonica*,

Fig. 1. Fern cladogram based on several sources (see text). Clade names follow Pryer et al. (2004), and representative taxa used in the current study are indicated in parentheses. Specific clades and branches mentioned in the text are labelled.
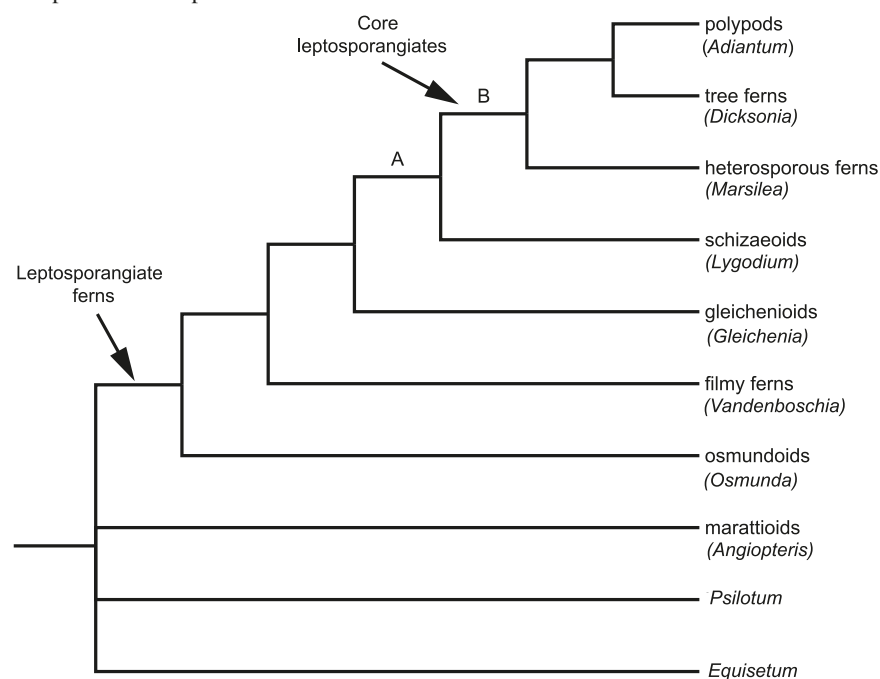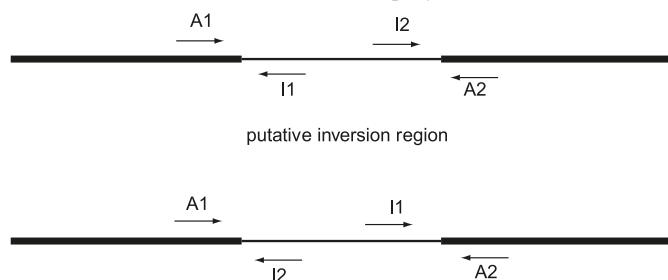

Fig. 2. Strategy and placement of primers to detect alternative genomic organizations in a putative inversion region. The anchored primers (prefix A) are outside the inversion region and the inversion primers (prefix I) are within the putative inversion. With one alternative pattern, primer pairs A1, I1 and I2, A2 should amplify, with other combinations failing. With the alternative inversion structure, A1, I2 and I1, A2 should amplify.

Lygodium japonicum, *Marsilea mutica*, and *Dicksonia antarctica* (Fig. 1; Table 2). DNA from *Angiopteris evecta* and *Adiantum capillus-veneris* was used to test primers and as positive and negative controls. Species and voucher information is available in Table 2.

PCR and sequencing

PCR amplification was performed using the Proofstart long PCR system (QIAGEN, Valencia, California, USA) consisting of a combination of QIAGEN *Taq* DNA Polymerase and QIAGEN Proofstart DNA Polymerase. We adjusted the extension time to approximately 1 min/kb. Length estimates were based on the distance between primers within the *A. capillus-veneris* or *A. evecta* genome, depending on the hypothesized gene order. For unknown arrangements we used a standard extension time of 8 min. Expected gene content was verified in each fragment by end sequencing.

Results

Here we focused on two regions of the fern plastome (one associated with the IR and one within the LSC) that we knew were variable in gene order. Our goal was to characterize the evolutionary changes. We compiled six new sequences totalling 136 932 bp. GenBank accession numbers for the new sequences are provided in Table 2. We sequenced the ends of most PCR fragments to verify inferences from primer combination results. Thus, in some regions we have sequence data, whereas in other regions we infer the gene order from the success or failure of different primer combinations as well as examination of the size of the PCR fragment. Figure 3 shows the inferred gene order for all taxa we sampled here as well as for the previously sequenced *A. evecta* and *A. capillus-veneris* plastid genomes. We denote regions inferred by PCR, those determined by partial sequencing of a feature, and those determined by complete sequencing of a feature. We also refer to a previous mapping study to infer some of the structure in *O. cinnamomea* (Stein et al. 1986). We first describe chloroplast genome structure in the IR region and then describe patterns with respect to the smaller inversions in the LSC.

By examining genome structure and referring to the phylogeny represented in Fig. 1, we are able to infer the order of genomic changes that gave rise to the unique gene order in the IR region of most ferns. *Angiopteris evecta*, *O. cinnamomea*, *V. radicans* (a filmy fern), and *G. japonica* (a gleichenioid fern) all have the same gene order in both regions examined. All the aforementioned taxa have essentially the gene order seen in most seed plants, which is ancestral in the context of the ferns. *Lygodium japonicum* represents the next major branch of the fern phylogeny: schizaeoid ferns. This taxon appeared to have both inver-

Table 1. Major PCR primers used.

Primer	Sequence
chlB-F	GAAGTTGSKGATACGATAGCTCGTA
chlL-F	CATCTTCGGGCCAWACATCTTCATA
matK-25R	CGTGTTGTACGGGGRTGTTTACARGC
matK-26R	GTTTACAGGCYAARGTTTTATCGC
matK-27F	TCAGCTCGGACTACTTTAACAGATG
matK-32F	ACAAGTCCGAGATGTTTCGTTC
ndhB-19F	GAACAGGATCCTTCATGTATAATCCG
ndhB-4R	GATCGAATGAAGCAACTACG
ndhB-9F	TTRTCAAGCTGAAGYTGTCC
ndhB15-F	CAGCTTTCCATCCATGCCAAAAGAG
ndhB20-F	GCGGGATTATACATGAAGGATCCTG
ndhF-2F	GGGAGTTAGTTGGAATGTGTT
ndhH-aF	TGGACGAATTTTCMATCTCCAAGGA
ndhH-bF	GGATGMTGAGGACCCATRCTTACTA
psbA-F	ACTAGAAGTTACCAAAGAACCATGC
psbA-R	GCATGGTTCTTTGGTAACTTCTAGT
rpl2-4R	GAATGCCCTACCCTTGAGTG
rpl2-aF	CATGKGGATGATCSACGGGATTCAT
rpl2-F	TCARAGGTAAGGCATTTCCTSYTGA
rpl2-R	TCARSAGGAAATGCCTTACCTYTGA
rpl32-F	TGGCMGTTCCGAARAAACGTACTTC
rpl32-R	GAAGTACGTTTYTTCGGAACKGCCA
rps16-F	TGATGCCTAATCGGGAGGTA
rps7-F	GAYYTTAAAGCCATACTTCGACCTG
rps7-R	CAGGTCGAAGTATGGCTTTAARRTC
rrn16-F	ACTGGAGTGAAGTCGTAACAAGGTA
rrn16-R	TACCTTGTTACGACTTCACTCCAGT
rrn5-F	ATCYCGAACTYGGTGGTGAAACTCT
rrn5-R	AGAGTTTCACCACCRAGTTCGRGAT
trnL-4F	TGGTGAACGATCTCCATCTG
trnL-F	TACCRTTTCACCATCAAGGC
trnL-R	GCCTTGATGGTGAAAYGGTA
trnN-F	CTGTTAACCGATTGGTCGTAGGTTC
trnN-R	GAACCTACGACCAATCGGTTAACAG
trnR-F	AGAGCACGTGGCTACGAACCACGGT
trnR-R	ACCGTGGTTCGTAGCCACGTGCTCT
ycf2-7F	GTCTCCTTGTATAGTATGGATTCA
ycf2-8R	CAAAGTAACAATTTTGGATCAGCTTC

Note: Additional specific sequencing primers were also used.

sions that we inferred in the region of the IR (Fig. 3). Thus, the IR region in *L. japonicum* appears similar to that of *A. capillus-veneris*, as it is in all remaining taxa sampled in our study (differing only in whether the 3′ end of *ndhF* is included in the IR). We have found no representative of any lineage that has only the first inversion, which would result in the "intermediate 1" gene order (Figs. 1 and 3).

Near the *petN* gene in the LSC is another region that varies in gene order among ferns. The inverted region is approximately 3.3 kb and two inversion steps are required to explain the differences between *A. evecta* and *A. capillus-veneris*. However, in this case there is no association with the ends of the IR, so that we cannot infer a single intermediate gene order; two are possible (Fig. 3). We observed the ancestral (*A. evecta*) gene order in *O. cinnamomea*, *V. radicans*, *G. japonica*, and *L. japonicum*. The *A. capillus-veneris* gene order was observed in all other taxa, and no sample had an "intermediate 2" gene order.

The gene *matK* encodes a maturase protein involved in splicing of group II introns (Ems et al. 1995; Vogel et al. 1997). In most land plants this gene sits within a group II intron of *trnK* in the LSC. However, in most ferns both *trnK* and its intron are lacking, but *matK* is present (Wolf et al. 2003). It has been assumed that the end point of the first of the large inversions involving the fern IR was responsible for the breakage and subsequent loss of *trnK* and its intron (Duffy et al. 2009). Here we examined sequence data from ferns with and without the large inversion for evidence of *matK*, *trnK*, the *trnK* intron, or remnants thereof. In *A. evecta* and *O. cinnamomea* the IR is oriented as in seed plants, and *matK* and *trnK* (with intron) are situated also as in seed plants. In *L. japonicum*, *M. mutica*, and *D. antarctica* the IR is reorganized as in *A. capillus-veneris*, and *trnK* and its intron are absent. Interestingly, however, we did not find *trnK* or its intron in *V. radicans* and *G. japonica*, even though these taxa have the same gene order as in seed plants (i.e., without the large inversions seen in the core leptosporangiate ferns). Our searches for *trnK* and its intron were conducted as follows. (1) Using sequences from *O. cinnamomea*, *V. radicans*, and *G. japonica*, we performed BLAST searches (with an expect threshold of 10 and wordsize of 11) of the region that should contain *trnK* against the *A. evecta* plastome. From this we identified all expected genes (*psbA*, *trnK* [including its intron], *matK*, *rps16*, and *chlB*) in *O. cinnamomea*. For *V. radicans* and *G. japonica* we got no BLAST hit for *trnK* and its intron. (2) We excised just the *matK–psbA* spacer and the *matK–rps16* spacer from *O. cinnamomea*, *V. radicans*, and *G. japonica*, and BLASTed each against the NCBI (nr database) and the plastomes of *A. evecta*, *A. capillus-veneris*, and *Psilotum*, getting no hits except for *O. cinnamomea*. (3) We also BLASTed 39 bp of the *trnK* intron domain 5 from *Cycas panzhihuaensis*, *Atropa belladonna*, *Sphagnum platyphyllum*, *Pellia borealis*, *Pinus thunbergii*, *Marchantia polymorpha*, *A. evecta*, and *Psilotum nudum* against the *matK–psbA* spacer. The nucleotide sequence of domain 5 is well conserved among land plants and identified the *trnK* intron in *A. evecta*, *O. cinnamomea*, and other taxa where it is present, but we could not detect the domain in *V. radicans* or *G. japonica*. (4) We BLASTed the *trnK* intron and exons from *A. evecta* against the *psbA–chlB* region and had significant hits in *O. cinnamomea*. The only one that returned any hits in *G. japonica* was the intron containing domains 1–4, but they were all 11-base hits with $E > 0.18$. This was repeated for *V. radicans*, which had the same weak hits for domains 1–4. There were also two weak (13 bp $E = 0.013$ and 12 bp $E = 0.051$) hits for domains 5–6 that are in the right area but in the wrong orientation and not spaced correctly relative to each other. One of them also hit in a second spot between *rps16* and *chlB*. These are AT-rich matches and probably are not meaningful, though it is possible that they represent degraded remnants of what was once a *trnK* exon. Thus, we infer that *trnK* and its intron are absent from both *V. radicans* and *G. japonica*, even though these taxa have the ancestral gene order usually associated with having functional *trnK* with an intron that contains *matK*. Plastomes of both these taxa contained *matK*, but it was not embedded in the *trnK* intron as it is in *A. evecta*, *O. cinnamomea*, and most other land plants.

Table 2. Taxon sampling information.

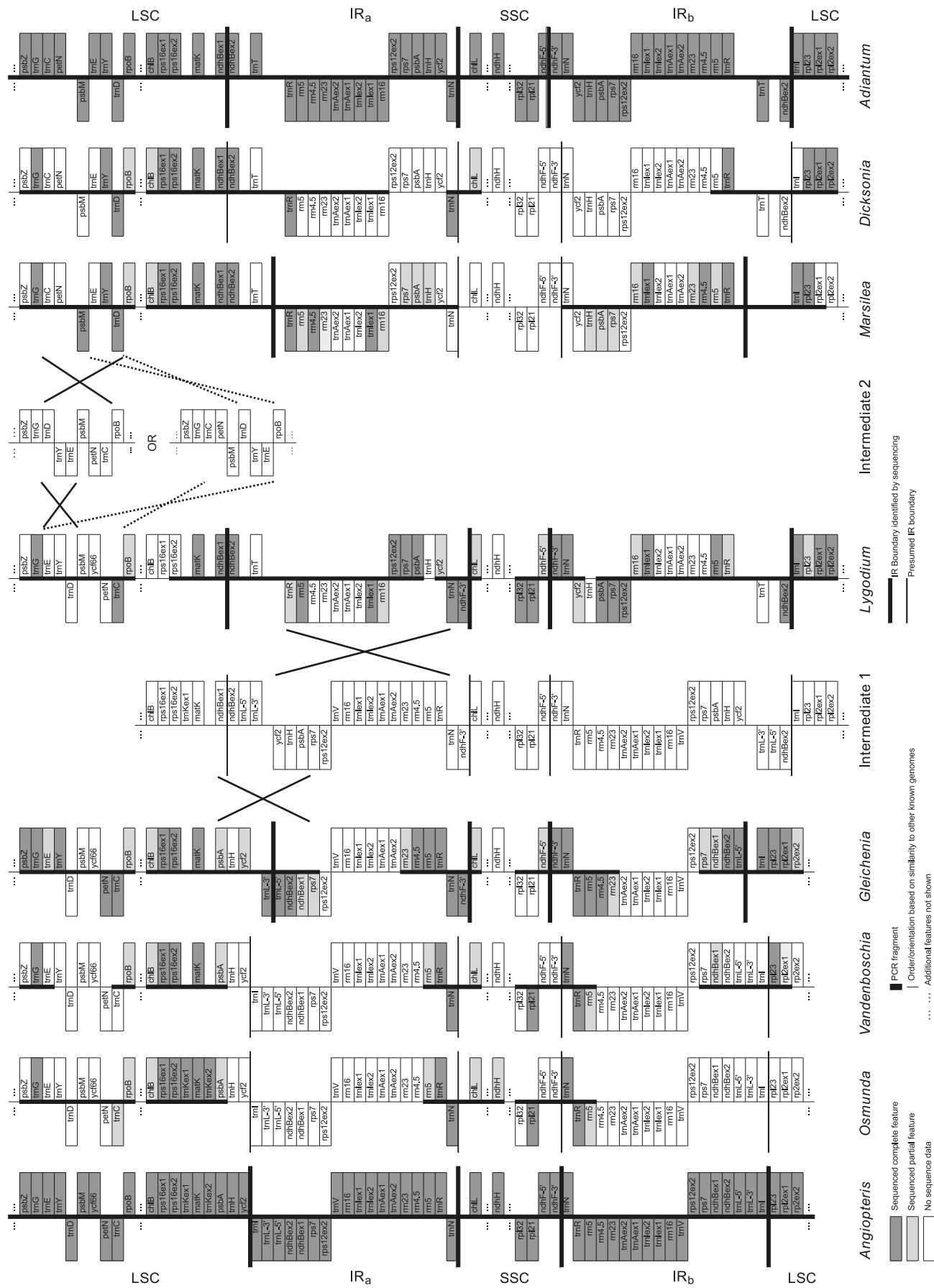| Lineage	Species	Collection information	Country	Accession No.	GenBank accession No.	Sequence (bp)
Marattioid ferns	*Angiopteris evecta*	Culture; Seattle, USA	USA	UC 1794629	DQ821119	Complete
Osmundaceous ferns	*Osmunda cinnamomea*	Wolf 745	USA	UC 1728349	HM021799	14 049
Filmy ferns	*Vandenboschia radicans*	A. Ebihara 051105-01	Japan	UTC 244420, also at TNS	HM021800	15 520
Gleichenioid ferns	*Gleichenia japonica*	M. Hasebe s.n.	Japan	UTC 244418	HM021798	28 958
Schizaeoid ferns	*Lygodium japonicum*	M. Hasebe s.n.	Japan	UTC 244417	HM021803	39 046
Heterosporous ferns	*Marsilea mutica*	Wolf 864 – and UC Botanical Garden	USA	UCBG97.0535	HM021801	24 134
Tree ferns	*Dicksonia antarctica*	A.R. Smith 2911	USA	UC	HM021802	15 225
Polypod ferns	*Adiantum capillus-veneris*				AY178864	Complete

Discussion

Ferns represent a major, diverse clade of land plants with approximately 11 000 species (Pryer et al. 2004). A clade that includes the ferns and horsetails is the sister group to the seed plants (Pryer et al. 2001). Thus, any comparative data on ferns are useful for studies within ferns and horsetails but are also potentially useful for studies of seed plants. Here we examined structural variation for the plastome of several major fern clades. We detected differences in gene order, which should provide useful information for botanists focusing on smaller groups within the clades we sampled. Below we discuss the implications of our results for understanding the dynamics of plastome evolution.

Our analysis of plastid genome structure in ferns indicates that two large and overlapping inversions (18 kb and 21 kb, respectively) both occurred on branch A of the fern phylogeny depicted in Fig. 1. Although these inversions were inferred long ago (Hasebe and Iwatsuki 1990), the phylogenetic placement was not known. We show here that the schizaeoid ferns, heterosporous ferns, tree ferns, and polypods all have the unusual gene order in the inverted repeat whereby the order of ribosomal genes is reversed relative to the order in other land plants. This group of ferns includes about 95% of fern species. The second region that we investigated, a 3.3 kb region of the LSC, also involved two inversions, both of which occurred on branch B (Fig. 1). Thus, we have two examples of major plastome reorganizations in which two overlapping inversions occurred on the same branch of the fern phylogeny. This suggests that there may be some form of selection pressure to maintain the strandedness of a particular gene. Once an inversion occurs, all genes within the region are transcribed from a different strand of the DNA. If certain transcript arrangements are more stable or functional, then there could be a disadvantage to the new organization, or an advantage to reverting to the old gene order. Thus, if inversions are often followed by a reversion, most will go undetected. Under this model, occasionally the reversion is not entirely aligned to the original inversion, resulting in a detectable reorganization. This hypothesis is supported by studies of suites of adjacent genes in plastomes indicating that strandedness is under selection (Cui et al. 2006). One possibility is that the position of plastome promoters may constrain structural evolution. An implication for overall plastome evolution is that most structures should be evolutionarily stable, which explains why *A. evecta* and tobacco have a very similar gene order despite being separated by more than 700 million years of evolutionary time (Pryer et al. 2004). However, if a lineage survives a change that destabilizes the structure, then it will quickly be followed by an event that restabilizes the structure, or the plastome may undergo very rapid structural evolution, as seen in several clades of angiosperms (Chumley et al. 2006; Haberle et al. 2008).

The curious case of the *trnK* intron

In land plants (except most ferns), *matK* is situated within the intron of another gene: *trnK*. Although a gene within another gene's intron is unique among land plant plastome genes, it is fairly typical for this class of intron (group II) in bacterial genomes and organellar genomes of most eukar-

Fig. 3. Plastome maps of all major lineages of ferns showing gene order in the region of the IR and flanking regions of the LSC and SSC. An additional region of the LSC is shown at the top. Hypothesized intermediate structures are also shown.

yotes (Hausner et al. 2006). Thus, the situation in ferns, whereby the maturase gene has lost its flanking intron, is the more unusual case (Duffy et al. 2009). Previously, we had assumed that loss of the *trnK* intron in ferns was caused by the first of the large plastome inversions, whereby the end point disrupted the *trnK* intron (Wolf et al. 2003; Duffy et al. 2009). However, our analysis here shows rather that the *trnK* intron had been lost prior to the inversion on the fern phylogeny. This raises the possibility that it was the earlier loss of the *trnK* intron, or *trnK* itself, that may have eventually allowed for the first large inversion. Regardless, we have shown here that the presence of *trnK* and its intron in ferns is independent of the large inversions associated with the IR. Furthermore, all ferns (except the small marattioid and osmundoid lineages) appear to lack *trnK* and its intron.

One reason that *matK* has been proposed as an ideal gene for barcoding is its high level of sequence variability. However, this makes it difficult to design PCR primers within the gene itself. This problem is circumvented by using primers within the more conserved flanking *trnK* gene in seed plants. This strategy cannot work in ferns, but now with our compilation of *matK* sequences from all major fern lineages, it may be possible to design primers, or alternative strategies, to characterize the gene in more fern taxa and perhaps eventually extend barcoding applications to ferns.

Inferring phylogenetic relationships among lineages that diverged hundreds of millions of years ago is inherently challenging. Although genomes contain much phylogenetic signal at the sequence level, it has been hoped that structural changes can provide better evidence than sequence data, because genomic rearrangements should be less susceptible to saturation and homoplasy. Our findings indicate some reason for optimism but there are limitations. By using a taxon sampling within a phylogenetic framework, we show that plastome inversions do not contradict phylogenies based on other types of data. However, we find evidence that plastomes undergo evolutionarily stable and destabilized periods — temporal hotspots — that are analogous to the spatial hotspots for inversions found in certain regions of genomes. This non-clocklike evolutionary tempo is in fact what is needed to characterize short but ancient phylogenetic branches. However, the rarity of such events means that an element of luck may be needed to detect a structural change on such a branch. The current rapid growth of complete genome sequences may provide an opportunity to test whether temporal hotspots are common in nuclear genomes of eukaryotes.

Acknowledgements

References

Chase, M.W., Cowan, R.S., Hollingsworth, P.M., van den Berg, C., Madrinan, S., Petersen, G., et al. 2007. A proposal for a standardised protocol to barcode all land plants. Taxon, **56**(2): 295–299.

Chumley, T.W., Palmer, J.D., Mower, J.P., Fourcade, H.M., Calie, P.J., Boore, J.L., and Jansen, R.K. 2006. The complete chloroplast genome sequence of *Pelargonium × hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol. Biol. Evol. **23**(11): 2175–2190. doi:10.1093/molbev/msl089. PMID:16916942.

Cui, L.Y., Leebens-Mack, J., Wang, L.S., Tang, J.J., Rymarquis, L., Stern, D.B., and dePamphilis, C.W. 2006. Adaptive evolution of chloroplast genome structure inferred using a parametric bootstrap approach. BMC Evol. Biol. **6**(1): 13. doi:10.1186/1471-2148-6-13.

Duffy, A.M., Kelchner, S.A., and Wolf, P.G. 2009. Conservation of selection on *matK* following an ancient loss of its flanking intron. Gene, **438**(1–2): 17–25. doi:10.1016/j.gene.2009.02.006. PMID:19236909.

Ems, S.C., Morden, C.W., Dixon, C.K., Wolfe, K.H., dePamphilis, C.W., and Palmer, J.D. 1995. Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*. Plant Mol. Biol. **29**(4): 721–733. doi:10.1007/BF00041163. PMID:8541499.

Gao, L., Yi, X., Yang, Y.X., Su, Y.J., and Wang, T. 2009. Complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*: insights into evolutionary changes in fern chloroplast genomes. BMC Evol. Biol. **9**(1): 130. doi:10.1186/1471-2148-9-130.

Graham, S.W., and Olmstead, R.G. 2000. Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. Am. J. Bot. **87**(11): 1712–1730. doi:10.2307/2656749. PMID:11080123.

Haberle, R.C., Fourcade, H.M., Boore, J.L., and Jansen, R.K. 2008. Extensive rearrangements in the chloroplast genome of *Trachelium caeruleum* are associated with repeats and tRNA genes. J. Mol. Evol. **66**(4): 350–361. doi:10.1007/s00239-008-9086-4. PMID:18330485.

Hasebe, M., and Iwatsuki, K. 1990. Chloroplast DNA from *Adiantum capillus-veneris* L., a fern species (Adiantaceae); clone bank, physical map and unusual gene localization in comparison with angiosperm chloroplast DNA. Curr. Genet. **17**(4): 359–364. doi:10.1007/BF00314885.

Hasebe, M., and Iwatsuki, K. 1992. Gene localization on the chloroplast DNA of the maiden hair fern; *Adiantum capillus-veneris*. Bot. Mag. Tokyo, **105**(3): 413–419. doi:10.1007/BF02497656.

Hausner, G., Olson, R., Simon, D., Johnson, I., Sanders, E.R., Karol, K.G., et al. 2006. Origin and evolution of the chloroplast *trnK* (*matK*) intron: a model for evolution of group II intron RNA structures. Mol. Biol. Evol. **23**(2): 380–391. doi:10.1093/molbev/msj047. PMID:16267141.

Helfenbein, K.G., and Boore, J.L. 2004. The mitochondrial genome of *Phoronis architecta* — comparisons demonstrate that phoronids are lophotrochozoan protostomes. Mol. Biol. Evol. **21**(1): 153–157. doi:10.1093/molbev/msh011. PMID:14595093.

Morton, B.R., and Clegg, M.T. 1993. A chloroplast DNA mutational hotspot and gene conversion in a noncoding region near *rbcL* in the grass family (Poaceae). Curr. Genet. **24**(4): 357–365. doi:10.1007/BF00336789. PMID:8252646.

Olmstead, R.G., and Palmer, J.D. 1994. Chloroplast DNA systematics: a review of methods and data analysis. Am. J. Bot. **81**(9): 1205–1224. doi:10.2307/2445483.

Palmer, J.D. 1985. Comparative organization of chloroplast genomes. Annu. Rev. Genet. **19**(1): 325–354. doi:10.1146/annurev.ge.19.120185.001545. PMID:3936406.

Palmer, J.D., Soltis, D.E., and Chase, M.W. 2004. The plant tree of life: an overview and some points of view. Am. J. Bot. **91**(10): 1437–1445. doi:10.3732/ajb.91.10.1437.

Pryer, K.M., Schneider, H., Smith, A.R., Cranfill, R., Wolf, P.G.,

Hunt, J.S., and Sipes, S.D. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature, **409**(6820): 618–622. doi:10.1038/35054555. PMID:11214320.

Pryer, K.M., Schuettpelz, E., Wolf, P.G., Schneider, H., Smith, A.R., and Cranfill, R. 2004. Phylogeny and evolution of ferns (monophylytes) with a focus on the early leptosporangiate divergences. Am. J. Bot. **91**(10): 1582–1598. doi:10.3732/ajb.91.10.1582.

Raubeson, L.A., and Jansen, R.K. 1992. Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science, **255**(5052): 1697–1699. doi:10.1126/science.255.5052.1697. PMID:17749424.

Raubeson, L.A., and Stein, D.B. 1995. Insights into fern evolution from mapping chloroplast genomes. Am. Fern J. **85**(4): 193–204. doi:10.2307/1547809.

Roper, J.M., Kellon Hansen, S., Wolf, P.G., Karol, K.G., Mandoli, D.F., Everett, K.D.E., et al. 2007. The complete plastid genome sequence of *Angiopteris evecta* (G. Forst.) Hoffm. (Marattiaceae). Am. Fern J. **97**(2): 95–106. doi:10.1640/0002-8444(2007)97[95:TCPGSO]2.0.CO;2.

Schuettpelz, E., Korall, P., and Pryer, K.M. 2006. Plastid *atpA* data provide improved support for deep relationships among ferns. Taxon, **55**(4): 897–906. doi:10.2307/25065684.

Stein, D.B., Palmer, J.D., and Thompson, W.F. 1986. Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus *Osmunda*. Curr. Genet. **10**(11): 835–841. doi:10.1007/BF00418530.

Stein, D.B., Conant, D.S., Ahearn, M.E., Jordan, E.T., Kirch, S.A., Hasebe, M., et al. 1992. Structural rearrangements of the chloroplast genome provide an important phylogenetic link in ferns. Proc. Natl. Acad. Sci. U.S.A. **89**(5): 1856–1860. doi:10.1073/pnas.89.5.1856. PMID:1542683.

Sugiura, C., Kobayashi, Y., Aoki, S., Sugita, C., and Sugita, M. 2003. Complete chloroplast DNA sequence of the moss *Physcomitrella patens*: evidence for the loss and relocation of *rpoA* from the chloroplast to the nucleus. Nucleic Acids Res. **31**(18): 5324–5331. doi:10.1093/nar/gkg726. PMID:12954768.

Vogel, J., Hübschmann, T., Börner, T., and Hess, W.R. 1997. Splicing and intron-internal RNA editing of *trnK–matK* transcripts in barley plastids: support for MatK as an essential splice factor. J. Mol. Biol. **270**(2): 179–187. doi:10.1006/jmbi.1997.1115. PMID:9236120.

Wolf, P.G., Rowe, C.A., Sinclair, R.B., and Hasebe, M. 2003. Complete nucleotide sequence of the chloroplast genome from a leptosporangiate fern, *Adiantum capillus-veneris* L. DNA Res. **10**(2): 59–65. doi:10.1093/dnares/10.2.59. PMID:12755170.

Wolf, P.G., Duffy, A.M., and Roper, J.M. 2009. Phylogenetic use of inversions in fern chloroplast genomes. Am. Fern J. **99**(2): 132–134.