Genes Translocated into the Plastid Inverted Repeat Show Decelerated Substitution Rates and Elevated GC Content

Fay-Wei Li[1,2,*], Li-Yaung Kuo[3], Kathleen M. Pryer[2], and Carl J. Rothfels[1]

[1]University Herbarium and Department of Integrative Biology, University of California, Berkeley

[2]Department of Biology, Duke University, Durham

[3]Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei

*Corresponding author: E-mail: fay.wei.li@duke.edu.

Accepted: July 1, 2016

Abstract

Plant chloroplast genomes (plastomes) are characterized by an inverted repeat (IR) region and two larger single copy (SC) regions. Patterns of molecular evolution in the IR and SC regions differ, most notably by a reduced rate of nucleotide substitution in the IR compared to the SC region. In addition, the organization and structure of plastomes is fluid, and rearrangements through time have repeatedly shuffled genes into and out of the IR, providing recurrent natural experiments on how chloroplast genome structure can impact rates and patterns of molecular evolution. Here we examine four loci (*psbA*, *ycf2*, *rps7*, and *rps12* exon 2–3) that were translocated from the SC into the IR during fern evolution. We use a model-based method, within a phylogenetic context, to test for substitution rate shifts. All four loci show a significant, 2- to 3-fold deceleration in their substitution rate following translocation into the IR, a phenomenon not observed in any other, nontranslocated plastid genes. Also, we show that after translocation, the GC content of the third codon position and of the noncoding regions is significantly increased, implying that gene conversion within the IR is GC-biased. Taken together, our results suggest that the IR region not only reduces substitution rates, but also impacts nucleotide composition. This finding highlights a potential vulnerability of correlating substitution rate heterogeneity with organismal life history traits without knowledge of the underlying genome structure.

Key words: GC content, genome structure, inverted repeat, plastome, rate heterogeneity.

Introduction

Rates of molecular evolution vary dramatically among organismal lineages and across genomes (Bromham and Penny 2003) and understanding what causes this rate variation is a fundamental topic in evolutionary biology (Lanfear et al. 2010). Past studies reporting on rate variation usually focused on establishing a correlation between substitution rates and those organismal traits that might potentially be affecting the supply of mutations (e.g., generation time or metabolic rate; Wu and Li 1985; Martin and Palumbi 1993; Smith and Donoghue 2008; Korall et al. 2010; Gaut et al. 2011; Lanfear et al. 2013), or the rate at which available mutations are fixed in the population (e.g., weakened purifying selection due to small population sizes; Woolfit and Bromham 2003, 2005). Establishing such correlations is important for understanding broad evolutionary processes in relation to certain life history traits. Few studies, on the other hand, have investigated whether there are other factors that might also influence evolutionary rates.

Plant chloroplast genomes (plastomes) generally comprise a pair of inverted repeat (IR) regions and two single-copy (SC) regions. The two IR copies are identical in sequence but run in opposite directions (fig. 1a), and their sequence identity is maintained by gene conversion (Birky and Walsh 1992). Biased gene conversion, whereby new mutations are preferentially corrected back to ancestral states (Birky and Walsh 1992), is hypothesized to be responsible for a significantly lower rate of nucleotide substitution in the IR compared to the SC region (Clegg et al. 1984; Wolfe et al. 1987; Wu and Chaw 2015; Zhu et al. 2016). If the IR region does indeed suppress nucleotide substitutions, it is expected that genes moving into or out of the IR (via expansion or contraction of the region) would experience corresponding substitution rate shifts—decelerating after entering the IR and accelerating after exiting it.

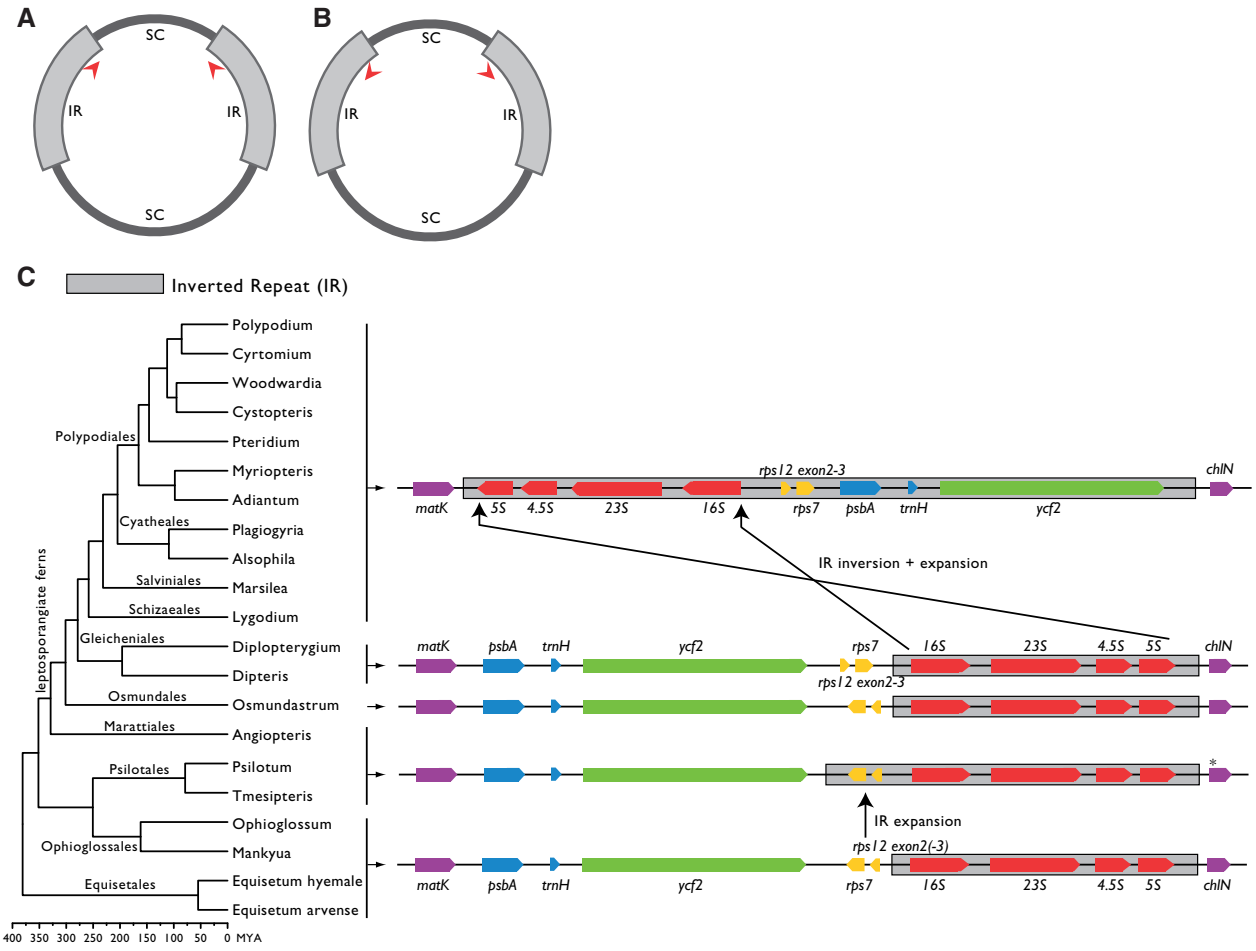

Fig. 1.—Chloroplast genome structure. (*A*) The typical plant chloroplast genome (plastome) comprises a pair of inverted repeat (IR) regions separating a single copy region (SC). The red arrowheads indicate the direction in which the rRNA genes in the IR are transcribed. (*B*) In some fern chloroplast genomes, the direction of transcription in the IR is inverted. (*C*) Genome rearrangements have resulted in changes in IR gene content. The phylogeny on the left shows the relationships among the sampled fern chloroplast genomes; a part of their genome organization is shown on the right. The tree topology and divergence times are derived from Rothfels et al. (2015). Gene lengths are not to scale, gene arrow tips indicate the direction of transcription, and a few genes are omitted for clarity. '*' indicates that *chlN* is not always present. Note that *rps12* in Ophioglossales, Psilotales, and Equisetales lacks the second intron, and therefore there is no exon 3.

Only a few studies have examined whether the IR region can influence gene evolutionary rates. Two of these focused on situations where genes in the IR were translocated to the SC. Perry and Wolfe (2002) discovered that in legumes, where the IR region has disappeared, the former IR genes experienced accelerated substitution rates. However, Lin et al. (2012) found that when *ycf2* was translocated out of the IR in *Ginkgo biloba*, its substitution rate did not increase (although it is possible that this event was too recent for a rate difference to be detectable). More recently, Zhu et al. (2016) tested whether genes translocated into the IR from the SC exhibited rate changes by examining such translocation events across a representative sample of angiosperms, gymnosperms, and ferns, and found that substitution rates did indeed decrease after gene translocation into the IR.

These three studies used pairwise synonymous substitutions to measure rate variation, which ignores the role of phylogenetic history and is unable to incorporate nucleotide substitution models to yield more accurate rate estimates. Here we estimate rate variation in a phylogenetic context, and employ a likelihood ratio test to identify significant rate shifts. In addition, we examine whether translocation into the IR alters the selective environment experienced by plastid loci, and whether plastome structure changes cause shifts in GC content. We focus on ferns (the sister group to seed plants), where plastome rearrangements have inverted the IR orientation (fig. 1b) and translocated five loci—*trnH*, *psbA*, *ycf2*, *rps7*, and exons 2–3 of *rps12* (henceforth simply "*rps12*")—from the SC region to the IR (Wolf et al. 2010). This expanded and inverted IR occurs in extant Schizaeales, Salviniales,

Table 1
Fern Plastomes Sampled in this Study

Taxon	Genbank Accession	Reference
Polypodium glycyrrhiza	KP136832	Wolf et al. (2015)
Cyrtomium falcatum	NC_028705	Choi and Park (unpublished)
Woodwardia unigemmata	NC_028543	Lu et al. (2015)
Cystopteris protrusa	KP136830	Wolf et al. (2015)
Pteridium aquilinum	NC_014348	Der (2010)
Myriopteris lindheimeri	NC_014592	Wolf et al. (2011)
Adiantum capillus-veneris	NC_004766	Wolf et al. (2003)
Plagiogyria glauca	KP136831	Wolf et al. (2015)
Alsophila spinulosa	NC_012818	Gao et al. (2009)
Marsilea crenata	NC_022137	Gao et al. (2013)
Lygodium japonicum	NC_022136	Gao et al. (2013)
Diplopterygium glaucum	NC_024158	Kim et al. (2014)
Dipteris conjugata	KP136829	Wolf et al. (2015)
Osmundastrum cinnamomeum	NC_024157	Kim et al. (2014)
Angiopteris evecta	NC_008829	Roper et al. (2007)
Psilotum nudum	NC_003386	Wakasugi et al. (1998)
Tmesipteris elongata	KJ569699	Zhong et al. (2014)
Ophioglossum californicum	NC_020147	Grewe et al. (2013)
Mankyua chejuensis	NC_017006	Kim and Kim (unpublished)
Equisetum arvense	NC_014699	Karol et al. (2010)
Equisetum hyemale	NC_020146	Grewe et al. (2013)

Cyatheales, and Polypodiales, whereas the ancestral IR configuration is found in Gleicheniales, Osmundales, Marattiales, Psilotales, Ophioglossales, and Equisetales, indicating that the IR reorganization occurred early in the evolution of leptosporangiate ferns (fig. 1c). Independent of this major IR inversion and expansion, IR boundaries also expanded in the eusporangiate fern orders Marattiales and Psilotales by incorporating *rps7* and *rps12* into the IR (fig. 1c). The only other locus that transitioned into and out of the IR is *ndhB* (Gao et al. 2013); however, because *ndhB* always resides next to, or right on, the IR boundary, its inclusion or exclusion from the IR appears to be labile and difficult to resolve phylogenetically. For this reason we did not analyze *ndhB*.

Results and Discussions

Rate Deceleration Following Translocation into IR

We compiled *psbA*, *ycf2*, *rps7*, and *rps12* sequence data from published plastomes (*trnH* and other tRNA genes were not included because of their short sequence lengths; table 1), and tested the fit of two clock models: a 1-rate model and a 2-rate model. The 1-rate model applies a single substitution rate across all branches of the phylogeny, whereas the 2-rate model allows one "in-IR" rate, when these loci are in the IR (blue branches in fig. 2*a* and *b*), and another "out-IR" rate, when they are in the SC (orange branches in fig. 2*a* and *b*). For all four translocated loci, the 1-rate model was rejected in favor of the 2-rate model ($P < 0.0000001$), and within the 2-rate model the "in-IR" rates were 2- to 3-fold slower than the "out-IR" rates (fig. 2c–*f*). This model-testing result indicates that substitution rates decreased following gene translocation into the IR, and is consistent with the findings of Zhu et al. (2016) and also a study by Li et al. (2011) that found the *trnH–psbA* intergenic spacer has reduced levels of sequence variation in Polypodiales (where it is located in the IR), compared to the other studied fern orders (where it is in the SC). Because the translocation of *rps7* and *rps12* into the IR occurred three times (fig. 2b) within ferns, we also tested a 4-rate model that allows three separate "in-IR" rates and one "out-IR" rate. For both loci, the 4-rate model fit significantly better than the 2-rate and 1-rate models ($P < 0.0000001$), and all three "in-IR" rates were slower than the "out-IR" rate (fig. 2d), a result that is again consistent with a rate deceleration for genes translocated into the IR.

To rule out the possibility that the rate deceleration observed in *psbA*, *ycf2*, *rps7*, and *rps12* is part of a plastome-wide phenomenon (e.g., a change in polymerase proof-reading efficiency; Parkinson et al. 2005), or that perhaps the improved fit is due simply to the extra parameter accommodating noise in the data rather than being related to the translocation (Lanfear 2011), we examined other chloroplast genes that have not translocated into or out of the IR. Based on the 1-rate model, we first confirmed that fern IR genes have significantly lower substitution rates than the non-IR genes ($P < 10^{-5}$; two-tailed *t*-test), a result consistent with past studies on other plant groups (Clegg et al. 1984; Wolfe et al. 1987; Wu and Chaw 2015; Zhu et al. 2016). Next we compared the 1-rate and 2-rate models as above, but for genes that did not move into or out of the IR. The majority of these genes showed no rate shift (i.e., the 1-rate model could not be rejected in favor of the 2-rate model), and if they did, the direction of the rate change was predominantly opposite to that seen in *psbA*, *ycf2*, *rps7*, and *rps12* (fig. 2c–*f* and supplementary table S1, Supplementary Material online). The only exception is *atpB*; although it experienced no translocation, it nonetheless showed a significant deceleration (fig. 2d). However, the rate change in *atpB* is only 1.2-fold, much lower than the rate change observed in the IR-translocated loci. This comparison demonstrates that the rate deceleration in our focal loci is exceptional among chloroplast genes, and is best explained by their translocation into the IR.

Rate Deceleration Is Not Likely Due to Selection at the Protein Level

Next, we tested whether the degree of rate deceleration differs between synonymous and nonsynonymous substitutions, which would suggest a change in selection pressures upon translocation into the IR. We used RELAX (Wertheim et al.

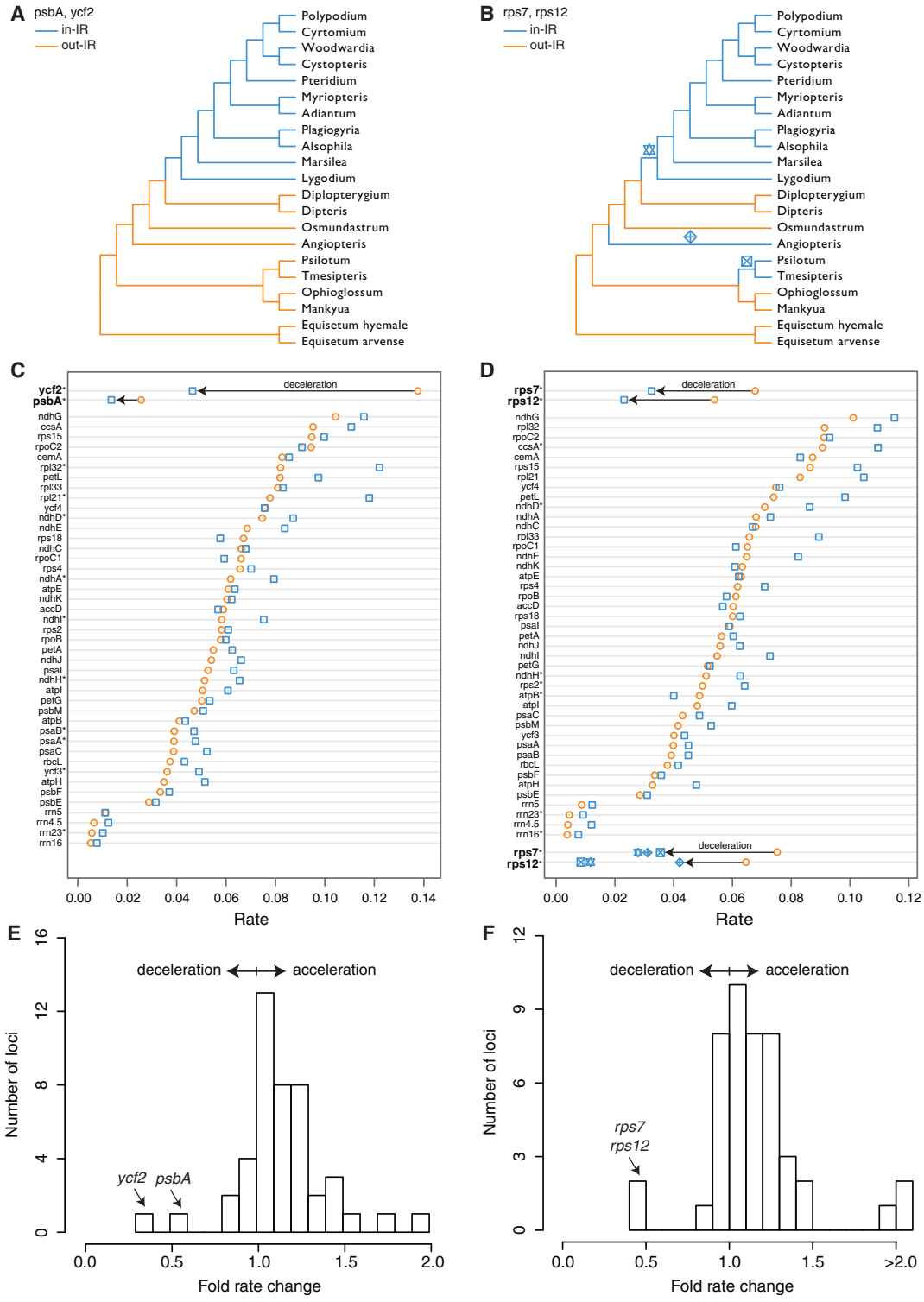

Fig. 2.—Substitution rates before and after translocation into the IR. (*A*, *B*) The resultant rate assignment schemes of the 2-rate models for *psbA* and *yfc2* (*A*) and *rps7* and *rps12* (*B*). The symbols on (*B*) branches mark the rate assignment scheme for the 4-rate model (see *D*). (*C*, *D*) Rate differences of all sampled chloroplast genes observed for the 2-rate or 4-rate model. "*" indicates genes in which the 1-rate model was rejected in favor of the 2-rate or 4-rate model ($P < 0.0000001$). The rate estimates from the 4-rate model of *rps7* and *rps12* are shown at the bottom of (*D*). The symbols correspond to those in (*B*). The rate unit is substitutions/site/million years. (*E*, *F*) The distributions of fold rate change. The rate deceleration that is observed in *ycf2*, *psbA*, *rps7*, and *rps12*, is unique among chloroplast genes, and is due to their translocation into the IR.

2015), a codon-based, branch-site random effects method, to test for any change in selective strength. We found that between the "in-IR" and "out-IR" branches, there is no significant difference in the nonsynonymous/synonymous substitution rate ratio (ω), suggesting that protein-level selection has little or no effect on the rate decrease. The exception is *ycf2*, in which selection relaxed along the "in-IR" branches ($P < 0.0001$; table 2). However, because relaxation of selection would tend to accelerate nucleotide substitution rates, and we still observed *ycf2* deceleration, this indicates its rate decrease is not due to selection.

GC Content Increases upon Translocation into the IR

Another hallmark of the IR is its high GC content relative to the SC region. Recently Wu and Chaw (2015) showed that in cycads there are more A/T to G/C substitutions in the IR (compared to the SC), and they proposed a GC-biased gene conversion mechanism for the IR. If this model is correct, we would expect to see an increase in GC content once a genomic segment translocates into the IR, particularly at the third-codon positions and in noncoding regions, which are under reduced selective constraint. Our analysis of GC content corroborates this prediction—the third codon positions of *psbA*, *ycf2*, *rps7*, and *rps12*, as well as the *trnH–psbA* intergenic spacer, have a higher GC content when they are in the IR than when in the SC region (fig. 3). Interestingly, GC variation is less obvious (or opposite, in *psbA*) in the first and second codon positions (fig. 3), implying that selection reduces this bias by limiting the substitutions that get fixed at these positions (since they usually result in amino acid changes). It should be noted that *ycf2* is the only locus with a significant GC increase at the second codon position, a result consistent with the relaxed selection we detected for it (table 2). To investigate whether the rate deceleration found above was an artifact of the GC content increase, we reanalyzed our focal loci with the third codon position removed. We found similar rate decreases, with no significant difference in the fold rate change ($P = 0.75$, paired two-tailed *t*-test; supplementary table S2, Supplementary Material online), suggesting that the shifts in rate and in GC content are decoupled in the IR.

Conclusion

In this study, we demonstrated that when genes are translocated into the IR, their nucleotide substitution rates dropped significantly (2- to 3-fold), and that this deceleration is not shared with other nontranslocated chloroplast genes. In addition to rate deceleration, GC content increases following translocation, indicating that the IR affects both substitution rates and GC content. Our finding also points to plastome rearrangements that can result in rate heterogeneity among lineages. This has important implications for studies trying to use genomic data to correlate rate heterogeneity together with life history traits, as well as for dating evolutionary events (e.g., Schuettpelz and Pryer 2006; Rothfels and Schuettpelz 2014). Without the knowledge of genome structure, or modeling for possible hidden rate shifts, the evolutionary inferences could be grossly misleading.

Table 2

Test for Shifts in Selective Pressures Using RELAX (Wertheim et al. 2015)

Locus	Selection Intensity *k*	*P* Value
psbA	1.02	0.8223
ycf2	0.42	<0.0001
rps7	1.96	0.2013
rps12	0.77	0.1178

Only *ycf2* showed a significant signature of selection relaxation ($k \neq 1$).

Materials and Methods

We sampled at least one plastome from each of the 11 fern orders (except for Hymenophyllales; table 1). The genome data were downloaded from Genbank, and individual gene sequences were extracted. Because some plastomes are incomplete, we focused on the genes that are present across all the sampled plastomes; a total of 48 loci were included. We inferred multiple sequence gene alignments using MUSCLE (Edgar 2004) followed by manual inspection and adjustment. We used baseml (implemented in PAML; Yang 2007) to estimate nucleotide substitution rates under the 1-rate (global clock) and 2-rate models (see fig. 2*a* and *b*). For the baseml runs, we used a GTR + G substitution model, with a fixed topology and divergence times derived from the 25-nuclear-locus study of Rothfels et al. (2015). The 2-rate models were tested against the 1-rate model using likelihood-ratio tests.

To test for differences in selection pressure among the "in-IR" and "out-IR" branches, we used a codon-based model-testing framework implemented in RELAX (Wertheim et al. 2015), available on Datamonkey (Delport et al. 2010). The key parameter in RELAX is *k*, which controls the degree of nonsynonymous/synonymous substitution rate ratio (ω) differences between the "in-IR" and "out-IR" branches ($\omega_{in} = \omega_{out}^{k}$). In the null model, k is fixed at 1 (i.e., no difference in selective pressures between the in-IR versus out-IR branches), whereas in the alternative model, *k* is a free parameter. Values of $k > 1$ make ω for the in-IR branches more extreme—stronger positive selection or stronger purifying selection—whereas values of $k < 1$ make ω closer to 1 (i.e., relaxation of selection). Both null and alternative models assume three site categories of ω. Likelihood ratio tests are then used to compare the fit of the free-*k* versus fixed-*k* models.

We calculated relative GC content for each sequence as the GC content of the focal sequence divided by the average GC content of the plastome, in order to control for genomic background variation. We used a two-tailed *t*-test to investigate

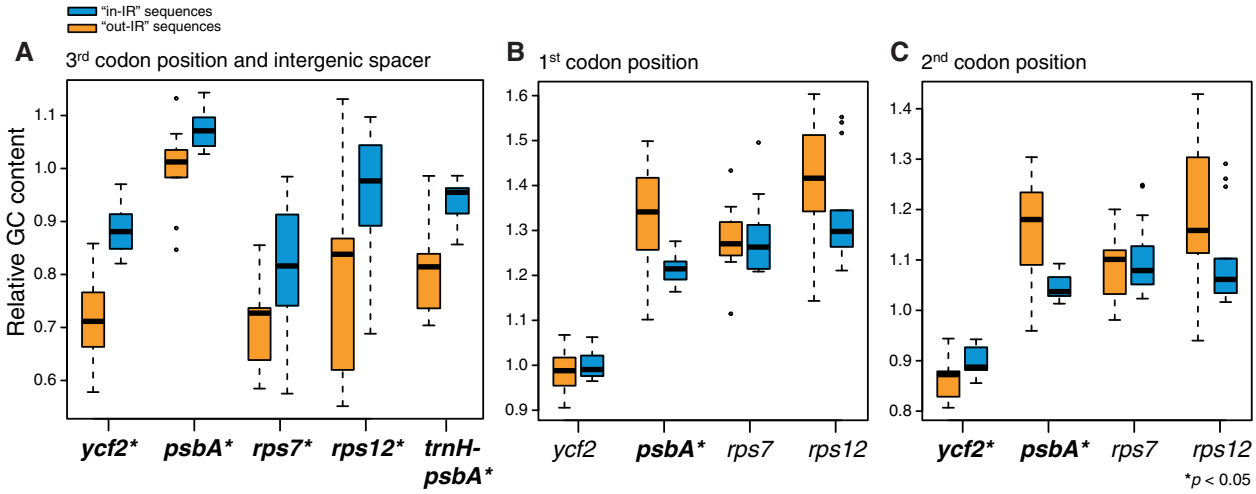

FIG. 3.—GC content before and after genes translocate into IR. (*A*) Third codon position and intergenic spacer. (*B*) Second codon position. (*C*) Third codon position. "*" denotes $P < 0.05$ in 2-tailed *t*-test. Relative GC content calculated as GC content of the locus divided by the whole-genome GC content.

whether there is any significant GC difference between the in-IR and out-IR sequences. All the alignments and the control files for running baseml are available on figshare (dx.doi.org/10.6084/m9.figshare.3483056.v2; dx.doi.org/10.6084/m9.figshare.3483059.v1).

Supplementary Material

Supplementary tables S1 and S2 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

We thank Paul Wolf for maintaining an online list of available fern plastomes (http://www.paulwolflab.com/data-protocols/fern_plastome_list) that helped to propel this study. This work was supported by National Science Foundation grants DEB-1145614 to K.M.P, DEB-1110767 to K.M.P. and C.J.R., and DEB-1407158 to K.M.P. and F.W.L.

Literature Cited

Birky CW, Walsh J. 1992. Biased gene conversion, copy number, and apparent mutation rate differences within chloroplast and bacterial genomes. Genetics 130:677–683.

Bromham L, Penny D. 2003. The modern molecular clock. Nat Rev Genet. 4:216–224.

Clegg MT, Rawson JR, Thomas K. 1984. Chloroplast DNA variation in pearl millet and related species. Genetics 106:449–461.

Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26:2455–2457.

Der JP. 2010. Genomic perspectives on evolution in bracken fern. PhD dissertation. Logan UT, USA: Utah State University.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Gao L, et al. 2013. Plastome sequences of *Lygodium japonicum* and *Marsilea crenata* reveal the genome organization transformation from basal ferns to core leptosporangiates. Genome Biol Evol. 5:1403–1407.

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

Gaut B, Yang L, Takuno S, Eguiarte LE. 2011. The patterns and causes of variation in plant nucleotide substitution rates. Annu Rev Ecol Evol Syst. 42:245–266.

Grewe F, Guo W, Gubbels EA, Hansen AK, Mower JP. 2013. Complete plastid genomes from *Ophioglossum californicum*, *Psilotum nudum*, and *Equisetum hyemale* reveal an ancestral land plant genome structure and resolve the position of Equisetales among monilophytes. BMC Evol Biol. 13:8.

Karol KG, et al. 2010. Complete plastome sequences of *Equisetum arvense* and *Isoetes flaccida*: implications for phylogeny and plastid genome evolution of early land plant lineages. BMC Evol Biol. 10:321.

Kim HT, Chung MG, Kim K-J. 2014. Chloroplast genome evolution in early diverged leptosporangiate ferns. Mol Cells. 37:372–382.

Korall P, Schuettpelz E, Pryer KM. 2010. Abrupt deceleration of molecular evolution linked to the origin of arborescence in ferns. Evolution 64:2786–2792.

Lanfear R. 2011. The local-clock permutation test: a simple test to compare rates of molecular evolution on phylogenetic trees. Evolution 65:606–611.

Lanfear R, et al. 2013. Taller plants have lower rates of molecular evolution. Nat Commun. 4:1879.

Lanfear R, Welch JJ, Bromham L. 2010. Watching the clock: studying variation in rates of molecular evolution between species. Trends Ecol Evol. 25:495–503.

Li F-W, et al. 2011. *rbcL* and *matK* earn two thumbs up as the core DNA barcode for ferns. PLoS One 6:e26597.

Lin CP, Wu CS, Huang YY, Chaw SM. 2012. The complete chloroplast genome of *Ginkgo biloba* reveals the mechanism of inverted repeat contraction. Genome Biol. Evol. 4:374–381.

Lu J-M, Zhang N, Du X-Y, Wen J, Li D-Z. 2015. Chloroplast phylogenomics resolves key relationships in ferns. J Syst Evol. 53:448–457.

Martin AP, Palumbi SR. 1993. Body size, metabolic rate, generation time, and the molecular clock. Proc Natl Acad Sci U S A. 90:4087–4091.
Parkinson CL, et al. 2005. Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. BMC Evol Biol. 5:73.
Perry AS, Wolfe KH. 2002. Nucleotide substitution rates in legume chloroplast DNA depend on the presence of the inverted repeat. J Mol Evol. 55:501–508.
Roper JM, et al. 2007. The complete plastid genome sequence of *Angiopteris evecta* (G. Forst.) Hoffm. (Marattiaceae). Am Fern J. 97:95–106.
Rothfels CJ, et al. 2015. The evolutionary history of ferns inferred from 25 low-copy nuclear genes. Am J Bot. 102:1089–1107.
Rothfels CJ, Schuettpelz E. 2014. Accelerated rate of molecular evolution for vittarioid ferns is strong and not driven by selection. Syst Biol. 63:31–54.
Schuettpelz E, Pryer KM. 2006. Reconciling extreme branch length differences: decoupling time and rate through the evolutionary history of filmy ferns. Syst Biol. 55:485–502.
Smith SA, Donoghue MJ. 2008. Rates of molecular evolution are linked to life history in flowering plants. Science 322:86–89.
Wakasugi T, Nishikawa A, Yamada K, Sugiura M. 1998. A complete nucleotide sequence of the plastid genome from a fern, *Psilotum nudum*. Endocyt Cell Res. 13 (Supp):147.
Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. 2015. RELAX: detecting relaxed selection in a phylogenetic framework. Mol Biol Evol. 32:820–832.
Wolf PG, et al. 2015. An exploration into fern genome space. Genome Biol Evol. 7:2533–2544.
Wolf PG, et al. 2011. The evolution of chloroplast genes and genomes in ferns. Plant Mol Biol. 76:251–261.
Wolf PG, Roper JM, Duffy AM. 2010. The evolution of chloroplast genome structure in ferns. Genome 53:731–738.
Wolf PG, Rowe CA, Sinclair RB, Hasebe M. 2003. Complete nucleotide sequence of the chloroplast genome from a leptosporangiate fern, *Adiantum capillus-veneris* L. DNA Res. 10:59–65.
Wolfe KH, Li W-H, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci U S A. 84:9054–9058.
Woolfit M, Bromham L. 2003. Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population sizes. Mol Biol Evol. 20:1545–1555.
Woolfit M, Bromham L. 2005. Population size and molecular evolution on islands. Proc Biol Sci. 272:2277–2282.
Wu C-I, Li W-H. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci U S A. 82:1741–1745.
Wu C-S, Chaw S-M. 2015. Evolutionary stasis in cycad plastomes and the first case of plastome GC-biased gene conversion. Genome Biol Evol. 7:2000–2009.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24:1586–1591.
Zhong B, Fong R, Collins LJ, McLenachan PA, Penny D. 2014. Two new fern chloroplasts and decelerated evolution linked to the long generation time in tree ferns. Genome Biol Evol. 6:1166–1173.
Zhu A, Guo W, Gupta S, Fan W, Mower JP. 2016. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. New Phytol. 209:1747–1756.

Associate editor: Shu-Miaw Chaw