

Chloroplast RNA editing going extreme: more than 3400 events of C-to-U editing in the chloroplast transcriptome of the lycophyte *Selaginella uncinata*

BASTIAN OLDENKOTT,¹ KAZUO YAMAGUCHI,^{2,3} SUMIKA TSUJI-TSUKINOKI,^{2,4} NILS KNIE,¹ and VOLKER KNOOP¹

¹IZMB—Institut für Zelluläre und Molekulare Botanik, Abteilung Molekulare Evolution, Universität Bonn, Kirschallee 1, 53115 Bonn, Germany

²Division of Functional Genomics, Advanced Science Research Center, Kanazawa University, Kanazawa 920-0934, Japan

ABSTRACT

RNA editing in chloroplasts and mitochondria of land plants differs significantly in abundance. For example, some 200–500 sites of cytidine-to-uridine RNA editing exist in flowering plant mitochondria as opposed to only 30–50 such C-to-U editing events in their chloroplasts. In contrast, we predicted significantly more chloroplast RNA editing for the protein-coding genes in the available complete plastome sequences of two species of the spike moss genus *Selaginella* (Lycopodiophyta). To evaluate these predictions we investigated the *Selaginella uncinata* chloroplast transcriptome. Our exhaustive cDNA studies identified the extraordinary number of 3415 RNA-editing events, exclusively of the C-to-U type, in the 74 mRNAs encoding intact reading frames in the *S. uncinata* chloroplast. We find the overwhelming majority (61%) of the 428 silent editing events leaving codon meanings unaltered directly neighboring other editing events, possibly suggesting a sterically more flexible RNA-editing deaminase activity in *Selaginella*. No evidence of RNA editing was found for tRNAs or rRNAs but we identified a total of 74 editing sites in cDNA sequences of four group II introns (petB16g2, petD18g2, ycf3i124g2, and ycf3i354g2) retained in partially matured transcripts, which strongly contribute to improved base-pairing in the intron secondary structures as a likely prerequisite for their splicing.

Keywords: chloroplast genome; group II introns; lycophyte; plastome; pyrimidine-exchange editing

INTRODUCTION

The modification of genetic information by the pyrimidine-exchange type of RNA editing is common in land plant mitochondria and chloroplasts (Chateigner-Boutin and Small 2011; Knoop 2011; Finster et al. 2012). Numerous cytidine-to-uridine conversions in the organelle mRNAs (and in some plant clades also reverse uridine-to-cytidine exchanges) primarily serve to convert codon meanings to encode evolutionarily conserved amino acids. In flowering plants some 200–500 sites of mitochondrial RNA editing exceed the generally much less abundant ~30–50 chloroplast editing sites in the angiosperms. “Editome” data for exhaustively analyzed organelle transcriptomes have recently been compiled as references for an update of the “Plant RNA-editing prediction and analysis computer tool” PREPACT (Lenz and Knoop 2013).

Making use of PREPACT we found that particularly frequent RNA editing was predicted for the chloroplast genome

sequence of *Selaginella uncinata* (Tsuji et al. 2007), prompting us to exhaustively investigate its chloroplast transcriptome. We here report on the identification of more than 3400 events of RNA editing in exon and intron sequences of *Selaginella uncinata* chloroplast (pre-)mRNAs, all of which are exclusively of the C-to-U type. Hence, chloroplast RNA editing in *Selaginella uncinata* is not only some 100-fold more abundant than in flowering plants but it now even surpasses the previous record as a first organelle transcriptome featuring more than 3000 sites of RNA editing.

RESULTS

RNA editing prediction and confirmation in *Selaginella uncinata* chloroplast mRNAs

We used the recently upgraded version 2.0 of PREPACT (Lenz and Knoop 2013) to scan several published land plant

³Present address: Setagaya Kamata 4-1-31-204, Tokyo 157-0077, Japan

⁴Present address: Laboratory of Plant Physiology, Department of Biology, Yamagata University, Yamagata 990-8560, Japan

Corresponding author: volker.knoop@uni-bonn.de

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.045575.114>.

© 2014 Oldenkott et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

chloroplast genome sequences for RNA-editing candidate sites. For the published lycophyte chloroplast genomes (Wolf et al. 2005; Tsuji et al. 2007; Smith 2009; Karol et al. 2010) we identified some 340 RNA editing candidate sites for *Huperzia lucidula* and about 460 for *Isoetes flaccida*, but some 1800 for *Selaginella moellendorffii* and an astonishing 2400 for *Selaginella uncinata* even when using stringent criteria (i.e., consensus prediction by at least 80% of the references implemented in PREPACT 2.0).

To check on these predictions and verify whether RNA editing would truly modify transcripts to such an extent we studied the *Selaginella uncinata* chloroplast transcriptome exhaustively by targeted RT-PCR cDNA analyses. Indeed, we found that the *S. uncinata* cp mRNAs are subject to RNA editing to an unparalleled extent. Altogether we found that the mRNAs of the 74 functional genes in the *Selaginella uncinata* chloroplast transcriptome were affected by 3415 events of RNA editing, all of which are exclusively of the C-to-U type. Some 12.6% of these editing events (428) are silent, all others change the meanings of the affected codons (Fig. 1; Supplemental Table 1; Supplemental Fig. 1). The extraordinary abundance of RNA editing is reflected by the numbers of start and stop codons created by threonine (ACG) to methionine (AUG) and arginine (CGA) or glutamine (CAR) to stop codon (UGA, UAA, UAG) editing, respectively, to properly define reading frame ends (Fig. 1). Methionine start codons are created in 52 cases and stop codons are introduced by editing in 31 of the transcripts, both at the same time are introduced in 19 of the genes. The highest numbers of editing sites were found in *rpoB* with 214 sites, *ndhF* with 164 sites and *ndhD* with 120 sites (Fig. 1; Supplemental Table 1). Only two chloroplast genes with intact and conserved reading frames were found to lack RNA editing altogether: *psbA* and *rps14*. Similarly we found that three pseudogenes ($\Psi accD$, $\Psi rpl20$, and $\Psi rpl21$) characterized by a lack of convincing overall similarity to other taxa including the sister lycophytes *Huperzia* and *Isoetes* are devoid of RNA editing, too. RNA editing could similarly not be identified for the small *psbN* locus where, however, 10 RNA-editing events including the creation of the start codon are strongly predicted to affect the otherwise conserved reading frame. Hence, the *psbN* locus likely represents a pseudogene in the *Selaginella uncinata* cpDNA, too, also supported by a unique proline-codon duplication of the coding sequence never seen elsewhere.

Finally, we checked three intergenic regions for transcription and RNA editing where genes present elsewhere were lost in *Selaginella uncinata* with only very minor or no pseudogene traces remaining: the *psaJ-rps18* spacer, where *rpl33* is located in other taxa, the *rps8-rpl36* spacer where *infA* is located elsewhere, and the *ccsA-ndhF* region, where *rpl21* and *rpl32* are present in other taxa. We retrieved 270, 132, and 737 bp, respectively, of cDNA amplicons from these regions and altogether identified five RNA-editing sites within this total of 1139 bp, hence amounting to 0.5% of edited nucleotides (nt) in those intergenic sequences. None of the five edits

served to increase similarities of recognizable pseudogene fragments with their functional counterparts in other taxa. Accordingly, we do not assume them to be an original leftover activity of proper editing in the formerly functional mRNAs but rather by-products of accidental misfiring of the editing machinery in the *Selaginella uncinata* chloroplast.

Silent editing: congruencies and deviations from previous observations

All theoretically possible C-to-U conversions possible among the 64 codons of the genetic code were found represented in the *Selaginella uncinata* chloroplast transcriptome (Supplemental Table 2). Single examples only were found for each of the silent editings in arginine (CGC>CGU), aspartate (GAC>GAU), cysteine (UGC>UGU), and glycine (GGC>GGU) codons and only two events of silent serine (AGC>AGU) codon editing. A general bias against purine nucleotides directly upstream of plant organelle RNA-editing sites has been observed previously and now receives yet more statistical support from the abundance of sites in *Selaginella uncinata* now available for analysis. With respect to silent third codon position editing we observed the distribution NUC (256 sites) > NCC (65 sites) > NAC (26 sites) > NGC (five sites).

While the above findings are fully congruent with previous findings, another observation concerning the occurrence of silent RNA-editing sites is strikingly divergent in the *S. uncinata* cp mRNAs. Whereas silent editing events normally occur randomly distributed within mRNAs, likely owing to similarities of upstream RNA recognition sequences for RNA-editing factors targeting relevant sites elsewhere, we here observe that most of the silent RNA-editing events occur in the immediate vicinity of other editing sites. Indeed, 261 of the 428 silent editing sites (i.e., 61%) in the *Selaginella uncinata* chloroplast are found immediately next to other edits (Fig. 2A). We henceforth refer to such sites as “NESIs,” i.e., “neighboring silents.” A high number of 5′-NESIs is also reflected by the high amount of silent first codon extra editing of proline CCR to leucine UUR codons. The 349 cases of proline (CCR) to leucine (CUR) editing in the second codon position are accompanied by 66 cases with additional silent first codon editing (i.e., to UUR) where the second codon position editing alone would have been sufficient for a change of codon meaning. These 66 cases outnumber the only 12 cases of independent silent CUR to UUR editing of leucine codons by more than a factor of five. A sterical flexibility of the deaminase activity operating on the ribonucleotide sequences in *Selaginella* may be an explanation for the high incidence of NESIs as we discuss below (Fig. 2B).

Checking RNA editing in rRNAs and tRNAs

A high frequency of messenger RNA editing may indicate that other types of RNAs could be affected, too. Accordingly, we checked all four *Selaginella uncinata* chloroplast ribosomal

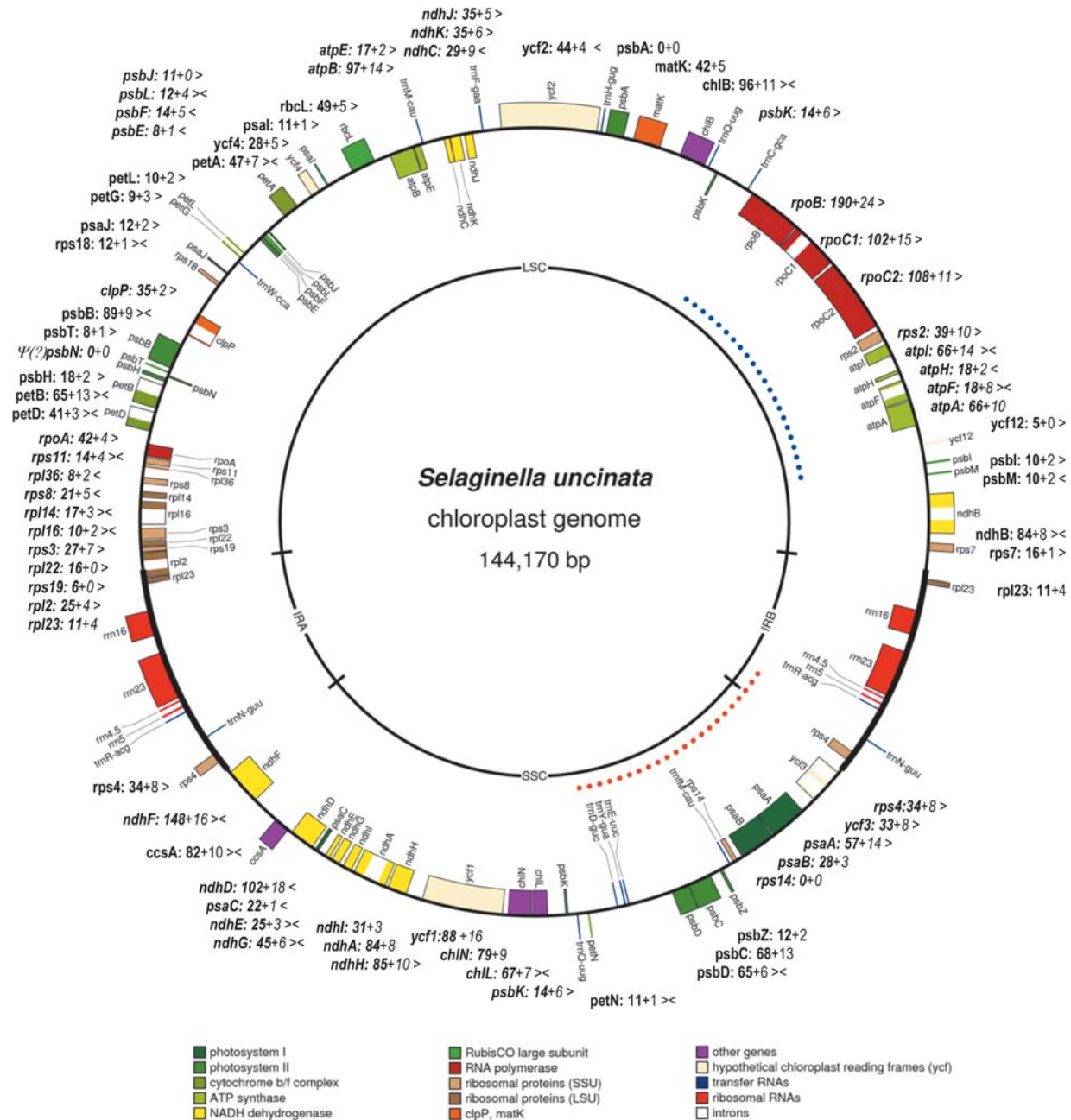


FIGURE 1. Map of the *Selaginella uncinata* chloroplast genome drawn with OGDRAW tool (Lohse et al. 2007) after manual editing of GenBank database entry AB197035 to properly annotate intron sequences and tRNA anticodons and to correct minor errors. Different colors indicate functional gene categories as indicated in the legend. The *S. uncinata* cpDNA is characterized by major plastome rearrangements and several genes present in bryophytes and/or in the sister lycophytes are entirely absent or remain as degenerated pseudogenes only (e.g., *accD*, *cemA*, *infA*, *psaM*, *rpl20*, *rpl21*, *rpl32*, *rpl33*, *rps12*, and *rps15*), very likely owing to their functional endosymbiotic gene transfer to the nucleus. Similarly, the *Selaginella* plastomes have a highly reduced set of only 12 tRNA genes. Numbers of nonsilent (bold) and silent (behind the plus sign) C-to-U RNA-editing sites identified in the respective mRNAs are indicated next to each protein-coding gene with italics indicating genes transcribed clockwise (*inner circle*) and normal font those transcribed counterclockwise (*outer circle*). Larger than (>) and smaller than (<) labels indicate the creation of start and stop codons via RNA editing, respectively. The *Selaginella* cpDNA is characterized by several deviations from the ancestral land plant chloroplast genome structure, which include (1) the translocation of the *rps4-trnD* region from the large (LSC) to the small single-copy (SSC) region (red dotted line), (2) the inversion of the *psbI-trnC* region (blue dotted line), (3) the inclusion of *rps4* and *rpl23* in the inverted repeat (IR) regions, (4) the loss of the *trnK* exons rendering the otherwise intron-borne maturase *matK* an orphan ORF, and (5) the duplication of the *psbK-trnQ* region in the LSC and SSC. Several genes present in other plant taxa including the related lycophytes are absent or frame-shifted pseudogenes. The *rpoA* reading frame is shortened by ~300 nt but subject to RNA editing, including the creation of a start codon at the end of the conserved sequence region.

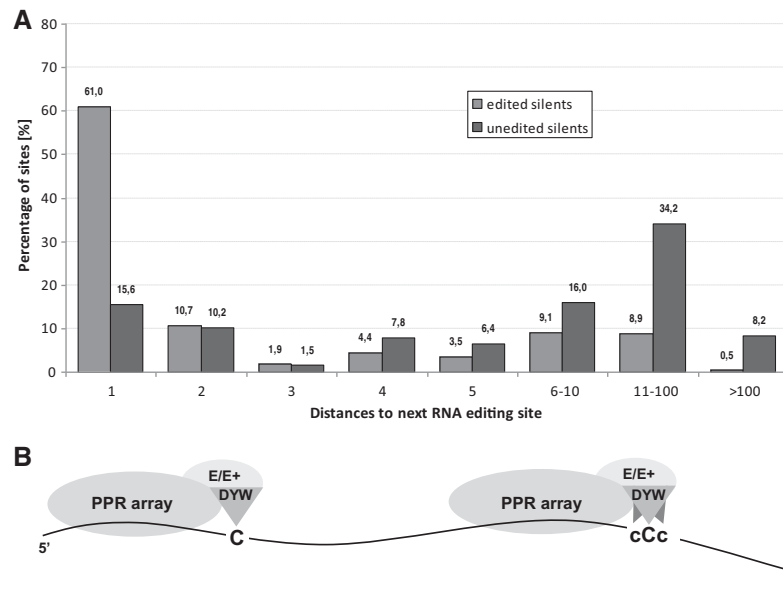


FIGURE 2. (A) A strong bias is found for silent RNA-editing events to occur in the immediate vicinity of other RNA-editing sites as “NESIs (neighboring silents)” in the *Selaginella uncinata* chloroplast transcriptome. Light gray bars indicate the respective percentages of the 428 silent edits observed sorted by distances to the respective next editing site as indicated in the categories on the x-axis. As a control, dark gray bars indicate the corresponding percentages for all 5738 additional potential silent cytidines remaining unedited within the 67,629 nt of analyzed cDNA, showing a peak in the 11–100 nt distance category. (B) RNA-binding pentatricopeptide repeat (PPR) proteins play a major role in identifying sequence targets in front of C-to-U RNA-editing sites. Several PPR proteins carrying a carboxy-terminal E/E+ domain linking into a “DYW domain” having cytidine deaminase similarity have been characterized as editing factors (left). In addition to their primary targets further cytidines may also be subject to editing when a similar RNA target is accidentally bound and the resulting change to uridine is tolerated, e.g., in a silent codon position. In the *Selaginella uncinata* chloroplast transcriptome the majority of silent edits is identified as “NESIs” (neighboring silents) in the immediate vicinity of other editing sites, possibly indicating a sterical flexibility of the DYW deaminase rather than spurious binding of the PPR array to non-target sites to be the underlying cause (right).

RNA species (*rrnL*, *rrnS*, *rrn5*, and *rrn4.5*) on transcript level. All cDNA sequences were found to be identical with the genomic DNA sequences and hence revealed no evidence for rRNA editing. We also reinspected the suggested cloverleaf tRNA secondary structure predictions for all *Selaginella uncinata* chloroplast tRNA genes (Lowe and Eddy 1997; Schattner et al. 2005) for the potential necessity of RNA editing to establish proper base pairings in the base-paired stem regions or to reconstitute highly conserved uridine positions (e.g., in tRNA consensus positions 8 and 33). Only one such candidate site was identified in the tRNA for aspartate (*trnD*), where a C-A mismatch in the anticodon stem could be converted to a proper U-A base pair. However, the candidate editing could not be verified on cDNA level. Hence, we could not find any evidence for RNA editing in either rRNAs or tRNAs in the *Selaginella uncinata* chloroplast.

RNA-editing sites in intron sequences

Organelle introns (mainly of the group II type in plant chloroplasts and mitochondria) are characterized by conserved

secondary and tertiary structures that have previously been shown to be affected by RNA editing in some cases (see e.g., Lippok et al. 1994; Börner et al. 1995; Carrillo and Bonen 1997; Carrillo et al. 2001; Castandet et al. 2010; Bégu et al. 2011; Farré et al. 2012). To detect RNA editing in intron sequences we used a strategy to aim for partially matured transcripts. To this end we selected two chloroplast regions encompassing two introns each where we could target cDNAs selecting for partially matured transcripts with one of the respective group II introns spliced out but the other one still retained with “bridge primers” covering the exon linkages. These were the *petB-petD* cotranscript containing introns *petBi6g2* and *petDi8g2* and the *ycf3* gene containing *ycf3i124g2* and *ycf3i354g2*, respectively (Fig. 3). Using the above strategy we were able to retrieve cDNAs for partially matured transcripts of these loci and indeed identified RNA editing events in all four intron sequences, again all of them exclusively of the C-to-U type: 11 sites in *ycf3i124g2*, 21 in *ycf3i354g2*, 14 in *petDi8g2*, and even 28 sites of RNA editing in *petBi6g2*. The majority of these editing events re-establish proper U-A base-pairing in many instances as becomes apparent in the highly conserved domains V of the intron carboxy-termini (Fig. 3). Likewise, a conserved unpaired

uridine is introduced by editing directly behind domain V in the case of *ycf3i354g2*, *petBi6g2*, and *petDi8g2*. Yet further editing events, such as, e.g., in the distal stem of domain V of *ycf3i124g2* (Fig. 3A), may be an ultimate prerequisite before functional splicing. The abundant editing also in the intron sequences makes their complete secondary structure modeling particularly complex and laborious. We nevertheless present a complete secondary structure model for *petBi6g2* for which we observed the highest incidence of RNA-editing sites (Supplemental Fig. 2). In this model, the majority of the 28 edits observed occur in essential base-pairing stem regions or in other crucial regions such as the exon binding site EBS1.

Notably, our strategy to investigate partially matured cotranscripts of the *petB-petD* locus with one of the two introns spliced out serendipitously identified seven events of RNA editing also in the transcribed spacer of 219 nt (Fig. 3; Supplemental Table 3). Hence, editing site abundance in this spacer between two heavily edited functional genes is characteristically higher than in the intergenic pseudogene regions described above.

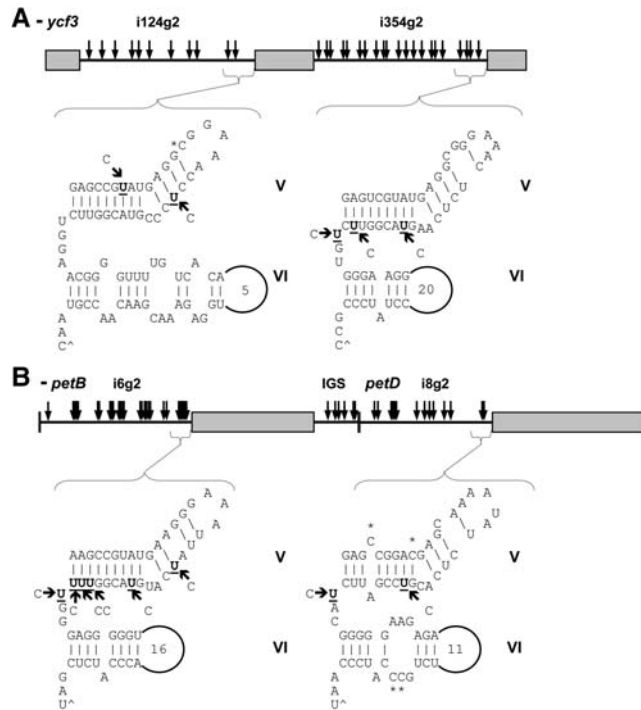


FIGURE 3. Partially matured transcripts of the *ycf3* (A) and of the *petB-petD* (B) locus with one of the respective introns spliced out were analyzed to investigate RNA editing in the respective other intron. Intron nomenclature follows an earlier proposal designating introns according to insertion site (Dombrowska and Qiu 2004) and intron type (Knoop 2004). Arrows indicate approximate positions of RNA editing events identified in *ycf3*i124g2 (11 sites), *ycf3*i354g2 (21 sites), *petB*i6g2 (28 sites), and *petD*i8g2 (14 sites), and in the *petB-petD* spacer (seven sites). The respective intron 3'-ends including the characteristic secondary domains V and VI structures of group II introns are shown together with the RNA-editing sites affecting these regions. Asterisks identify potential additional candidate sites of C-to-U editing in further matured introns. A complete secondary structure model for *petB*i6g2 is shown in Supplemental Figure 2.

We consider the number of intron editing sites we have identified a lower bound since these were seen in unspliced introns and yet more editing events (such as e.g., the ones labeled in Fig. 3) may be ultimate prerequisites for splicing. Such editing events may only be identified in spliced-out introns to be possibly identified in the future with sophisticated primers designed to target the 5'-2' lariat formed at the adenosine branch-points in domain VI. Extrapolating from the average of about 18 editing sites per intron identified in the four introns for which we succeeded to obtain cDNA sequences and assuming similar numbers of editing sites in the seven remaining introns (*atpFi145g2*, *clpPi71g2*, *ndhAi556g2*, *ndhBi726g2*, *rpl2i397g2*, *rpl16i9g2*, and *rpoCi432g2*) would already now bring the total number of RNA editing events in the 11 *Selaginella uncinata* chloroplast introns to some 200.

Loss of an intron and of RNA-editing sites

The second *clpP* intron (*clpPi363g2*) conserved in other taxa including the sister lycophytes *Huperzia lucidula* and *Isoetes*

flaccida is lost in the *Selaginella* cpDNAs, precluding the “one-of-two” strategy to obtain intron cDNA sequences outlined above. On the other hand, the *clpP* intron loss warranted a reinspection of flanking RNA-editing patterns. We observe an intriguing absence of RNA editing over 98 nt around the former *clpPi363g2* insertion site between editing site *clpPeU320PL* and *clpPeU418RC* (Fig. 4). At the same time, the upstream intron *clpPi71g2* still present in *Selaginella* is flanked by an editing event immediately upstream of the 5' splice site and in nucleotide +9 behind the 3' splice site. Similarly, editing events are consistently present in distances of maximally 9 nt from intron insertion sites in the *Selaginella uncinata* genes in all other cases (Supplemental Fig. 1). We conclude that loss of *clpPi363g2* was accompanied by a recombination event involving regions of a mature *clpP* transcript/cDNA leading to a concomitant loss of editing events in its vicinity, similar to previously reported cases of organelle intron loss (e.g., Geiss et al. 1994; Itchoda et al. 2002; Lopez et al. 2007; Ran et al. 2010; Grewe et al. 2011). Alternative explanations such as horizontal gene transfer (Hepburn et al. 2012) or a pressure on intron retention exerted by RNA editing (Kagami et al. 2012) must be kept in mind, however. Whereas average distances between editing sites are 18.3 nt, a total of 62 mRNA regions of at least 96 nt exist in the *Selaginella uncinata* chloroplast transcriptome that are free of RNA-editing sites, similar to the *clpP* case (Supplemental Fig. 1). Importantly, a total of only 18 events of RNA editing are reasonably predicted in those regions of altogether >14,600 nt, which supports the idea that cDNA-mediated recombination has caused a loss of editing there.

DISCUSSION

The unparalleled amount of more than 3500 chloroplast RNA-editing sites now identified in *Selaginella uncinata* is all the more extraordinary since it even surpasses the record amounts of 2152 RNA-editing sites previously documented for the *Selaginella moellendorffii* mitochondrion (Hecht et al. 2011). First preliminary data (B Oldenkott and V Knoop, unpubl.) suggest that *S. uncinata* may possibly exceed *S.*

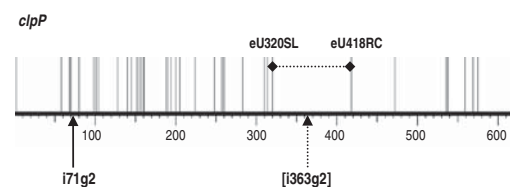


FIGURE 4. The loss of the second group II intron (*clpPi363g2*) lacking in the *Selaginella uncinata* *clpP* gene but conserved in other taxa appears to have been accompanied by a loss of editing sites around the former intron insertion site. The editing sites closest to the former intron insertion site are located 43 nt upstream (*clpPeU320SL*) and 55 nt downstream (*clpPeU418RC*), respectively, whereas the upstream intron *clpPi71g2* retained in *S. uncinata* is flanked by editing sites *clpPeU70HY* directly upstream and editing site *clpPeU80PL* only 9 nt downstream.

moellendorffii in mitochondrial RNA-editing abundance and we predicted somewhat lower chloroplast editing in *S. moellendorffii* than in *S. uncinata*. The ultimate mitochondrial vs. chloroplast editing ratios in the same *Selaginella* species would ultimately be only in the range from 1:1 to maximally 2:1.

The huge numbers of RNA-editing sites in both *Selaginella* organelles are all the more intriguing given that both the mitochondrial and chloroplast genomes are at the same time among the gene-poorest exemplars of their kind among land plants. Many protein-coding genes are lost from both organelles obviously owing to endosymbiotic gene transfer to the nucleus and even the tRNA gene sets are extremely reduced in the chloroplast (Tsuiji et al. 2007; Smith 2009) or even lost altogether in the mitochondria (Hecht et al. 2011). Paradoxically, the highly gene-reduced organelle genomes of *Selaginella* harbor the current record amounts of plant C-to-U RNA editing and their comparatively high GC content is likely a part of the explanation.

Intriguingly, all of the now altogether more than 5500 identified sites of RNA editing in the *Selaginella* organelle transcriptomes are exclusively of the C-to-U type. It is particularly noteworthy that neither the *Selaginella* chloroplasts nor the mitochondria revealed but a single type of U-to-C editing, which is so abundant not only in ferns and hornworts but also in the sister lycophyte *Isoetes engelmannii* (Grewe et al. 2011). Hence, a complete loss of reverse U-to-C editing—i.e., loss of a likely transamination activity as opposed to the cytidine deaminase activity of C-to-U editing—is the obvious phylogenetic explanation and a highly interesting issue, all the more given that no U-to-C editing factor has as yet been characterized, mainly because the phenomenon is represented by “nonmodel” plant taxa.

In contrast, after the initial discovery of the first chloroplast (Kotera et al. 2005) and mitochondrial (Zehrman et al. 2009) C-to-U RNA-editing factors in *Arabidopsis thaliana*, the prime model organism of molecular botany, many other such factors have been identified. Its low number of RNA-editing sites and the unique model organism features have recently made the mitochondrion of the model moss *Physcomitrella patens* the first plant organelle with a full assignment of eight nuclear-encoded C-to-U editing protein factors to all of its 11 mitochondrial editing sites (Ichinose et al. 2013; Schallenberg-Rüdinger et al. 2013; Sugita et al. 2013). These factors are RNA-binding pentatricopeptide repeat (PPR) proteins of the so-called “DYW-type” (Fig. 2B). The PPR arrays of such proteins mediate sequence-specific RNA-binding in a predictable manner (Barkan et al. 2012; Takenaka et al. 2013; Yagi et al. 2013), whereas the carboxy-terminal DYW domain of these proteins has cytidine deaminase similarity (Salone et al. 2007; Iyer et al. 2011), although a direct biochemical proof for the latter is still lacking. Among land plants, RNA-editing abundance and the diversity of DYW-type PPR gene families appear to correlate well (Rüdinger et al. 2008, 2012). Expectedly, the *Selaginella moellendorffii* genome (Banks et al. 2011) features an extended DYW-type

PPR protein gene family—however, its ~200 members are far below numbers of some 4000 to 5000 organellar RNA-editing events.

All of the above observations plus the fact that the genus *Selaginella* is highly variable and rich in number of species (~700), the nuclear genome sizes of which appear to be generally very small (i.e., ~100 MBp) (Little et al. 2007) make *Selaginella* species particularly interesting additional taxa for future studies of the plant RNA-editing machinery. Other than the mysterious absence of U-to-C editing, a vast field is opened to test and improve the PPR-binding code targeting specific RNA sequences both bioinformatically and ideally also with transgenic approaches in the future. Similarly, the RNA-editing activity in *Selaginella* may prove to be more prone to affect additional cytidine nucleotides in the immediate vicinity of proper RNA targets than in the other lineages (Fig. 2B).

Finally, the abundant intron editing described here is particularly noteworthy. Whereas the PPR-RNA-binding code hitherto considers linear RNA sequences the additional layer of RNA secondary structures comes into play and the succession and mutual influence of RNA editing and RNA folding processes offer an amazing field of complexity in the RNA metabolism in plant organelles.

MATERIALS AND METHODS

Sequence handling and RNA-editing analyses

Sequence handling and analysis was done using the MEGA alignment feature (Tamura et al. 2011). The PREPACT server under www.prepact.de (Lenz et al. 2010; Lenz and Knoop 2013) was used to predict and analyze RNA editing. Potential RNA-editing events predicted by at least 80% of the references in PREPACT 2.0 revealing BLASTX target hits were considered as strong candidate sites (scored as “1”). Lower predictions of ultimately confirmed editings sites were scored with their respective percentage values (0 to 0.8) to obtain cumulative sensitivity values. Specificity values were scored as the true number of edits E divided by E + the number strong false positive candidate sites. Overall values of sense-changing edits averaged over all *Selaginella uncinata* chloroplast coding regions with RNA editing were 82.5% sensitivity at a specificity of 95.1%. The cDNA mode of PREPACT was used to obtain graphic displays of editing sites as shown in Supplemental Figure 1 and to obtain the editing site list as shown in Supplemental Table 1.

Plant material and molecular work

Selaginella uncinata material in the Japanese laboratory was provided by Kunihiko Ueda (Tsuiji et al. 2007) and independently obtained from a commercial source and expert-verified by Dr. Christian Schulz (Bochum) in the German laboratory and further cultivated under high humidity conditions in the laboratory. No case of sequence conflict was observed for the different biological source materials. Total RNA was isolated via the TRI Reagent Protocol (Sigma-Aldrich) or using the RNeasy Plant Mini Kit (Qiagen) followed by on-column DNA digestion using the RNase-free DNase set.

Synthesis of cDNA was accomplished using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific/Fermentas) in the presence of random hexamer primers or with gene-specific primers using SuperScript III reverse transcriptase (Invitrogen). A complete list of oligonucleotides used for RT-PCR amplification is available upon request from the investigators. PCR products were sequenced directly as described previously (Tsuji et al. 2007) or amplified RT-PCR products were alternatively recovered from a 0.8%–2.0% agarose gel using the NucleoSpin Extract II Kit (Macherey Nagel) and cloned into the pGEM-T Easy vector (Promega).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Prof. K. Ueda for providing *S. uncinata*, Dr. Christian Schulz (Bochum) for expert examination of *Selaginella* materials, Monika Polsakiewicz (Bonn) for expert technical assistance, and Dr. Henning Lenz for providing a software tool to identify unedited silent cytidines. We also thank two anonymous reviewers for their helpful remarks to improve our manuscript.

Received March 28, 2014; accepted July 8, 2014.

REFERENCES

- Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, Depamphilis C, Albert VA, Aono N, Aoyama T, Ambrose BA, et al. 2011. The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* **332**: 960–963.
- Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, Small I. 2012. A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet* **8**: e1002910.
- Bégu D, Castandet B, Araya A. 2011. RNA editing restores critical domains of a group I intron in fern mitochondria. *Curr Genet* **57**: 317–325.
- Börner GV, Mörl M, Wissinger B, Brennicke A, Schmelzer C. 1995. RNA editing of a group II intron in *Oenothera* as a prerequisite for splicing. *Mol Gen Genet* **246**: 739–744.
- Carrillo C, Bonen L. 1997. RNA editing status of *nad7* intron domains in wheat mitochondria. *Nucleic Acids Res* **25**: 403–409.
- Carrillo C, Chapdelaine Y, Bonen L. 2001. Variation in sequence and RNA editing within core domains of mitochondrial group II introns among plants. *Mol Gen Genet* **264**: 595–603.
- Castandet B, Choury D, Bégu D, Jordana X, Araya A. 2010. Intron RNA editing is essential for splicing in plant mitochondria. *Nucleic Acids Res* **38**: 7112–7121.
- Chateigner-Boutin AL, Small I. 2011. Organellar RNA editing. *Wiley Interdiscip Rev RNA* **2**: 493–506.
- Dombrowska E, Qiu YL. 2004. Distribution of introns in the mitochondrial gene *nad1* in land plants: phylogenetic and molecular evolutionary implications. *Mol Phylogenet Evol* **32**: 246–263.
- Farré JC, Akinin C, Araya A, Castandet B. 2012. RNA editing in mitochondrial *trans*-introns is required for splicing. *PLoS One* **7**: e52644.
- Finster S, Legen J, Qu Y, Schmitz-Linneweber C. 2012. Land plant RNA editing or: Don't be fooled by plant organellar DNA sequences. In *Genomics of Chloroplasts and Mitochondria* (ed. Bock R, Knoop V), pp. 293–321. Springer, Netherlands.
- Geiss KT, Abbas GM, Makaroff CA. 1994. Intron loss from the NADH dehydrogenase subunit 4 gene of lettuce mitochondrial DNA: evidence for homologous recombination of a cDNA intermediate. *Mol Gen Genet* **243**: 97–105.
- Grewe F, Herres S, Viehöver P, Polsakiewicz M, Weisshaar B, Knoop V. 2011. A unique transcriptome: 1782 positions of RNA editing alter 1406 codon identities in mitochondrial mRNAs of the lycophyte *Isoetes engelmannii*. *Nucleic Acids Res* **39**: 2890–2902.
- Hecht J, Grewe F, Knoop V. 2011. Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of *Selaginella moellendorffii* mitochondria: the root of frequent plant mtDNA recombination in early tracheophytes. *Genome Biol Evol* **3**: 344–358.
- Hepburn NJ, Schmidt DW, Mower JP. 2012. Loss of two introns from the *Magnolia tripetala* mitochondrial *cox2* gene implicates horizontal gene transfer and gene conversion as a novel mechanism of intron loss. *Mol Biol Evol* **29**: 3111–3120.
- Ichinose M, Sugita C, Yagi Y, Nakamura T, Sugita M. 2013. Two DYW subclass PPR proteins are involved in RNA editing of *ccmFc* and *atp9* transcripts in the moss *Physcomitrella patens*: first complete set of PPR editing factors in plant mitochondria. *Plant Cell Physiol* **54**: 1907–1916.
- Itchoda N, Nishizawa S, Nagano H, Kubo T, Mikami T. 2002. The sugar beet mitochondrial *nad4* gene: an intron loss and its phylogenetic implication in the Caryophyllales. *Theor Appl Genet* **104**: 209–213.
- Iyer LM, Zhang D, Rogozin IB, Aravind L. 2011. Evolution of the deaminase fold and multiple origins of eukaryotic editing and mutagenic nucleic acid deaminases from bacterial toxin systems. *Nucleic Acids Res* **39**: 9473–9497.
- Kagami H, Nagano H, Takahashi Y, Mikami T, Kubo T. 2012. Is RNA editing implicated in group II intron survival in the angiosperm mitochondrial genome? *Genome* **55**: 75–79.
- Karol KG, Arumuganathan K, Boore JL, Duffy AM, Everett KD, Hall JD, Hansen SK, Kuehl JV, Mandoli DF, Mishler BD, 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.
- Knoop V. 2004. The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr Genet* **46**: 123–139.
- Knoop V. 2011. When you can't trust the DNA: RNA editing changes transcript sequences. *Cell Mol Life Sci* **68**: 567–586.
- Kotera E, Tasaka M, Shikanai T. 2005. A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* **433**: 326–330.
- Lenz H, Knoop V. 2013. PREPACT 2.0: predicting C-to-U and U-to-C RNA editing in organelle genome sequences with multiple references and curated RNA editing annotation. *Bioinform Biol Insights* **7**: 1–19.
- Lenz H, Rüdinger M, Volkmar U, Fischer S, Herres S, Grewe F, Knoop V. 2010. Introducing the plant RNA editing prediction and analysis computer tool PREPACT and an update on RNA editing site nomenclature. *Curr Genet* **56**: 189–201.
- Lippok B, Brennicke A, Wissinger B. 1994. Differential RNA editing in closely related introns in *Oenothera* mitochondria. *Mol Gen Genet* **243**: 39–46.
- Little DP, Moran RC, Brenner ED, Stevenson DW. 2007. Nuclear genome size in *Selaginella*. *Genome* **50**: 351–356.
- Lohse M, Drechsel O, Bock R. 2007. Organellar Genome DRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet* **52**: 267–274.
- Lopez L, Picardi E, Quagliariello C. 2007. RNA editing has been lost in the mitochondrial *cox3* and *rps13* mRNAs in Asparagales. *Biochim* **89**: 159–167.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**: 955–964.
- Ran JH, Gao H, Wang XQ. 2010. Fast evolution of the retroprocessed mitochondrial *rps3* gene in Conifer II and further evidence for the phylogeny of gymnosperms. *Mol Phylogenet Evol* **54**: 136–149.
- Rüdinger M, Polsakiewicz M, Knoop V. 2008. Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat (PPR) proteins in jungermanniid but not in marchantiid liverworts. *Mol Biol Evol* **25**: 1405–1414.

- Rüdinger M, Volkmar U, Lenz H, Groth-Malonek M, Knoop V. 2012. Nuclear DYW-type PPR gene families diversify with increasing RNA editing frequencies in liverwort and moss mitochondria. *J Mol Evol* **74**: 37–51.
- Salone V, Rüdinger M, Polsakiewicz M, Hoffmann B, Groth-Malonek M, Szurek B, Small I, Knoop V, Lurin C. 2007. A hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Lett* **581**: 4132–4138.
- Schallenberg-Rüdinger M, Kindgren P, Zehrmann A, Small I, Knoop V. 2013. A DYW-protein knockout in *Physcomitrella* affects two closely spaced mitochondrial editing sites and causes a severe developmental phenotype. *Plant J* **76**: 420–432.
- Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* **33**: W686–W689.
- Smith DR. 2009. Unparalleled GC content in the plastid DNA of *Selaginella*. *Plant Mol Biol* **71**: 627–639.
- Sugita M, Ichinose M, Ide M, Sugita C. 2013. Architecture of the PPR gene family in the moss *Physcomitrella patens*. *RNA Biol* **10**: 1256–1262.
- Takenaka M, Zehrmann A, Brennicke A, Graichen K. 2013. Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. *PLoS One* **8**: e65343.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.
- Tsuji S, Ueda K, Nishiyama T, Hasebe M, Yoshikawa S, Konagaya A, Nishiuchi T, Yamaguchi K. 2007. The chloroplast genome from a lycophyte (microphylophyte), *Selaginella uncinata*, has a unique inversion, transpositions and many gene losses. *J Plant Res* **120**: 281–290.
- Wolf PG, Karol KG, Mandoli DF, Kuehl J, Arumuganathan K, Ellis MW, Mishler BD, Kelch DG, Olmstead RG, Boore JL. 2005. The first complete chloroplast genome sequence of a lycophyte, *Huperzia lucidula* (Lycopodiaceae). *Gene* **350**: 117–128.
- Yagi Y, Hayashi S, Kobayashi K, Hirayama T, Nakamura T. 2013. Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PLoS One* **8**: e57286.
- Zehrmann A, Verbitskiy D, van der Merwe JA, Brennicke A, Takenaka M. 2009. A DYW domain-containing pentatricopeptide repeat protein is required for RNA editing at multiple sites in mitochondria of *Arabidopsis thaliana*. *Plant Cell* **21**: 558–567.