SCIENCE CHINA
Life Sciences

THEMATIC ISSUE: Evolution of genes and genomes
•RESEARCH PAPER•

April 2019 Vol.62 No.4: 498–506
https://doi.org/10.1007/s11427-018-9450-1

Complete loss of RNA editing from the plastid genome and most highly expressed mitochondrial genes of *Welwitschia mirabilis*

Weishu Fan[1,2,3], Wenhu Guo[2,4], Lexis Funk[2,4], Jeffrey P. Mower[2,3*] & Andan Zhu[1*]

[1]*Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China;*
[2]*Center for Plant Science Innovation, University of Nebraska, Lincoln, NE 68588, USA;*
[3]*Department of Agronomy and Horticulture, University of Nebraska, Lincoln, NE 68583, USA;*
[4]*School of Biological Sciences, University of Nebraska, Lincoln, NE 68588, USA*

Received September 10, 2018; accepted October 2, 2018; published online March 7, 2019

Comparative genomics among gymnosperms suggested extensive loss of mitochondrial RNA editing sites from *Welwitschia mirabilis* based on predictive analysis. However, empirical or transcriptome data to confirm this massive loss event are lacking, and the potential mechanisms of RNA site loss are unclear. By comparing genomic sequences with transcriptomic and reverse-transcription PCR sequencing data, we performed a comprehensive analysis of the pattern of RNA editing in the mitochondrial and plastid genomes (mitogenome and plastome, respectively) of *W. mirabilis* and a second gymnosperm, *Ginkgo biloba*. For *W. mirabilis*, we found only 99 editing sites located in 13 protein-coding genes in the mitogenome and a complete loss of RNA editing from the plastome. The few genes having high editing frequency in the *Welwitschia* mitogenome showed a strong negative correlation with gene expression level. Comparative analyses with *G. biloba*, containing 1,405 mitochondrial and 345 plastid editing sites, revealed that the editing loss from *W. mirabilis* is mainly due to the substitution of editable cytidines to thymidines at the genomic level, which could be caused by retroprocessing. Our result is the first study to uncover massive editing loss from both the mitogenome and plastome in a single genus. Furthermore, our results suggest that gene expression level and retroprocessing both contributed to the evolution of RNA editing in plant organellar genomes.

RNA editing, massive loss, expression levels, organelle genomes, *Welwitschia*

Citation: Fan, W., Guo, W., Funk, L., Mower, J.P., and Zhu, A. (2019). Complete loss of RNA editing from the plastid genome and most highly expressed mitochondrial genes of *Welwitschia mirabilis*. Sci China Life Sci 62, 498–506. https://doi.org/10.1007/s11427-018-9450-1

INTRODUCTION

In the mitochondrial genome (mitogenome) and plastid genome (plastome) of land plants, the posttranscriptional and site-specific conversion of cytidines to uridines (C-to-U) is named RNA editing, with some lineages also performing U-to-C editing (Covello and Gray, 1989; Hein et al., 2016; Hiesel et al., 1989; Shikanai, 2006). C-to-U RNA editing has been observed in all major land plant groups (Chaw et al., 2008; Grewe et al., 2011; Oda et al., 1992; Unseld et al., 1997), suggesting that it originated in the common ancestor of land plants. The mechanism of this process is still unclear, although it requires a large number of site-specificity factors, including PPR proteins, some of which may also provide catalytic activity (Schallenberg-Rüdinger and Knoop, 2016).

The prevalence of mitochondrial RNA editing varies substantially across lineages. It is rare in the moss *Physcomitrella patens*, which harbors only 11 C-to-U edit sites (Rüdinger et al., 2009), and completely absent from the liverwort *Marchantia polymorpha* (Rüdinger et al., 2008), whereas it is most abundant in lycophytes, with over 2,152 sites in the spike moss *Selaginella moellendorffii* (Hecht et al., 2011) and 1,782 positions in the quillwort *Isoetes en-*

*Corresponding authors (Jeffrey P. Mower, email: jpmower@unl.edu; Andan Zhu, email: zhuandan@mail.kib.ac.cn)

gelmannii (Grewe et al., 2011). In gymnosperms, predictive analysis identified 1,214 sites in *Cycas taitungensis* and 1,306 sites in *Ginkgo biloba* (Guo et al., 2016), which was largely corroborated by experimental analysis of a subset of sites via RT-PCR (Chaw et al., 2008; Guo et al., 2016). RNA editing occurs somewhat less frequently in ferns, with 1,014 C-to-U and 58 U-to-C sites in *Ophioglossum californicum*, and 965 C-to-U and 19 U-to-C sites in *Psilotum nudum* (Guo et al., 2017). Editing is even lower and solely C-to-U in angiosperms, from 189 editing sites in *Silene noctiflora* (Sloan et al., 2010) to 835 sites in *Amborella trichopoda* (Rice et al., 2013).

RNA editing frequency in the plastome is generally less, and often substantially so, than in the mitogenome, but it also shows great disparity among different lineages (Chateigner-Boutin and Small, 2011; Kugita, 2003). In lycophytes, an extraordinary number of 3,415 C-to-U RNA editing sites was ascertained from *Selaginella uncinata* (Oldenkott et al., 2014). Among ferns, editing content ranges from hundreds of sites in *Adiantum capillus-veneris* and *O. californicum* to only 27 sites in *P. nudum* and an apparent absence of any RNA editing in *Equisetum hyemale* (Guo et al., 2015; Knie et al., 2016; Wolf et al., 2004). Plastid editing among gymnosperms is also variable, ranging from 26 to 255 sites (Chen et al., 2011; Hecht et al., 2011; Wolf et al., 2004). In flowering plants, 20–60 C-to-U editing events are usually identified in most plastomes (Hirose et al., 1999; Ichinose and Sugita, 2016), although recent reports in *Phalaenopsis aphrodite* (Chen et al., 2017) and *Amborella trichopoda* (Hein et al., 2016) have identified over 100.

The underlying mechanistic cause of variation in RNA editing frequency has been often attributed to lineage-specific loss, either via retroprocessing, accelerated substitution rates, or selection against the retention of edit sites (Lynch, 2006; Mower, 2008; Parkinson et al., 2005). *Welwitschia mirabilis*, the sole member in the Welwitschiaceae (hereafter referred to as *Welwitschia*), is an intriguing candidate to explore this issue because its mitogenome and plastome have accelerated substitution rates (Drouin et al., 2008; Guo et al., 2016; McCoy et al., 2008), and the mitogenome was predicted to have lost most RNA editing sites relative to other gymnosperms (Guo et al., 2016). To address these questions, we examined the RNA editing in mitochondrial and plastid genomes of *Welwitschia*, and compared these results to those from a second gymnosperm, *G. biloba* (hereafter referred to as *Ginkgo*), which has slow organellar substitution rates and evidence for substantial RNA editing.

RESULTS

Transcriptional capacity of *Welwitschia* organelle genomes

We constructed a strand-specific RNA-seq library to obtain expression signatures of *Welwitschia* organelle genomes, generating approximately 55.4 million clean reads. After applying a customized filter, a total of 48.2 M clean reads were used for mapping to the *Welwitschia* mitogenome (NC_029130) and plastome (KT347148) that we previously reported, of which 2.3 M (4.77%) read pairs mapped to the *Welwitschia* mitogenome and 16.9 M pairs (35.06%) mapped to the plastome. Coverage for both genomes was extensive, particularly at genic regions, consistent with functional expectations (Figure S1 in Supporting Information).

Mitochondrial RNA editing is very limited in *Welwitschia* but abundant in *Ginkgo*

Based on the RNA-seq data mapped to the *Welwitschia* mitogenome, a total of 134 C-to-U editing sites were identified (Table S1 in Supporting Information). Among these, 99 sites (73.9%) are distributed in 13 protein-coding genes, while the remaining editing sites either distributed within 1 kb of coding regions (11.9%) or located in intergenic regions (14.2%). We found that 88 amino acids were changed and one stop codon was created by editing. Of the 29 annotated protein-coding genes in the *Welwitschia* mitogenome (Guo et al., 2016), more than half are completely devoid of any editing and another four genes contain only a single editing site in their coding regions (Table 1). To validate the editing sites in the *Welwitschia* mitogenome, RT-PCR was also used to amplify all coding regions. This approach identified 45 editing sites in nine genes, which is comparatively less than that of the RNA-seq based method (Table S2 in Supporting Information). In total, 57 of the 99 (57.6%) coding-located editing sites identified by RNA-Seq method were not found via RT-PCR, presumably due to the lower editing efficiency of these sites that may hinder detection by RT-PCR.

To compare the *Welwitschia* mitogenome editing pattern with another gymnosperm, we mapped RNA-seq data from *Ginkgo* to its mitogenome. A total of 1,405 C-to-U editing sites were detected in the *Ginkgo* mitogenome, approximately 10-fold more than that in the *Welwitschia* mitogenome (Figure 1). The majority of sites (1,162/1,405) affect coding regions, and RNA editing was found in all 41 protein-coding genes (Table 1; Table S3 in Supporting Information). Annotation of these editing sites indicated that 73.2%, 8.7% and 1.0% C-to-U editing events caused non-synonymous, synonymous and nonsense changes, respectively.

Interestingly, the *Welwitschia* and *Ginkgo* mitogenomes only share 44 editing sites in the protein-coding regions, representing 44.4% of the total editing sites in *Welwitschia* (Figure 2A). For the non-overlapped portion of the other 55 editing sites from *Welwitschia*, the corresponding sites in *Ginkgo* included 38 C, 15 T and two A sites, implying that most of the unique editing sites in *Welwitschia* were retained as unedited cytidines in *Ginkgo* (Figure 2B). In contrast, the

Table 1 Summary of *Ginkgo* and *Welwitschia* mitogene editing sites

	Gene	*Ginkgo*			*Welwitschia*		
		Length	Editing number	Frequency	Length	Editing number	Frequency
ATP synthase F1 subunits	*atp1*	1,527	23	0.0151	1,569	0	0
	atp4	594	27	0.0455	594	0	0
	atp6	768	63	0.0820	768	0	0
	atp8	489	19	0.0389	510	0	0
	atp9	225	21	0.0933	225	0	0
	Subtotal	3,603	153	0.0425	3,666	0	0
Cytochrome c biogenesis	*ccmB*	639	21	0.0329	636	19	0.0299
	ccmC	711	45	0.0633	741	3	0.0040
	ccmFc	1,386	14	0.0101	1,566	11	0.0070
	ccmFn	1,815	35	0.0193	1,893	0	0
	Subtotal	4,551	115	0.0253	4,836	33	0.0068
Cytochrome b	*cob*	1,197	3	0.0025	1,143	0	0
	Subtotal	1,197	3	0.0025	1,143	0	0
Cytochrome c oxidase subunits	*cox1*	1,563	3	0.0019	1,572	0	0
	cox2	765	27	0.0353	801	0	0
	cox3	798	46	0.0576	801	0	0
	Subtotal	3,126	76	0.0243	3,174	0	0
Other proteins	*matR*	2,034	31	0.0152	2,676	14	0.0052
	mttB	780	37	0.0474	798	10	0.0125
	Subtotal	2,814	68	0.0242	3,474	24	0.0069
NADH dehydrogenase subunits	*nad1*	978	38	0.0389	981	2	0.0020
	nad2	1,473	57	0.0387	1,479	0	0
	nad3	357	7	0.0196	351	0	0
	nad4	1,488	103	0.0692	1,488	1	0.0007
	nad4L	303	20	0.0660	372	0	0
	nad5	2,004	114	0.0569	2,070	2	0.0010
	nad6	606	47	0.0776	636	22	0.0346
	nad7	1,185	45	0.0380	1,185	9	0.0076
	nad9	594	25	0.0421	612	0	0
	Subtotal	8,988	456	0.0507	9,174	36	0.0039
Large ribosomal proteins	*rpl2*	1,440	13	0.0090	–	–	–
	rpl5	570	17	0.0298	–	–	–
	rpl10	483	12	0.0248	549	1	0.0018
	rpl16	411	11	0.0268	–	–	–
	Subtotal	2,904	53	0.0183	549	1	0.0018
Small ribosomal proteins	*rps1*	594	27	0.0455	–	–	–
	rps2	696	9	0.0129	–	–	–
	rps3	1,737	33	0.0190	1,704	1	0.0006
	rps4	1,041	47	0.0451	615	4	0.0065
	rps7	621	18	0.0290	–	–	–
	rps10	324	2	0.0062	–	–	–
	rps11	519	19	0.0366	–	–	–
	rps12	372	18	0.0484	369	0	0
	rps13	381	2	0.0052	–	–	–
	rps14	303	11	0.0363	–	–	–
	rps19	282	13	0.0461	–	–	–
	Subtotal	6,870	199	0.0290	2,688	5	0.0019
Complex II	*sdh3*	402	23	0.0572	–	–	–
	sdh4	396	16	0.0404	426	0	0
	Subtotal	798	39	0.0489	426	0	0
Total		34,851	1,162	0.0333	29,130	99	0.0034

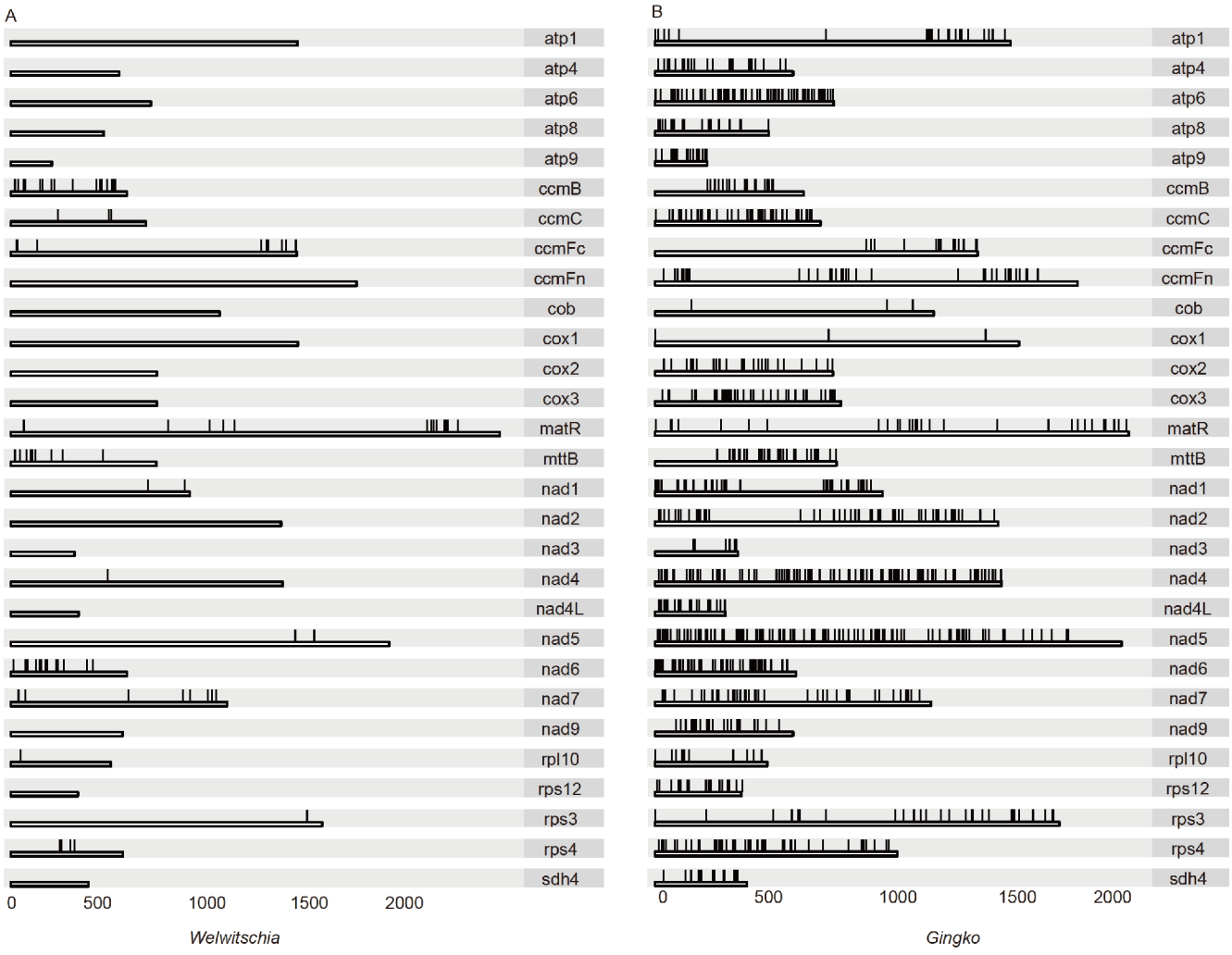

Figure 1 RNA editing site distribution in 29 shared protein-coding genes in *Welwitschia* and *Ginkgo* mitogenomes. Vertical lines indicate RNA editing sites and horizontal lines represent gene length. Selected gene names are shown on the right and the length bar shown at the bottom.

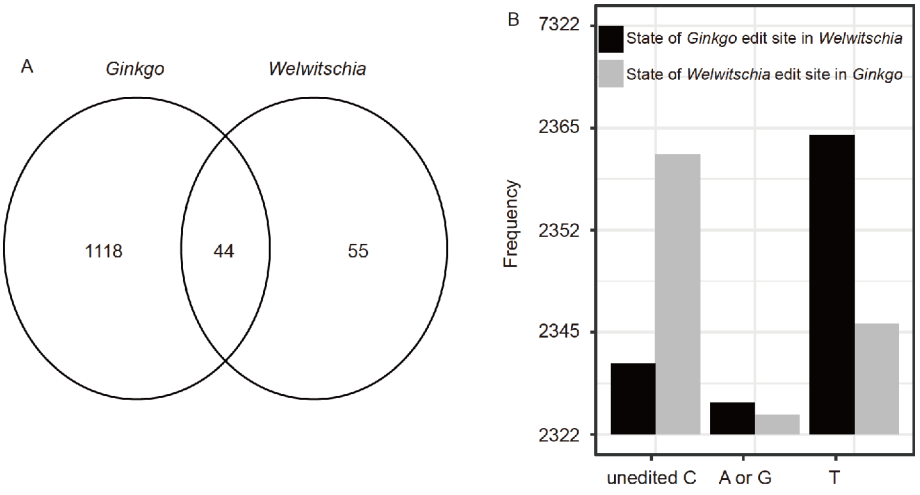

Figure 2 Comparative evolutionary analysis of editing sites in *Welwitschia* and *Ginkgo* mitochondrial genes. A, Venn diagram depicts the numbers of conserved and unique editing sites. B, State of the unique *Ginkgo* editing sites in *Welwitschia* and the unique *Welwitschia* editing sites in *Ginkgo*. The relative frequency of retaining cytosine (C-C) and thymidine (C-T) and others (C-R, R=A, G) substitutions are shown.

unique editing sites in *Ginkgo* (1,118 sites) revealed the opposite pattern, in which the majority of the unique edit sites in *Ginkgo* (74.3%) were present as thymidines in the *Welwitschia* mitogenome. The result indicated that the massive loss of RNA editing sites in the *Welwitschia* mitogenome is mainly attributed to the substitution of editing sites by thymidines.

The substantial variance of the mitochondrial editing number between *Welwitschia* and *Ginkgo* is not only due to the loss of many mitochondrial protein-coding genes from *Welwitschia* (Guo et al., 2016), but also due to the reduced editing frequency among genes. For each mitochondrial protein-coding gene, *Ginkgo* shows higher editing frequency compared with *Welwitschia* (Table 1). Considering the total length of each gene, the highest edited frequency occurred in *nad* genes in the *Ginkgo* mitogenome with a frequency of 0.0507, which is more than 10 times higher than that in *Welwitschia*. Additionally, the *atp*, *cox* and *sdh* genes, which also have relatively high editing frequency in *Ginkgo* (0.0425, 0.0243 and 0.0489, respectively), are completely unedited in *Welwitschia*. The total number of editing sites in the *Welwitschia* mitogenome is the lowest of all reported seed plants.

Retention of RNA editing sites in *Welwitschia* low-expressed mitochondrial genes

RNA-Seq data can be used to distinguish whether a specific site has been fully or partially edited. Since the majority of the editing sites in *Welwitschia* were partially edited (Table S1 in Supporting Information), we predicted that the retention of editing sites could be intrinsically tied to gene expression level. We measured the number of RNA-Seq reads mapped to the gene per million reads mapped to the whole genome, and studied the relationship between gene expression level and editing frequency. Regression analysis indicated these two features were significantly (P<0.05) negatively correlated (Figure 3). Genes with relatively high editing frequency tend to be low expressed in the mitogenome. For example, the *ccmB* gene exhibited the second highest editing frequency and the lowest expression level among all the protein-coding genes.

Complete elimination of RNA editing from the *Welwitschia* plastome

The dramatic reduction of RNA editing in the *Welwitschia* mitogenome prompted us to examine the level of editing in the *Welwitschia* plastome. In fact, we did not find any evidence of plastid RNA editing in *Welwitschia*. In contrast, previous studies identified 255 editing sites in *Ginkgo* chloroplast protein-coding genes (He et al., 2016) and at least two more in untranslated regions (Kudla and Bock, 1999),

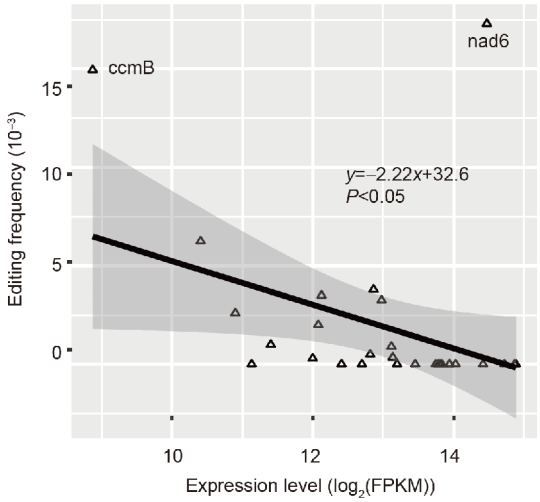

Figure 3 The relationship between gene expression level and editing frequency in the *Welwitschia* mitogenome. Triangles represent protein-coding genes. The shade area represents the 95% confidence interval.

suggesting a relatively high level of editing. To cross-check these RT-PCR based results and obtain a more comprehensive picture of RNA editing distribution in this genus, we mapped Illumina RNA-Seq reads to the *Ginkgo* plastome. This reanalysis identified a total of 345 edited sites (Table S3 in Supporting Information), confirming a widespread distribution of editing sites in the *Ginkgo* plastome. Among these, 286 (82.9%) editing sites were within protein-coding regions, essentially altering 258 amino acids, while seven editing sites restored or generated stop codons. Outside of the coding regions, another 47, four and one editing site were distributed in UTR and intergenic regions, respectively (Table S3 in Supporting Information).

The *Welwitschia* plastome is very compact, containing only 66 protein-coding genes (McCoy et al., 2008). Assuming similar conserved editing sites were present in the *Ginkgo-Welwitschia* ancestor, the loss of all *ndh* and *chl* genes as well as various other genes could account for approximately one-third (35.5%) of ancestral plastid editing sites from *Welwitschia*. Of the retained genes in *Welwitschia*, edited sites in *Ginkgo* plastid genes were found to be converted at the genomic level to T (84.5%), A (3.3%) or G (1.7%), while the remaining sites were present as an unedited C (8.9%) in the *Welwitschia* plastome (Figure S2 in Supporting Information). To our knowledge, this is the first case of complete elimination of RNA editing from a seed plant organelle genome.

DISCUSSION

The post-transcriptional process of RNA editing is a prevalent phenomenon in land plant organellar transcripts. Although it has been observed in all major plant lineages, the

number of C-to-U RNA editing sites is variable. Even within the gymnosperm clade, editing frequency shows substantial divergence. A previous study on the mitogenomes of gymnosperms predicted a dramatic reduction of RNA editing in *Welwitschia* relative to the ancestrally high level of editing in other gymnosperms (Guo et al., 2016). In this study, we used both RNA-Seq and RT-PCR data to verify that RNA editing frequency is extremely low in *Welwitschia* compared with *Ginkgo*. Within the 29 functional protein-coding genes in the *Welwitschia* mitogenome, RNA editing sites were detected from only 13 of them, which implied that more than half (55.17%) of the genes lost RNA editing completely. RNA editing loss from a single gene has been reported in seed plants, such as *cox3* and *rps13* that lacked RNA editing in Iridaceae and Amaryllidaceae (Lopez et al., 2007). However, such a massive RNA editing loss from most of the genes in a seed plant mitogenome has not been reported yet. The RT-PCR results were largely consistent but much less sensitive compared with RNA-Seq data, as approximately 40% of the editing sites were not detected through RT-PCR. This is presumably related to editing efficiency such that the RT-PCR method may fail to identify partial editing sites that are edited at low efficiency (Mower and Palmer, 2006; Schuster et al., 1990).

Although previous studies showed a diverse number of editing sites across gymnosperms, such as *Ginkgo* (284 sites), *Pinus thunbergii* (26 sites in 12 transcripts) and *Cycas taitungensis* (85 sites in 25 transcripts) (Chen et al., 2011; He et al., 2016; Wakasugi et al., 1996), we were surprised to discover that the *Welwitschia* plastid transcriptome contained no editing sites. Our recent study on a large scale of gymnosperm samplings also found RNA editing in all other representative clades within gymnosperms (Fan et al., unpublished). Therefore, the loss of editing from the *Welwitschia* plastome is an extraordinary case among gymnosperms.

Our study showed that the RNA editing sites in *Welwitschia* are dramatically reduced or completely lost compared with *Ginkgo* in both the mitogenome and plastome. A reduction of editing during angiosperm evolution has also been reported for many angiosperm mitogenomes (Edera et al., 2018; Mower, 2008; Shields and Wolfe, 1997; Sloan et al., 2010). Given that other gymnosperms were predicted or experimentally confirmed to have high levels of RNA editing, the absence of so many sites from the *Welwitschia* mitogenome is most consistent with extensive loss. In our study, comparative analyses uncovered that the majority of the editing cytidine in *Ginkgo* were present as thymidine in *Welwitschia* (representing 73.2% in mitogenome and 84.5% in plastome), indicating that most editing loss occurred by a genomic mutation from C to T (Figure 2; Figure S2 in Supporting Information). The reasons for these specific site substitutions could be point mutation or retroprocessing (Grewe et al., 2011; Mower, 2008; Ran et al., 2010; Sloan et al., 2010). An RNA-mediated gene conversion (retroprocessing) model has been proposed to explain some of the RNA editing loss, resulting in preferential C-to-T substitution at RNA editing sites. If the retroprocessing is the driving force for RNA editing loss, genes with increased transcript level would facilitate retroprocessing events. In this study, by applying the deep sequencing method, we tested if the transcription level is correlated with editing frequency. Our result provided evidence in support of this hypothesis because genes with high expression level contain fewer editing sites (Figure 3; Figure S1 in Supporting Information). This is also the possible reason for complete editing loss in *Welwitschia* plastome as the plastid genes are expressed in a very high level (Figure S1 in Supporting Information). The retention of RNA editing sites in low-expressed genes also suggests the nonadaptive role of the process, which is consistent with the point of view that RNA editing is a source of genetic variation, instead of restoring DNA mutations (Landweber and Gilbert, 1993).

Considering that RNA editing is supposed to be a correction mechanism to correct T-to-C mutations in organelle genomes, our results are consistent with this hypothesis since most of the RNA editing sites in *Ginkgo* were converted to thymidine in *Welwitschia* (Figure 1B; Figure S2 in Supporting Information). In contrast, the *Welwitschia* mitogenome and plastome seem to have an unusual editing pattern, given that most of the corresponding sites in *Ginkgo* are not edited or a thymidine. This suggests that many of these edited sites in *Welwitschia* are not evolutionarily conserved and may not be functionally important. These non-conserved edit sites may suggest a change of site-specific recognition factors, such as PPR (pentatricopeptide repeat) proteins that selectively bind upstream of the cytidines to be edited in the transcript (Cheng et al., 2016; Lurin et al., 2004). Loss of PPR proteins may also play a role in driving the loss of so many RNA editing sites in the *Welwitschia* organelles.

MATERIALS AND METHODS

Plant material and RNA extraction

The *Welwitschia mirabilis* plant used for RNA extraction was grown in the Beadle Center greenhouse at the University of Nebraska-Lincoln, which is the same individual previously used for mitochondrial and chloroplast genome sequencing (Guo et al., 2016). Fresh leaf tissues were collected right before RNA extraction. Total RNA was isolated using TRIzol (Life Technologies Corporation, USA) according to procedures suggested by the manufacturer. The isolated RNA was stored at –80°C until use for either transcriptome sequencing or RT-PCR analysis.

Strand-specific RNA-Seq library construction and transcriptome sequencing

To capture genome-wide transcriptional signals in *Welwitschia* organellar genomes, a strand-specific RNA library was constructed with the NEBNext Ultra Directional RNA Library Prep Kit (New England Biolabs) following manufacturer instructions and sequenced with an Illumina HiSeq 2500 system (BGI, Hong Kong) using 150-bp paired-end reads. Raw reads were deposited in Genbank under SAMN10352752.

Identification of putative editing sites

To perform quality and adapter trimming for *Welwitschia* paired end transcriptome data, the Trimmomatic software (Bolger et al., 2014) was used with the following parameters: -phred33, LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:145. Trimmed data were mapped to the *Welwitschia* mitogenome (accession number: NC_029130) and plastome (KT347148) using Bowtie2 version 2.2.8, allowing as much as nine mismatches (Langmead and Salzberg, 2012). The SAMtools mpileup version 1.3 (Li et al., 2009) was used to identify the mismatched sites between mapped transcript reads and the reference genome sequences. Further gene-based and region-based annotations of SNPs were implemented by using ANNOVAR (Wang et al., 2010). All C:T or G:A (in the antisense strand) mismatches were accepted as putative editing sites for false-positive filtering steps.

For comparison, a deeply sequenced *Ginkgo* RNA-Seq library (SRX2527274) was downloaded from the NCBI sequence read archive (SRA) database. Clean reads were mapped to the *Ginkgo* mitogenome (NC_027976) and plastome (KP099648) using the procedures described above.

Filtering false positives and identifying fully or partially edited sites

To reduce false positives due to amplification bias, sequencing and mapping errors, the following filters were applied during mapping and annotation steps, requiring: (i) reads to map uniquely in paired mode; (ii) a variant call quality $q>20$; (iii) depth of variant-supporting bases $\geqslant 4$; (iv) editing sites in the coding and up-/downstream regions of the same gene belong to one strand. The editing efficiency was determined by ANNOVAR annotation. We considered homozygous variants (1/1; $\geqslant 80\%$) to be highly edited, and heterozygous variants (0/1; $<80\%$) to be partially edited.

Validation of editing sites using reverse-transcription PCR

Identified RNA editing sites in *Welwitschia* mitochondrial genes were further verified by RT-PCR. Briefly, first-strand cDNA synthesis was prepared from isolated RNA primed with random hexamers. RT-PCR assays were performed using first-strand cDNA as template. Primers used to amplify all the protein-coding genes of *Welwitschia* mitogenome were designed, excluding regions that contain editing sites (Table S4 in Supporting Information) (Mower and Palmer, 2006). Additional primer sets were designed for genes longer than 1 kb (*ccmFn*, *matR*, *rps3*), and alternative primers were designed as the upstream or downstream sequence if no best optimized sequence was found (*mttB*) (Table S4 in Supporting Information). The RT-PCR program settings were used as described before (Hepburn et al., 2012) with optimized annealing temperature.

Sequence alignments and comparative analyses of RNA editing sites

To analyze the evolution of editing sites in gymnosperms, organellar protein-coding genes from *Ginkgo* and *Welwitschia* were individually aligned at the amino acid level with MUSCLE version 3.8.31 (Edgar, 2004) and constrained to their nucleotide alignments.

Correlation between expression level and editing frequency

RNA-Seq mapping data were used to determine gene-specific expression levels in fragments per kilobase per million (FPKM) using cufflinks v2.2.1 (Trapnell et al., 2012), and to calculate per-base coverage across transcriptional regions with BEDTools genomecov (Quinlan, 2014). The editing frequency per gene was calculated by normalizing the number of editing sites with gene length. The correlation between expression levels and editing frequencies were calculated using "cor.test" function in R, and plotted with ggplot2 (Wickham, 2016).

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *We thank Amy Hilske and Samantha Link for care of Welwitschia plants in the Beadle Center greenhouse. This research was supported by the National Natural Science Foundation of China (31801051 to WF), by the Large-scale Scientific Facilities of the Chinese Academy of Sciences (2017-LSF-GBOWS-02) and by CAS Pioneer Hundred Talents Program (to AZ).*

References

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.
Chateigner-Boutin, A.L., and Small, I. (2011). Organellar RNA editing. WIREs RNA 2, 493–506.
Chaw, S.M., Chun-Chieh Shih, A., Wang, D., Wu, Y.W., Liu, S.M., and Chou, T.Y. (2008). The mitochondrial genome of the gymnosperm

Cycas taitungensis contains a novel family of short interspersed elements, Bpu sequences, and abundant RNA editing sites. Mol Biol Evol 25, 603–615.
Chen, H., Deng, L., Jiang, Y., Lu, P., and Yu, J. (2011). RNA editing sites exist in protein-coding genes in the chloroplast genome of *Cycas taitungensis*. J Integr Plant Biol 53, 961–970.
Chen, T.C., Liu, Y.C., Wang, X., Wu, C.H., Huang, C.H., and Chang, C.C. (2017). Whole plastid transcriptomes reveal abundant RNA editing sites and differential editing status in *Phalaenopsis aphrodite* subsp. *formosana*. Bot Stud 58, 38.
Cheng, S., Gutmann, B., Zhong, X., Ye, Y., Fisher, M.F., Bai, F., Castleden, I., Song, Y., Song, B., Huang, J., et al. (2016). Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. Plant J 85, 532–547.
Covello, P.S., and Gray, M.W. (1989). RNA editing in plant mitochondria. Nature 341, 662–666.
Drouin, G., Daoud, H., and Xia, J. (2008). Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. Mol PhyloGenets Evol 49, 827–831.
Edera, A.A., Gandini, C.L., and Sanchez-Puerta, M.V. (2018). Towards a comprehensive picture of C-to-U RNA editing sites in angiosperm mitochondria. Plant Mol Biol 97, 215–231.
Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792–1797.
Grewe, F., Herres, S., Viehöver, P., Polsakiewicz, M., Weisshaar, B., and 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.
Guo, W., Grewe, F., Fan, W., Young, G.J., Knoop, V., Palmer, J.D., and Mower, J.P. (2016). *Ginkgo* and *Welwitschia* mitogenomes reveal extreme contrasts in gymnosperm mitochondrial evolution. Mol Biol Evol 33, 1448–1460.
Guo, W., Grewe, F., and Mower, J.P. (2015). Variable frequency of plastid RNA editing among ferns and repeated loss of uridine-to-cytidine editing from vascular plants. PLoS ONE 10, e0117075.
He, P., Huang, S., Xiao, G., Zhang, Y., and Yu, J. (2016). Abundant RNA editing sites of chloroplast protein-coding genes in *Ginkgo biloba* and an evolutionary pattern analysis. BMC Plant Biol 16, 257.
Hecht, J., Grewe, F., and 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.
Hein, A., Polsakiewicz, M., and Knoop, V. (2016). Frequent chloroplast RNA editing in early-branching flowering plants: pilot studies on angiosperm-wide coexistence of editing sites and their nuclear specificity factors. BMC Evol Biol 16, 23.
Hepburn, N.J., Schmidt, D.W., and Mower, J.P. (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.
Hiesel, R., Wissinger, B., Schuster, W., and Brennicke, A. (1989). RNA editing in plant mitochondria. Science 246, 1632–1634.
Hirose, T., Kusumegi, T., Tsudzuki, T., and Sugiura, M. (1999). RNA editing sites in tobacco chloroplast transcripts: editing as a possible regulator of chloroplast RNA polymerase activity. Mol Gen Genets MGG 262, 462–467.
Ichinose, M., and Sugita, M. (2017). RNA editing and its molecular mechanism in plant organelles. Genes 8, 5.
Knie, N., Grewe, F., and Knoop, V. (2016). Monilophyte mitochondrial *rps1* genes carry a unique group II intron that likely originated from an ancient paralog in *rpl2*. RNA 22, 1338–1348.
Kudla, J., and Bock, R. (1999). RNA editing in an untranslated region of the *Ginkgo* chloroplast genome. Gene 234, 81–86.
Kugita, M. (2003). RNA editing in hornwort chloroplasts makes more than half the genes functional. Nucleic Acids Res 31, 2417–2423.
Landweber, L.F., and Gilbert, W. (1993). RNA editing as a source of genetic variation. Nature 363, 179–182.
Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.
Lopez, L., Picardi, E., and Quagliariello, C. (2007). RNA editing has been lost in the mitochondrial *cox3* and *rps13* mRNAs in Asparagales. Biochimie 89, 159–167.
Lurin, C., Andrés, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyère, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., et al. (2004). Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16, 2089–2103.
Lynch, M. (2006). Streamlining and simplification of microbial genome architecture. Annu Rev Microbiol 60, 327–349.
McCoy, S.R., Kuehl, J.V., Boore, J.L., and Raubeson, L.A. (2008). The complete plastid genome sequence of *Welwitschia mirabilis*: an unusually compact plastome with accelerated divergence rates. BMC Evol Biol 8, 130.
Mower, J.P. (2008). Modeling sites of RNA editing as a fifth nucleotide state reveals progressive loss of edited sites from angiosperm mitochondria. Mol Biol Evol 25, 52–61.
Mower, J.P., and Palmer, J.D. (2006). Patterns of partial RNA editing in mitochondrial genes of *Beta vulgaris*. Mol Genet Genomics 276, 285–293.
Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y., Kohchi, T., et al. (1992). Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. J Mol Biol 223, 1–7.
Oldenkott, B., Yamaguchi, K., Tsuji-Tsukinoki, S., Knie, N., and Knoop, V. (2014). Chloroplast RNA editing going extreme: more than 3400 events of C-to-U editing in the chloroplast transcriptome of the lycophyte *Selaginella uncinata*. RNA 20, 1499–1506.
Parkinson, C.L., Mower, J.P., Qiu, Y.L., Shirk, A.J., Song, K., Young, N.D., DePamphilis, C.W., and Palmer, J.D. (2005). Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. BMC Evol Biol 5, 73.
Quinlan, A.R. (2014). BEDTools: The Swiss-Army tool for genome feature analysis. Curr Protocols BioInf 47, 11.12.1–11.12.34.
Ran, J.H., Gao, H., and Wang, X.Q. (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.
Rice, D.W., Alverson, A.J., Richardson, A.O., Young, G.J., Sanchez-Puerta, M.V., Munzinger, J., Barry, K., Boore, J.L., Zhang, Y., dePamphilis, C.W., et al. (2013). Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. Science 342, 1468–1473.
Rüdinger, M., Funk, H.T., Rensing, S.A., Maier, U.G., and Knoop, V. (2009). RNA editing: only eleven sites are present in the *Physcomitrella patens* mitochondrial transcriptome and a universal nomenclature proposal. Mol Genet Genomics 281, 473–481.
Rüdinger, M., Polsakiewicz, M., and Knoop, V. (2008). Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat proteins in jungermanniid but not in marchantiid liverworts. Mol Biol Evol 25, 1405–1414.
Schallenberg-Rüdinger, M., and Knoop, V. (2016). Chapter Two - Coevolution of Organelle RNA Editing and Nuclear Specificity Factors in Early Land Plants. In Advances in Botanical Research. Vol. 78, S.A. Rensing, ed. (Academic Press), pp. 37-93.
Schuster, W., Wissinger, B., Unseld, M., and Brennicke, A. (1990). Transcripts of the NADH-dehydrogenase subunit 3 gene are differentially edited in *Oenothera* mitochondria. EMBO J 9, 263–269.
Shields, D.C., and Wolfe, K.H. (1997). Accelerated evolution of sites undergoing mRNA editing in plant mitochondria and chloroplasts. Mol Biol Evol 14, 344–349.

Shikanai, T. (2006). RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci 63, 698–708.
Sloan, D.B., MacQueen, A.H., Alverson, A.J., Palmer, J.D., and Taylor, D. R. (2010). Extensive loss of RNA editing sites in rapidly evolving Silene mitochondrial genomes: selection vs. retroprocessing as the driving force. Genetics 185, 1369–1380.
Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7, 562–578.
Unseld, M., Marienfeld, J.R., Brandt, P., and Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nat Genet 15, 57–61.
Wakasugi, T., Hirose, T., Horihata, M., Tsudzuki, T., Kossel, H., and Sugiura, M. (1996). Creation of a novel protein-coding region at the RNA level in black pine chloroplasts: the pattern of RNA editing in the gymnosperm chloroplast is different from that in angiosperms. Proc Natl Acad Sci USA 93, 8766–8770.
Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38, e164.
Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis, (Springer-Verlag New York).
Wolf, P.G., Rowe, C.A., and Hasebe, M. (2004). High levels of RNA editing in a vascular plant chloroplast genome: analysis of transcripts from the fern *Adiantum capillus-veneris*. Gene 339, 89–97.

SUPPORTING INFORMATION

Figure S1 RNA-Seq read coverages for mitogenome and plastome.

Figure S2 Editable site substitution pattern in the *Welwitschia* plastome.

Table S1 RNA editing sites distribution in *Welwitschia* mitogenome

Table S2 RT-PCR validation of *Welwitschia* mitogenes

Table S3 RNA editing sites distribution in *Ginkgo* mitogenome and plastome

Table S4 Gene specific primers for *Welwitschia* mitochondrial genes

The supporting information is available online at http://life.scichina.com and http://link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.