New Phytologist

Review

Tansley review

The evolution of RNA editing and pentatricopeptide repeat genes

Author for correspondence:
Ian Small
Tel: +61 08 6488 4499
Email: iansmall@cyllene.uwa.edu.au

Received: *23 February 2011*
Accepted: *22 March 2011*

Sota Fujii and Ian Small

Australian Research Council Centre of Excellence in Plant Energy Biology, University of Western Australia, 35 Stirling Highway, Crawley 6009, WA, Australia

Contents

New Phytologist (2011) **191**: 37–47
doi: 10.1111/j.1469-8137.2011.03746.x

Key words: chloroplasts, evolution, mitochondria, pentatricopeptide repeat, RNA editing.

Summary

The pentatricopeptide repeat (PPR) is a degenerate 35-amino-acid structural motif identified from analysis of the sequenced genome of the model plant *Arabidopsis thaliana*. From the wealth of sequence information now available from plant genomes, the PPR protein family is now known to be one of the largest families in angiosperm species, as most genomes encode 400–600 members. As the number of PPR genes is generally only *c.* 10–20 in other eukaryotic organisms, including green algae, the family has obviously greatly expanded during land plant evolution. This provides a rare opportunity to study selection pressures driving a 50-fold expansion of a single gene family. PPR proteins are sequence-specific RNA-binding proteins involved in many aspects of RNA processing in organelles. In this review, we will summarize our current knowledge about the evolution of PPR genes, and will discuss the relevance of the dramatic expansion in the family to the functional diversification of plant organelles, focusing primarily on RNA editing.

I. Introduction

Organellar function is essential for eukaryotic life, and depends upon activities that are maintained by proteins either internally synthesized within the organelle (organelle-encoded) or imported from cytosol (nucleus-encoded). Many of these functions are carried out by enzyme complexes with both organelle-encoded and nucleus-encoded subunits, and therefore coordination of the expression of organelle and nuclear genes is a critical matter. For example, the plant mitochondrial proteome can be estimated to consist of *c.* 2000 gene products (Millar *et al.*, 2005). Of these, only *c.* 40 proteins are encoded within the mitochondrial genome, most of which encode essential

subunits of oxidative phophorylation enzymatic complexes and ribosomal proteins (Kubo *et al.*, 2000; Notsu *et al.*, 2002; Handa, 2003; Ogihara *et al.*, 2005; Sugiyama *et al.*, 2005; Tian *et al.*, 2006; Allen *et al.*, 2007; Kubo & Newton, 2008; Fujii *et al.*, 2010). Although these represent only a small proportion of all mitochondrial proteins, their importance to mitochondrial function means that incorrect regulation of these mitochondrial genes would severely affect the whole system. As plant mitochondria encode no machinery to manage their own RNA expression and post-transcriptional RNA modification processes, these essential steps are totally reliant on nuclear-encoded gene products (Binder & Brennicke, 2003).

RNA metabolism plays a particularly important role in organelle gene expression (Stern *et al.*, 2010) and a wide array of different RNA binding proteins are found in organelles. Pentatricopeptide repeat (PPR) proteins are the most numerous of these. The first PPR protein to be described was the *Saccharomyces cerevisiae* mitochondrial protein Pet309, found to participate in translation of *cox1* (Manthey & McEwen, 1995; Manthey *et al.*, 1998; Tavares-Carreon *et al.*, 2008). Subsequently, Pet309, as well as the protein P67 implicated in transcription in *Triticum aestivum* mitochondria (Ikeda & Gray, 1999) and CRP1 in *Zea mays* involved in translation of photosynthesis genes (Fisk *et al.*, 1999; Schmitz-Linneweber *et al.*, 2005), were recognized to be members of a large family of related proteins following the systematic analysis of the *Arabidopsis thaliana* (thale cress) genome (Aubourg *et al.*, 2000; Small & Peeters, 2000). Since the discovery and definition of the PPR consensus sequence (Small & Peeters, 2000), predicted to form an antiparallel double alpha-helical motif, many studies have been conducted on PPR proteins covering biochemistry, molecular functions, cellular functions and roles in development (Schmitz-Linneweber & Small, 2008). The general picture that results from these studies is that PPR proteins form sequence-specific associations with RNA, and that these associations affect folding, processing and/or translation of the RNA, thus manipulating expression of the transcript. For example, CRP1 in maize is shown to associate with the 5′UTR region of photosynthetic genes *psaC* and *petA*, and mutants disrupted in *CRP1* gene lack translation of these genes, leading to the defects in photosynthesis (Fisk *et al.*, 1999; Schmitz-Linneweber *et al.*, 2005). *Arabidopsis* CRR4 was the first gene found in plants to be directly involved in cytosine to uridine RNA editing (Kotera *et al.*, 2005), since followed by many other PPR editing factors (Table 1). These are only a couple of examples, and for more discussion of the molecular functions of PPR proteins readers are redirected to reviews elsewhere that treat this topic comprehensively (Andres *et al.*, 2007; Delannoy *et al.*, 2007; Saha *et al.*, 2007; Schmitz-Linneweber & Small, 2008; Chateigner-Boutin & Small, 2010). The PPR family is subdivided into two major classes, P and PLS. Whereas P-class PPR proteins consist of an orthodox tandem alignment of 35-amino-acid PPR (P) motifs, PLS-class proteins contain, in addition, slightly longer (L) or shorter (S) variant PPR motifs in tandem arrays of characteristic triplets, P-L-S (Lurin *et al.*, 2004; Rivals *et al.*, 2006; O'Toole *et al.*, 2008). PLS-class proteins can be divided into two further groups, the E subclass and the DYW subclass, based on their C-terminal domains (Lurin *et al.*, 2004; Rivals *et al.*, 2006; O'Toole *et al.*, 2008).

The aim of this review is to integrate our current knowledge about the evolution of the PPR gene family and connect this with the changes thought to have taken place in the sequence and expression of the organelle genomes during the history of land plants. Given the rapid expansion of PPR proteins in land plants, there is an interest in understanding what selective forces might have operated on them to increase their numbers. Recent progress in genome sequencing and the consequent enrichment of comparative genomic databases are starting to allow us to understand how and why the plant-specific expansion of the PPR family occurred.

II. Expansion of the PPR family

Outside land plants, the organism with largest set of known PPR genes is the parasitic protozoan *Trypanosoma brucei*, commonly known as the causative agent for sleeping sickness, with 28 PPR members (Pusnik *et al.*, 2007). This compares with the 450 and 477 PPR genes in *Arabidopsis thaliana* and *Oryza sativa* (rice), model species for eudicots and monocots, respectively (Lurin *et al.*, 2004; O'Toole *et al.*, 2008). The green alga *Chlamydomonas reinhardtii* (Merchant *et al.*, 2007) contains only 12 PPR genes, whilst the moss *Physcomitrella patens* (Rensing *et al.*, 2008) contains 103 (O'Toole *et al.*, 2008; Schmitz-Linneweber & Small, 2008). These data suggest that the initial expansion of PPR genes took place after the separation of the land plant lineage from green algae such as *Chlamydomonas* but before the divergence of seed plants from bryophytes.

At present, sufficient information is available to conduct whole-genome comparative genomic studies for 18 plant species (from websites such as Phytozome: http://www.phytozome.net/). The distribution of PPR genes in these species is shown in Fig. 1. It is now certain that a wide range of angiosperms belonging to the core-eudicots or monocots contain a very large number of PPR genes, and that considering variations in genome size and total gene content between these species, the number of PPR genes is relatively constant. The least number of PPR genes is found in *Medicago truncatula* (365) and the most in *Glycine max* (629). The partially assembled *Selaginella moellendorffii* (spikemoss, lycophyte) genome contains > 1000 PPR genes (Fig. 1), although the numbers given here should be treated as

Table 1 Known editing factors from plant organelles

Organelle	AGI	Subclass	Name	Edited gene	Reference
Chloroplast	At1g08070	DYW	OTP82	*ndhG*, *ndhB*	Hammani *et al.* (2009)
Chloroplast	At1g11290	DYW	CRR22	*ndhB*, *ndhD*, *rpoB*	Okuda *et al.* (2009)
Chloroplast	At1g15510	DYW	AtECB2/VAC1	*accD*	Yu *et al.* (2009), Tseng *et al.* (2010)
Chloroplast	At1g59720	DYW	CRR28	*ndhB*, *ndhD*	Okuda *et al.* (2009)
Chloroplast	At2g02980	DYW	OTP85	*ndhD*	Hammani *et al.* (2009)
Chloroplast	At2g29760	DYW	OTP81	*rps12 intron*	Hammani *et al.* (2009)
Chloroplast	At3g22690	DYW	YS1	*rpoB*	Zhou *et al.* (2008)
Chloroplast	At3g57430	DYW	OTP84	*ndhF*, *psbZ*, *ndhB*	Hammani *et al.* (2009)
Chloroplast	At3g63370	DYW	OTP86	*rps14*	Hammani *et al.* (2009)
Chloroplast	At5g13270	DYW	RARE1	*accD*	Robbins *et al.* (2009)
Chloroplast	At5g48910	DYW	LPA66	*psbF*	Cai *et al.* (2009)
Chloroplast	At2g45350	E	CRR4	*ndhD*	Kotera *et al.* (2005)
Chloroplast	At5g59200	E	OTP80	*rpl23*	Hammani *et al.* (2009)
Chloroplast	At1g05750	E	CLB19	*rpoA*, *clpP*	Chateigner-Boutin *et al.* (2008)
Chloroplast	At5g55740	E	CRR21	*ndhD*	Okuda *et al.* (2007)
Mitochondria	At2g03880	DYW	REME1	*nad2*, *orfX*	Bentolila *et al.* (2010)
Mitochondria	At2g25580	DYW	MEF8	*nad5*	Takenaka *et al.* (2010)
Mitochondria	At3g12770	DYW	MEF22	*nad3*	Takenaka *et al.* (2010)
Mitochondria	At4g14850	DYW	LOI1/MEF11	*cox3*, *nad4*, *ccb203*	Tang *et al.* (2010), Verbitskiy *et al.* (2010)
Mitochondria	At5g52630	DYW	MEF1	*rps4*, *nad7*, *nad2*	Zehrmann *et al.* (2009)
Mitochondria	*Oryza sativa*	DYW	OGR1	*cox2*, *cox3*, *ccmC*, *nad2*, *nad4*	Kim *et al.* (2009)
Mitochondria	*Physcomitrella patens*	DYW	PpPPR_56	*nad3*, *nad4*	Ohtani *et al.* (2010)
Mitochondria	*Physcomitrella patens*	DYW	PpPPR_71	*ccmF*	Tasaki *et al.* (2010)
Mitochondria	*Physcomitrella patens*	DYW	PpPPR_77	*cox2*, *cox3*	Ohtani *et al.* (2010)
Mitochondria	*Physcomitrella patens*	DYW	PpPPR_91	*nad5*	Ohtani *et al.* (2010)
Mitochondria	At1g62260	E	MEF9	*nad7*	Takenaka (2010)
Mitochondria	At2g22410	E	SLO1	*nad4*, *nad9*	Sung *et al.* (2010)
Mitochondria	At3g05240	E	MEF19	*ccb206*	Takenaka *et al.* (2010)
Mitochondria	At3g18970	E	MEF20	*rps4*	Takenaka *et al.* (2010)
Mitochondria	At5g19020	E	MEF18	*nad4*	Takenaka *et al.* (2010)
Mitochondria	At2g20540	E	MEF21	*cox3*	Takenaka *et al.* (2010)

preliminary and are likely to change with improved coverage and gene annotation.

The mechanism of expansion is likely to have involved retrotransposition. As many as 80% of the PPR genes in *Arabidopsis* and *Oryza* are intron-less (Lurin *et al.*, 2004; Rivals *et al.*, 2006; O'Toole *et al.*, 2008), whereas > 75% of *P. patens* PPR genes contain introns. Evolutionarily older PPR genes in rice and *A. thaliana* contain more introns (O'Toole *et al.*, 2008) and some of the putative orthologues to the *P. patens* intron-containing genes lack introns in angiosperms (O'Toole *et al.*, 2008). These data are consistent with amplification by retrotransposition.

Gene families can also increase in size by genome duplication events. Angiosperm genomes have undergone multiple large genome duplication events (Bowers *et al.*, 2003; De Bodt *et al.*, 2005). These duplication events have large effects on gene number in gene families that are already present in large numbers. However, as shown in Fig. 2, abundant genes containing short repeat motifs such as ankyrin, tetratricopeptide repeat (TPR) and WD40 motifs have not increased in number during the evolution of land plants, implying that, for these families, gene loss has balanced the creation of new genes by duplication. This contrasts with the large increase in number of leucine-rich repeat (LRR) genes and PPR genes (Fig. 2). This suggests there may have been some form of selection for retention of duplicated LRR and PPR genes early in land plant evolution. On the other hand, it was shown that PPR gene content is well conserved between *A. thaliana* and *Oryza* (O'Toole *et al.*, 2008) and that the PPR genes duplicated in independent whole-genome duplication events occurring in these angiosperm lineages were rarely retained. This suggests that whatever the selective forces acting to increase PPR gene number in early land plant evolution, they appear no longer to act in this way in extant angiosperms.

In the case of LRR genes, their mode of action in the defence against pathogens can explain selection for increasing diversity. What might have caused selective pressures acting to retain large numbers of PPR genes following retrotransposition or genome duplications? In the next two

New Phytologist

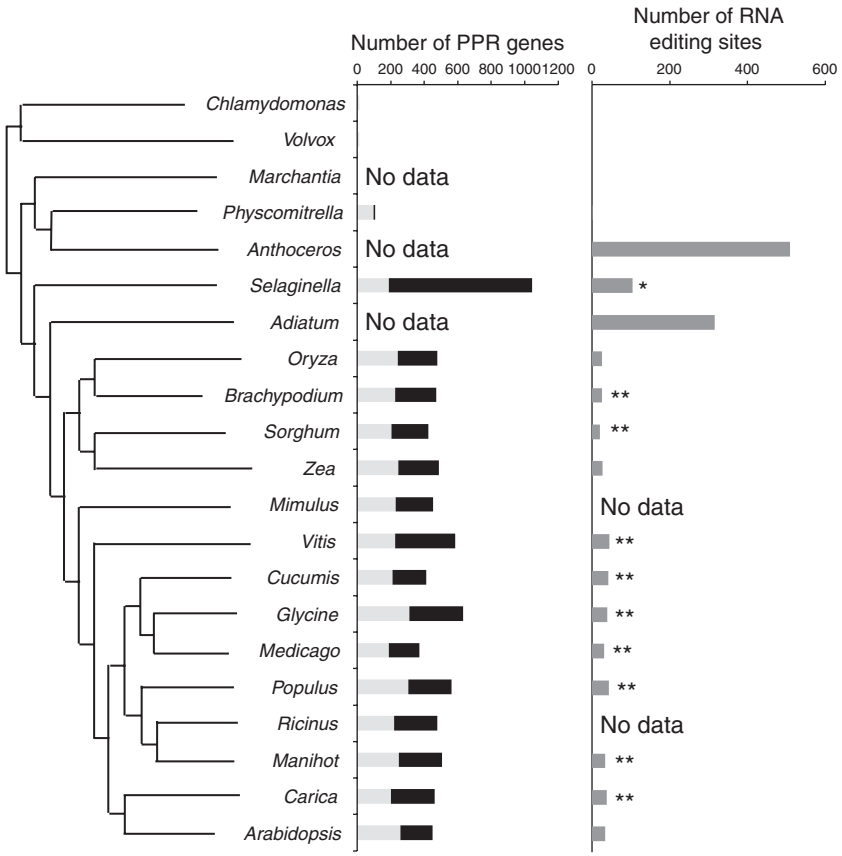

Fig. 1 Numbers of pentatricopeptide repeat (PPR) genes and plastid RNA editing sites in various plant species. Left histogram: light bars, P-class PPR genes; dark bars, PLS-class PPR genes. Right histogram: *, number of RNA editing sites detected from expressed sequence tags (Smith, 2009); **, number of RNA editing sites predicted by PREP (Mower, 2009).

sections, we will explore the growth of PPR family, focusing on the connection to one of its main molecular functions, RNA editing.

III. RNA editing and the evolution of PPR genes

RNA editing in plant mitochondria or chloroplasts commonly involves post-transcriptional substitution of cytidine by uridine (C-to-U) or, less commonly, uridine by cytidine (Gray, 1996, 2009; Maier *et al.*, 1996; Mulligan *et al.*, 1999; Shikanai, 2006; Takenaka *et al.*, 2008). Many RNA editing events play important roles in gene expression by generating start codons at ACG sites, correcting codons to encode conserved amino acids and generating required stop codons.

Beginning with the identification of CRR4 (Kotera *et al.*, 2005), many PPR proteins have been identified as factors required for C-to-U RNA editing in chloroplasts (summarized in Table 1). All of these PPR proteins are members of the plant-specific PLS class, one of two major subgroups within the PPR family. A possible exception is the *Arabidopsis* P-class protein PPR596, which acts to suppress RNA editing in mitochondrial *rps3* transcripts (Doniwa *et al.*, 2010). It remains to be seen whether PPR596 acts in a mechanistically similar way to other editing factors or whether the effect on editing is an indirect result of alterations in RNA processing or turnover.

The majority of PLS-class members may be dedicated to RNA editing, given the large numbers of editing mutants affected in expression of these genes (Table 1). So far, the only exceptions are the DYW-subclass protein CRR2 involved in RNA cleavage of a chloroplast transcript (Hashimoto *et al.*, 2003; Okuda *et al.*, 2009) and the E-subclass protein OTP70 involved in splicing of chloroplast *rpoC1* transcripts (Chateigner-Boutin *et al.*, 2011). The presence of DYW domains is strictly correlated in plant evolution with the presence of RNA editing (Salone *et al.*, 2007; Rudinger *et al.*, 2008). *Marchantia polymorpha*, one of the earliest branching land plants, lacks both DYW-subclass PPR genes and RNA editing, although the molecular phylogeny suggests it is likely that *Marchantia* has lost these properties rather than never had them (Groth-Malonek *et al.*, 2007; Salone *et al.*, 2007; Rudinger *et al.*, 2008). The origin of the DYW domain is unclear but there is a possibility that horizontal gene transfer has taken place between a plant and a heterolobosean protist (Knoop &

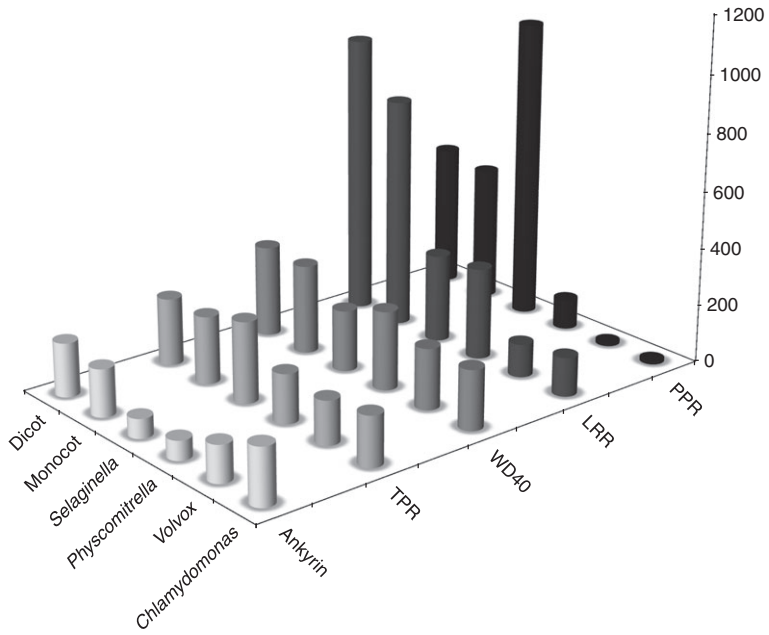

Fig. 2 Numbers of short tandem repeat motif genes present in plant taxa. The pentatricopeptide repeat (PPR) family has expanded more than other families of structurally similar genes that were probably present in greater numbers in the ancestors of land plants. TPR, tetratricopeptide repeat; LRR, leucine-rich repeat; PPR, pentatricopeptide repeat.

Rudinger, 2010). PLS-class PPR proteins had not previously been found outside the plant kingdom, and even within it, they are not present in algae (O'Toole *et al.*, 2008; Knoop & Rudinger, 2010). However, the recently sequenced genome of a protist species, *Naegleria gruberi* (Fritz-Laylin *et al.*, 2010), contains 12 PLS-class PPR genes, with 10 of them carrying a DYW domain (Knoop & Rudinger, 2010). Additional investigation is required to map the phylogenetic distribution of these proteins in protists and settle the direction of any gene transfer.

Both RNA editing (of the type found in plant organelles) and PLS-class PPR genes first appeared in the earliest land plants. The number of C-to-U RNA editing sites in organelles is variable between plants, ranging from only one in *Physcomitrella* chloroplasts to > 500 in the chloroplasts of the bryophyte *Anthoceros*, and even more in most land plant mitochondria (Fig. 1) (Kugita *et al.*, 2003; Tillich *et al.*, 2006; Jobson & Qiu, 2008). Unfortunately, comprehensive RNA editing information and complete nuclear genome sequences (i.e. information on PPR gene numbers) only rarely coexist. Species with both sets of data available are limited to *Physcomitrella*, *Selaginella*, *Oryza*, *Zea* (maize) and *Arabidopsis* (Fig. 1). Only 10 organellar RNA editing sites are known in *P. patens* (Miyata & Sugita, 2004; Tasaki & Sugita, 2010) and 103 PPR genes (including 10 DYW-subclass PPR genes; Fig. 1). There are at least 104 RNA editing sites in *S. moellendorffii* chloroplasts (Smith, 2009), many more than in angiosperms (20–40 sites). Mitochondrial RNA editing information is not available for *Selaginella*; however, a high number of RNA editing sites is expected from the extremely high GC content of the mitochondrial genome in this species because of the positive correlation between GC content and RNA editing (Malek *et al.*, 1996; Smith, 2009). This species is estimated to possess over 1000 PPR genes, whereas angiosperms, *Oryza*, *Zea* and *Arabidopsis* possess *c.* 450 (Fig. 1). Furthermore, it is the number of PLS-class PPR genes that correlates best (Pearson's correlation = 0.96) with the number of editing sites (Fig. 1) rather than the P subclass (Pearson's correlation = 0.5), in accordance with the mechanistic links between PLS-class proteins and RNA editing. The correlation between numbers of RNA editing sites and number of PPR genes implies that any selective pressures acting to increase RNA editing in early land plants would also have selected for increasing numbers of PPR genes.

IV. Sequence drift in the organelle genomes of early land plants

Algal organelle genomes are generally AT-rich, including those of the nearest relatives to the land plant lineage (Smith, 2009). By contrast, many land plant plastid genomes are relatively GC-rich (see Fig. 3), implying a sustained sequence drift over hundreds of millions of years. Phylogenetic reconstructions show that the rate of genomic T → C transitions must have far exceeded the rate of reverse mutations during this long period (Fig. 4). The potential link with RNA editing is evident, given that it reverses these DNA mutations at the RNA level. Indeed, several studies have conclusively shown that the number of RNA editing sites is strongly positively correlated with

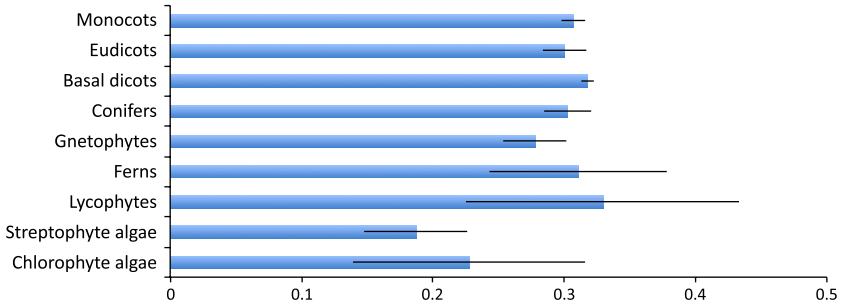

Fig. 3 Proportion of GC nucleotides at the third codon position in chloroplast genomes. Coding sequences from green algal and land plant chloroplast genomes were aligned. The nucleotide at the third position of each aligned codon was scored; the proportion of G + C is shown for each group. The standard deviation of the values for each group is indicated by a horizontal line. The number of genomes in each group is as follows: monocots (27), eudicots (79), basal dicots (10), conifers (10), gnetophytes (three), ferns (seven), lycophytes (three), streptophyte algae (six), chlorophyte algae (20).

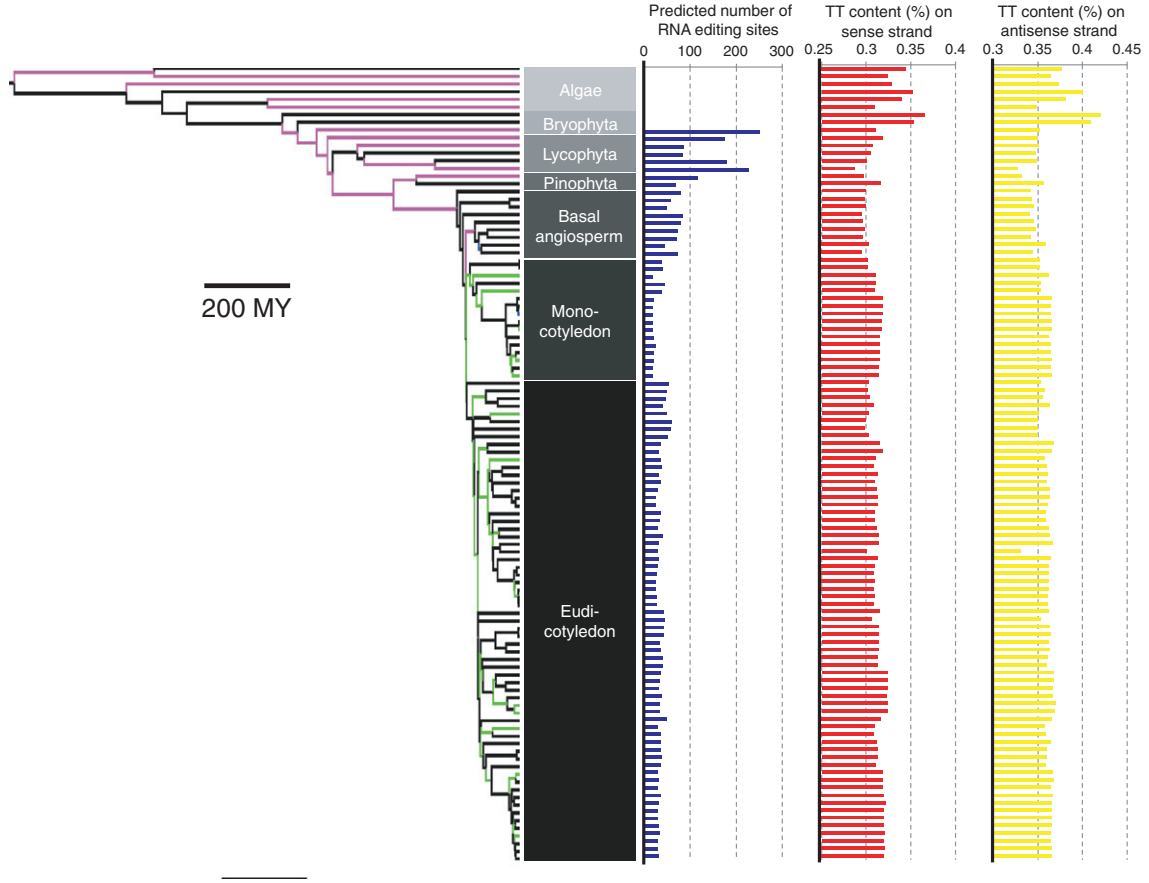

Fig. 4 Genome sequence trends during land plant evolution. The topology of the dendrogram was reconstructed from the concatenated multiple alignment of 51 chloroplast genes deduced from 104 plant species, and by using the maximum-likelihood method (Stamatakis, 2006). Numbers of synonymous substitutions (dS) were estimated by PAML (Yang, 2007), and absolute branch length was calculated using the molecular divergence estimation program r8s (Sanderson, 2003). Divergence time of the root was fixed as 1200 million yr ago (MYA), as described in Lang *et al.* (2008), and the minimum divergence time of angiosperms was set as 132 MYA from the oldest fossil record (Crane *et al.*, 1995). Magenta branches indicate when the T-to-C substitution rate was more than twice that of the C-to-T substitution rate, and green edges indicate when the C-to-T substitution rate was more than twice that of T-to-C substitutions. The histogram on the left indicates the number of RNA editing sites predicted by PREP (Mower, 2009) for each extant species. The histograms in the middle and to the right show the proportion of thymines in TT pairs within the sense strand or antisense (transcribed strand) of 51 chloroplast genes.

genome GC content (Malek *et al.*, 1996; Smith, 2009; Yura *et al.*, 2009). The reason for this sustained mutation bias is not clear. One explanation that has been suggested is that increased exposure to DNA-damaging UV light in land plants might have led to a relative suppression of adjacent pyrimidines (especially adjacent thymines) as a result of the formation of cyclobutane pyrimidine dimers and other photoproducts that can block transcription or replication, and which can be mutagenic (Singer & Ames, 1970; Maier *et al.*, 1996; Rozema *et al.*, 2002; Yura *et al.*, 2009). The support for this in extant sequences is weak, as judged by strand-specific dinucleotide frequencies (Fig. 4), but it remains a plausible hypothesis for changes in early land plants when UV exposure was probably considerably higher, because of both lower stratospheric ozone concentrations and the fact that land plants had not yet evolved the full set of UV-protecting traits possessed by extant plants (Takahashi & Ohnishi, 2004; Hader & Sinha, 2005).

An important point when considering the evolution of RNA editing is that phylogenetic reconstructions show very little long-term conservation of editing sites. There is little overlap between those present in extant ferns, mosses, spikeworts or angiosperms (Freyer *et al.*, 1995, 1997; Malek *et al.*, 1996; Tillich *et al.*, 2009). This has two ramifications. First, it means we are unable to state with confidence how much RNA editing was present in the common ancestors of these various lineages, as almost all the extant sites arose after the lineages split. Second, it shows that many land plant lineages continue to show a high rate of appearance of new editing sites.

V. Loss of RNA editing in angiosperms

Over 100 chloroplast editing sites are known in early-branching land plants such as *Anthoceros*, *Alsophila*, *Adiantum* or *Selaginella* (Freyer *et al.*, 1997; Kugita *et al.*, 2003; Wolf *et al.*, 2004; Tillich *et al.*, 2006; Smith, 2009), but fewer sites are edited in angiosperm chloroplast RNAs. From this, several authors have proposed that angiosperm chloroplasts are losing editing sites faster than they gain new ones (Freyer *et al.*, 1997; Tillich *et al.*, 2006; Jobson & Qiu, 2008). To demonstrate this, we have plotted the number of predicted RNA editing sites in each species against its evolutionary distance from the common ancestor of all angiosperms (Fig. 5). The two parameters are clearly anti-correlated, implying a steady overall loss of RNA editing sites in angiosperm chloroplasts. By using known divergence times to estimate phylogenetic branch length, an RNA editing site is estimated to be lost in 2.43 million yr, on average (Fig. 5). This loss of RNA editing sites is again linked to a concomitant change in DNA mutation rates; in angiosperm plastids, the rate of C → T transitions exceeds that of T → C transitions, a reversal of the situation in early land plants (Fig. 4).

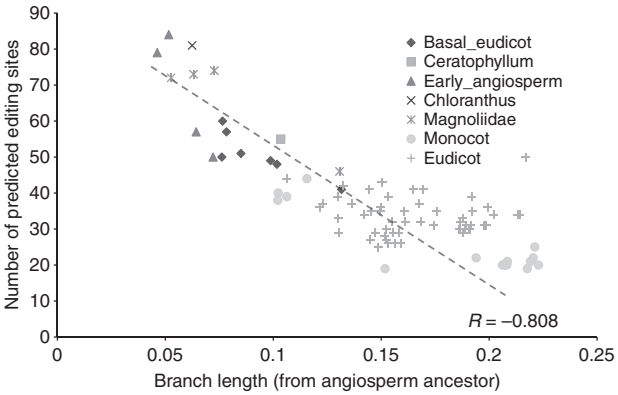

Fig. 5 Loss of RNA editing during angiosperm evolution. The vertical axis plots the number of plastid RNA editing sites predicted in each angiosperm species. The horizontal axis shows the cumulative edge length (number of synonymous substitutions per site, dS) from the common ancestor of angiosperms to each extant angiosperm species.

An interesting question is whether the RNA editing process is advantageous to the plant, or whether, on the contrary, it carries an associated cost. Advantages might include permitting a finer control over gene expression or providing a form of 'heterosis' by permitting the production of proteins with slightly different sequence from the same gene. Costs include the cost of producing the editing machinery (i.e. large numbers of PPR proteins) and the cost of producing and disposing of incorrect proteins translated from incompletely edited transcripts. We compared the conservation of C sites in angiosperm chloroplast coding sequences by summing cumulative branch length from when they appeared to when they reverted to a T (Fig. 6). We then compared the conservation of those C sites that are known or predicted to be edited at the RNA level with those that are not (Fig. 6). The evidence is that sites subject to RNA editing revert to a T more quickly than sites that are not edited.

Several explanations are possible for why edited cytidines might revert more frequently than unedited cytidines. Sequence context affects nucleotide substitution frequency, and the pronounced bias in sequence context around editing sites might lead to more frequent reversion. Alternatively, retrotranscription of edited organelle transcripts is a likely explanation for reversion of multiple closely linked sites and may explain a general high rate of reversion for editing sites (Cuenca *et al.*, 2010; Sloan *et al.*, 2010). Finally, it is possible that there is a sufficient cost associated with editing (the cost of synthesizing and importing editing factors, plus the costs of dealing with the consequences of incompletely edited transcripts and off-target effects of editing factors) to drive selection of reversion events and the loss of superfluous editing factors. It has been demonstrated recently that reversion of an editing site can lead to loss of the relevant editing factor (Hayes &

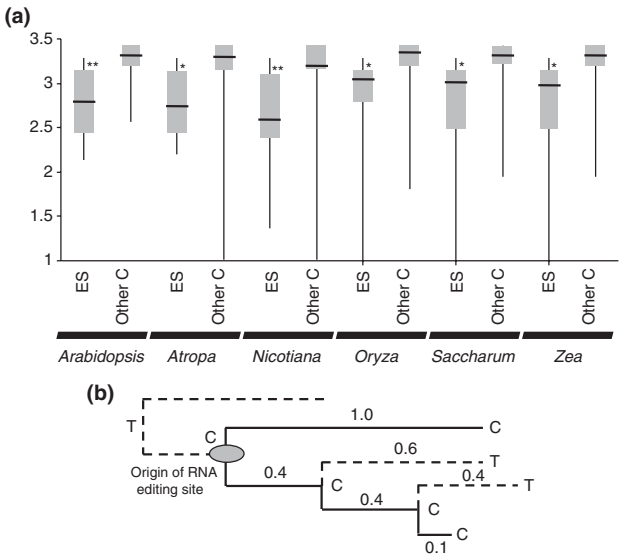

Fig. 6 Conservation of RNA editing sites (ES), in comparison to other unedited C sites (Other C). (a) Values are an indication of the duration that these sites persisted as a C during angiosperm evolution. Boxes include 50% of distributions. Horizontal bars indicate means. *, $P < 0.05$; **, $P < 0.01$, Student's *t*-test. (b) Schematic diagram of the values in the upper panel were calculated. In this example, solid edges indicate the period that a particular C was conserved. Thus the cumulative edge length for this site is 1.9 (values are relative to the average number of synonymous substitutions per site, dS).

Mulligan, 2011). It seems likely that angiosperm chloroplasts are on their way to ridding themselves of RNA editing. In this context it is worth noting that the number of RNA editing sites reduces quite rapidly in organelle genomes with accelerated substitution rates (Parkinson *et al.*, 2005; Sloan *et al.*, 2010).

VI. Other functions besides editing

Not all PPR proteins are implicated in RNA editing. Assuming as a first approximation that all PLS-class proteins are editing factors and all P-class proteins are not, then roughly half of the PPR proteins in angiosperms (*c.* 200–250) have other functions. This is still a huge increase in numbers over the 10–20 homologous proteins in algae and other eukaryotes, and this expansion probably occurred before, and independently of, the expansion in PLS-class PPR genes (O'Toole *et al.*, 2008). These P-class PPR proteins have roles in many other RNA-processing activities, including splicing, RNA turnover and translation, with positive or negative effects depending on the target and the context (Schmitz-Linneweber & Small, 2008). Many of these activities probably passively result from tight and sequence-specific binding, and, unlike RNA editing, do not require additional catalytic activities (Pfalz *et al.*, 2009; Prikryl *et al.*, 2011). The selection pressures acting to drive expansion of these PPR genes are less evident than those encoding RNA editing factors, but are likely to be similar, involving correction at the RNA level of genome-level mutations and rearrangements (Maier *et al.*, 2008). This is most evident in a particular clade of P-class proteins related to fertility restorer genes associated with cytoplasmic male sterility (CMS). These *restorer-to-fertility-like* (*RFL*) genes are numerous in angiosperms and show particularly rapid rates of evolutionary change (Geddy & Brown, 2007; Fujii *et al.*, 2011); at least some of them are required to block expression of the recombinant mitochondrial genes that induce CMS (Bentolila *et al.*, 2002; Brown *et al.*, 2003; Desloire *et al.*, 2003; Kazama & Toriyama, 2003; Koizuka *et al.*, 2003; Komori *et al.*, 2004; Wang *et al.*, 2006; Kato *et al.*, 2007; Kazama *et al.*, 2008; Uyttewaal *et al.*, 2008; Barr & Fishman, 2010).

VII. Conclusion

The current understanding of PPR proteins is that they are modular sequence-specific binding factors (Schmitz-Linneweber & Small, 2008). This modularity, coupled with highly specific target recognition, is perhaps what has made this protein family such an attractive resource for evolution to make use of in various cellular activities for plants. Amongst these various activities, RNA editing is one of the most enigmatic processes in terms of understanding how it came about, how it could have expanded to the extent it has in some land plant lineages, and to what extent it is advantageous. The discoveries described in this review are helping to make sense of these questions and have shown us that the evolution of land plant organelle genomes is inextricably linked to the evolution of the nuclear PPR gene family, and vice versa. It now seems probable that a rapid expansion in numbers of PPR genes permitted RNA editing on a large scale, and that this in turn allowed genome sequence drift. If (and it is a big 'if') this sequence drift was advantageous to early terrestrial plants (by conferring increased tolerance to temperature extremes, UV or oxidative stress induced by conditions on land) then PPR proteins and RNA editing may have had an important role in accelerating the colonization of the land.

References

Allen JO, Fauron CM, Minx P, Roark L, Oddiraju S, Lin GN, Meyer L, Sun H, Kim K, Wang C *et al.* 2007. Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. *Genetics* **177**: 1173–1192.

Andres C, Lurin C, Small I. 2007. The multifariuous roles of PPR proteins in plant mitochondrial gene expression. *Physiologia Plantarum* **129**: 14–22.

Aubourg S, Boudet N, Kreis M, Lecharny A. 2000. In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. *Plant Molecular Biology* **42**: 603–613.

Barr CM, Fishman L. 2010. The nuclear component of a cytonuclear hybrid incompatibility in mimulus maps to a cluster of pentatricopeptide repeat genes. *Genetics* **184**: 455–465.

Bentolila S, Alfonso AA, Hanson MR. 2002. A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. *Proceedings of the National Academy of Sciences, USA* **99**: 10887–10892.

Bentolila S, Knight W, Hanson M. 2010. Natural variation in arabidopsis leads to the identification of REME1, a pentatricopeptide repeat-DYW protein controlling the editing of mitochondrial transcripts. *Plant Physiology* **154**: 1966–1982.

Binder S, Brennicke A. 2003. Gene expression in plant mitochondria: transcriptional and post-transcriptional control. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **358**: 181–188; discussion 188-189.

Bowers JE, Chapman BA, Rong J, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**: 433–438.

Brown GG, Formanova N, Jin H, Wargachuk R, Dendy C, Patil P, Laforest M, Zhang J, Cheung WY, Landry BS. 2003. The radish *RFO* restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. *Plant Journal* **35**: 262–272.

Cai W, Ji D, Peng L, Guo J, Ma J, Zou M, Lu C, Zhang L. 2009. Lpa66 is required for editing *psbf* chloroplast transcripts in arabidopsis. *Plant Physiology* **150**: 1260–1271.

Chateigner-Boutin A, Colas des Francs-Small C, Delannoy E, Kahlau S, Tanz S, Falcon de Longevialle A, Fujii S, Small I. 2011. OTP70 is a pentatricopeptide repeat protein of the E subgroup involved in splicing of the plastid transcript *rpoC1*. *Plant Journal* **65**: 532–542.

Chateigner-Boutin AL, Ramos-Vega M, Guevara-Garcia A, Andres C, de la Luz Gutierrez-Nava M, Cantero A, Delannoy E, Jimenez LF, Lurin C, Small I *et al.* 2008. CLB19, a pentatricopeptide repeat protein required for editing of *rpoA* and *clpP* chloroplast transcripts. *Plant Journal* **56**: 590–602.

Chateigner-Boutin AL, Small I. 2010. Plant RNA editing. *RNA Biology* **7**: 213–219.

Crane PR, Friis EM, Pederson KR. 1995. The origin and early diversification of angiosperms. *Nature* **374**: 27–33.

Cuenca A, Petersen G, Seberg O, Davis JI, Stevenson DW. 2010. Are substitution rates and RNA editing correlated? *BMC Evolutionary Biology* **10**: 349.

De Bodt S, Maere S, Van de Peer Y. 2005. Genome duplication and the origin of angiosperms. *Trends in Ecology & Evolution* **20**: 591–597.

Delannoy E, Stanley WA, Bond CS, Small ID. 2007. Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. *Biochemical Society Transactions* **35**: 1643–1647.

Desloire S, Gherbi H, Laloui W, Marhadour S, Clouet V, Cattolico L, Falentin C, Giancola S, Renard M, Budar F *et al.* 2003. Identification of the fertility restoration locus, *RFO*, in radish, as a member of the pentatricopeptide-repeat protein family. *EMBO Reports* **4**: 588–594.

Doniwa Y, Ueda M, Ueta M, Wada A, Kadowaki K, Tsutsumi N. 2010. The involvement of a PPR protein of the P subfamily in partial RNA editing of an arabidopsis mitochondrial transcript. *Gene* **454**: 39–46.

Fisk DG, Walker MB, Barkan A. 1999. Molecular cloning of the maize gene *CRP1* reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO Journal* **18**: 2621–2630.

Freyer R, Kiefer-Meyer MC, Kossel H. 1997. Occurrence of plastid RNA editing in all major lineages of land plants. *Proceedings of the National Academy of Sciences, USA* **94**: 6285–6290.

Freyer R, Lopez C, Maier RM, Martin M, Sabater B, Kossel H. 1995. Editing of the chloroplast *ndhB* encoded transcript shows divergence between closely related members of the grass family (poaceae). *Plant Molecular Biology* **29**: 679–684.

Fritz-Laylin LK, Prochnik SE, Ginger ML, Dacks JB, Carpenter ML, Field MC, Kuo A, Paredez A, Chapman J, Pham J *et al.* 2010. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. *Cell* **140**: 631–642.

Fujii S, Bond CS, Small ID. 2011. Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. *Proceedings of the National Academy of Sciences, USA* **108**: 1723–1728.

Fujii S, Kazama T, Yamada M, Toriyama K. 2010. Discovery of global genomic re-organization based on comparison of two newly sequenced rice mitochondrial genomes with cytoplasmic male sterility-related genes. *BMC Genomics* **11**: 209.

Geddy R, Brown GG. 2007. Genes encoding pentatricopeptide repeat (PPR) proteins are not conserved in location in plant genomes and may be subject to diversifying selection. *BMC Genomics* **8**: 130.

Gray MW. 1996. RNA editing in plant organelles: a fertile field. *Proceedings of the National Academy of Sciences, USA* **93**: 8157–8159.

Gray MW. 2009. RNA editing in plant mitochondria: 20 years later. *IUBMB Life* **61**: 1101–1104.

Groth-Malonek M, Wahrmund U, Polsakiewicz M, Knoop V. 2007. Evolution of a pseudogene: exclusive survival of a functional mitochondrial *nad7* gene supports *Haplomitrium* as the earliest liverwort lineage and proposes a secondary loss of RNA editing in marchantiidae. *Molecular Biology and Evolution* **24**: 1068–1074.

Hader DP, Sinha RP. 2005. Solar ultraviolet radiation-induced DNA damage in aquatic organisms: potential environmental impact. *Mutation Research* **571**: 221–233.

Hammani K, Okuda K, Tanz SK, Chateigner-Boutin AL, Shikanai T, Small I. 2009. A study of new arabidopsis chloroplast RNA editing mutants reveals general features of editing factors and their target sites. *Plant Cell* **21**: 3686–3699.

Handa H. 2003. The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Research* **31**: 5907–5916.

Hashimoto M, Endo T, Peltier G, Tasaka M, Shikanai T. 2003. A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in arabidopsis. *Plant Journal* **36**: 541–549.

Hayes ML, Mulligan RM. 2011. Pentatricopeptide repeat proteins constrain genome evolution in chloroplasts. *Molecular Biology and Evolution*. doi: 10.1093/molbev/msr023

Ikeda TM, Gray MW. 1999. Characterization of a DNA-binding protein implicated in transcription in wheat mitochondria. *Molecular and Cellular Biology* **19**: 8113–8122.

Jobson RW, Qiu YL. 2008. Did RNA editing in plant organellar genomes originate under natural selection or through genetic drift? *Biology Direct* **3**: 43.

Kato H, Tezuka K, Feng YY, Kawamoto T, Takahashi H, Mori K, Akagi H. 2007. Structural diversity and evolution of the *rf-1* locus in the genus *Oryza*. *Heredity* **99**: 516–524.

Kazama T, Nakamura T, Watanabe M, Sugita M, Toriyama K. 2008. Suppression mechanism of mitochondrial ORF79 accumulation by RF1 protein in Bt-type cytoplasmic male sterile rice. *Plant Journal* **55**: 619–628.

Kazama T, Toriyama K. 2003. A pentatricopeptide repeat-containing gene that promotes the processing of aberrant *atp6* RNA of cytoplasmic male-sterile rice. *FEBS Letters* **544**: 99–102.

Kim SR, Yang JI, Moon S, Ryu CH, An K, Kim KM, Yim J, An G. 2009. Rice *OGR1* encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. *Plant Journal* **59**: 738–749.

Knoop V, Rudinger M. 2010. DYW-type PPR proteins in a heterolobosean protist: plant RNA editing factors involved in an ancient horizontal gene transfer? *FEBS Letters* **584**: 4287–4291.

Koizuka N, Imai R, Fujimoto H, Hayakawa T, Kimura Y, Kohno-Murase J, Sakai T, Kawasaki S, Imamura J. 2003. Genetic characterization of a pentatricopeptide repeat protein gene, ORF687, that restores fertility in the cytoplasmic male-sterile kosena radish. *Plant Journal* **34**: 407–415.

Komori T, Ohta S, Murai N, Takakura Y, Kuraya Y, Suzuki S, Hiei Y, Imaseki H, Nitta N. 2004. Map-based cloning of a fertility restorer gene, RF-1, in rice (*Oryza sativa* L.). *Plant Journal* **37**: 315–325.

Kotera E, Tasaka M, Shikanai T. 2005. A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* **433**: 326–330.

Kubo T, Newton KJ. 2008. Angiosperm mitochondrial genomes and mutations. *Mitochondrion* **8**: 5–14.

Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T. 2000. The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA(Cys)(GCA). *Nucleic Acids Research* **28**: 2571–2576.

Kugita M, Yamamoto Y, Fujikawa T, Matsumoto T, Yoshinaga K. 2003. RNA editing in hornwort chloroplasts makes more than half the genes functional. *Nucleic Acids Research* **31**: 2417–2423.

Lang D, Zimmer AD, Rensing SA, Reski R. 2008. Exploring plant biodiversity: the *Physcomitrella* genome and beyond. *Trends in Plant Science* **13**: 542–549.

Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere 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.

Maier RM, Zeltz P, Kossel H, Bonnard G, Gualberto JM, Grienenberger JM. 1996. RNA editing in plant mitochondria and chloroplasts. *Plant Molecular Biology* **32**: 343–365.

Maier UG, Bozarth A, Funk HT, Zauner S, Rensing SA, Schmitz-Linneweber C, Borner T, Tillich M. 2008. Complex chloroplast RNA metabolism: Just debugging the genetic programme? *BMC Biology* **6**: 36.

Malek O, Lattig K, Hiesel R, Brennicke A, Knoop V. 1996. RNA editing in bryophytes and a molecular phylogeny of land plants. *EMBO Journal* **15**: 1403–1411.

Manthey GM, McEwen JE. 1995. The product of the nuclear gene *PET309* is required for translation of mature mRNA and stability or production of intron-containing RNAs derived from the mitochondrial *cox1* locus of *Saccharomyces cerevisiae*. *EMBO Journal* **14**: 4031–4043.

Manthey GM, Przybyla-Zawislak BD, McEwen JE. 1998. The *Saccharomyces cerevisiae* PET309 protein is embedded in the mitochondrial inner membrane. *European Journal of Biochemistry* **255**: 156–161.

Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Marechal-Drouard L *et al.* 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**: 245–250.

Millar AH, Heazlewood JL, Kristensen BK, Braun HP, Moller IM. 2005. The plant mitochondrial proteome. *Trends in Plant Science* **10**: 36–43.

Miyata Y, Sugita M. 2004. Tissue- and stage-specific RNA editing of *rps14* transcripts in moss (*Physcomitrella patens*) chloroplasts. *Journal of Plant Physiology* **161**: 113–115.

Mower JP. 2009. The PREP suite: predictive RNA editors for plant mitochondrial genes, chloroplast genes and user-defined alignments. *Nucleic Acids Research* **37**: W253–W259.

Mulligan RM, Williams MA, Shanahan MT. 1999. RNA editing site recognition in higher plant mitochondria. *Journal of Heredity* **90**: 338–344.

Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K. 2002. The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Molecular Genetics and Genomics* **268**: 434–445.

Ogihara Y, Yamazaki Y, Murai K, Kanno A, Terachi T, Shiina T, Miyashita N, Nasuda S, Nakamura C, Mori N *et al.* 2005. Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucleic Acids Research* **33**: 6235–6250.

Ohtani S, Ichinose M, Tasaki E, Aoki Y, Komura Y, Sugita M. 2010. Targeted gene disruption identifies three PPR-DYW proteins involved in RNA editing for five editing sites of the moss mitochondrial transcripts. *Plant and Cell Physiology* **51**: 1942–1949.

Okuda K, Chateigner-Boutin AL, Nakamura T, Delannoy E, Sugita M, Myouga F, Motohashi R, Shinozaki K, Small I, Shikanai T. 2009. Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in arabidopsis chloroplasts. *Plant Cell* **21**: 146–156.

Okuda K, Myouga F, Motohashi R, Shinozaki K, Shikanai T. 2007. Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proceedings of the National Academy of Sciences, USA* **104**: 8178–8183.

O'Toole N, Hattori M, Andres C, Iida K, Lurin C, Schmitz-Linneweber C, Sugita M, Small I. 2008. On the expansion of the pentatricopeptide repeat gene family in plants. *Molecular Biology and Evolution* **25**: 1120–1128.

Parkinson CL, Mower JP, Qiu YL, Shirk AJ, Song K, Young ND, DePamphilis CW, Palmer JD. 2005. Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. *BMC Evolutionary Biology* **5**: 73.

Pfalz J, Bayraktar OA, Prikryl J, Barkan A. 2009. Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. *EMBO Journal* **28**: 2042–2052.

Prikryl J, Rojas M, Schuster G, Barkan A. 2011. Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proceedings of the National Academy of Sciences, USA* **108**: 415–420.

Pusnik M, Small I, Read LK, Fabbro T, Schneider A. 2007. Pentatricopeptide repeat proteins in *Trypanosoma brucei* function in mitochondrial ribosomes. *Molecular and Cellular Biology* **27**: 6876–6888.

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y *et al.* 2008. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **319**: 64–69.

Rivals E, Bruyere C, Toffano-Nioche C, Lecharny A. 2006. Formation of the arabidopsis pentatricopeptide repeat family. *Plant Physiology* **141**: 825–839.

Robbins JC, Heller WP, Hanson MR. 2009. A comparative genomics approach identifies a PPR-DYW protein that is essential for C-to-U editing of the arabidopsis chloroplast *accD* transcript. *RNA* **15**: 1142–1153.

Rozema J, Bjorn LO, Bornman JF, Gaberscik A, Hader DP, Trost T, Germ M, Klisch M, Groniger A, Sinha RP *et al.* 2002. The role of UV-b radiation in aquatic and terrestrial ecosystems–an experimental and functional analysis of the evolution of UV-absorbing compounds. *Journal of Photochemistry and Photobiology. B, Biology* **66**: 2–12.

Rudinger M, Polsakiewicz M, Knoop V. 2008. Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat proteins in jungermanniid but not in marchantiid liverworts. *Molecular Biology and Evolution* **25**: 1405–1414.

Saha D, Prasad AM, Srinivasan R. 2007. Pentatricopeptide repeat proteins and their emerging roles in plants. *Plant Physiology and Biochemistry* **45**: 521–534.

Salone V, Rudinger 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 Letters* **581**: 4132–4138.

Sanderson MJ. 2003. R8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**: 301–302.

Schmitz-Linneweber C, Small I. 2008. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends in Plant Science* **13**: 663–670.

Schmitz-Linneweber C, Williams-Carrier R, Barkan A. 2005. RNA immunoprecipitation and microarray analysis show a chloroplast pentatricopeptide repeat protein to be associated with the 5' region of mRNAs whose translation it activates. *Plant Cell* **17**: 2791–2804.

Shikanai T. 2006. RNA editing in plant organelles: machinery, physiological function and evolution. *Cellular and Molecular Life Sciences* **63**: 698–708.

Singer CE, Ames BN. 1970. Sunlight ultraviolet and bacterial DNA base ratios. *Science* **170**: 822–825.

Sloan DB, MacQueen AH, Alverson AJ, Palmer JD, Taylor DR. 2010. Extensive loss of RNA editing sites in rapidly evolving *Silene* mitochondrial genomes: selection vs. retroprocessing as the driving force. *Genetics* **185**: 1369–1380.

Small ID, Peeters N. 2000. The PPR motif – a TPR-related motif prevalent in plant organellar proteins. *Trends in Biochemical Sciences* **25**: 46–47.

Smith DR. 2009. Unparalleled GC content in the plastid DNA of selaginella. *Plant Molecular Biology* **71**: 627–639.

Stamatakis A. 2006. Raxml-vi-hpc: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.

Stern DB, Goldschmidt-Clermont M, Hanson MR. 2010. Chloroplast RNA metabolism. *Annual Review of Plant Biology* **61**: 125–155.

Sugiyama Y, Watase Y, Nagase M, Makita N, Yagura S, Hirai A, Sugiura M. 2005. The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. *Mol Genet Genomics* **272**: 603–615.

Sung TY, Tseng CC, Hsieh MH. 2010. The SLO1 PPR protein is required for RNA editing at multiple sites with similar upstream sequences in arabidopsis mitochondria. *Plant Journal* **63**: 499–511.

Takahashi A, Ohnishi T. 2004. The significance of the study about the biological effects of solar ultraviolet radiation using the exposed facility on the international space station. *Biological Sciences in Space* **18**: 255–260.

Takenaka M. 2010. MEF9, an E-subclass pentatricopeptide repeat protein, is required for an RNA editing event in the *nad7* transcript in mitochondria of arabidopsis. *Plant Physiology* **152**: 939–947.

Takenaka M, Verbitskiy D, van der Merwe JA, Zehrmann A, Brennicke A. 2008. The process of RNA editing in plant mitochondria. *Mitochondrion* **8**: 35–46.

Takenaka M, Verbitskiy D, Zehrmann A, Brennicke A. 2010. Reverse genetic screening identifies five E-class PPR proteins involved in RNA editing in mitochondria of *Arabidopsis thaliana*. *Journal of Biological Chemistry* **285**: 27122–27129.

Tang J, Kobayashi K, Suzuki M, Matsumoto S, Muranaka T. 2010. The mitochondrial PPR protein LOVASTATIN INSENSITIVE 1 plays regulatory roles in cytosolic and plastidial isoprenoid biosynthesis through RNA editing. *Plant Journal* **61**: 456–466.

Tasaki E, Hattori M, Sugita M. 2010. The moss pentatricopeptide repeat protein with a DYW domain is responsible for RNA editing of mitochondrial *ccmFc* transcript. *Plant Journal* **62**: 560–570.

Tasaki E, Sugita M. 2010. The moss *Physcomitrella patens*, a model plant for the study of RNA editing in plant organelles. *Plant Signaling and Behavior* **5**: 727–729.

Tavares-Carreon F, Camacho-Villasana Y, Zamudio-Ochoa A, Shingu-Vazquez M, Torres-Larios A, Perez-Martinez X. 2008. The pentatricopeptide repeats present in pet309 are necessary for translation but not for stability of the mitochondrial *cox1* mRNA in yeast. *Journal of Biological Chemistry* **283**: 1472–1479.

Tian X, Zheng J, Hu S, Yu J. 2006. The rice mitochondrial genomes and their variations. *Plant Physiology* **140**: 401–410.

Tillich M, Lehwark P, Morton BR, Maier UG. 2006. The evolution of chloroplast RNA editing. *Molecular Biology and Evolution* **23**: 1912–1921.

Tillich M, Sy VL, Schulerowitz K, von Haeseler A, Maier UG, Schmitz-Linneweber C. 2009. Loss of *matK* RNA editing in seed plant chloroplasts. *BMC Evolutionary Biology* **9**: 201.

Tseng CC, Sung TY, Li YC, Hsu SJ, Lin CL, Hsieh MH. 2010. Editing of *accD* and *ndhF* chloroplast transcripts is partially affected in the arabidopsis *vanilla cream 1* mutant. *Plant Molecular Biology* **73**: 309–323.

Uyttewaal M, Arnal N, Quadrado M, Martin-Canadell A, Vrielynck N, Hiard S, Gherbi H, Bendahmane A, Budar F, Mireau H. 2008. Characterization of *Raphanus sativus* pentatricopeptide repeat proteins encoded by the fertility restorer locus for Ogura cytoplasmic male sterility. *Plant Cell* **20**: 3331–3345.

Verbitskiy D, Zehrmann A, van der Merwe JA, Brennicke A, Takenaka M. 2010. The PPR protein encoded by the *LOVASTATIN INSENSITIVE 1* gene is involved in RNA editing at three sites in mitochondria of *Arabidopsis thaliana*. *Plant Journal* **61**: 446–455.

Wang Z, Zou Y, Li X, Zhang Q, Chen L, Wu H, Su D, Chen Y, Guo J, Luo D *et al.* 2006. Cytoplasmic male sterility of rice with Boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. *Plant Cell* **18**: 676–687.

Wolf PG, Rowe CA, 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.

Yang Z. 2007. Paml 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**: 1586–1591.

Yu QB, Jiang Y, Chong K, Yang ZN. 2009. AtECB2, a pentatricopeptide repeat protein, is required for chloroplast transcript *accD* RNA editing and early chloroplast biogenesis in *Arabidopsis thaliana*. *Plant Journal* **59**: 1011–1023.

Yura K, Sulaiman S, Hatta Y, Shionyu M, Go M. 2009. RESOPS: a database for analyzing the correspondence of RNA editing sites to protein three-dimensional structures. *Plant and Cell Physiology* **50**: 1865–1873.

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.

Zhou W, Cheng Y, Yap A, Chateigner-Boutin AL, Delannoy E, Hammani K, Small I, Huang J. 2008. The arabidopsis gene *YS1* encoding a dyw protein is required for editing of *rpoB* transcripts and the rapid development of chloroplasts during early growth. *Plant Journal* **58**: 82–96.