The C-terminal DYW domain of a PPR protein from the protist *Naegleria gruberi* can act as cytidine deaminase in moss mitochondria

Yingying Yang, Bastian Oldenkott, Elena Lesch, Volker Knoop and Mareike Schallenberg-Rüdinger*

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

*Corresponding author
mareike.ruedinger@uni-bonn.de, Phone: +49 228 73-6464

ORCID-IDs
Yingying Yang: 0000-0002-5518-0307
Bastian Oldenkott: 0000-0002-2513-2443
Elena Lesch: 0000-0002-3449-3929
Volker Knoop: 0000-0002-8485-9423
Mareike Schallenberg-Rüdinger: 0000-0002-6874-4722

Key words*: Naegleria gruberi*, *Physcomitrium patens,* mitochondrial RNA editing factor, PLS-type pentatricopeptide repeat (PPR) protein, DYW-type cytidine deaminase, PPR-RNA recognition code, horizontal gene transfer

Abstract

Organellar cytidine-to-uridine RNA editing driven by RNA-binding DYW-type pentatricopeptide repeat proteins is largely restricted to land plants. One of the rare exceptions is the heterolobosean protist *Naegleria gruberi*, which encodes ten DYW-type PPR proteins in its nuclear genome and features two C-to-U RNA editing sites in its mitochondrial transcriptome. Bioinformatic analyses favored NgPPR45 and NgPPR51 as top candidates for mitochondrial import and targeting those two RNA editing sites. Here, we tested the cytidine deaminase functionality of their C-terminal DYW domains making use of the model moss *Physcomitrium patens*. We designed protein chimeras combining the RNA-binding region of moss editing factor PpPPR78 and different portions of the putative cytidine deaminase domain of the two *N. gruberi* DYW-type PPR proteins, respectively. Chimeras were introduced into a KO plant line of PpPPR78, aiming to restore RNA editing at its assigned sites. Indeed, cox1eU755SL was edited in up to 82 % of transcripts by the chimera of PpPPR78 and the E1-E2-DYW domain of *N. gruberi* NgPPR45. In contrast, neither chimeras with smaller portions of the catalytic C-terminal domain of NgPPR45 nor chimeras using NgPPR51 domains were functional *P. patens*. We conclude that the C-terminal DYW domain of NgPPR45 is a functional cytidine deaminase and that the match of its PPR array to the putative binding regions upstream of the two mitochondrial editing targets of *Naegleria gruberi* make this factor the prime candidate to edit the latter in the protist.

Introduction

Genetic information can be changed on transcript level by RNA editing. Depending on the species, nuclear or organellar transcripts are edited by insertion, deletion or conversion of nucleotides (reviewed in Knoop, 2011; Knoop, 2022). In Metazoa, A-to-I RNA editing of nuclear transcripts is the dominant type of RNA editing (Nishikura, 2010) and only few cases of C-to-U of RNA editing are known (Meier et al., 2005). In plants, mitochondrial and chloroplast transcripts are affected by pyrimidine RNA editing (C-U or U-C) which typically leads to the correction of conserved codon identities or reading frames (reviewed in Knoop, 2022; Small et al., 2020). No single case of C-to-U RNA editing is known in algae by now, leading to the assumption, that the mechanism evolved in the common ancestor of land plants as consequence of colonizing land (Fujii and Small, 2011). In all land plants, except the marchantiid liverworts (Rüdinger et al., 2012), a few sites (moss *Funaria hygrometrica*, with two sites in the chloroplast and eight sites in the mitochondria, (Rüdinger et al., 2011b)) up to several thousands (*Selaginella* with 2700 sites in mitochondria (Hecht et al., 2011), and 3400 sites in the chloroplast (Oldenkott et al., 2014)) need to be edited to ensure translation of the correct protein sequence. In liverworts, mosses and seed plants only cytidines are converted into uridines in organellar transcripts (Dong et al., 2019; Rüdinger et al., 2012). In hornworts, some lycophytes and ferns classic C-to-U RNA editing is also accompanied by the reverse type of editing changing uridines to cytidines (Duff, 2006; Grewe et al., 2011; Gutmann et al., 2020; Knie et al., 2016).

C-to-U RNA editing in land plants is mediated by members of the large family of RNA binding pentatricopeptide repeat (PPR) proteins (Cheng et al., 2016; Gutmann et al., 2020; Ichinose et al., 2022). These alpha solenoid proteins consist of a PPR stretch with two to 30 circa 35 amino acid long PPR repeats (Knoop and Rüdinger, 2010; Lurin et al., 2004). Plant PPR proteins involved in RNA editing belong to the "PLS" subfamily. These proteins were only found to be encoded in species with RNA editing sites identified in their organellar transcriptomes (Gutmann et al., 2020; Rüdinger et al., 2008; Salone et al., 2007). Their

PPR arrays differ from those of “P-type” PPR proteins by exhibiting long (L, 35–36 aa) and short (S, 31–32 aa) PPR variants alternating with the canonical P-type PPRs of 35 amino acids (Cheng et al., 2016; Lurin et al., 2004). P- and S-type PPRs recognize and bind specific nucleotides in an one-repeat-one-nucleotide manner following a PPR-RNA binding code, which was proposed in 2012 and refined in several follow-up publications (Figure 1, Barkan et al., 2012; Takenaka et al., 2013; Yan et al., 2019). The 5th and the last (L) amino acid of each P- and S-type repeat form hydrogen bonds with the assigned nucleotides which places the PPR protein to edit the cytidine four nucleotides downstream of the nucleotide corresponding to the last PPR repeat (Figure 1, Barkan et al., 2012). The PPR array is followed by TPR-like E1 and E2 helix-turn-helix motifs and a DYW domain, named after its conserved C-terminal aspartate, tyrosine and tryptophan tripeptide (Cheng et al., 2016; Lurin et al., 2004).

The DYW domain features the typical $HxE(x)_nCxxC$ deaminase signature (Iyer et al., 2011; Salone et al., 2007) and was recently unequivocally assigned to exert the cytidine deaminase function (Hayes and Santibanez, 2020; Oldenkott et al., 2019; Takenaka et al., 2021). Single moss DYW-type PPR proteins were shown to edit provided targets, when transferred into the bacterium *Escherichia coli* (Oldenkott et al., 2019), *in vitro* (Hayes and Santibanez, 2020) or even in the cytoplasm of human cells (Ichinose et al., 2022; Lesch et al., 2022; Takenaka et al., 2021). Single amino acid changes in the cytidine deaminase signature of the DYW domain of such editing factors led to complete loss of editing, also *in planta* (Boussardon et al., 2014; Ichinose and Sugita, 2018; Wagoner et al., 2015).

Flowering plants encode some complete, but also many C-terminally truncated DYW-type PPR proteins lacking parts of the DYW domain (E+ subgroup), the complete DYW domain (E2 subgroup) or the complete E2 and DYW domain (E1 subgroup; Cheng et al., 2016). They need to recruit a DYW-domain *in trans* to obtain editing functionality (Andrés-Colás et al., 2017; Boussardon et al., 2012; Diaz et al., 2017; Gerke et al., 2020; Guillaumot et al., 2017; Wang et al., 2022). In angiosperms, additional factors like Multiple organellar RNA editing factors (MORFs), also called RNA-editing factor Interacting Protein (RIPs), organelle

RNA recognition motif-containing (ORRM) proteins, organellar zinc finger (OZ) proteins or P-type PPR proteins were also identified to participate in RNA editing of particular editing sites, building a complex editosome (Andrés-Colás et al., 2017; Bentolila et al., 2012; Gipson et al., 2022; Guillaumot et al., 2017; Sun et al., 2016; Sun et al., 2015; Takenaka et al., 2012, reviewed in Small et al., 2023).

The model moss *Physcomitrium patens* encodes only ten complete DYW-type PPR proteins in its nuclear genome (O'Toole et al., 2008), which correlates well with only eleven editing sites identified in its mitochondrial transcriptome and two sites in its plastid transcriptome (Miyata and Sugita, 2004; Rüdinger et al., 2009). Nine of the ten DYW-type PPR proteins were assigned to one up to two editing sites in the chloroplast or mitochondrion, respectively (Ichinose et al., 2014; Schallenberg-Rüdinger et al., 2013a), while the tenth DYW type PPR protein was shown to act in group II intron splicing of *cox1* instead (Ichinose et al., 2012). No RNA editing helper proteins were identified in *P. patens* so far (Uchiyama et al., 2018).

With more genomic and transcriptomic sequence data becoming available, rare cases of DYW domains encoded in genomes of species outside of land plants were discovered (Gutmann et al., 2020; Schallenberg-Rüdinger et al., 2013b). Besides identification of single DYW domain sequences in genomes of fungi (Schallenberg-Rüdinger et al., 2013b), dinoflagellates (Mungpakdee et al., 2014) or algae (Gutmann et al., 2020; Schallenberg-Rüdinger et al., 2013b), DYW type protein families of 10 up to 20 proteins were found to be encoded in the genomes of the slime mold *Physarum* (20, Schaap et al., 2015), the heterolobosean protists *Acrasis* (12, Fu et al., 2014) and *Naegleria gruberi* (10, Knoop and Rüdinger, 2010). In parallel, C-to-U RNA editing sites were found in the mitochondrial transcripts of these species. In *Physarum* four C-to-U RNA editing sites in the mitochondrial transcriptome are accompanied by many other RNA editing events, including insertion of single Cs, Us and dinucleotides as well as deletions (Bundschuh et al., 2011; Gott et al., 2005). In the protists *Acrasis* and *Naegleria* two mitochondrial C-to-U RNA editing sites, cox1eU1120HY and cox3eU787RW in *Naegleria* and atp6eU722SL and cobeU409HY in *Acrasis*, were identified. Editing sites are labeled as proposed in Rüdinger et al. (2009) starting with the

gene name of the affected transcript, eU for editing from C to U, the position of the C to be edited in the coding sequence and the resulting change of the codon identity.

As knockout studies with these heterolobosean protists are not feasible to date (Faktorová et al., 2020), we developed an alternative approach to test the functionality of the deaminase domains of the DYW-type PPR proteins of *Naegleria gruberi*. We used the model moss *Physcomitrium patens* KO line of PpPPR78 (Rüdinger et al., 2011b), which lacks editing of the two mitochondrial editing sites cox1eU755SL and rps14eU137SL. Chimeric protein constructs of PpPPR78 with parts of its C-terminal domain replaced by the counterparts of one of the two selected *N. gruberi* DYW-type PPR proteins NgPPR45 and NgPPR51 were introduced in the KO line to test their ability to restore RNA editing.

Results

Selection of DYW-type PPR proteins from *Naegleria gruberi* for complementation studies

The complete genome sequence of the heterolobosean protist *Naegleria gruberi* was published 2010 (NCBI GenBank entry: NW_003163326.1) and subsequently ten DYW-type PPR proteins and one PLS protein lacking the C-terminal extensions were found to be encoded in the nuclear genome (Fritz-Laylin et al., 2010; Knoop and Rüdinger, 2010).

The predicted encoded proteins vary in the number of PPR repeats and only one DYW-type PPR protein Ng51788 (NgPPR51) and the pure PLS protein (Ng45423) have signal peptides predicted for mitochondrial localization encoded N-terminally of their PPR arrays (Rüdinger et al., 2011a). For nine of the ten DYW-type PPR proteins (except Ng32401), PPR protein models with predicted PPR repeats and respective amino acids five (5) and last (L) for target recognition were already available on the PPR plantenergy webpage ((https://ppr.plantenergy.uwa.edu.au/ppr/). With the tool “Search for PPRs” available on the same webpage PPR repeats can be searched for in input sequences with slightly different parameters. Predicted models for most of the ten DYW type PPR proteins showed differences in comparison to the models available on the webpage, especially in terms of biased PPR (Cheng et al., 2016). To give one example,

PPR repeat P-6 of Ng45424, Ng76708 and Ng69406 showed insertions in the region between the two alpha helixes of the PPR repeat and were only predicted as a PPR repeat using the "Search for PPR" option. We manually inspected the different predictions and decided for the more reliable model with the highest number of PPR repeats predicted but the avoidance of overlapping PPR repeats for each DYW-type PPR protein to be presented in Figure 1 (all predictions are found in Supplementary table 1 and protein sequences are presented in supplementary data 2).

We aligned the PPR repeat region of each DYW-type PPR protein of N. gruberi to the sequence upstream of the editing targets cox1eU1120HY and cox3eU787RW, respectively, to identify the DYW-type PPR proteins, which are most likely to bind to the target RNAs. Interestingly, not the protein NgPPR51, which is predicted to be localized to the mitochondria, but the proteins Ng45424 and Ng32041 were identified to match the respective target sequence of cox1eU1120HY and cox3eU787RW best (Figure 1), respectively. Ng45424 showed six perfectly matching PPR-nucleotide pairs with the cis elements of cox1eU1120HY and five with target cox3eU787RW, respectively. Ng32041 showed seven matches with target cox3eU787RW, but also three mismatches in the C-terminal region of the PPR stretch, which was shown to be more relevant for recognizing the target than the N-terminal region in earlier studies (Oldenkott et al., 2020; Oldenkott et al., 2019; Takenaka et al., 2013). In our predicted gene model, however, the coding sequences of Ng45424 and Ng32041 lack an N-terminal organellar signal peptide.

One single DYW-type PPR protein sequence is buildup of predicted genes Naegr45423 and Naegr45424

To investigate, whether our current gene models are incomplete, we examined the upstream regions of both predicted genes Naegr45424 and Naegr32041 in the available scaffolds of the *N. gruberi* genome.

We indeed found a sequence in the 5' region of the predicted Naegr32041 gene that could be translated into 11 additional PPR repeats, linked to the predicted Ng32041 protein via a 29 aa long linker (Supplementary table 1). However, no clear signal peptide for mitochondrial localization was predicted with the WoLFPSORT tool (Supplementary table 4).

The predicted gene Naegr45424 was found to be located only 154 nucleotides downstream of Naegr45423. The latter is predicted to encode for the only pure PLS-type protein and includes an intron of 324 bp. As both predicted genes show numerous repetitive regions (Supplementary figure 1), we re-checked the gene makeup via PCR using primers with unique binding sites outside of the repetitive elements. Sequencing revealed that the coding regions are not separated by an intergenic region and the intron sequence in gene Naegr45423 could not be confirmed (Supplementary data 1, Figure 2). However, the amplicon confirming the connection of Naegr45423 and Naegr45424 was shorter than the calculated size of the combination of the two CDS. The size reduction can be explained by the highly repetitive character of the region between Naegr45423 and Naegr45424, which might have resulted in scaffolding/assembly mistakes in the initial genome annotation.

An RT-PCR was performed to prove that Naegr45423 and Naegr45424 are linked by the repetitive region and transcribed jointly (Figure 2). The complete gene will be named NgPPR45 in the following. The N-terminus of NgPPR45 contains a clear signal peptide for mitochondrial localization (predicted with WoLFPSORT and TargetP 2.0, Supplementary table 4). The combined PPR array of NgPPR45 encompasses 19 PPR repeats and matches the target sequences of both editing sites in the mitochondrial transcripts of *N. gruberi* equally well as the PPR protein encoded by the original Naegr45424 gene model (Figure 2).

Differences in the conservation of the C-terminal domains of NgPPR45 and NgPPR51 of *Naegleria gruberi* and the nine DYW type PPR editing factors of *Physcomitrium patens*

NgPPR45 and NgPPR51 both contain E1, E2 and DYW domains with a conserved cytidine deaminase signature which, however, show particular amino acid differences in comparison to the protein sequences of the PPR editing factors of the moss *Physcomitrium patens* (Figure 3). The most prominent difference is the lack of conservation of the „SW“ within the PgxSWiEv motif in the so-called PG-Box, which was shown to be important for proper RNA editing function in flowering plants (Okuda et al., 2007; Takenaka et al.,

178 2021) and just recently identified to participate in forming the cytidine binding pocket (Toma-Fukai et al.,
179 2022).

180 The C-terminal DYW tripeptide is also modified in both PPR proteins. The change of the Y to an N in
181 NgPPR51 might not have a strong impact, as several DYW proteins, for example PPR65, PPR98 and PPR91
182 of *P. patens* ending with DFW (Figure 3) and OTP86 of *Arabidopsis thaliana* with DSW (Takenaka et al.
183 2021), also show different amino acids in the second position of the tripeptide. The G instead of the D in
184 NgPPR45 might have an influence on editing capacity, as most of the characterized DYW-type PPR proteins
185 have a D at that position and an earlier mutational study with DYW1, a short DYW type PPR protein of
186 *Arabidopsis thaliana* interacting with E+ protein CRR4 on chloroplast editing site ndhDeU2TM (Boussardon
187 et al., 2012), already showed that a change of D into A reduced RNA editing capacity of that editing factor
188 (Boussardon et al., 2014).

189 Other differences from the plant consensus E1E2DYW are either found in the C-terminal domains of
190 NgPPR45 or of NgPPR51, as both proteins only share 46% sequence identity within their E1E2DYW domain.
191 Whereas NgPPR51 has the HSE motif within the zinc binding region $HxE(x)_nCxxC$, which is highly conserved
192 in most of the 400 plant C-U DYW-type PPR editing factors characterized so far (PREPACT Edifacts (Lenz et
193 al., 2018)), NgPPR45 exhibits an HAE motif, which can predominately be found in the only recently
194 identified putative reverse editing factors of hornworts and ferns (Gerke et al., 2020; Gutmann et al., 2020;
195 Ichinose et al., 2022), but also in other bacterial, fungal or mammalian cytidine deaminases for example
196 ADAR 1 and 2 (Iyer et al., 2011; Salone et al., 2007; Takenaka et al., 2021). In four other DYW type PPR
197 proteins of *Naegleria gruberi,* the HAE motif can be identified as well, while five, including NgPPR51, show
198 the dominant triplet HSE. Ng76525 exhibits an HCE instead of HxE, but displays several other mutations
199 including the deletion of the PG box and loss of functionality is likely (Supplementary data 2). NgPPR51 in
200 contrast to NgPPR45 also lacks conservation of the WGAL motif at the start of the E1 domain, also
201 conserved in most, but not all plant type PPR editing factors.

Complementation studies in PpPPR78 KO plant line

The characteristic differences of the DYW domains of the *N. gruberi* DYW type PPR proteins, the lack of information about the functionality of DYW domains outside of land plants and the question of functional compatibility of these domains with plant counterparts separated in evolution by 1.5 billion years brought us to the initial design of our experiments.

We generated different chimeras of PpPPR78 of *P. patens* with C-terminal domains exchanged with those of the *N. gruberi* DYW-type PPR proteins NgPPR45 and NgPPR51. Three different fusion points were chosen based on the assumed functionality of the catalytic C-terminal domain (Figure 4). The first set of chimeras consisted of the PPR stretch of PpPPR78 and the complete C-terminal extensions (E1E2DYW) of NgPPR45 or NgPPR51, respectively. In the second set of chimeras, the PPR region and the E1 and E2 motifs of PpPPR78 were combined with the DYW domain of the two *Naegleria* proteins, respectively. The last set of chimeras was built of the PPR region, E1, E2 and the N-terminal part of the DYW domain of PpPPR78 with the C-terminal part of the DYW domain of the two *Naegleria* proteins, respectively. This C-terminal part of the DYW domain was initially defined as core DYW domain (Cheng et al., 2016; Lurin et al., 2004) and some fusion proteins of different plant editing factors with exchanged „short“ DYW domain (DYW*) were proven to be functional (Ichinose and Sugita, 2018; Schallenberg-Rüdinger et al., 2017).

The different chimeras were transformed into the KO PpPPR78 line using protoplast transformation and inserted into the *P. patens* intergenic (PIG) region via homologous recombination (Okuda et al., 2009; Oldenkott et al., 2020). For each construct a minimum of three stable lines were selected and expression of the transgene, as well as RNA editing ratios, were evaluated (Supplementary table 2).

None of the plant lines expressing chimeras with the N-terminal part of PpPPR78 and the C-terminal part of NgPPR51 showed any editing at the two editing sites of PpPPR78. A different picture emerged for the chimeras with C-terminal domains of NgPPR45. The chimera harboring the complete E1E2DYW domain of NgPPR45 edited cox1eU755SL with 24% - 82 % efficiency (Figure 5), depending on the stable plant line

investigated. The second editing site assigned to PpPPR78, rps14eU137SL, however, was not edited in any of the plant lines. Chimeras of PpPPR78 and shorter parts of the C-terminal domain of NgPPR45 likewise did not show editing in any of the generated stable plant lines investigated.

Equal expression levels of different complementation constructs do not lead to same editing levels

Expression of chimeric proteins was driven by the strong actin promoter (XY). We investigated the expression levels of the different chimeras in stable complementation lines in comparison to the native expression of PpPPR78 in *P. patens* wild type Gransden, which is in general low (Schallenberg-Rüdinger et al., 2017), via quantitative real time PCR. In plants expressing PpPPR78-NgPPR45EDYW, the different editing efficiencies of *cox1* correlated well with the corresponding expression of the fusion protein itself (Figure 5). The plant line with only 2.7fold transgene expression compared to the wild type expression, exhibited 28 % editing of cox1eU755SL. With an increase of expression to up to 1000fold, the editing efficiency increased to 82 %, but did not reach >99 % like in the wild type Gransden.

Other chimeras, which did not show any editing, were expressed to comparable ratios (Figure 5). In case of PpPPR78-NgPPR45DYW even a 1,000-fold higher expression than the PpPPR78 expression in the wild type did not lead to any editing of one of the targets. Thus, different expression levels between lines expressing the same construct do influence editing rates, but differing expression levels are not the driving force for diverging editing capacities of different constructs.

Discussion

With our here presented study, we confirm that the C-terminal domain of DYW-type PPR protein NgPPR45 of the protist *Naegleria gruberi* has cytidine deaminase activity and is able to edit specific cytidines when fused to the upstream PPR array of PpPPR78 for RNA binding *in planta*.

Compatibility of the DYW domain of NgPPR45 of *Naegleria gruberi* with moss editing factor PpPPR78

This is to some extent surprising, as meanwhile studies have shown that the interchangeability of domains between different DYW-type PPR proteins is limited (Ichinose & Sugita, 2018; Maeda et al., 2022; Yang et al., 2023#2). Even the exchange of C-terminal domains of different DYW-type PPR proteins of the same species did not always result in functional editing factors *in planta*. *P. patens'* editing factor PPR56, for example, edits its main target nad4eU272SL with lower efficiencies (35 % instead of 95 %), when its DYW domain is replaced by the DYW domain of the chloroplast editing factor PpPPR45, but not at all when it´s replaced by one of the mosses other eight editing factors (Ichinose and Sugita, 2018, Figure 2).

Ichinose and colleagues identified the conservation of residues 37–42 of the DYW* domain (in Figure 2 positions 166-171) to be important for successful exchangeability, but it remained open, if the motif participates in the recognition of the -3 to -1 region upstream of assigned editing sites (Ichinose and Sugita, 2018; Okuda et al., 2014), or if that region interacts with upstream regions of the PPR protein. Tin contrast, the fusion of the PPR region of PpPPR78 with different portions of the C-terminal domain of editing factor PpPPR79 of *Physcomitrium patens* was shown to result in functional editing factor chimeras (Schallenberg-Rüdinger et al., 2017). This motivated us to use PpPPR78 in this study. In contrast to PpPPR78 and PpPPR79 sharing the same amino acids in position 37-42 of the DYW domain, NgPPR45 and NgPPR51 present another motif than PpPPR78 or any other *P. patens* editing factor (Figure 2). This could possibly be one explanation for the lack of functionality of PpPPR78 being fused to the DYW domain or portions of the DYW domain of NgPPR45 or NgPPR51 only.

The complete C-terminal domain (E1E2DYW domain) of PpPPR78, however, can be replaced by the one of NgPPR45 of *Naegleria gruberi* and led to editing of up to 82% of the cox1eU755SL site in *P. patens*. The E1 and E2 motifs, sharing similarities to TPR repeats which mediate protein-protein interaction (Blatch and Lässle, 1999), might interact with the DYW domain of NgPPR45 and an exchange of the complete C-terminal domain might therefore be successful in contrast to the DYW domain only. Differences of the E1,

E2 and DYW domain of NgPPR45 in comparison to the consensus of the plant E1, E2 and DYW domains might additionally hinder the interaction of the protists E1 and E2 domain with the plant DYW domain or *vice versa*.

Restricted influence of editing factor expression levels on editing rates

Natively DYW-type PPR proteins are only lowly expressed in land plants and the number of individual editing factors in mitochondria was found to be likewise low (Fuchs et al., 2020; Lurin et al., 2004). The more surprising is the high editing efficiency of most editing sites *in planta* (Bentolila et al., 2013; Small et al., 2023). Cox1eU755SL is also one of the sites, which is fully edited *in planta* in different tissues investigated (Rüdinger et al., 2009; Rüdinger et al., 2011b; Uchida et al., 2011), although PpPPR78 is only expressed to low levels under standard growth conditions (Figure 5, Schallenberg-Rüdinger et al., 2017). The second editing site assigned to PpPPR78, rps14eU137SL, is edited in 60-80% of transcripts in *P. patens* (Rüdinger et al., 2009; Rüdinger et al., 2011b). Intriguingly, overexpression of PpPPR78 in the KO PPR78 background of *P. patens* could increase editing of rps14eU137SL to 100%. A chimera of PpPPR78 and PpPPR79 could likewise fully edit the *cox1* site, but rps14eU137SL in only 26 % up to 63 % of transcripts (Schallenberg-Rüdinger et al., 2017). We postulated that the PPR protein RNA interaction is the limiting factor and not the abundance of the protein. The same holds true for our different PpPPR78-NgPPR45 and PpPPR78-NgPPR51 chimeras. Even expressions 1,000-fold higher than PpPPR78 in the wild type did not lead to RNA editing by the chimera with replaced DYW domain only. Highest expression of PpPPR78 with the E1E2DYW domain of PpPPR45 also did not lead to any editing of rps14eU137SL, revealing once more that this site is more difficult to be addressed than its counterpart in the *cox1* transcript.

The expression level of the PpPPR78-NgPPR45EDYW, however, influences the editing efficiency of cox1eU755SL. Five independent stable complementation lines show editing of 24% up to 82% and the editing investigated in each line correlates well with the PPR protein expression levels tested via quantitative RT-PCR (Figure 5).

Other studies also confirmed that once a PPR protein or PPR chimera can edit its assigned target, the expression of transcript and subsequently the amount of editing factors in the organelle affects the editing efficiencies at particular sites (Loiacono et al., 2022; Oldenkott et al., 2020).

The DYW domain of NgPPR45 shares similarity with reverse editing enzymes, but still acts in C-to-U RNA editing

DYW-type PPR proteins are mainly restricted to land plants and sporadic appearance of such proteins in species outside of the plant kingdom can most likely be explained by horizontal gene transfer (HGT, (Fu et al., 2014; Knoop and Rüdinger, 2010; Schallenberg-Rüdinger et al., 2013b). Given the divergent sequence conservation of the DYW-type PPR proteins of *Naegleria gruberi* and land plants, an HGT very early in plant evolution some 500 million years ago has most likely seeded the DYW-type PPR protein genes in a protist related to *Naegleria* (Knoop and Rüdinger, 2010). While the cytidine deaminase signature $HxE(x)_nCxxC$ is highly conserved in all DYW domains encoded in *N. gruberi*, other motifs which were also identified as key motifs for a fully functional DYW domain were found to be modified in NgPPR45 and NgPPR51 in comparison to the DYW-type PPR proteins of *P. patens* (Figure 2).

Whereas the E1E2DYW domain of NgPPR51 indeed did not show cytidine deaminase activity when fused to the PPR array of PpPPR78, the E1E2DYW domain of NgPPR45 did.

Interestingly, the C-terminal domain of NgPPR45 shows a relaxed conservation of certain motifs, which are also degenerated in C-terminal domains of the KPAxA PLS-type PPR proteins. This is a subtype of DYW-type PPR proteins exclusively found in hornworts, ferns and lycophytes and suspected to catalyze the reverse RNA editing reaction from U-to-C, which only appears in the organelles of these land plant clades (Gerke et al., 2020; Gutmann et al., 2020).

The SW of the PG box is degenerated, as well as the SHP motif, which is completely missing in many of the KPAxA PLS-type PPR proteins (Gerke et al., 2020; Gutmann et al., 2020; Takenaka et al., 2021). The C-terminal DYW triplet is modified into a GYW and the HSExLA motif conserved in all so far characterized C-

to-U editing factors in land plants (citation missing) is modified into a HAExLA, what is also dominantly found in the KPAxA PLS-type PPR proteins.

That this C-terminal domain is now able to edit a cytidine into a uridine *in planta* is surprising, but might also point towards the hypothesis that in early land plant evolution, C-to-U and U-to-C RNA editing was present (Small et al., 2020), with initial RNA editing enzymes that operated in both directions (Knoop, 2022; Small et al., 2023).

Recent phylogenetic concepts assume bryophytes (hornworts, mosses, liverworts) as one monophyletic group (Puttick et al., 2018; Su et al., 2021), what would consequently mean, that U-to-C RNA editing was gained together with C-to-U RNA editing in first land plants, possible linked to the conquest of land (Fujii et al., 2013).

A recent study with synthetic KPAxA PLS-type PPR proteins successfully tested in the two heterologous systems of *Escherichia coli* and humans, also confirmed that these proteins can indeed act in both directions (Ichinose et al., 2022).

Successful transfer of editing factors between different genetic systems

With the proof of functionality of the E1E2DYW domain of NgPPR45 of *Naegleria gruberi*, we have shown that an evolutionary early transfer of a functional editing factor had been successful. This is a further example of the transferability of these editing factors, with huge families of up to several 1000 members in land plants (Banks et al., 2011; Gerke et al., 2020), but with only single exceptional cases in species of other kingdoms accompanied by only low numbers of editing sites identified in their mitochondrial transcriptomes (Bundschuh, 2015; Fu et al., 2014; Rüdinger et al., 2011b).

In recent studies it was shown that DYW-type PPR proteins can be transferred into other genetic systems. Transferred moss editing factors PpPPR56 and PpPPR65 were shown to edit there delivered targets in the bacterium *Escherichia coli*, in human cell cytosol and *in vitro* (Oldenkott et al., 2019; Lesch et al., 2022; Hayes et al., 2020). Another moss editing factor PpPPR79 edits matching editing sites, when transferred

into flowering plant *Arabidopsis thaliana* (Oldenkott et al., 2020). Editing factors of *Arabidopsis thaliana*, however, were not functional in *P. patens* or *E. coli* yet (Maeda et al., 2022; Oldenkott et al., 2020). One reason for this could be the lack of additional helper proteins, which are needed for efficient RNA editing of many sites in flowering plants (Maeda et al., 2022). Indeed, a synthetic editing factor engineered on the basis of DYW-type PPR protein CLB19 of *A. thaliana* was shown to increase its activity in *E. coli* when co-expressed with the seed plant specific RNA editing helper protein MORF9 (Gutmann et al., 2020; Royan et al., 2021). Other PPR protein chimeras with the PPR stretch of PpPPR56 combined with DYW domains of different *Arabidopsis thaliana* editing factors were also tested in bacteria and most of them were shown to be inactive (Maeda et al., 2022; Takenaka et al., 2021).

We also tested PpPPR78 with the E1E2DYW domain of NgPPR45 in the *E. coli* system. Neither the co-transcribed cox1eU755SL nor rps14eU137SL were edited by the chimera in the heterologous system. It remains to be seen, if bacterial factors hinder the editing activity of that protein in bacteria or if other unknown mediators are needed to support the editing in plant organelles.

The reduced editing of cox1eU755SL and the lack of editing of rps14eU137SL by chimera PpPPR78-NgPPR45EDYW could at the end also be a consequence of a reduced compatibility of the C-terminal domain of NgPPR45 with the upstream *cis* element of the *rps14* site (Takenaka et al., 2022).

NgPPR45 might be the best candidate to edit the two endogenous editing sites in *Naegleria gruberi* mitochondria

As NgPPR45 shows a clear target signal for mitochondrial localization and fits best to cox1eU1120HY in the mitochondrial transcriptome of *Naegleria gruberi* in comparison to the other identified DYW-type PPR proteins in *Naegleria,* participation of this editing factor in editing of that site is likely. The PPR stretch of NgPPR45 also moderately fits to the second editing site cox3eU780RW with five matches and only one mismatch in the core region of recognition. Such mismatches were even found to be relevant for proper editing of targets like shown for *P. patens* PPR65. Upon removal of the mismatch by modifying the target

sequence, ccmFCeU103PS is edited less efficiently than the original target in the heterologous *E. coli* system (Oldenkott et al., 2019).

Anyway, the final proof, that this protein edits one or even both editing sites in the mitochondrial transcriptome of Naegleria gruberi is still lacking. The high number of repetitive elements within the PPR stretch (Supplementary figure 1) hindered us to synthesize or amplify the complete protein gene for expression in an heterologous system (Lesch et al., 2022; Oldenkott et al., 2019) or at least to perform electromobility shift assays to test the binding to the appropriate targets (Kindgren et al., 2015; Matsuda et al., 2020; Schallenberg-Rüdinger et al., 2013a).

Unknown functions of the other nine DYW-type PPR proteins in *Naegleria gruberi*

The function of the other DYW-type PPR proteins encoded in the genome of the protist, however, remains puzzling. Protein models for genes Naegr76525, Naegr46207 and Naegr66503 consist of up to six PPR repeats and a degenerated C-terminal domain only. These genes might be pseudogenes or cryptic truncated PPR protein genes as also found sporadically in diverse species outside of land plants such as chlorophyte algae (Gutmann et al., 2020) or in the charophyte algae *Nitella hyalina* (Schallenberg-Rüdinger et al., 2013b). Naegr76708 shows a deletion of the E1 and E2 domain and an incomplete DYW domain and can be excluded as functional RNA editing factor as well. For NgPPR51, a function in mitochondria is likely due to the clear mitochondrial signal peptide of the protein, but no editing target matches the PPR array of that protein. If the protein has another function in RNA processing via interaction with a particular intergenic RNA region like chloroplast DYW-type PPR protein CRR2 in *A. thaliana* (Hashimoto et al., 2003; Ruwe et al., 2018) or participating in splicing like PpPPR43 in *P. patens* (Ichinose et al., 2012) cannot be answered yet.

For the other four DYW-type PPR proteins, a function in RNA editing cannot be ruled out to date. Naegr70351 has 14 PPR repeats which are likely not arranged in the PLS triplet manner, typically found in editing factors. DYW-type PPR proteins Naegr69406, Naegr76708 and Naegr32041, when the N-terminal

elongation is included, present long PPR binding regions with 25, 23 and 28 PPR repeats, respectively. These are PPR repeat regions longer than the stretch of most editing factors identified in land plants like *Arabidopsis thaliana* or *Physcomitrium patens* (editing factor overview available via the PREPACT search tool Edifacts (Lenz et al., 2018)). As only NgPPR51 and NgPPR45 possess a clear signal peptide for mitochondrial localization, one may speculate on a function in RNA editing of nuclear-cytosolic transcripts. In nature, no single case of a DYW-type PPR protein acting on a nuclear transcript is known so far, but it was shown recently that moss editing factor PpPPR56 upon expression in human cells not only edits its endogenous delivered target but also numerous off-targets in the cytosolic transcriptome (Lesch et al., 2022). This proves the functionality of DYW-type editing factors in the cytosol in principle.

Materials and Methods

Amplification and sequence analysis of NgPPR45423 and NgPPR45424

Nucleic acids were prepared from *Naegleria gruberi* strain NEG-M as described in Rüdinger et al. (2011a) and kindly provided by Dr. Lillian Fritz-Laylin (Fritz-Laylin et al., 2010). RNA was treated with DNase I (Thermofisher) to remove vestiges of DNA. First strand cDNA was synthesized using the NucleoSpin RNA kit (Macherey-Nagel) and oligodT_{18} primers. Different primers (Supplementary table 3) were used to amplify NgPPR45423 and NgPPR45424 on DNA and cDNA level with Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific). PCR products were gel-purified (BLIRT kit), sequenced (Macrogen Europe) and aligned with the *Naegleria gruberi* genome sequence (Fritz-Laylin et al., 2010), 2010, NCBI Genebank entry NW_003163326.1) using MEGA7 (Kumar et al., 2016) and manual adjustment (Supplementary data 1 and 3).

Plant material and growth conditions

Physcomitrium patens (Hedw.) Bruch & Schimp., Gransden (Rensing et al., 2020; Rensing et al., 2008) wild type and KO PpPPR78 Gransden (Rüdinger et al., 2011b) were cultivated following Oldenkott et al. (2020). Gametophores were cultivated on modified Knop medium plates (250 mg/L KH_2PO_4, 250 mg/L KCl, 250 mg/L $MgSO_4x7H_2O$, 1000 mg/ $Ca(NO_3)_2x4H_2O$, 12.5 mg/L $FeSO_4x7H_2O$, 0.22 mM $CuSO_4$, 0.19 mM $ZnSO_4$,

10 mM H_3BO_3, 0.1 mM Na_2MoO_4, 2 mM $MnCl_2$, 0.23 mM $CoCl_2$, and 0.17 mM KI, pH 5.8, 1% [w/v] agar; (Rüdinger et al., 2011b) at 21°C, with a 16-h-light (photosynthetic photon flux density of 65 mmol/m^2/s, neon tubes, Osram HO 39W/865 Lumilux Cool Daylight)/8-h-dark cycle.

Complementation of *Physcomitrium patens* ppr78 KO line

Physcomitrium patens DNA was prepared using the CTAB preparation method (Doyle and Doyle, 1987). Respective primers were used to amplify parts of PpPPR78 and NgPPR45 or NgPPR51 coding sequences in initial PCRs to be fused subsequently in an overlap extension PCR (Higuchi et al., 1988), primers see supplementary table 2). Fusion PCR products were gel-purified and inserted into plasmid PIG_AN between the rice actin1 promoter and the nos terminator. The constructs are flanked by regions homologous to the *P. patens* intergenic (PIG) region (Okuda et al., 2009) to be inserted into the *P. patens* genome via homologous recombination (Schallenberg-Rüdinger et al., 2017). Constructs were introduced into KO PpPPR78 ecotype Gransden protoplasts using polyethylene glycol–mediated transformation (Hohe et al., 2004) as described in Oldenkott et al. (2020). For selection, mutant lines were cultivated on Knop agar (1% w/v) with 30 µg/ml hygromycin B.

For detection of stable integration of the gene construct, DNA was prepared using the quick extraction method described in Edwards et al. 1991. Genotyping PCRs using TaqNova DNA Polymerase (Blirt) or Q5 High-Fidelity DNA Polymerase (New England Biolabs) were performed following Schallenberg-Rüdinger et al. (2017). Primers Act1Pfor and NosTrev2 were used to confirm the presence of the transgene. The correct orientation of the construct and the insertion into the PIG region was tested with primer combination PpPIG1gen_for and Act1Pfor (for primers see Supplementary table 3). Transgenic lines, which showed expected PCR products with both primer combinations were considered as stable lines.

RNA editing detection and transgene expression

Three to five single gametophores of each stable plant line and control lines (KO PpPPR78, OE PpPPR78 DYW domain truncated, OE PpPPR78, OE PpPPR78 fused with PpPPR79 (E1E2 and) DYW domain, published in Schallenberg-Rüdinger et al. (2017)) were transferred to fresh Knop plates for standardized growth under conditions as described above. After 42 days, equal amounts of plants were harvested for RNA extraction by using the NucleoSpin RNA kit (Macherey-Nagel), followed by DNase I treatment (Thermo Fisher Scientific).

To detect RNA editing, cDNA was synthesized from DNase treated RNA by using random hexamer primer (Roth) and RevertAid reverse transcriptase (Thermo Fisher Scientific). Primer pair 11altfor/16rev and PPrps14for/PPrps14rev was used to amplify the target sequence containing editing positions cox1eU755SL and rps14eU137SL, respectively. PCR assays included cDNA corresponding to 4 ng of RNA, 0.2 µM of each primer, 0.2 mM dNTPs, 1 unit of Taq polymerase with 1x supplied PCR buffer in double-distilled water in total volume of 25 µl. Amplification went through 5 min of initiation followed by 35 cycles, each including 30 sec degeneration at 96 °C, 30 sec annealing at 45 °C, 1 min elongation at 72 °C, and a 5 min finishing step at 72 °C. Purified PCR products were sequenced and RNA editing was detected by using BioEdit 7.0.5.3 (Hall, 1997). The editing efficiency was quantified by the ratio of the thymidine peak height to the sum of thymidine and cytidine peak heights in the chromatogram position, corresponding to the editing site.

Quantitative real time PCR

To quantify the expression of the inserted PPR chimera in the different generated plant lines, 62 ng of DNase treated RNA were used for cDNA synthesis per 20 µl assay with oligodT_{18} primers. The real-time PCR was performed using the SYBR green master mix (Invitrogen) with cDNA corresponding to 3.1 ng initial total RNA per 20 µl assay. cDNA was analyzed on a Bio-Rad CFX96 Real-Time system with the following program: 95°C for 10 min, 50 cycles of 95°C for 15 sec, 58°C for 20 sec and 72°C for 20 sec, finishing with the melt stage: 95°C for 30 sec, 48°C to 95°C with 0.5°C increase each 5 sec. Primer qpcr78-for-all1 and

460 qpcr78-rev-all1 were used to amplify PpPPR78 and PpPPR78 chimeras. Reference gene *Ade PRT*
461 (Phypa_443007) was used for normalization as recommended by Bail et al. (2013). Triplicate
462 measurements were performed for each of three biological replicates. Measurements were analyzed using
463 the delta-delta-Ct method. Melting curves were analyzed to ensure product specificity.

464 Acknowledgements

465 We wish to thank Dr. Lilian Fritz-Laylin for providing DNA and RNA of *Naegleria gruberi* and Monika
466 Polsakiewicz and Sarah Brenner for technical assistance. The authors gratefully acknowledge funding by
467 Deutsche Forschungsgemeinschaft (DFG) SCHA 1952/2-2 to M.S.-R.

468 Author contributions

469 B.O and M.S-R. cloned the constructs. B.O. and E.L. introduced them into *P. patens* KO PPR78. Y.Y, B.O and
470 E.L. performed the genotyping and RNA editing analysis. Y.Y. did the genome locus analysis and performed
471 the qRT-PCR experiments with initial help of E.L.. V.K and M.S-R designed and supervised the study
472 program and contributed to data analyses. Y.Y. prepared the figures. M.S-R wrote the manuscript, which
473 was read, edited and ultimately approved by all co-authors.

Figures

Figure 1. The ten DYW-type PPR proteins of *Naegleria gruberi*

Coding sequences of the ten DYW-type PPR proteins encoded in the genome of *Naegleria gruberi* were subscribed based on Fritz-Laylin et al. (2010) and motifs and amino acids at binding positions 5th and Last (L) of each PPR repeat were identified using the PPR finder (https://ppr.plantenergy.uwa.edu.au, Cheng et al 2016). N-terminal sequences upstream of the PLS array, which were predicted as signal peptides for mitochondrial localization are colored in orange (Naegr51788), other sequences and spacer sequences within the PPR protein that were not recognized as PPR motifs or C-terminal E1/E2(green)/DYW(blue) domains are displayed in grey (size adjusted by sequence length). The DYW domain of Naegr76525 and Naegr76708 showing amino acid deletions are colored in light blue. 5th and L amino acids of each PLS repeat are shown and P and S motifs, which are important for binding, are shaded in grey. Binding fit of S and P motifs to the corresponding nucleotides upstream of the two mitochondrial editing sites (bold, underlined) of *Naegleria gruberi*, cox1eU1120HY and cox3eU780RW, are highlighted in green for a match, in yellow for the less favored match and in red for a mismatch, based on the PPR-RNA binding code: T/S + N/S: A>G, T/S + D:G>A, N + S: C>U, N + D: U>C, N + N: C/U (Barkan et al., 2012), respectively..

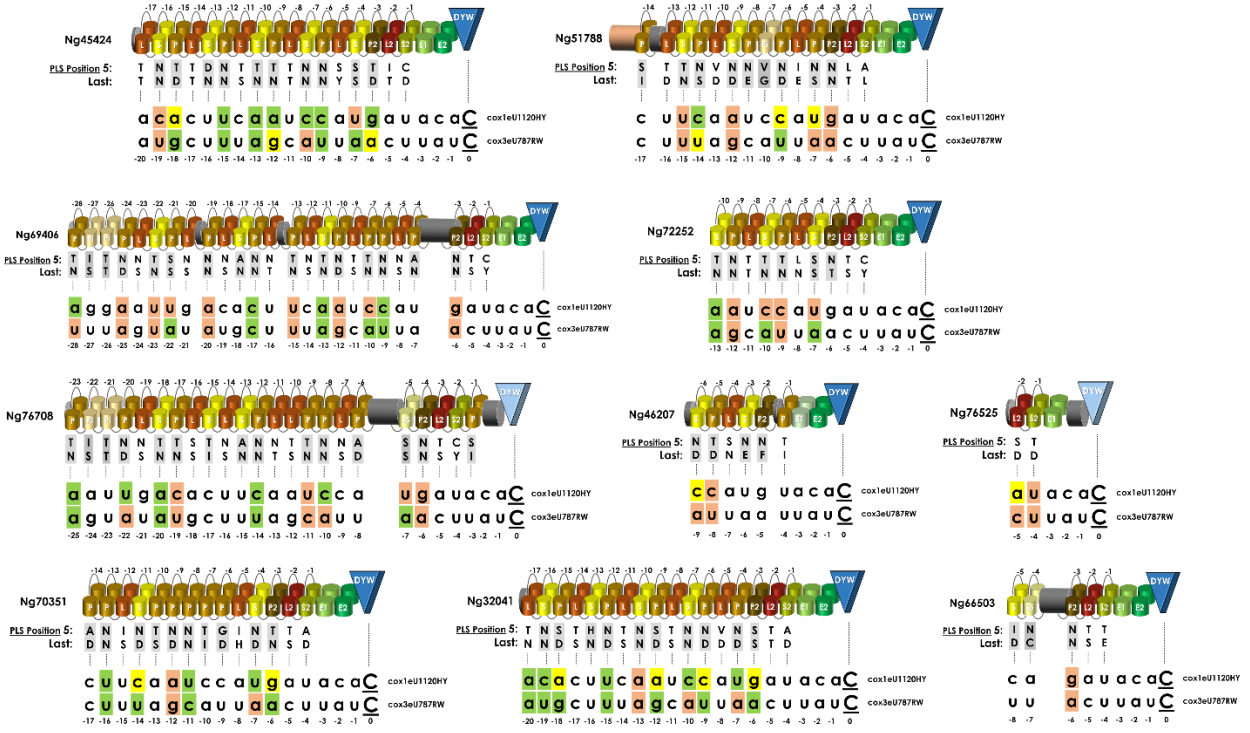
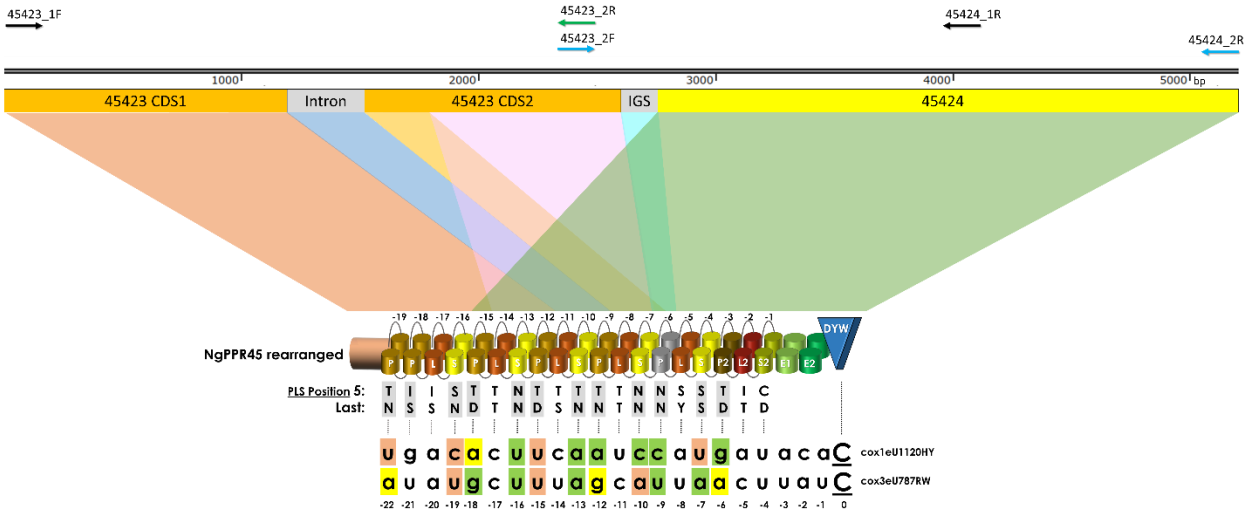


Figure 2. Combination of predicted genes Naegr45423 and Naegr45424 to the gene encoding editing factor NgPPR45.

A. Different primer pairs (arrows) binding to unique parts of predicted genes Nagegr45423 and Naegr45424 were used to amplify the investigated genome region. Sanger sequencing could identify the predicted intron of Naegr45423 and the intergenic region to be part of the coding region. As a result, Naegr45423 and Naegr45424 belong to one reading frame (Supplementary data 1). Sequences within Naegr45423 and Naegr45424 are highly repetitive (shown in different shading), which might have caused difficulties within the initial assembly. The co-transcription of the two predicted genes was finally proven on transcript level. IGS=predicted intergenic region. **B**. Gel electrophoresis of PCR products confirming the proximity of predicted gene regions Naegr45423 and Naegr45424 on genomic DNA level (left) with primer pair 45423_1F and 45423_2R (D1) and 45423_2F and 45424_2R (D2), respectively and on cDNA level with primer pair 45423_1F and 45424_1R (C1).



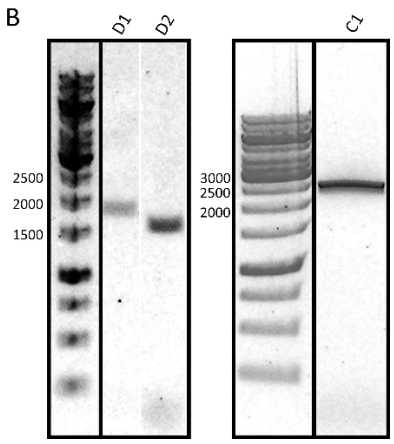


Figure 3. Alignment of C-terminal E1, E2 and DYW domain of NgPPR45 and NgPPR51 of the protist *Naegleria gruberi* and the nine DYW-type PPR editing factors of the moss *Physcomitrium patens*.

Light blue, blue and dark blue indicate amino acid conservation higher than 30%, 50% and 80%, respectively. Conserved motifs associated with the catalytic function of the DYW domain (PG box, zinc binding motifs, DYW motif) or the regulation of activity (gating domain consisting of α1 and ß3 and ß4, (Takenaka et al., 2021) are labeled. Start of the short DYW domain (DYW*) defined by Lurin et al. (2004) is indicated with an arrow, amino acids 37-42 of DYW* suggested to be important for domain compatibility (Ichinose and Sugita, 2018) are highlighted as well. The figure was prepared with Jarview 2.11.2.6.

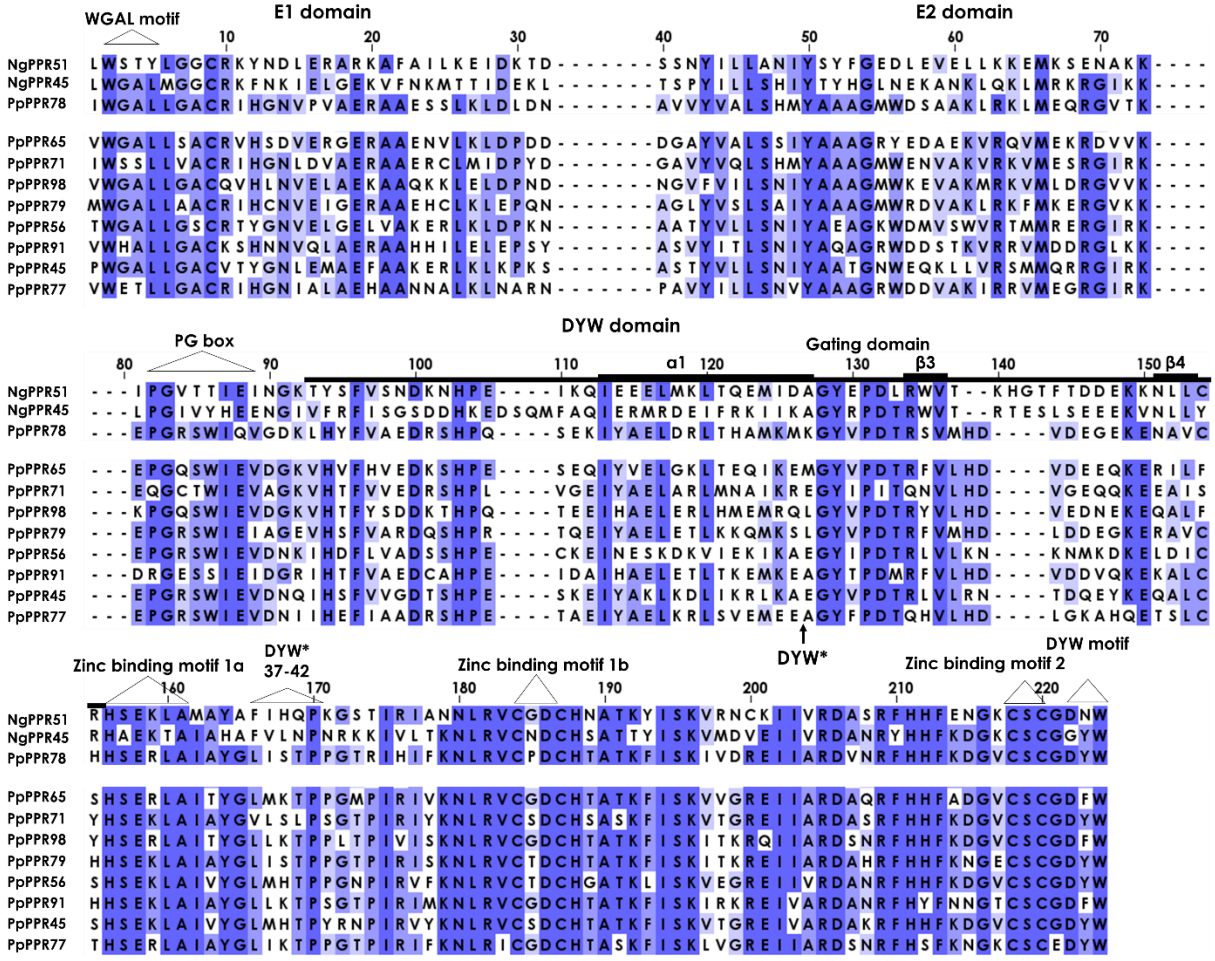


Figure 4. Functional complementation of *P. patens* KO PpPPR78 plants using PpPPR78 chimeras with C-termini of NgPPR45 or NgPPR51 of *Naegleria gruberi*.

Chimeric constructs of PpPPR78 and NgPPR45 with three different fusion points were introduced into the *ppr78* ko line. In wild type *P. patens,* PpPPR78 edits mitochondrial sites cox1eU755SL and rps14eU137SL, with >99% and 60-80% editing efficiency, respectively (shown on top, and >99% presented as 100%, as no C peak is recognizable in the Sanger sequencing chromatogram). Chimeras of PpPPR78 and NgPPR51 did not complement the *ppr78* ko line and editing was not regained in any generated line (see chromatograms of selected lines in the blue box and Supplementary table 2 for the complete data set). (green box) Complementation was successful when the complete C-terminal domain of PpPPR78 including the E1, E2 and DYW domain was replaced by the one of NgPPR45. This resulted in editing of 24% - 82% of cox1eU755SL (shown is the chromatogram of a line with moderate editing of 59%). The rps14eU137SL site was not edited in these complementation lines either.

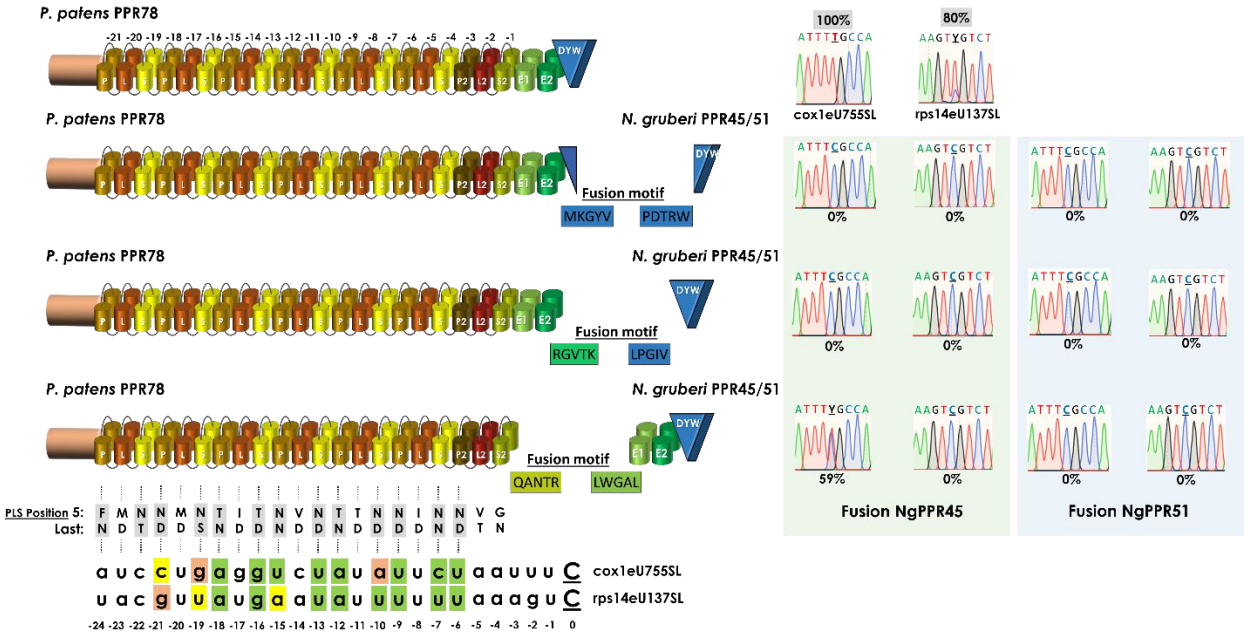
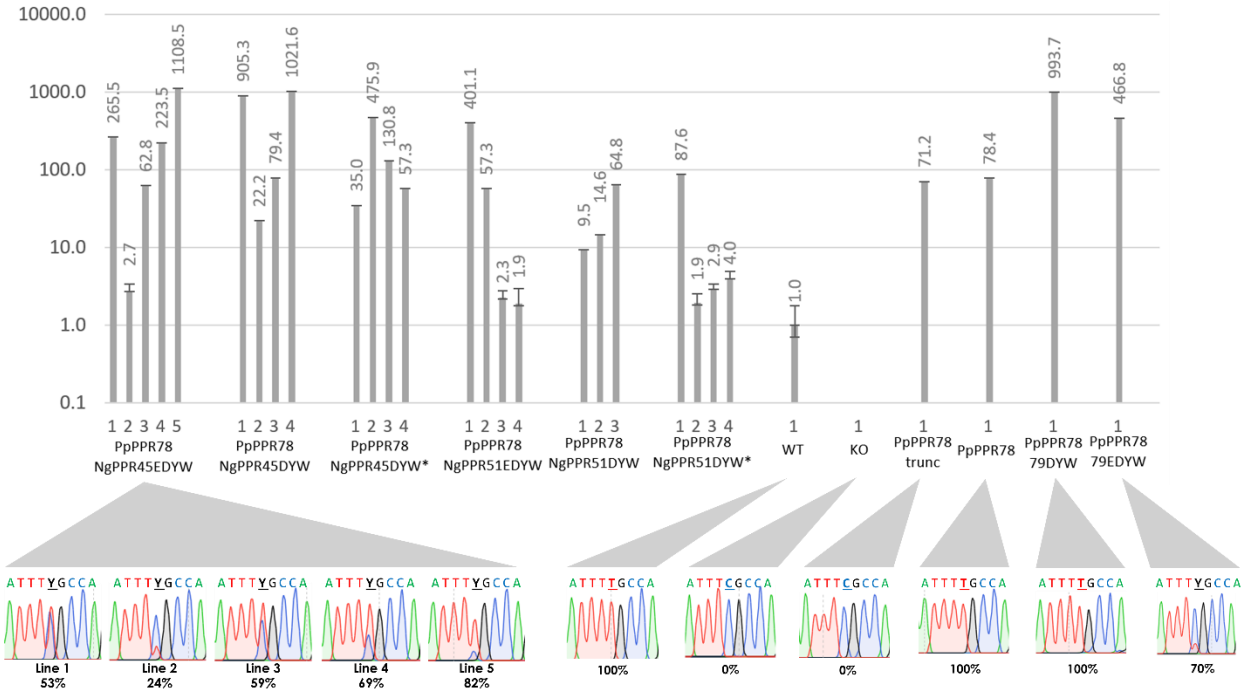


Figure 5. Quantitative real time PCR analysis of complementation lines of *P. patens* KO PpPPR78.

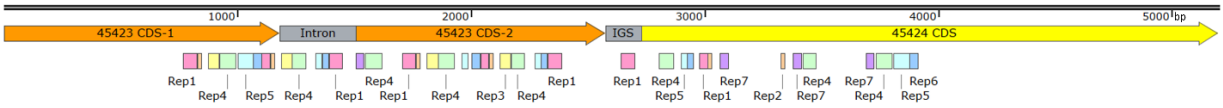
The expression of different PpPPR78 chimeras in each stable overexpression line was analyzed by quantitative RT-PCR in comparison to wildtypic PpPPR78 expression and other complementation lines of an earlier study (right side, (Schallenberg-Rüdinger et al., 2017). KO PpPPR78 was used as the negative control. The values are means of 3 biological replicates (error bars indicate SD). The regain of editing of site cox1eU755SL correlates with the expression levels of introduced PpPPR78-NgPPR45EDYW in the investigated complementation lines.



Supplementary information

Supplementary figure 1. Annotated genome region of Naegr45423 and Naegr45424 contains highly repetitive regions.

Repetitive regions ("Rep") are indicated below in different colored bars. From the end of Naegr45423 CDS-1 until mid of the Naegr45424 CDS, seven different repetitive regions were identified, which are also appearing within the putative intron region of Naegr45423 and the putative intergenic region (IGS). Figure prepared with Snapgene viewer V6.2.1. For details see Supplementary data 1.



Supplementary table 1: Output of PPR Finder for the different PLS-type PPR proteins of *Naegleria gruberi*

PPR models presented on the PPR plantenergy webpage subtool "PPR" (https://ppr.plantenergy.uwa.edu.au/ppr/) and predicted via search tool option "Search for PPR" (https://ppr.plantenergy.uwa.edu.au/fasta/) are displayed for each of the DYW-type PPR proteins from *N. gruberi*, if available. Detected PPR motifs and C-terminal extensions are listed with the region of the motif (start-end), length, score, type, amino acids associated with nucleotide recognition (2nd, 5th and last) and sequence of the motifs detected. End and start of motifs with gaps in between or with an overlap are highlighted in red (column "start-end"). Length number of motifs with unusual length are colored in red (column "length").

Supplementary table 2. *Physcomitrium* KO PPR78 complementation lines

Constructs inserted into KO PPR78 are given with the transgenic line number (Line No) and editing efficiencies (Ed) for both editing sites cox1eU755SL and rps14eU137SL measured for three independent biological replicates (REP) for each line investigated. Average (Ave) and standard deviation (SD) is calculated. Expression levels in relation to WT expression of PPR78 are displayed as well.

Supplementary table 3. Oligonucleotides

Oligonucleotides used in this study. All oligonucleotides were synthesized by IDT (Integrated DNA technologies Europe, BVBA, Leuven, Belgium).

Supplementary table 4. Signal peptide prediction of assembled NgPPR45 and N-terminal extended Ng32041

Results of the prediction of localization of NgPPR45 and N-terminal extended Ng32041 performed with TargetP 2.0 (https://services.healthtech.dtu.dk/service.php?TargetP-2.0) and WoLFPSORT (https://wolfpsort.hgc.jp/).

Supplementary Data

Supplementary data 1. Alignment of sequences to rearrange the annotation of the gene encoding for NgPPR45

The genome region of Naegr45423 and Naegr45424 is aligned with Sanger sequenced PCR products based on DNA and cDNA amplification.

Supplementary data 2. Alignment of the ten DYW-type PPR proteins of *Naegleria gruberi*

Presented are the ten DYW-type PPR proteins of *Naegleria gruberi* aligned with the ten DYW-type PPR proteins of *Physcomitrium patens* for comparison. The rearranged NgPPR45 is presented as well as the N-terminally elongated Naegr32401.

Supplementary data 3. Sanger sequencing file for NgPPR45 assembly and editing analysis of plants

Collection of Sanger sequencing results of PCR products used for NgPPR45 assembly and Sanger sequencing results of PCR products to evaluate the editing of cox1eU755SL and rps14eU137SL in the different transgenic lines generated within this study.

References

Andrés-Colás, N., Q. Zhu, M. Takenaka, B.D. Rybel, D. Weijers, and D.V.D. Straeten. 2017. Multiple PPR protein interactions are involved in the RNA editing system in Arabidopsis mitochondria and plastids. *PNAS*.

Bail, A.L., S. Scholz, and B. Kost. 2013. Evaluation of Reference Genes for RT qPCR Analyses of Structure-Specific and Hormone Regulated Gene Expression in Physcomitrella patens Gametophytes. *PLoS One*. 8.

Banks, J.A., T. Nishiyama, M. Hasebe, J.L. Bowman, M. Gribskov, ..., and I.V. Grigoriev. 2011. The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science*. 332:960-963.

Barkan, A., M. Rojas, S. Fujii, A. Yap, Y.S. Chong, C.S. Bond, and I. Small. 2012. A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet*. 8:e1002910.

Bentolila, S., A.M. Babina, A. Germain, and M.R. Hanson. 2013. Quantitative trait locus mapping identifies REME2, a PPR-DYW protein required for editing of specific C targets in Arabidopsis mitochondria. *RNA Biol*. 10:1520-1525.

Bentolila, S., W.P. Heller, T. Sun, A.M. Babina, G. Friso, K.J. van Wijk, and M.R. Hanson. 2012. RIP1, a member of an Arabidopsis protein family, interacts with the protein RARE1 and broadly affects RNA editing. *Proc Natl Acad Sci U S A*. 109:E1453-1461.

Blatch, G.L., and M. Lässle. 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays*. 21:932-939.

Boussardon, C., A. Avon, P. Kindgren, C.S. Bond, M. Challenor, C. Lurin, and I. Small. 2014. The cytidine deaminase signature HxE(x)n CxxC of DYW1 binds zinc and is necessary for RNA editing of ndhD-1. *New Phytol*. 203:1090-1095.

Boussardon, C., V. Salone, A. Avon, R. Berthome, K. Hammani, K. Okuda, T. Shikanai, I. Small, and C. Lurin. 2012. Two interacting proteins are necessary for the editing of the NdhD-1 site in Arabidopsis plastids. *Plant Cell*. 24:3684-3694.

Bundschuh, R. 2015. The search for the determinants of insertional RNA editing in Physarum polycephalum mitochondrial RNAs. *In* The First International Workshop on Physarum Transport Networks 2015.

Bundschuh, R., J. Altmüller, C. Becker, P. Nürnberg, and J.M. Gott. 2011. Complete characterization of the edited transcriptome of the mitochondrion of Physarum polycephalum using deep sequencing of RNA. *Nucleic Acids Research*. 39:6044-6055.

Cheng, S., B. Gutmann, X. Zhong, Y. Ye, M.F. Fisher, F. Bai, I. Castleden, Y. Song, B. Song, J. Huang, X. Liu, X. Xu, B.L. Lim, C.S. Bond, S.M. Yiu, and I. Small. 2016. Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. *Plant J*. 85:532-547.

Diaz, M.F., S. Bentolila, M.L. Hayes, M.R. Hanson, and R.M. Mulligan. 2017. A protein with an unusually short PPR domain, MEF8, affects editing at over 60 Arabidopsis mitochondrial C targets of RNA editing. *Plant J*. 92:638-649.

Dong, S., C. Zhao, S. Zhang, H. Wu, W. Mu, T. Wei, N. Li, T. Wan, H. Liu, J. Cui, R. Zhu, B. Goffinet, and Y. Liu. 2019. The Amount of RNA Editing Sites in Liverwort Organellar Genes Is Correlated with GC Content and Nuclear PPR Protein Diversity. *Genome Biology and Evolution*. 11:3233-3239.

Doyle, J.J., and J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*. 19:11-15.

Duff, R.J. 2006. Divergent RNA editing frequencies in hornwort mitochondrial nad5 sequences. *Gene*. 366:285-291.

Faktorová, D., R.E.R. Nisbet, J.A.F. Robledo, E. Casacuberta, L. Sudek, ..., and J. Lukeš. 2020. Genetic tool development in marine protists: emerging model organisms for experimental cell biology. *Nature Methods*. 17:pages481–494.

Fritz-Laylin, L.K., S.E. Prochnik, M.L. Ginger, J.B. Dacks, M.L. Carpenter, M.C. Field, A. Kuo, A. Paredez, J. Chapman, J. Pham, S. Shu, R. Neupane, M. Cipriano, J. Mancuso, H. Tu, A. Salamov, E. Lindquist, H. Shapiro, S. Lucas, I.V. Grigoriev, W.Z. Cande, C. Fulton, D.S. Rokhsar, and S.C. Dawson. 2010. The genome of Naegleria gruberi illuminates early eukaryotic versatility. *Cell*. 140:631-642.

Fu, C.-J., S. Sheikh, W. Miao, S.G.E. Andersson, and S.L. Baldauf. 2014. Missing Genes, Multiple ORFs, and C-to-U Type RNA Editing in Acrasis kona (Heterolobosea, Excavata) Mitochondrial DNA. *Genome Biol Evol*. 6:2240-2257.

Fuchs, P., N. Rugen, C. Carrie, M. Elsässer, I. Finkemeier, J. Giese, T.M. Hildebrandt, K. Kühn, V.G. Maurino, C. Ruberti, M. Schallenberg-Rüdinger, J. Steinbeck, H.-P. Braun, H. Eubel, E.H. Meyer, S.J. Müller-Schüssele, and M. Schwarzländer. 2020. Single organelle function and organization as estimated from Arabidopsis mitochondrial proteomics. *Plant J*. 101:420-441.

Fujii, S., N. Sato, and T. Shikanai. 2013. Mutagenesis of Individual Pentatricopeptide Repeat Motifs Affects RNA Binding Activity and Reveals Functional Partitioning of Arabidopsis PROTON GRADIENT REGULATION3. *The Plant Cell*. 25.

Fujii, S., and I. Small. 2011. The evolution of RNA editing and pentatricopeptide repeat genes. *New Phytol*. 191:37-47.

Gerke, P., P. Szövényi, A. Neubauer, H. Lenz, B. Gutmann, R. McDowell, I. Small, M. Schallenberg-Rüdinger, and V. Knoop. 2020. Towards a plant model for enigmatic U-to-C RNA editing: the organelle genomes, transcriptomes, editomes and candidate RNA editing factors in the hornwort Anthoceros agrestis. *New Phytol*. 225:1976-1992.

Gipson, A.B., M.R. Hanson, and S. Bentolila. 2022. The RanBP2 zinc finger domains of chloroplast RNA editing factor OZ1 are required for protein-protein interactions and conversion of C to U. *Plant J*. 109:215-226.

Gott, J.M., N. Parimi, and R. Bundschuh. 2005. Discovery of new genes and deletion editing in Physarum mitochondria enabled by a novel algorithm for finding edited mRNAs. *Nucleic Acids Res*. 33:5063-5072.

Grewe, F., S. Herres, P. Viehover, M. Polsakiewicz, B. Weisshaar, and V. Knoop. 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.

Guillaumot, D., M. Lopez-Obando, K. Baudry, A. Avon, G. Rigaill, A. Falcon de Longevialle, B. Broche, M. Takenaka, R. Berthome, G. De Jaeger, E. Delannoy, and C. Lurin. 2017. Two interacting PPR proteins are major Arabidopsis editing factors in plastid and mitochondria. *Proc Natl Acad Sci U S A*. 114:8877-8882.

Gutmann, B., S. Royan, M. Schallenberg-Rüdinger, H. Lenz, I.R. Castleden, R. McDowell, M.A. Vacher, J. Tonti-Filippini, C.S. Bond, V. Knoop, and I.D. Small. 2020. The Expansion and Diversification of Pentatricopeptide Repeat RNA-Editing Factors in Plants. *Molecular Plant*. 13.

Hall, T. 1997. Biological sequence alignment editor for Win95/98/NT/2K/XP.

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

Hayes, M.L., and P.I. Santibanez. 2020. A plant pentatricopeptide repeat protein with a DYW-deaminase domain is sufficient for catalyzing C-to-U RNA editing in vitro. *J Biol Chem*. 295:3497-3505.

Hecht, J., F. Grewe, and V. Knoop. 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.

Higuchi, R., B. Krummel, and R.K. Saiki. 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res*. 16:7351-7367.

Hohe, A., T. Egener, J.M. Lucht, H. Holtorf, C. Reinhard, G. Schween, and R. Reski. 2004. An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, Physcomitrella patens. *Curr Genet*. 44:339-347.

Ichinose, M., M. Kawabata, Y. Akaiwa, Y. Shimajiri, I. Nakamura, T. Tamai, T. Nakamura, Y. Yagi, and B. Gutmann. 2022. U-to-C RNA editing by synthetic PPR-DYW proteins in bacteria and human culture cells. *Communications Biology*. 5.

Ichinose, M., and M. Sugita. 2018. The DYW Domains of Pentatricopeptide Repeat RNA Editing Factors Contribute to Discriminate Target and Non-Target Editing Sites. *Plant Cell Physiol*. 59:1652-1659.

Ichinose, M., E. Tasaki, C. Sugita, and M. Sugita. 2012. A PPR-DYW protein is required for splicing of a group II intron of cox1 pre-mRNA in Physcomitrella patens. *Plant J*. 70.

Ichinose, M., M. Uchida, and M. Sugita. 2014. Identification of a pentatricopeptide repeat RNA editing factor in Physcomitrella patens chloroplasts. *FEBS Lett*. 588:4060-4064.

Iyer, L.M., D. Zhang, I.B. Rogozin, and L. Aravind. 2011. Evolution of the deaminase fold and multiple origins of eukaryotic editing and mutagenic nucleic acid deaminases from bacterial toxin systems. *Nucleic Acids Res*. 39:9473-9497.

Kindgren, P., A. Yap, C.S. Bond, and I. Small. 2015. Predictable alteration of sequence recognition by RNA editing factors from Arabidopsis. *Plant Cell*. 27:403-416.

Knie, N., F. Grewe, S. Fischer, and V. Knoop. 2016. Reverse U-to-C editing exceeds C-to-U RNA editing in some ferns - a monilophyte-wide comparison of chloroplast and mitochondrial RNA editing suggests independent evolution of the two processes in both organelles. *BMC Evolutionary Biology*. 16.

Knoop, V. 2011. When you can't trust the DNA: RNA editing changes transcript sequences. *Cell Mol Life Sci*. 68:567-586.

Knoop, V. 2022. C-to-U and U-to-C: RNA editing in plant organelles and beyond. *Journal of Experimental Botany*. erac488.

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

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 33:1870-1874.

Lenz, H., A. Hein, and V. Knoop. 2018. Plant organelle RNA editing and its specificity factors: enhancements of analyses and new database features in PREPACT 3.0. *BMC Bioinformatics*. 19:255.

Lesch, E., M.T. Schilling, S. Brenner, Y. Yang, O.J. Gruss, V. Knoop, and M. Schallenberg-Rüdinger. 2022. Plant mitochondrial RNA editing factors can perform targeted C-to-U editing of nuclear transcripts in human cells. *Nucleic Acids Res*. 50:9966-9983.

Loiacono, F.V., D. Walther, S. Seeger, W. Thiele, I. Gerlach, D. Karcher, M.A. Schöttler, R. Zoschke, and R. Bock. 2022. Emergence of Novel RNA-Editing Sites by Changes in the Binding Affinity of a Conserved PPR Protein. *Molecular Biology and Evolution*. 39.

Lurin, C., C. Andres, S. Aubourg, M. Bellaoui, F. Bitton, C. Bruyere, M. Caboche, C. Debast, J. Gualberto, B. Hoffmann, A. Lecharny, M. Le Ret, M.L. Martin-Magniette, H. Mireau, N. Peeters, J.P. Renou, B. Szurek, L. Taconnat, and I. Small. 2004. Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*. 16:2089-2103.

Maeda, A., S. Takenaka, T. Wang, B. Frink, T. Shikanai, and M. Takenaka. 2022. DYW deaminase domain has a distinct preference for neighboring nucleotides of the target RNA editing sites. *The plant Journal*. 111:756-767.

Matsuda, T., M. Sugita, and M. Ichinose. 2020. The L motifs of two moss pentatricopeptide repeat proteins are involved in RNA editing but predominantly not in RNA recognition. *PLOS ONE*.

718 Meier, J.C., C. Henneberger, I. Melnick, C. Racca, R.J. Harvey, U. Heinemann, V. Schmieden, and R. Grantyn.
719 2005. RNA editing produces glycine receptor α3P185L, resulting in high agonist potency. *Nature*
720 *Neuroscience*. 8:736-744.
721 Miyata, Y., and M. Sugita. 2004. Tissue- and stage-specific RNA editing of rps 14 transcripts in moss
722 (Physcomitrella patens) chloroplasts. *J Plant Physiol*. 161:113-115.
723 Mungpakdee, S., C. Shinzato, T. Takeuchi, T. Kawashima, R. Koyanagi, K. Hisata, M. Tanaka, H. Goto, M.
724 Fujie, S. Lin, N. Satoh, and E. Shoguchi. 2014. Massive gene transfer and extensive RNA editing of
725 a symbiotic dinoflagellate plastid genome. *Genome Biol Evol*. 6:1408-1422.
726 Nishikura, K. 2010. Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem*.
727 79:321-349.
728 O'Toole, N., M. Hattori, C. Andres, K. Iida, C. Lurin, C. Schmitz-Linneweber, M. Sugita, and I. Small. 2008.
729 On the expansion of the pentatricopeptide repeat gene family in plants. *Mol bio Evol.* 25:1120-
730 1128.
731 Okuda, K., A.L. Chateigner-Boutin, T. Nakamura, E. Delannoy, M. Sugita, F. Myouga, R. Motohashi, K.
732 Shinozaki, I. Small, and T. Shikanai. 2009. Pentatricopeptide repeat proteins with the DYW motif
733 have distinct molecular functions in RNA editing and RNA cleavage in Arabidopsis chloroplasts.
734 *Plant Cell*. 21:146-156.
735 Okuda, K., F. Myouga, R. Motohashi, K. Shinozaki, and T. Shikanai. 2007. Conserved domain structure of
736 pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proc Natl Acad Sci U S A*.
737 104:8178-8183.
738 Okuda, K., H. Shoki, M. Arai, T. Shikanai, I. Small, and T. Nakamura. 2014. Quantitative analysis of motifs
739 contributing to the interaction between PLS-subfamily members and their target RNA sequences
740 in plastid RNA editing. *Plant J*. 80:870-882.
741 Oldenkott, B., M. Burger, A.-C. Hein, A. Jörg, J. Senkler, H.-P. Braun, V. Knoop, M. Takenaka, and M.
742 Schallenberg-Rüdinger. 2020. One C-to-U RNA Editing Site and Two Independently Evolved Editing
743 Factors: Testing Reciprocal Complementation with DYW-Type PPR Proteins from the Moss
744 Physcomitrium (Physcomitrella) patens and the Flowering Plants Macadamia integrifolia and
745 Arabidopsis. *Plant Cell*. 32:2997-3018.
746 Oldenkott, B., K. Yamaguchi, S. Tsuji-Tsukinoki, N. Knie, and V. Knoop. 2014. Chloroplast RNA editing going
747 extreme: more than 3400 events of C-to-U editing in the chloroplast transcriptome of the
748 lycophyte Selaginella uncinata. *RNA*. 20:1499-1506.
749 Oldenkott, B., Y. Yang, E. Lesch, V. Knoop, and M. Schallenberg-Rüdinger. 2019. Plant-type
750 pentatricopeptide repeat proteins with a DYW domain drive C-to-U RNA editing in Escherichia coli.
751 *Communications Biology*. 2.
752 Puttick, M.N., J.L. Morris, T.A. Williams, C.J. Cox, D. Edwards, P. Kenrick, S. Pressel, C.H. Wellman, H.
753 Schneider, D. Pisani, and P.C.J. Donoghue. 2018. The Interrelationships of Land Plants and the
754 Nature of the Ancestral Embryophyte. *Curr Biol*. 28:733-745.
755 Rensing, S.A., B. Goffinet, R. Meyberg, S.-Z. Wu, and M. Bezanilla. 2020. The Moss Physcomitrium
756 (Physcomitrella) patens: A Model Organism for Non-Seed Plants. *Plant Cell*. 32:1361-1376.
757 Rensing, S.A., D. Lang, A.D. Zimmer, A. Terry, A. Salamov, H. Shapiro, T. Nishiyama, P.F. Perroud, E.A.
758 Lindquist, Y. Kamisugi, T. Tanahashi, K. Sakakibara, T. Fujita, K. Oishi, I.T. Shin, Y. Kuroki, A. Toyoda,
759 Y. Suzuki, S. Hashimoto, K. Yamaguchi, S. Sugano, Y. Kohara, A. Fujiyama, A. Anterola, S. Aoki, N.
760 Ashton, W.B. Barbazuk, E. Barker, J.L. Bennetzen, R. Blankenship, S.H. Cho, S.K. Dutcher, M. Estelle,
761 J.A. Fawcett, H. Gundlach, K. Hanada, A. Heyl, K.A. Hicks, J. Hughes, M. Lohr, K. Mayer, A.
762 Melkozernov, T. Murata, D.R. Nelson, B. Pils, M. Prigge, B. Reiss, T. Renner, S. Rombauts, P.J.
763 Rushton, A. Sanderfoot, G. Schween, S.H. Shiu, K. Stueber, F.L. Theodoulou, H. Tu, Y. Van de Peer,
764 P.J. Verrier, E. Waters, A. Wood, L. Yang, D. Cove, A.C. Cuming, M. Hasebe, S. Lucas, B.D. Mishler,
765 R. Reski, I.V. Grigoriev, R.S. Quatrano, and J.L. Boore. 2008. The Physcomitrella genome reveals
766 evolutionary insights into the conquest of land by plants. *Science*. 319:64-69.

Royan, S., B. Gutmann, C.C.d. Francs-Small, S. Honkanen, J. Schmidberger, A. Soet, Y.K. Sun, L.V.P. Sanglard, C.S. Bond, and I. Small. 2021. A synthetic RNA editing factor edits its target site in chloroplasts and bacteria. *Communications Biology*. 4.

Rüdinger, M., L. Fritz-Laylin, M. Polsakiewicz, and V. Knoop. 2011a. Plant-type mitochondrial RNA editing in the protist Naegleria gruberi. *RNA*. 17:2058-2062.

Rüdinger, M., H.T. Funk, S.A. Rensing, U.G. Maier, and V. Knoop. 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., M. Polsakiewicz, and V. Knoop. 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.

Rüdinger, M., P. Szovenyi, S.A. Rensing, and V. Knoop. 2011b. Assigning DYW-type PPR proteins to RNA editing sites in the funariid mosses Physcomitrella patens and Funaria hygrometrica. *Plant J*. 67:370-380.

Rüdinger, M., U. Volkmar, H. Lenz, M. Groth-Malonek, and V. Knoop. 2012. Nuclear DYW-type PPR gene families diversify with increasing RNA editing frequencies in liverwort and moss mitochondria. *J Mol Evol*. 74:37-51.

Ruwe, H., B. Gutmann, C. Schmitz-Linneweber, I. Small, and P. Kindgren. 2018. The E-domain of CRR2 participates in sequence-specific recognition of RNA in plastids. *New Phytol*.

Salone, V., M. Rüdinger, M. Polsakiewicz, B. Hoffmann, M. Groth-Malonek, B. Szurek, I. Small, V. Knoop, and C. Lurin. 2007. A hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Lett*. 581:4132-4138.

Schaap, P., I. Barrantes, P. Minx, N. Sasaki, R.W. Anderson, M. Bénard, K.K. Biggar, N.E. Buchler, R. Bundschuh, X. Chen, C. Fronick, L. Fulton, G. Golderer, N. Jahn, V. Knoop, L.F. Landweber, C. Maric, D. Miller, A.A. Noegel, R. Peace, G. Pierron, T. Sasaki, M. Schallenberg-Rüdinger, M. Schleicher, R. Singh, T. Spaller, K.B. Storey, T. Suzuki, C. Tomlinson, J.J. Tyson, W.C. Warren, E.R. Werner, G. Werner-Felmayer, R.K. Wilson, T. Winckler, J.M. Gott, G. Glöckner, and W. Marwan. 2015. The Physarum polycephalum Genome Reveals Extensive Use of Prokaryotic Two-Component and Metazoan-Type Tyrosine Kinase Signaling. *Genome Biol Evol*. 8:109-125.

Schallenberg-Rüdinger, M., P. Kindgren, A. Zehrmann, I. Small, and V. Knoop. 2013a. A DYW-protein knockout in Physcomitrella affects two closely spaced mitochondrial editing sites and causes a severe developmental phenotype. *Plant J*. 76:420-432.

Schallenberg-Rüdinger, M., H. Lenz, M. Polsakiewicz, J.M. Gott, and V. Knoop. 2013b. A survey of PPR proteins identifies DYW domains like those of land plant RNA editing factors in diverse eukaryotes. *RNA Biol*. 10:1549-1556.

Schallenberg-Rüdinger, M., B. Oldenkott, M. Hiss, P.L. Trinh, V. Knoop, and S.A. Rensing. 2017. A Single-Target Mitochondrial RNA Editing Factor of Funaria hygrometrica Can Fully Reconstitute RNA Editing at Two Sites in Physcomitrella patens. *Plant Cell Physiol*. 58:496-507.

Small, I., J. Melonek, A.-V. Bohne, J. Nickelsen, and C. Schmitz-Linneweber. 2023. Plant Organellar RNA Maturation. *The Plant Cell*.

Small, I.D., M. Schallenberg-Rüdinger, M. Takenaka, H. Mireau, and O. Ostersetzer-Biran. 2020. Plant organellar RNA editing: what 30 years of research has revealed. *The Plant journal: for cell and moelecular biology*. 101:1040-1056.

Su, D., L. Yang, X. Shi, X. Ma, X. Zhou, S.B. Hedges, and B. Zhong. 2021. Large-Scale Phylogenomic Analyses Reveal the Monophyly of Bryophytes and Neoproterozoic Origin of Land Plants. *Molecular Biology and Evolution*. 38:3332–3344.

Sun, T., S. Bentolila, and M.R. Hanson. 2016. The Unexpected Diversity of Plant Organelle RNA Editosomes. *Trends Plant Sci*. 21:962-973.

Sun, T., X. Shi, G. Friso, K. Van Wijk, S. Bentolila, and M.R. Hanson. 2015. A zinc finger motif-containing protein is essential for chloroplast RNA editing. *PLoS Genet*. 11:e1005028.

Takenaka, M., S. Takenaka, T. Barthel, B. Frink, S. Haag, D. Verbitskiy, B. Oldenkott, M. Schallenberg-Rüdinger, C.G. Feiler, M.S. Weiss, G.J. Palm, and G. Weber. 2021. DYW domain structures imply an unusual regulation principle in plant organellar RNA editing catalysis. *Nature Catalysis*. 4:510-522.

Takenaka, M., A. Zehrmann, A. Brennicke, and K. Graichen. 2013. Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. *PLoS One*. 8:e65343.

Takenaka, M., A. Zehrmann, D. Verbitskiy, M. Kugelmann, B. Hartel, and A. Brennicke. 2012. Multiple organellar RNA editing factor (MORF) family proteins are required for RNA editing in mitochondria and plastids of plants. *Proc Natl Acad Sci U S A*. 109:5104-5109.

Toma-Fukai, S., Y. Sawada, A. Maeda, H. Shimizu, T. Shikanai, M. Takenaka, and T. Shimizu. 2022. Structural insight into the activation of an Arabidopsis organellar C-to-U RNA editing enzyme by active site complementation. *The Plant Cell*.

Uchida, M., S. Ohtani, M. Ichinose, C. Sugita, and M. Sugita. 2011. The PPR-DYW proteins are required for RNA editing of rps14, cox1 and nad5 transcripts in Physcomitrella patens mitochondria. *FEBS Lett*. 585:2367-2371.

Uchiyama, H., M. Ichinose, and M. Sugita. 2018. Chloroplast ribonucleoprotein-like proteins of the moss Physcomitrella patens are not involved in RNA stability and RNA editing. *Photosynthetica*. 56:62-66.

Wagoner, J.A., T. Sun, L. Lin, and M.R. Hanson. 2015. Cytidine deaminase motifs within the DYW domain of two pentatricopeptide repeat-containing proteins are required for site-specific chloroplast RNA editing. *J. Biol. Chem.* 290:2957-2968.

Wang, Y., H. Li, Z.-Q. Huang, B. Ma, Y.-Z. Yang, Z.-H. Xiu, L. Wang, and B.-C. Tan. 2022. Maize PPR-E proteins mediate RNA C-to-U editing in mitochondria by recruiting the trans deaminase PCW1. *The Plant Cell*. 35:529-551.

Yan, J., Y. Yao, S. Hong, Y. Yang, C. Shen, Q. Zhang, D. Zhang, T. Zou, and P. Yin. 2019. Delineation of pentatricopeptide repeat codes for target RNA prediction. *Nucleic Acids Res*. 47:3728-3738.