


Review

An Overview of Pentatricopeptide Repeat (PPR) Proteins in the Moss *Physcomitrium patens* and Their Role in Organellar Gene Expression

Mamoru Sugita 

Graduate School of Informatics, Nagoya University, Chikusa-ku, Nagoya 464-8601, Japan; sugita@gene.nagoya-u.ac.jp; Tel.: +81-52-789-5376

Abstract: Pentatricopeptide repeat (PPR) proteins are one type of helical repeat protein that are widespread in eukaryotes. In particular, there are several hundred PPR members in flowering plants. The majority of PPR proteins are localized in the plastids and mitochondria, where they play a crucial role in various aspects of RNA metabolism at the post-transcriptional and translational steps during gene expression. Among the early land plants, the moss *Physcomitrium* (formerly *Physcomitrella*) *patens* has at least 107 PPR protein-encoding genes, but most of their functions remain unclear. To elucidate the functions of PPR proteins, a reverse-genetics approach has been applied to *P. patens*. To date, the molecular functions of 22 PPR proteins were identified as essential factors required for either mRNA processing and stabilization, RNA splicing, or RNA editing. This review examines the *P. patens* PPR gene family and their current functional characterization. Similarities and a diversity of functions of PPR proteins between *P. patens* and flowering plants and their roles in the post-transcriptional regulation of organellar gene expression are discussed.

Keywords: pentatricopeptide repeat (PPR) protein; moss; *Physcomitrium patens*; plastids; mitochondria; post-transcriptional regulation; RNA processing; RNA splicing; RNA editing; RNA stabilization



Citation: Sugita, M. An Overview of Pentatricopeptide Repeat (PPR) Proteins in the Moss *Physcomitrium patens* and Their Role in Organellar Gene Expression. *Plants* **2022**, *11*, 2279. <https://doi.org/10.3390/plants11172279>

Academic Editor: Riccardo Velasco

Received: 1 August 2022

Accepted: 29 August 2022

Published: 31 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Plastids and mitochondria, which are intracellular organelles in plants and algae, evolved from ancestral prokaryotic organisms, cyanobacteria and α -proteobacteria, respectively [1]. Chloroplasts (one type of plastid) carry out oxygen-evolving photosynthesis while mitochondria perform aerobic respiration, processes that are essential for almost all cellular functions [1]. Plastids and mitochondria have retained their own genome and transcription-translation machinery, which is composed of nuclear- and organellar-encoded proteins and RNAs [2,3]. Most organellar genes are organized in clusters and are often co-transcribed as polycistronic precursor RNAs that are then post-transcriptionally processed into multiple intermediate or mature RNAs [4–7]. Precursor, intermediate and mature RNAs are relative stable and accumulate at respective steady-state levels. Such post-transcriptional RNA processing, including RNA splicing, RNA cleavage and translational initiation, represents an important step in the control of organellar gene expression (Figure 1). The stabilization of RNA is also important for efficient translation. This complex processing is accomplished mostly by nuclear-encoded RNA-binding proteins such as pentatricopeptide repeat (PPR) proteins [8–10].

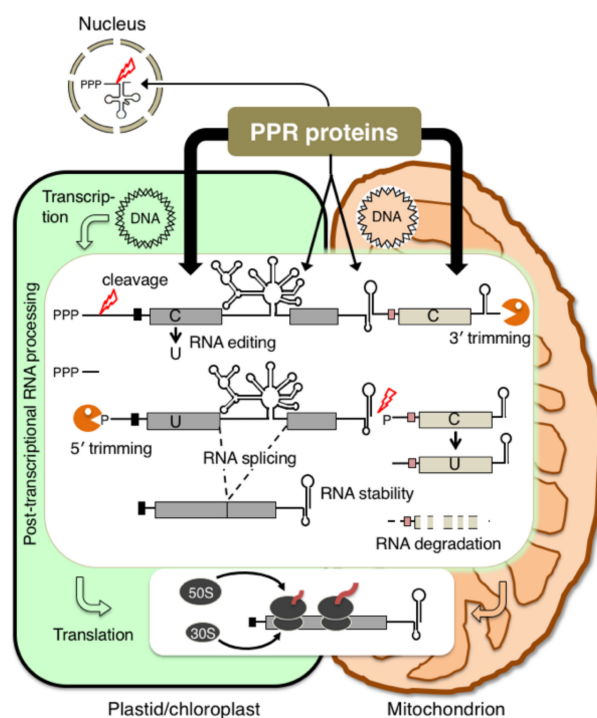


Figure 1. The complexity of post-transcriptional RNA processing events in plastids/chloroplasts and plant mitochondria. Plastid/chloroplast and plant mitochondrial genes are often transcribed as long polycistronic RNA precursors. Several post-transcriptional RNA processing events are necessary to produce the mature mRNAs: 5' end processing by endonucleolytic cleavage, 5' and 3' ends of pre-mRNAs are trimmed by exonucleases (shown by the Pacman icons), specific cytidines (C) are edited to uridines (U), introns are spliced, and intergenic regions are cleaved. Translation initiation is also an important step in the regulation of plant organellar gene expression. All these steps proceed via the participation of numerous nucleus-encoded RNA-binding PPR proteins. Almost all PPR proteins are imported into either plastids/chloroplasts or mitochondria and some members are dually targeted to both organelles. Small members of PPR proteins are imported to the nucleus, where they are involved in pre-tRNA maturation.

The large PPR gene family was first characterized in 2000 by the presence of tandem arrays of a degenerate 35-amino-acid (aa) repeat, which folds into a pair of α -helices and is widely distributed in fungi, animals and plants [11]. The same gene family, named the plant combinatorial and modular protein or PCMP family, was independently reported as a novel family that was unique to plants by Aubourg et al. [12]. They are categorized into a helical repeat protein family, including Pumilio (Puf repeat), β -catenin (Arm repeat) and pp2A (HEAT repeat) [13]. Plant PPR proteins are grouped into two subfamilies, P and PLS. The P subfamily (or P type) consists of canonical 35 aa PPR (P) motifs while the PLS subfamily is organized by repeated units of the P motif, long PPR (L) and short PPR (S) motif variants [14]. In the original definition by Lurin et al. [14], the PLS subfamily was further divided into four types, PLS, E, E+ and DYW, based on their PPR motif composition and characteristic C-terminal domain structures. Thereafter, based on a structural modeling approach, Cheng et al. [15] defined 10 different variants of the PPR motif (P, P1, P2, L1, L2, S1, S2, SS, E1 and E2) and divided the PPR proteins into six types, P, PLS, E1, E2, E+ and DYW (Figure 2). The DYW domain, based on the original definition, was composed of 106 aa [14] while the DYW domain, based on the new definition, becomes expanded to about 136 aa. Unlike vascular plants, only three types (P, PLS and DYW) are found in the moss *Physcomitrium patens*.

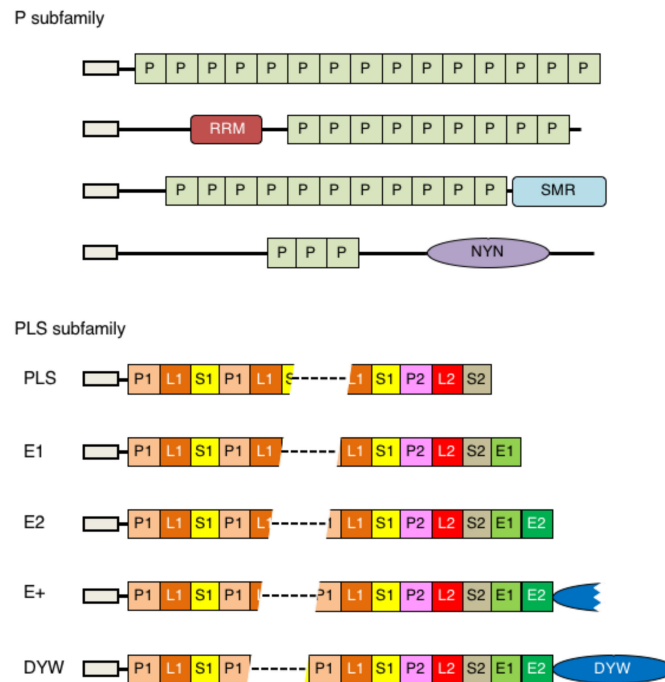


Figure 2. Architecture of pentatricopeptide repeat (PPR) proteins in land plants. Schematic structures of PPR proteins in different subfamilies and types are shown, according to Cheng et al. [15]. N-terminal boxes indicate a transit peptide targeting to organelles. The number of PPR motifs in each protein varies from two to more than thirty, and the first motif can be any of P, P1, L1 or S1. The P subfamily consists of P motifs only or P motifs and additional functional domain(s) such as RRM, SMR or NYN, as is described in the text. The E+ type consists of proteins with a degenerate or truncated DYW domain. The PLS subfamily is composed of 10 to 28 repeated motifs in *Physcomitrium*.

Nucleus-encoded PPR proteins constitute an extraordinarily large family in land plants, comprising 500 members in *Arabidopsis thaliana* (*Arabidopsis*) and 5,000 members in the whisk fern, *Selaginella moellendorffii* [15]. The majority of plant PPR proteins are localized in plastids or mitochondria [14,16] where they play important roles in a wide range of physiological and developmental functions [8–10]. Using forward genetics, various mutants that are defective in photosynthesis, respiration, cytoplasmic male sterility or embryogenesis were isolated and many PPR genes were then identified as being responsible for the mutated phenotypes [8–10,14]. Most PPR proteins that have been investigated are required for various post-transcriptional steps that are associated with RNA in plant organelles [17–24]. Thus, in the early 2000s, functional analyses of PPR proteins were performed using flowering plants, namely *Arabidopsis*, rice and maize. In contrast to studies performed in flowering plants, knowledge regarding PPR proteins in early land plants was limited [25]. In this review, the current knowledge of the function of *P. patens* PPR proteins in plastids and mitochondria is summarized, and attempts are made to highlight the differences and similarities of PPR proteins between mosses and angiosperms.

2. *P. patens* Is a Model Plant for Studying the Molecular Function of PPR Genes

The past two decades, my colleagues and I have engaged in the study of the functional analyses of the moss, *P. patens* PPR (*PpPPR*) proteins with an evolutionary point of view, relative to the PPR proteins in land plants. *P. patens* (hereafter *Physcomitrium*) has emerged as a powerful model system in plant functional genomics studies [26,27]. The genome sequences of its plastid, mitochondrion and nucleus have already been disclosed [28–30]. Gene targeting is feasible in the nuclear and chloroplast genomes [31,32]. Reverse genetics is usually applied to this moss and stable transformants can be easily generated via

homologous recombination [32,33]. Furthermore, the *Physcomitrium* gametophyte, the haploid phase of the life cycle, is dominant, making it possible to study the phenotypes of the knockouts directly without further crosses. Moreover, the *PPR* gene family is rather compact in *Physcomitrium*. Therefore, this moss provides researchers with a good tool for the study of *PPR*, in comparison to flowering plants. To characterize the function of *Physcomitrium* *PPR* proteins, my laboratory has constructed gene-targeted knockout or knockdown mutant lines via homologous recombination. To date, 42 *PpPPR* genes were tagged using an antibiotic resistant gene cassette and the characterization of their mutants is in progress. This reverse-genetics approach has thus far revealed the function of 22 *PpPPR* genes.

3. The Compact *PPR* Gene Family in *Physcomitrium*

Moss *PPR* genes were firstly identified in *Physcomitrium* in 2004 [25]. Thereafter, the whole genome sequence of this moss was disclosed in 2008 [30], and the genome database has been frequently updated. Accordingly, at least 107 *PPR* genes were identified in *Physcomitrium*, and their encoded proteins were numbered sequentially from *PpPPR_1* to *PpPPR_107* (Table 1). The *PPR* gene family is rather small in *Physcomitrium* when it is compared to the large *PPR* gene families in vascular plants [15]. The liverwort, *Marchantia polymorpha*, and the charophytic alga, *Chara braunii*, also possess a compact *PPR* gene family composing between 74 and 57 members, respectively [34,35].

The structure of *Physcomitrium* *PPR* genes has diverged somewhat from the *PPR* genes in *Arabidopsis* and rice. Intron-containing *PPR* genes represent three-fourths of all *PPR* genes in *Physcomitrium* but only one-fourth of them in *Arabidopsis* and rice [36]. Thus, *Physcomitrium* *PPR* genes are generally characterized as intron-rich. The gene structure and encoded amino acid sequence of many *PPR* proteins are conserved in *Physcomitrium* and *Arabidopsis* plants. This conservation suggests that such homologous *PPR* proteins have the same or similar function in mosses and flowering plants. Presumably, intron-rich *PPR* genes may represent “ancient” *PPR* genes that pre-dated the occurrence of the retrotransposition-mediated expansion of the *PPR* gene family in land plants [36]. Among the 107 *PpPPR* proteins, at least 18 paralogous pairs have been found (Table 2) and these may have occurred via gene family expansion due to recent genome duplications [30]. Respective paralogous pairs may have redundant or diverse functions.

Table 1. *Physcomitrium patens* pentatricopeptide repeat (PpPPR) proteins.

Protein Name	Gene Locus ID ¹	Type ²	PPR Motif Bead Patterns and Additional Domain/Motif in Parenthesis ³	Subcellular Localization ⁴		Phenotype of Moss Colony of Gene Knockout Lines	Function Identified in <i>P. patens</i>	Arabidopsis Homologs	Function Identified in Arabidopsis	Refs.
				Pred	Exp					
PpPPR_1	Pp3c3_19290	P	P-P-P-P-P-P-P-P-P-P-P-P	m				At5g50280 (EMB1006)	required for embryo development	[14]
PpPPR_2	Pp3c16_9240	P	P-P-P-P-P-P-P-P-P-P	c	C	Smaller than WT				
PpPPR_3	Pp3c5_10110	P	(RRM)-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	other	C	WT-like		At5g04810 (AtPPR4)	<i>rps12</i> -intron 1 <i>trans</i> -splicing	[37,38]
PpPPR_4	Pp3c17_11510	P	P-P-Pi-P-P-P-P-P-P-P	c	C	Very small colony	RNA splicing of group II intron in pre-tRNA ^{Ile}			[39]
PpPPR_5	Pp3c21_11730	P	P-P-P-P-P-P-P-P	m						
PpPPR_6	Pp3c5_21760	P	P-P-P-P	c	C					
PpPPR_7	Pp3c25_10050	P	P-P-P-P-P-P-Pi-P-P-P-(LAGLIDADG)	c				At2g15820 (OTP51)	<i>ycf3</i> -intron 2 splicing	[40]
PpPPR_8	Pp3c14_19310	P	P-P-P-Pi-P-P-P-P-P-P-P	m		Smaller than WT				
PpPPR_9	Pp3c24_14870	PLS	P1-L1-S1-P1-L1-S1-P1-L1-S1-P2	m	M	Smaller than WT	RNA splicing of <i>cox1</i> intron 3			[41]
PpPPR_10	Pp3c21_550	P	P-P	c				At4g21190 (EMB1417)	chloroplast localized, required for embryo development	[16]
PpPPR_11	Pp3c3_2440	P	P-P-P-P-P	other	M	Smaller than WT	stabilization of <i>nad7</i> mRNA			Unpublished
PpPPR_12	Pp3c7_22430	P	P-Pi-P-P-P-P-P-P-P-P	c						
PpPPR_13	Pp3c7_17210	P	P-P-P-P-P-P-P-P-P-P-P-P-P	c						

Table 1. Cont.

Protein Name	Gene Locus ID ¹	Type ²	PPR Motif Bead Patterns and Additional Domain/Motif in Parenthesis ³	Subcellular Localization ⁴		Phenotype of Moss Colony of Gene Knockout Lines	Function Identified in <i>P. patens</i>	Arabidopsis Homologs	Function Identified in Arabidopsis	Refs.
				Pred	Exp					
PpPPR_24	Pp3c2_3210	P	P-P-P-P-P-P-P-P-P	m						
PpPPR_25	Pp3c16_5160	PLS	P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1	m	M					
PpPPR_26	Pp3c10_24680	P	P-P-P-P-P-P-P-P	m						
PpPPR_27	Pp3c20_15110	P	P-P-P-P-P-P-P-P-P	other	C	WT-like		At3g53170	chloroplast nucleoid	[47]
PpPPR_28	Pp3c6_24160	P	P-P-P-P-P-P-P-P-P-P	c	C	Smaller than WT				
PpPPR_29	Pp3c5_2770	P	P-P-P-P-P-P-P-P-P	m						
PpPPR_30	Pp3c7_10060	P	P-P-P-P-P-P-P-(SMR)	m				At1g18900	mitochondrial localized, function unknown	[16]
PpPPR_31	Pp3c6_23550	PLS	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1	m	M	Smaller than WT	RNA splicing of <i>atp9</i> intron 1 and <i>nad5</i> intron 3			[41]
PpPPR_32	Pp3c8_15500	P	P-P-P-P-P-P-P-P-P-P-P	c	C	Smaller than WT	<i>psaC</i> mRNA accumulation			[48]
PpPPR_33	Pp3c1_37670	P	P-P-P-P-P-P-Pi-P-P	m						
PpPPR_34	Pp3c22_1710	PLS	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1	other	C					
PpPPR_35	Pp3c8_13830	P	P-P-P-P-P-P-P-P-P	other						
PpPPR_36	Pp3c4_14490	P	P-P-P-P-P	m						
PpPPR_37	Pp3c4_14140	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	m	M	Bigger than WT				

Table 1. Cont.

Protein Name	Gene Locus ID ¹	Type ²	PPR Motif Bead Patterns and Additional Domain/Motif in Parenthesis ³	Subcellular Localization ⁴		Phenotype of Moss Colony of Gene Knockout Lines	Function Identified in <i>P. patens</i>	Arabidopsis Homologs	Function Identified in Arabidopsis	Refs.
				Pred	Exp					
PpPPR_38	Pp3c6_26920	P	P-P-P-P-P-P-P-P-P-P	c	C	Very small colony	RNA processing of <i>clpP-5'-rps12</i> mRNA			[49,50]
PpPPR_39	Pp3c4_3090	P	P-P-P-P-P-Pi-P-P	c				At3g42630 (PBF2)	RNA splicing of <i>ycf3</i>	[51]
PpPPR_40	Pp3c23_13490	P	P-P-P-P-P-P-P-P-P	c	C			At5g67570 (DG1)	involved in the regulation of early chloroplast development	[52]
PpPPR_41	Pp3c13_18720	P	P-Pi-P-P-P-P	other	C					
PpPPR_42	Pp3c21_12360	P	P-P-P-P-P-P-P-P-P-P-(SMR)	m				At5g02830	chloroplast nucleoid	[47]
PpPPR_43	Pp3c3_24770	DYW	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	m	M	Very small colony	RNA splicing of <i>cox1</i> intron 3			[53]
PpPPR_44	Pp3c17_24090	P	P-P-P-P-P-P-Pi-P-P-P	m						
PpPPR_45	Pp3c11_7720	DYW	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	c	C		RNA editing at <i>rps14-C2</i>			[54]
PpPPR_46	Pp3c3_5760	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	c				At5g39980 (PDM3)	essential for chloroplast development	[55]
PpPPR_47	Pp3c4_17980	P	P-Pi-P-P-P-P-P-P-P-P	other						

Table 1. Cont.

Protein Name	Gene Locus ID ¹	Type ²	PPR Motif Bead Patterns and Additional Domain/Motif in Parenthesis ³	Subcellular Localization ⁴		Phenotype of Moss Colony of Gene Knockout Lines	Function Identified in <i>P. patens</i>	Arabidopsis Homologs	Function Identified in Arabidopsis	Refs.
				Pred	Exp					
PpPPR_48	Pp3c24_4430	P	P-P-P-P-P-P-P-P-P-P-Pi-P-P	c	M	WT-like		At1g10910 (PDM2)	involved in the regulation of chloroplast development	[56]
PpPPR_49	Pp3c1_6490	P	P-P-P-P-P-P-P-P-Pi-P-P-P-P-P-P-P-P-P-P	m	M	WT-like				
PpPPR_50	Pp3c11_2930	P	P-P-P-Pi-P-Pi-Pi-P-P-P-P-P-P-P-P-P	m						
PpPPR_51	Pp3c12_14320	P	P-P-P-P-P-P-P-P-P	c	C			At4g34830 (MRL1)	stabilizes <i>rbcL</i> 5' end	[44]
PpPPR_52	Pp3c25_7340	P	P-P-P-P-P	c	C					
PpPPR_53	Pp3c13_17120	P	P-P-P-P-P-P-P-P-P	c	C	WT-like		At5g02860	chloroplast nucleoid	[47]
PpPPR_54	Pp3c12_26140	P	P-P	m	C/M	Bigger than WT				
PpPPR_55	Pp3c6_14920	P	P-P-P	c				At3g46870 (THA8-like)	<i>ycf3</i> -intron 2 and <i>trnA</i> splicing	[57]
PpPPR_56	Pp3c19_930	DYW	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	m	M	WT-like	RNA editing at <i>nad3</i> -C230, <i>nad4</i> -C272			[58]
PpPPR_57	Pp3c12_4690	P	P-P-P-P-P-P-P-P	m						
PpPPR_58	Pp3c1_21850	P	P-P-P-P-Pi-Pi-P-P-P-P	m				At4g35850	present in mitochondrial complexome	[59]
PpPPR_59	Pp3c22_3070	P	P-P-P-P-P-P-P-P-P-P-P (SMR)	c	C	WT-like		At5g02830	chloroplast nucleoid	[47]
PpPPR_60	Pp3c16_9420	P	P-P-P-P-P-P-P-P	m						

Table 1. Cont.

Protein Name	Gene Locus ID ¹	Type ²	PPR Motif Bead Patterns and Additional Domain/Motif in Parenthesis ³	Subcellular Localization ⁴		Phenotype of Moss Colony of Gene Knockout Lines	Function Identified in <i>P. patens</i>	Arabidopsis Homologs	Function Identified in Arabidopsis	Refs.
				Pred	Exp					
PpPPR_61	Pp3c14_7210	P	P-P-P-P-Pi-Pi-P-P-P-P	m				At4g35850	present in mitochondrial complexome	[59]
PpPPR_62	Pp3c16_9880	P	P-P-P-P-(SMR)	m				At2g17033	chloroplast localized, function unknown	[16]
PpPPR_63	Pp3c7_17100	P	P-Pi-P-(NYN)	other	NUC	Very small colony	5'-end processing of pre-tRNA	At2g16650 (PRORP2)-At4g21900 (PRORP3)	5'-end processing of pre-tRNA	[60,61]
PpPPR_64	Pp3c11_11830	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-Pi-P-(SMR)	other	C	Smaller than WT	Expression of <i>psaA-psaB-rps14</i>	At1g74850 (pTAC2)	involved in transcription by PEP	[62,63]
PpPPR_65	Pp3c4_16600	DYW	S2-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	m	M	Very small colony	RNA editing at <i>ccmFc-C103</i>			[64,65]
PpPPR_66	Pp3c16_5890	P	P-P-P-P-P-P-P-P-P-P-P	c	C	WT-like	RNA splicing of <i>ndhA</i>	At2g35130 (AtPPR66L)	RNA splicing of <i>ndhA</i>	[66]
PpPPR_67	Pp3c2_30390	P	P-P-P-(NYN)	other	C/M	WT-like	5'-end processing of pre-tRNA	At2g32230 (RPORP1)	5'-end processing of pre-tRNA	[60,61]
PpPPR_68	Pp3c2_27580	P	P-P-P-P-P-P-P-P-P-P-P-P	m						
PpPPR_69	Pp3c17_5040	PLS	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2	sp	C	WT-like		At4g18520 (PDM1/SEL1)	RNA editing of <i>accD</i> , <i>trnK</i> -intron splicing, RNA processing of <i>rpoA</i>	[67]

Table 1. Cont.

Protein Name	Gene Locus ID ¹	Type ²	PPR Motif Bead Patterns and Additional Domain/Motif in Parenthesis ³	Subcellular Localization ⁴		Phenotype of Moss Colony of Gene Knockout Lines	Function Identified in <i>P. patens</i>	Arabidopsis Homologs	Function Identified in Arabidopsis	Refs.
				Pred	Exp					
PpPPR_70	Pp3c8_4280	P	P-P-P-P-P-P-P-P-(CBS)	c	C/NUC			At5g10690 (CBSPPR1)	chloroplast nucleoid, possibly play a role in regulating transcription or replication	[47,68]
PpPPR_71	Pp3c14_16110	DYW	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	m	M	Very small colony	RNA editing at <i>ccmFc-C122</i>			[69]
PpPPR_72	Pp3c6_26210	P	P-P-P-P-P-P-P-P-P-P	c	C/M		function unknown	At2g35130 (AtPPR66L)	RNA splicing of <i>ndhA</i>	[66]
PpPPR_73	Pp3c3_31860	P	P-P-P-P-P-Pi-Pi-P	m	C/M					
PpPPR_74	Pp3c17_13520	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	c	C	Smaller than WT				
PpPPR_75	Pp3c5_16950	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-(SMR)	c	C			At2g31400 (GUN1)	involved in retrograde signaling to the nucleus.	[70]
PpPPR_76	Pp3c16_280	P	(RRM)-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	other	C			At5g04810 (AtPPR4)	<i>rps12</i> -intron 1 <i>trans</i> -splicing	[37,38]
PpPPR_77	Pp3c5_15090	DYW	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	other	M	Very small colony	RNA editing at <i>cox2-C370</i> , <i>cox3-C733</i>			[71]
PpPPR_78	Pp3c2_12230	DYW	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	other	M	WT-like	RNA editing at <i>rps14-C137</i> , <i>cox1-C755</i>			[71,72]

Table 1. Cont.

Protein Name	Gene Locus ID ¹	Type ²	PPR Motif Bead Patterns and Additional Domain/Motif in Parenthesis ³	Subcellular Localization ⁴		Phenotype of Moss Colony of Gene Knockout Lines	Function Identified in <i>P. patens</i>	Arabidopsis Homologs	Function Identified in Arabidopsis	Refs.
				Pred	Exp					
PpPPR_79	Pp3c5_7610	DYW	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	m	M	Smaller than WT	RNA editing at <i>nad5</i> -C598			[58]
PpPPR_80	Pp3c4_9690	P	P-P-P-P-P-P-P-P-P-P	c	C			At4g39620 (AtPPR5)	<i>trnG</i> splicing and intron stabilization	[43]
PpPPR_81	Pp3c14_15490	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-SMR	other	C	Bigger than WT				
PpPPR_82	Pp3c9_6880	P	P-P-P-P-P-P-P-P-P-P	other	C					
PpPPR_83	Pp3c2_1940	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-Pi-P-P-P	other				At2g41720 (EMB2654)	<i>rps12</i> -intron 1 trans-splicing	[73,74]
PpPPR_84	Pp3c15_12010	P	P-P-P-P-P-P-P-P-P-P	c						
PpPPR_85	Pp3c6_15900	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-SMR	c	C	WT-like		At2g31400 (GUN1)	involved in retrograde signaling to the nucleus	[70]
PpPPR_86	Pp3c17_2130	P	P-P-Pi-P-Pi-P-P-P-P-P-P-P-P-P-P-P-P-P-Pi-P	c						
PpPPR_87	Pp3c8_15040	P	P-P-P-P-P-P-P-P-P-P-P-P-P	m						
PpPPR_88	Pp3c18_8600	P	P-P-P-P-P-P-P-P-P-P	m						
PpPPR_89	Pp3c1_28760	P	P-P-P-P-P-P-P	m	M					
PpPPR_90	Pp3c11_21930	15P	P-P-Pi-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	other				At5g42310 (AtCRP1)	stabilizes 5' and 3' ends in <i>petB-petD</i> intergenic region; activates <i>petA</i> , <i>psaC</i> , and <i>petD</i> translation	[17,18,75,76]

Table 1. Cont.

Protein Name	Gene Locus ID ¹	Type ²	PPR Motif Bead Patterns and Additional Domain/Motif in Parenthesis ³	Subcellular Localization ⁴		Phenotype of Moss Colony of Gene Knockout Lines	Function Identified in <i>P. patens</i>	Arabidopsis Homologs	Function Identified in Arabidopsis	Refs.
				Pred	Exp					
PpPPR_91	Pp3c17_23250	DYW	P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-P1-L1-S1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	m	M	Very small colony	RNA editing at <i>nad5-C730</i>			[58]
PpPPR_92	Pp3c5_2530	P	P-P-P-P-P-P-P-Pi-P-P-P-P-P-P-P-P-P-P-P-P	c	C	Smaller than WT		At4g30825 (BFA2)	accumulation of the <i>atpH/F</i> transcript	[77]
PpPPR_93	Pp3c6_3910	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	other						
PpPPR_94	Pp3c16_4140	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	c	C			At4g30825 (BFA2)	accumulation of the <i>atpH/F</i> transcript	[77]
PpPPR_95	Pp3c5_26320	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	c						
PpPPR_96	Pp3c4_4900	P	P-P-P-P-P-P-P-P-P-P-Pi-P-(SMR)	c	C					
PpPPR_97	Pp3c3_19780	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	m	M					
PpPPR_98	Pp3c27_5540	DYW	P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-L2-S2-E1-E2-DYW	m	M	WT-like	RNA editing at <i>atp9-C92</i>			[64,65]
PpPPR_99	Pp3c12_2390	P	P-P-P-P-P-P-P-P-P-P-P-P-P-P-P-P	c	C			At3g09650 (HCF152)	stabilizes 5' and 3' ends in <i>psbH-petB</i> intergenic region, also stimulates <i>petB</i> splicing	[20,78]
PpPPR_100	Pp3c12_8330	P	P-P-P-P-Pi-Pi-P-P-P	c				At2g30100	chloroplast localized, function unknown	[16]

Table 1. Cont.

Protein Name	Gene Locus ID ¹	Type ²	PPR Motif Bead Patterns and Additional Domain/Motif in Parenthesis ³	Subcellular Localization ⁴		Phenotype of Moss Colony of Gene Knockout Lines	Function Identified in <i>P. patens</i>	Arabidopsis Homologs	Function Identified in Arabidopsis	Refs.
				Pred	Exp					
PpPPR_101	Pp3c2_36070	P	P1-P1-P2-L2-P1-P1-P2-P1-P1-P1-P1-L2-P1-P1	m						
PpPPR_102	Pp3c6_11500	P	P-P-P-P	c	C	WT-like				
PpPPR_103	Pp3c17_4890	P	P1-L1-P2	other						
PpPPR_104	Pp3c10_16850	P	P-P-Pi-P-(NYN)	m	C/M	Smaller than WT	5'-end processing of pre-tRNA	At2g32230 (RPORP1)	5'-end processing of pre-tRNA	[60,61]
PpPPR_105	Pp3c24_8560	PLS	L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P1-L1-S1-P2-S1-P1-L1-S2	m	C	WT-like				
PpPPR_106	Pp3c22_3080	P	P-P-(SMR)	sp						
PpPPR_107	Pp3c1_8170	P	P-P-P-P-P-(SAP)-P	m	C			At3g04260 (pTAC3)	light-dependent transcription	[79]

¹ Gene locus ID is from *P. patens* genome release v3.3 (<https://phytozome-next.jgi.doe.gov> (accessed on 1 July 2022)). ² PPR types are according to Lurin et al. [14] and Cheng et al. [15]; “P” for the P type, “PLS” for the PLS type containing P1, P2, L1, L2, S1, S2 and “DYW” for the DYW type containing E1, E2 and DYW domain. ³ Motif bead patterns are according to Cheng et al. [15] (<https://ppr.plantenergy.uwa.edu.au> (accessed on 1 July 2022)). Additional domains or motifs are RRM (RNA recognition motif), LAGLIDADG (LAGLIDADG RNA motif), SMR (small Mut-related), NYN (Nedd4-BP1, YacP nucleases domain), CBS (cystathionine β-synthase domain) and SAP (SAF-A/B, Acinus and PIAS domain). ⁴ Subcellular localization. Predicted localizations (Pred) were provided by the TargetP server v1.1 (<http://www.cbs.dtu.dk/services/TargetP-1.1/index.php> (accessed on 1 July 2022)). The prediction is noted in lowercases as follows: m, mitochondria; c, chloroplasts; sp, signal peptide; other, any other location. Experimental localizations (Exp) of fluorescent proteins were from published studies and unpublished studies of our laboratory. The conclusion is indicated in uppercases as follows: M, mitochondria; C, chloroplasts; C/M, chloroplasts and mitochondria; NUC, nucleus; C/NUC, chloroplasts and nucleus.

Table 2. Paralogous pairs of Physcomitrium PPR proteins.

Protein Name	Type	Additional Domain	Amino Acid Length (aa)	Amino Acid Identity (%)	Arabidopsis Homolog																																																																																																																																						
PpPPR_3	P	RRM	939	62	At5g04810 (AtPPR4)																																																																																																																																						
PpPPR_76			947			PpPPR_5	P		611	79.2		PpPPR_88	611	PpPPR_6	P		451	62.5		PpPPR_102	395	PpPPR_7	P	LAGLIDADG	1195	71.5	At2g15820 (OTP51)	PpPPR_22	883	PpPPR_13	P		657	75.4		PpPPR_101	670	PpPPR_16	P		637	70.8		PpPPR_89	630	PpPPR_17	P		489	58.5	At4g39620 (AtPPR5)	PpPPR_80	482	PpPPR_19	P		958	65.8	At4g34830 (MRL1)	PpPPR_51	982	PpPPR_26	P		1110	71.2		PpPPR_40	961	PpPPR_27	P		487	68.2	At3g53170	PpPPR_35	532	PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104	993	PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010
PpPPR_5	P		611	79.2																																																																																																																																							
PpPPR_88			611			PpPPR_6	P		451	62.5		PpPPR_102	395	PpPPR_7	P	LAGLIDADG	1195	71.5	At2g15820 (OTP51)	PpPPR_22	883	PpPPR_13	P		657	75.4		PpPPR_101	670	PpPPR_16	P		637	70.8		PpPPR_89	630	PpPPR_17	P		489	58.5	At4g39620 (AtPPR5)	PpPPR_80	482	PpPPR_19	P		958	65.8	At4g34830 (MRL1)	PpPPR_51	982	PpPPR_26	P		1110	71.2		PpPPR_40	961	PpPPR_27	P		487	68.2	At3g53170	PpPPR_35	532	PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000
PpPPR_6	P		451	62.5																																																																																																																																							
PpPPR_102			395			PpPPR_7	P	LAGLIDADG	1195	71.5	At2g15820 (OTP51)	PpPPR_22	883	PpPPR_13	P		657	75.4		PpPPR_101	670	PpPPR_16	P		637	70.8		PpPPR_89	630	PpPPR_17	P		489	58.5	At4g39620 (AtPPR5)	PpPPR_80	482	PpPPR_19	P		958	65.8	At4g34830 (MRL1)	PpPPR_51	982	PpPPR_26	P		1110	71.2		PpPPR_40	961	PpPPR_27	P		487	68.2	At3g53170	PpPPR_35	532	PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000								
PpPPR_7	P	LAGLIDADG	1195	71.5	At2g15820 (OTP51)																																																																																																																																						
PpPPR_22			883			PpPPR_13	P		657	75.4		PpPPR_101	670	PpPPR_16	P		637	70.8		PpPPR_89	630	PpPPR_17	P		489	58.5	At4g39620 (AtPPR5)	PpPPR_80	482	PpPPR_19	P		958	65.8	At4g34830 (MRL1)	PpPPR_51	982	PpPPR_26	P		1110	71.2		PpPPR_40	961	PpPPR_27	P		487	68.2	At3g53170	PpPPR_35	532	PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																
PpPPR_13	P		657	75.4																																																																																																																																							
PpPPR_101			670			PpPPR_16	P		637	70.8		PpPPR_89	630	PpPPR_17	P		489	58.5	At4g39620 (AtPPR5)	PpPPR_80	482	PpPPR_19	P		958	65.8	At4g34830 (MRL1)	PpPPR_51	982	PpPPR_26	P		1110	71.2		PpPPR_40	961	PpPPR_27	P		487	68.2	At3g53170	PpPPR_35	532	PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																								
PpPPR_16	P		637	70.8																																																																																																																																							
PpPPR_89			630			PpPPR_17	P		489	58.5	At4g39620 (AtPPR5)	PpPPR_80	482	PpPPR_19	P		958	65.8	At4g34830 (MRL1)	PpPPR_51	982	PpPPR_26	P		1110	71.2		PpPPR_40	961	PpPPR_27	P		487	68.2	At3g53170	PpPPR_35	532	PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																
PpPPR_17	P		489	58.5	At4g39620 (AtPPR5)																																																																																																																																						
PpPPR_80			482			PpPPR_19	P		958	65.8	At4g34830 (MRL1)	PpPPR_51	982	PpPPR_26	P		1110	71.2		PpPPR_40	961	PpPPR_27	P		487	68.2	At3g53170	PpPPR_35	532	PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																								
PpPPR_19	P		958	65.8	At4g34830 (MRL1)																																																																																																																																						
PpPPR_51			982			PpPPR_26	P		1110	71.2		PpPPR_40	961	PpPPR_27	P		487	68.2	At3g53170	PpPPR_35	532	PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																
PpPPR_26	P		1110	71.2																																																																																																																																							
PpPPR_40			961			PpPPR_27	P		487	68.2	At3g53170	PpPPR_35	532	PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																								
PpPPR_27	P		487	68.2	At3g53170																																																																																																																																						
PpPPR_35			532			PpPPR_39	P		582	58.3	At3g42630 (PBF2)	PpPPR_73	603	PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																																
PpPPR_39	P		582	58.3	At3g42630 (PBF2)																																																																																																																																						
PpPPR_73			603			PpPPR_42	P	SMR	936	56.1	At5g02830	PpPPR_59	935	PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																																								
PpPPR_42	P	SMR	936	56.1	At5g02830																																																																																																																																						
PpPPR_59			935			PpPPR_58	P		530	74.2	At4g35850	PpPPR_61	522	PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																																																
PpPPR_58	P		530	74.2	At4g35850																																																																																																																																						
PpPPR_61			522			PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)	PpPPR_72	578	PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																																																								
PpPPR_66	P		578	77.3	At2g35130 (AtPPR66L)																																																																																																																																						
PpPPR_72			578			PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)	PpPPR_67	791	PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																																																																
PpPPR_63	P	NYN	655	60.7–79.1	At2g32230 (PRORP1), At2g16650 (PRORP2), At4g21900 (PRORP3)																																																																																																																																						
PpPPR_67			791																																																																																																																																								
PpPPR_104			993			PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)	PpPPR_85	871	PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																																																																														
PpPPR_75	P	SMR	871	91.5	At2g31400 (GUN1)																																																																																																																																						
PpPPR_85			871			PpPPR_82	P		717	76.8		PpPPR_84	717	PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																																																																																						
PpPPR_82	P		717	76.8																																																																																																																																							
PpPPR_84			717			PpPPR_92	P		1010	73.7	At4g30825 (BFA2)	PpPPR_94	1000																																																																																																																														
PpPPR_92	P		1010	73.7	At4g30825 (BFA2)																																																																																																																																						
PpPPR_94			1000																																																																																																																																								

4. Subcellular Localization of PPR Proteins in Physcomitrium

In silico and in vivo analyses have shown that 97 out of 107 PpPPR proteins are presumably localized in either the plastids or mitochondria, or both (Table 1). An in vivo analysis using a transient assay or transgenic Physcomitrium plants expressing the PpPPR-green fluorescent protein (GFP) fusion protein demonstrated the subcellular localization of 60 PpPPR proteins, wherein 38 of which are chloroplast-targeted, 17 are localized in mitochondria, and four are targeted to both. PpPPR_63 is localized in the nucleus and

its paralogs (PpPPR₆₇ and 104) are localized in both the chloroplasts and mitochondria [60]. Thirteen PpPPR proteins are predicted to be localized in the plastids and 24 in the mitochondria, while the location of the remaining 10 has not yet been predicted.

5. P-Type of PPR Proteins in Physcomitrium

More than half (55%) of the Arabidopsis PPR genes encode P-type proteins while most (85%) of the Physcomitrium PPR proteins are of the P type. About 40% of Physcomitrium P-type proteins show high aa identities (30–50%) with Arabidopsis PPR proteins, such as EMBRYO DEFECTIVE (EMB)1006 [14], AtPPR4 [37,38], AtPPR5 [43], Maturation/stability of *RbcL* 1 (MRL1) [44], THYLAKOID ASSEMBLY 8 (THA8)-like [57], GENOME UNCOUPLED 1 (GUN1) [70], plastid Transcriptionally Active Chromosome 2 (pTAC2) [62], ORGANELLE TRANSCRIPT PROCESSING 51 (OTP51) [40], and Proteinaceous RNase P (PRORP) 1, 2 and 3 [61] and others [42,46,47,51,52,55,56,59,67,68,73–77,79] (Table 1). The remaining 60% of the P-type PpPPR proteins may be unique to this moss but their functions are totally unknown.

The P-type PPR proteins are mostly involved in intergenic RNA processing, RNA splicing, RNA stabilization, as well as translation initiation [8–10,43,80]. In all these cases, PPR proteins bind in a gene-specific manner to target RNAs. Bioinformatic analyses of PPR proteins and target RNA sequences have proposed an RNA recognition code for PPR proteins involving a combination of amino acid residues at two or three positions in a PPR motif that determines the base preference [81–83]. The crystal structures of THA8 and PPR10 in an RNA-bound state were determined [84,85]. The molecular basis for the specific and modular recognition of RNA bases, A, G and U was revealed. The structural elucidation of RNA recognition by PPR proteins provides an important framework for the potential biotechnological applications of PPR proteins in RNA-related research areas [86–88].

5.1. P-Type Proteins Involved in RNA Processing/Stabilization

5.1.1. PpPPR₃₈

In Physcomitrium, the first functional analysis of a PPR protein was achieved for PpPPR₃₈ which has 10 P motifs [49]. PpPPR₃₈ is localized in the chloroplasts and is responsible for the intergenic RNA cleavage between *clpP* and 5'-*rps12*, encoding the ATP-dependent protease proteolytic subunit and the ribosomal protein S12, respectively. The *clpP* gene, which contains two introns, is co-transcribed with the 5'-*rps12* and *rpl20* genes, and a primary transcript of 3.2 kb is produced. The primary transcript is then cleaved at the intergenic spacer between *clpP* and 5'-*rps12* and is spliced to produce a 0.6-kb mature *clpP* mRNA. This post-transcriptional event proceeds rapidly and therefore the primary transcript accumulates in small amounts. In the PpPPR₃₈ knockout (KO) mutants, the primary transcript accumulates at a substantial level while mature *clpP* mRNA is significantly decreased [49]. PpPPR₃₈ binds specifically to the intergenic region of *clpP*-5'-*rps12* dicistronic mRNA and may stabilize a processed *clpP* 3'-end [50]. The binding of the PpPPR₃₈ protein could recruit endonucleases and thus may lead to RNA cleavage. Interestingly, the processed *rps12* 5'-ends in moss and Arabidopsis were mapped at analogous positions and the sequence corresponding to the overlap between the processed *clpP* and *rps12* mRNAs shows a striking conservation between angiosperms and moss [78]. These findings suggest that an unidentified conserved protein, possibly a PpPPR₃₈ ortholog, binds to the *clpP*-*rps12* intergenic region and blocks the RNA degradation in angiosperms.

5.1.2. PpPPR₂₁

PpPPR₂₁ is comprised of 19 P motifs and is chloroplast-localized. The KO mutants of PpPPR₂₁ grow slowly and exhibit a significant reduction of the photosystem II (PSII) core protein D1 (PsbA) level, and a concomitantly poor level of the PSII supercomplexes [45]. Similar phenotypic features were reported in the green alga, *Chlamydomonas reinhardtii* (*Chlamydomonas*) and tobacco, *psbI* gene KO mutants [89,90]. Tobacco, *psbI*, KO mutants are autotrophically viable, but PsbA and PsbO levels are reduced to 50% of the level of

the wild type (WT) while the PSII complexes are poorly formed, suggesting that *PsbI* is essential for the stability of dimeric PSII and supercomplexes [45]. Notably, the *psbI-ycf12* dicistronic mRNA (encoding PSII I-subunit and YCF12) is lost in the *PpPPR_21* KO mutants while other PSII gene transcript levels are not altered in these mutants [45]. An RNA electrophoresis mobility shift assay (REMSA) showed that the recombinant PpPPR_21 binds efficiently to the 5' untranslated region (UTR) of *psbI* mRNA [45]. These observations suggest that PpPPR_21 is involved in the accumulation of a *psbI-ycf12* mRNA.

PpPPR_21 homologs, which are referred to as PPR21L, are widely distributed in bryophytes, ferns and seed plants, but not in *Chlamydomonas*. PpPPR_21 shows 45% aa identity and 85% similarity to Arabidopsis At5g02860 (*AtPPR21L*), which is predicted to be localized in the chloroplasts [16]. A predicted target sequence of Arabidopsis At5g02860 is found in the 5'-UTR of the Arabidopsis *psbI* transcript. Unfortunately, no homozygous mutants of T-DNA-tagged *At5g02860* lines have been obtained, so the loss of function of *AtPPR21L* may lead to embryonic lethality. To confirm whether *AtPPR21L* is a functional ortholog of PpPPR_21, other *AtPPR21L* mutants will be generated and characterized, for example, by generating knockdown mutants.

5.1.3. PpPPR_32

PpPPR_32 consists of 13 P motifs and is localized in the chloroplasts [48]. *PpPPR_32* homologs are not found in seed plants and are unique to early land plants. *PpPPR_32* KO mutants grow autotrophically but they have reduced protonema growth and the poor formation of photosystem I (PSI) complexes. In addition, a significant reduction in the transcript level of the *psaC* gene encoding the iron sulfur protein of PSI was observed in the KO mutants, but the transcript levels of other PSI genes were not altered. This indicates that PpPPR_32 is essential for the accumulation of *psaC* transcript and PSI complexes. In land plants, *psaC* is organized as a gene cluster with *ndhH*, *A*, *I*, *G*, *E* and *D* genes and is located between *ndhE* and *D*. These genes are co-transcribed as a long primary transcript (>7 kb) and then they are post-transcriptionally processed to multiple and overlapping intermediates and mature transcripts in seed plants [91,92]. The monocistronic *psaC* transcript is produced from a *psaC-ndhD* dicistronic mRNA by the cleavage of its intergenic region, and then it represents the translatable mRNA [91,93]. Based on the stoichiometry of accumulating PSI and NDH complexes [94], the expression of *psaC* is two orders higher than that of *ndhD* in higher plants. Likewise, the monocistronic *psaC* mRNA is extremely stable, more so than *ndh* transcripts in *Physcomitrium* chloroplasts.

PsaC is an essential component, both for photochemical activity and for the stable assembly of PSI in *Chlamydomonas* [95] and higher plants [96]. The loss of PsaC leads to defects in PSI activity and the rapid degradation of other subunits within the core complex. A reduction in the *psaC* transcript level may cause a reduction in the PSI subunits in *PpPPR_32* KO mutants. An in silico analysis predicted that PpPPR_32 binds to one of two stem-loop structures in the 3'-UTR of *psaC* mRNA [48]. The binding of PpPPR_32 to the stem-loop may facilitate the stabilization of *psaC* mRNA. The binding of PpPPR_32 could prevent endonucleolytic cleavage by blocking the access of RNases, thus stabilizing the RNA. This binding might occur in the vicinity of secondary structure elements, such as hairpins, that are known targets for endonucleases.

5.2. P-Type Proteins Involved in RNA Splicing

Several P-type PPR proteins that are involved in the splicing of plastid group II introns have been identified. Maize PPR4, with an N-terminal RNA recognition motif (RRM), facilitates *rps12* trans-splicing through direct interaction with intron RNA [37]. Maize PPR5 is involved in the splicing or the stability of pre-tRNA^{Gly} [43]. Arabidopsis OTP51 with two C-terminal LAGLIDADG motifs is required for the cis-splicing of the *ycf3-2* intron [40]. Maize and Arabidopsis THA8 is essential for the splicing of both *ycf3-2* and tRNA^{Ala} introns [57]. PPR4, PPR5, OTP51 and THA8 homologs were found in *Physcomitrium*

(Table 1) but their function has yet to be elucidated. To date, PpPPR_4 and PpPPR_66 were identified as P-type proteins that are required for RNA splicing in plastid transcripts [39,66].

5.2.1. PpPPR_4

The ten P motif-containing PpPPR_4, which is not related to maize PPR4, is required for the group II intron splicing of pre-tRNA^{Ile}-GAU [39]. The tRNA^{Ile} gene is co-transcribed as a long primary transcript from a *rrn16-trnI(GAU)-trnA(UGC)-rrn23-rrn4.5* gene cluster, then an intron-containing pre-tRNA^{Ile} is produced and spliced to produce a mature tRNA^{Ile}. PpPPR_4 KO mutants display severe growth retardation, a reduced effective quantum yield of PSII and a strongly reduced level of plastid-encoded proteins, such as PSII D1 (PsbA) protein, the β subunit of ATP synthase, and the stromal enzyme, Rubisco. Analyses of the chloroplast transcriptome revealed that the disruption of PpPPR_4 resulted in a significant reduction of the mature tRNA^{Ile} (GAU) and the aberrant accumulation of intron-containing pre-tRNA^{Ile}. The recombinant PpPPR_4 was shown to bind preferentially to domain III of the tRNA^{Ile} group II intron [39]. Thus, the PpPPR_4 gene KO specifically impaired the splicing of tRNA^{Ile} (GAU), but it did not affect tRNA^{Ala} (UGC) splicing. This is distinct from maize and Arabidopsis *tha8* mutants, where tRNA^{Ala} splicing was strongly inhibited, whereas tRNA^{Ile} splicing was not affected [57].

In angiosperms, the splicing of pre-tRNA^{Ile} is known to require at least four nuclear-encoded factors, ribonuclease III domain-containing protein RNC1 [97], WTF1 (a protein harboring a “domain of unknown function 860”, assigned the name WTF1 “what’s this factor?”) [98], DEAD box RNA helicase RH3 [99], mitochondrial transcription termination factor 4 (mTERF4) [100] and the chloroplast *trnK* intron-encoded maturase (MatK) [101]. These are thought to be major splicing factors and constitute a spliceosome-like complex that is for splicing chloroplast group II introns [102]. The Physcomitrium genome encodes an RNC1 homolog, a WTF1 homolog, four RH3 homologs and three mTERF4 homologs. It will be important in the future to examine whether these homologs are involved in the splicing of tRNA^{Ile} (GAU) in Physcomitrium. Intriguingly, a PpPPR_4 homolog was found in a lycophyte, *Selaginella moellendorffii*, but not in Arabidopsis and rice. Presumably, the Arabidopsis PPR protein(s) involved in the splicing of pre-tRNA^{Ile} (GAU) may have lost their importance during evolution, whereas other proteins may have replaced PpPPR_4.

5.2.2. PpPPR_66

PpPPR_66 is a chloroplast-localized protein with 11 P motifs. PpPPR_66 KO mutants exhibited a WT-like growth phenotype, but they showed aberrant chlorophyll fluorescence due to defects in the chloroplast NADH dehydrogenase-like (NDH) activity [66]. In addition, the chloroplast NDH complex was completely lost in the KO mutants. Among the 11 *ndh* genes in the chloroplast genome, *ndhA* expression was specifically affected in the KO mutants. The chloroplast *ndhA* gene contains a group II intron and its intron splicing was almost completely blocked in the PpPPR_66 KO mutants. PpPPR_66 binds to the 5' half of domain I of the *ndhA* group II intron [66].

PpPPR_66 shows 44% aa identity and 81% similarity to the Arabidopsis PPR protein (At2g35130). The At2g35130 gene is interrupted by seven introns and their positions are identical to the intron positions of the PpPPR_66 gene, suggesting that At2g35130 is an ortholog of PpPPR_66. In fact, among the Arabidopsis At2g35130 KO null mutant lines, SALK_043507 and SALK_065137 exhibited defects in chloroplast NDH activity and *ndhA* splicing [66]. PpPPR_66-like homologs (PPR66L) are widely distributed in land plants but not in green algae, *Chlamydomonas*, *Volvox* or *Chlorella*. Streptophyta (charophytes and land plants) have *ndhA* genes with a group II intron, while green algae have an intron-less *ndhA* gene in the chloroplast genome. Thus, there is likely coevolution of PPR66L and the *ndhA* intron in the land plant lineage. Although Arabidopsis At2g35130 (AtPPR66L) was involved in *ndhA* splicing, AtPPR66L cDNA did not rescue the *ndhA* splicing deficiency in the PpPPR_66 KO mutant [66]. Thus, PpPPR_66 and AtPPR66L are required for *ndhA*

splicing but their mode of action for splicing the *ndhA* transcript might differ slightly between Physcomitrium and Arabidopsis. This possibility still needs to be assessed.

In seed plants, the splicing of the *ndhA* transcript is known to require several nuclear-encoded factors, including CHLOROPLAST RNA SPLICING 2 (CRS2) [103], CRS2-ASSOCIATED FACTOR 1 and 2 (CAF1 and CAF2) [104] and CHLOROPLAST RNA SPLICING AND RIBOSOME MATURATION (CRM) FAMILY MEMBER 2 (CFM2) [105]. These splicing-related factors have not been identified in Physcomitrium.

5.3. P-Type Proteins with an NYN Domain

RNase P is a ubiquitous endonuclease that removes the 5' leader sequence from pre-tRNAs in all organisms [106]. In Arabidopsis, RNA-free proteinaceous RNase P1 (PRORP1), PRORP2 and PRORP3 were shown to be enzyme(s) for pre-tRNA 5'-end processing in organelles and the nucleus [61,107]. PRORPs contain three P motifs and an NYN (Nedd4-binding protein 1 and the bacterial YacP-like metallo-nuclease nucleases) domain [108]. PRORP1 is localized in both the chloroplasts and the mitochondria, whereas PRORP2 and 3 are localized in the nucleus. In Physcomitrium, three PRORP-like proteins, PpPPR₆₃, 67 and 104, were identified [60]. PpPPR₆₃ was localized in the nucleus, while PpPPR₆₇ and 104 were found in both the chloroplasts and the mitochondria. The three proteins exhibited pre-tRNA 5'-end processing activity *in vitro* [60]. PpPPR₆₃ KO mutants display the growth retardation of protonema colonies, indicating that unlike Arabidopsis nuclear PRORP2 and 3, the moss nuclear PpPPR₆₃ is not essential for viability, but it is involved in the growth and development of protonema filaments. In the KO mutants, the level of nuclear-encoded tRNA^{Asp} (GUC) decreased slightly, whereas the levels of most nuclear-encoded tRNAs were not altered. This indicates that most of the cytosolic mature tRNAs are produced normally, without proteinaceous RNase P-like PpPPR₆₃. Single PpPPR₆₇ or 104 gene KO mutants display different phenotypes of protonema growth and plastid tRNA^{Arg} (ACG) accumulation. However, the levels of all other tRNAs are not altered in the KO mutants. In addition, *in vitro* RNase P assays showed that PpPPR₆₇ and PpPPR₁₀₄ efficiently cleaved plastid pre-tRNA^{Arg} (CCG) and pre-tRNA^{Arg} (UCU), but they cleaved pre-tRNA^{Arg} (ACG) with different efficiency [60]. This suggests that the two proteins have an overlapping function, whereas their substrate specificity is not identical.

5.4. P-Type Proteins with an SMR Domain

DNA-damage repair systems, including the mismatch-repair pathway, operate in prokaryotes and eukaryotes. The mismatch-recognizing protein, MutS, is central to the mismatch-repair pathway because it recognizes and binds to mismatched nucleotides. The *mutS* and *mutS2* genes exist in a variety of bacterial species and their eukaryotic homologues form a multigene family. Novel, small proteins containing the MutS2 C-terminus were also found in bacteria and fungi and they were named as small MutS-related (SMR) proteins [109]. SMR domain-containing PPR proteins, referred as PPR-SMR proteins, constitute a small family in higher plants, with eight PPR-SMR proteins in Arabidopsis and maize [110]. Arabidopsis PPR-SMR protein GUN1 is involved in plastid-to-nucleus retrograde signaling [70], and SUPPRESSOR OF VARIATION7 (SVR7) could be required for FtsH-mediated chloroplast biogenesis [111]. Like SVR7, Arabidopsis SUPPRESSOR OF THYLAKOID FORMATION 1 (SOT1) facilitates the correct processing of plastid 23S–4.5S rRNA precursors [112]. Maize PPR53, an ortholog of Arabidopsis SOT1, plays a role in the expression of the *ndhA* and 5' end processing of the 23S rRNA precursor [113]. The SMR domain of SOT1 conferred RNA endonucleolytic activity [114]. Thus, PPR-SMR proteins play essential roles in embryo development, chloroplast biogenesis and gene expression [110,115].

In Physcomitrium, 10 PPR-SMR proteins were identified, of which PpPPR₅₉, 64, 75, 81, 85 and 96 targeted the chloroplasts. PpPPR₇₅ and 85 were previously assigned as a GUN1 homolog, and PpPPR₆₄ was assigned as a pTAC2 homolog [116] (Table 1). The loss of *pTAC2* (*ptac2*) results in pale yellow-green primary leaves and a seedling-lethal

phenotype in Arabidopsis [62]. Plastid-encoded RNA polymerase (PEP) and its accessory polypeptides are either absent or strongly reduced in the *ptac2* mutant. Similar phenotypes were shown in the rice mutant, *OspTAC2* [117]. In contrast, *PpPPR_64* KO mutants grow autotrophically but more slowly than the WT do [63]. Even though the levels of PSI and PSII are considerably reduced in the KO mosses, unlike in Arabidopsis *ptac2*, most PEP-dependent plastid transcripts, including *psbA*, accumulate at similar levels in the WT and KO mosses. However, the transcript level of the *psaA-psaB-rps14* gene cluster was significantly reduced in the KO mutants. Thus, *PpPPR_64* is not likely a functional orthologue of pTAC2.

Single KO mutants of *PpPPR_59* and *85* exhibited no obvious external phenotypic differences in their protonema growth when compared to that of WT moss plants. *PpPPR_75* is highly conserved (91.5% aa identity) with *PpPPR_85* (Table 2), so it may complement with the loss-of-function of *PpPPR_85*. Interestingly, KO mutants of *PpPPR_81* have a larger protonema colony than WT do.

6. PLS-Type PPR Proteins in Physcomitrium

PLS-type PPR proteins represent a small group in the PLS-subfamily in Arabidopsis and rice [15]. The function of most PLS-type proteins remains elusive. Arabidopsis chloroplast-localized PLS-type protein Pigment-Deficient Mutant 1 (PDM1)/SEEDLING LETHAL1 (SEL1) is reported to influence the RNA editing of the *accD* transcript encoding the acetyl-CoA carboxylase β subunit [67,118] and the splicing of *trnK* transcript encoding tRNA^{Lys} and *ndhA* transcripts [67]. Rice PLS-type protein PALE GREEN LEAF12 (PGL12) is also a chloroplast-localized protein that was involved in the processing of 16S rRNA and the splicing of the *ndhA* transcript [119].

The Physcomitrium nuclear genome encodes six PLS-type proteins, *PpPPR_9*, 25, 31, 34, 69 and 105 (Table 1). *PpPPR_34*, 69 and 105 are targeted to chloroplasts while the others are mitochondria-localized. *PpPPR_69* seems to be a PDM1/SEL1 homolog, but the remaining five proteins are structurally unrelated to the Arabidopsis and rice PLS-type proteins. KO mutants of *PpPPR_69* or *105* exhibit no obvious external phenotypic differences when they are compared to WT. In contrast, the KO mutants of *PpPPR_9* and *31* encoding the mitochondrial localized PLS-type protein display slower protonema growth than the WT does [41]. The *PpPPR_31* KO mutants show a considerable reduction in the splicing of *nad5* intron 3 and *atp9* intron 1 (Figure 3). The *PpPPR_9* KO mutants display a severe reduction in the splicing of *cox1* intron 3 (Figure 3). Their intron splicing efficiency in the KO mutants was reduced to less than 50% relative to that in the WT. In Physcomitrium, the splicing of *cox1* intron 3 required DYW-type *PpPPR_43* [53]. The splicing of *cox1* intron 3 was completely abolished in the *PpPPR_43* KO mutants while it occurred—but at a lower efficiency—in the *PpPPR_9* KO mutants. The loss of *PpPPR_43* resulted in severe growth retardation while the knockout of *PpPPR_9* only led to the generation of smaller protonema colonies when it is compared to WT. This phenotypic difference can be explained by the different splicing efficiencies of *cox1* intron 3 in both mutants. *PpPPR_43* may contribute as a major factor to the splicing of *cox1* intron 3, and *PpPPR_9* may assist its splicing as an auxiliary factor.

In Arabidopsis mitochondria, nine genes are interrupted by 23 group II introns that are removed out, as 18 undergo *cis*-splicing and five undergo *trans*-splicing. In Physcomitrium, there are 16 intron-containing genes that are interrupted by 25 group II and two group I introns [29]. All of these introns are spliced out by *cis*-splicing. The mitochondrial genes, *cox1* and *atp9*, have four and three introns in Physcomitrium but no introns in Arabidopsis. The mitochondrial *nad5* gene is interrupted by three introns in Physcomitrium. The first and third introns are group II introns, and the second intron is a group I intron. In contrast, the Arabidopsis *nad5* gene is interrupted by two *cis*- and two *trans*-group II introns. Thus, the positions of the intron insertion sites in the mitochondrial genes are completely different between Physcomitrium and Arabidopsis. Hence, *PpPPR_9*, 31 and 43 are Physcomitrium-

specific PPR splicing factors. PLS-type PPR proteins that are involved in the splicing of other introns of either plastids or mitochondrial genes have yet to be identified.

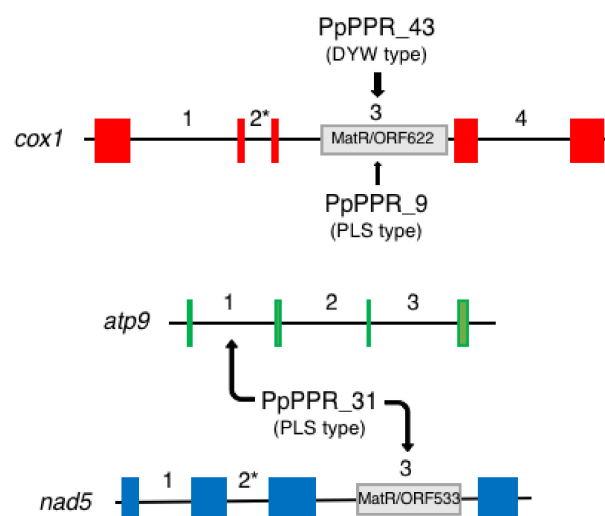


Figure 3. PPR splicing factors and their intron targets in *Physcomitrium*. The *cox1* gene is interrupted by four introns and the *atp9* and *nad5* genes by three introns in *Physcomitrium*. The second introns (marked with asterisks) of *cox1* and *nad5* genes are group I introns. The other introns are group II introns. MatR/ORF622 or ORF533 indicates an intron-encoded maturase-like protein. PpPPR_31 is involved in the splicing of the first intron of *atp9* and the third intron of *nad5*. The splicing of the third intron of *cox1* requires two PPR proteins, PpPPR_43 and PpPPR_31. PpPPR_43 may be a major factor in the splicing of *cox1* intron 3, and PpPPR_9 may assist its splicing as an auxiliary factor.

7. DYW-Type PPR Proteins in *Physcomitrium*

Physcomitrium has 10 DYW-type PPR proteins, PpPPR_43, 45, 56, 65, 71, 77, 78, 79, 91 and 98 (Table 1). In *Arabidopsis*, 82 DYW domain-containing proteins were identified and many of them are involved in RNA editing [15]. The DYW domain was named after its characteristic C-terminal tripeptide, Asp (D)-Tyr (Y)-Trp (W) and is not found in any other proteins or organisms apart from land plants, except for a heterolobosean protist, *Naegleria gruberi* [120]. The protist DYW-type proteins are hypothesized to derive from horizontal gene transfer from plants in very early land plant evolution. DYW-type proteins are closely linked to RNA editing in plant organellar transcripts, as it is explained next.

RNA editing is a post-transcriptional modification to nuclear, plastid and mitochondrial genome-encoded transcripts, and occurs in a wide range of organisms [121–123]. In seed plants, 30 to 60 cytidine (C)-to-uridine (U) RNA editing sites are found in plastids and 300 to 600 sites are found in mitochondria [122]. To date, more than 150 PLS-subfamily proteins, including the DYW type, have been identified as editing site-recognition factors [123–125]. The binding of these PPR proteins to short *cis*-elements immediately upstream of editing sites is required for C-to-U processing [123].

In *Physcomitrium*, RNA editing occurs at only two sites in chloroplasts [126] and at 11 sites in mitochondria [69,127]. Among these 13 C-to-U editing sites, only three editing events at the *ccmFc*-C103, *ccmFc*-C122 and *nad5*-C598 sites also occur in *Arabidopsis* mitochondria. Nine out of ten *Physcomitrium* DYW-type proteins were identified as editing site-recognition factors [54,58,64,65,69,71,72], indicating that each DYW-type protein participates in one or two editing events (Table 1). This is the first full assignment of DYW-type editing protein factors to all their organellar editing sites in a plant species. Interestingly, the moss, *Funaria hygrometrica*, a closely related species of *Physcomitrium*, lacks both the PpPPR_56 ortholog and its target *nad3*-C230 and *nad4*-C272 editing sites. *F. hygrometrica* has nine DYW-type proteins but lacks the PpPPR_56 ortholog [65]. This suggests that DYW-type genes and their cognate editing sites were mutually constrained during their evolution [65,128].

In *Physcomitrium* chloroplasts, C-to-U RNA editing at the *rps14*-C2 site occurs at a high efficiency (80%) and creates a translation initiation codon AUG. In addition, the *rps14*-1C site in the 5'-UTR is edited at a low efficiency (5%) [129]. These editing sites also exist in the related moss, *F. hygrometrica*, but are not found in the chloroplasts of higher plants. The knockdown of the *PpPPR_45* gene reduced the extent of RNA editing at the *rps14*-C2 site, whereas the over-expression of *PpPPR_45* increased the levels of RNA editing at both the *rps14*-C2 site and its neighboring -C1 site. This suggests that the expression level of *PpPPR_45* may affect the extent of RNA editing at the two neighboring sites. The efficiency of RNA editing at the *rps14*-C2 site was 70–80% in the young protonemata and decreased to 20% in old protonemata and the fully developed leafy shoots [129]. Thus, the RNA editing of this site is regulated in a tissue- and stage-specific manner and it may affect the efficiency of *rps14* mRNA translation.

8. Molecular Basis of RNA Editing in *Physcomitrium*

Although DYW-type proteins play a role in RNA editing as site-recognition factors in land plants, including *Physcomitrium*, the role of the DYW domain in RNA editing has long been elusive. In 2007, Salone et al. [130] proposed a hypothesis that the DYW domains catalyze C-to-U RNA editing. This is because the DYW domain contains a conserved region that includes invariant residues that match the active site of cytidine deaminases (C/HxE...PCxxC) from various organisms. Moreover, there was a correlation between the presence of nuclear DYW genes and the occurrence of organelle RNA editing among land plants [130,131]. Recently, this hypothesis was proved by in vivo orthogonal RNA editing assays in *Escherichia coli* and in vitro assays with purified proteins from *Physcomitrium* DYW-type proteins [132,133]. The crystal structure of the DYW domain of *Arabidopsis* OTP86 was recently determined, showing the potential RNA path on the DYW domain and identifying key residues required for the regulation and catalysis to occur [134]. Thus, a repeated PLS motifs-tract recognizes an immediate upstream sequence from a target editing site and a DYW domain catalyzes the C-to-U RNA editing reaction.

The chloroplast ribonucleoprotein (cpRNP) family containing two RRM associates with large transcript pools and influences multiple plastid RNA processing steps [135–138]. In particular, cpRNPs also are involved in RNA editing in tobacco [139] and *Arabidopsis* [140]. CP31A, a member of the cpRNP family, influenced the efficiency of editing at 13 sites in *Arabidopsis* chloroplasts [140]. In *Physcomitrium*, cpRNP-like proteins, PpRBP2a and PpRBP2b, are present in the chloroplasts [141,142]. KO mutants of either one or two *PpRBP2* genes exhibited a WT-like phenotype and the efficiency of RNA editing at the *rps14* sites was not altered in the KO mutants. This suggests that PpRBP2a and 2b are functionally distinct from *Arabidopsis* cpRNPs and might not be required for RNA editing in mosses. Organellar RRM-containing protein, ORRM1, was reported to be required for plastid RNA editing at multiple sites in *Arabidopsis* and maize [143] but is not found in *Physcomitrium*.

Several other protein factors that are involved in organellar RNA editing were identified in flowering plants. Multiple organellar RNA-editing factors (MORFs)/RNA editing interacting proteins (RIPs) are required for plastid and mitochondrial RNA editing in flowering plants [144,145]. Ten members of the MORF/RIP family were identified in *Arabidopsis*. They interacted with each other and also with some PPR editing factors and formed specific homo- and heteromeric interactions [146]. These factors are organized in a higher ordered editing complex (~200 kDa, called the “editosome”) [145]. The protein components of the editosome vary depending on each target site in either plastids or mitochondria. Two additional proteins, protoporphyrinogen IX oxidase 1 (PPO1) and organelle zinc finger (OZ), were also characterized as general editing factors [147,148]. Notably, PPO1, a critical enzyme for the tetrapyrrole biosynthetic pathway, plays an unexpected role in chloroplast editing at multiple sites in *Arabidopsis* [147]. PPO1 interacts with three chloroplast MORF proteins but not with the PPR proteins, suggesting that PPO1 controls the level of chloroplast editing via the stabilization of the MORFs. The OZ family contains four members, OZ1–4, in *Arabidopsis* and the disruption of *OZ1* led to an alteration in the

level of editing of most sites in chloroplasts [148]. OZ1 interacts with PPR editing factors and ORRM1, but not with MORFs. The OZ family is present in many plant lineages but not in algae. Physcomitrium and Selaginella encode OZ-like proteins but they have minimal similarity. A new imprinted gene, *NUWA* (*At3g49240*), encoding a P-type PPR protein has influenced RNA editing in plastids and mitochondria [149,150]. *NUWA* enhanced the interaction of E+-type PPR protein and DYW2, a short, atypical DYW protein [150]. The overexpression of cationic peroxidase 3 (OCP3) also affected the editing of multiple sites in the chloroplast *ndhB* transcript [151]. Porphobilinogen deaminase HEMC interacted with the PPR protein, *AtECB2*, to achieve chloroplast RNA editing [152]. Curiously, most of these RNA editing-related proteins are not found in Physcomitrium, suggesting that the RNA editing machinery is largely different between mosses and seed plants. At present, DYW-type PPR proteins are a sole key player required for RNA editing in Physcomitrium. In contrast, in Arabidopsis, four types of PPR proteins are involved in C-to-U RNA editing: DYW type (many cases), E+ type (CRR4, SLO2 and CWM1, etc.), P type (*NUWA*) and a short, atypical DYW protein (DYW1, DYW2 and MEF8, etc.) [153]. RNA editing occurs in the editosome complex, and it is specific to a respective editing site (Figure 4). However, the possibility that unidentified factors are necessary for site-recognition, and the efficiency of RNA editing events together with DYW-type proteins in Physcomitrium cannot be excluded. Unlike the complex editosome of seed plants, RNA editing may occur in a simpler editing complex that is composed of a single DYW-type PPR editing protein and a few other unidentified non-PPR editing factors, at least in mosses (Figure 4).

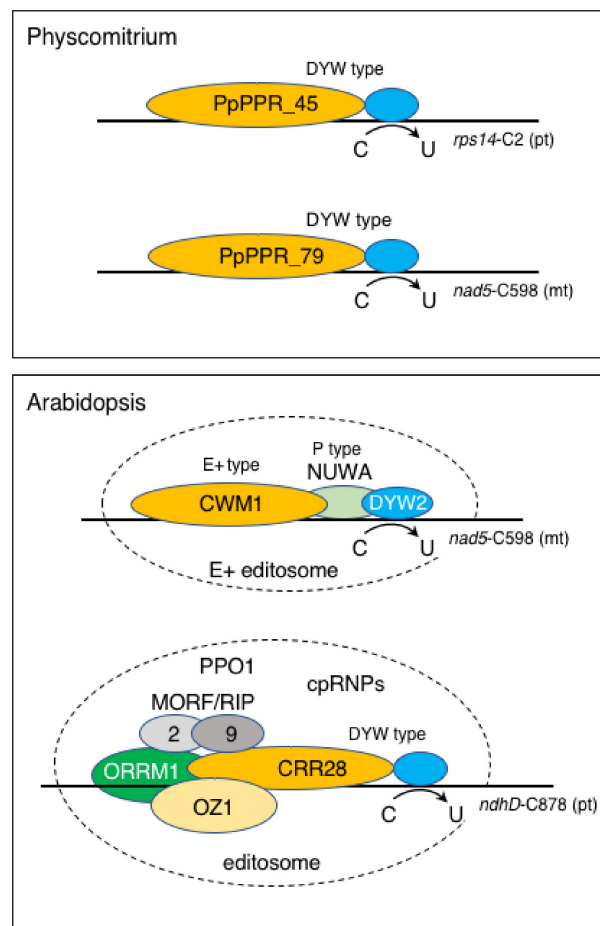


Figure 4. Models of RNA editing machinery in Physcomitrium and Arabidopsis. In Physcomitrium, nine DYW-type PPR proteins are responsible for editing at all 13 sites. They are a sole key player required for RNA editing. A PLS repeats-tract recognizes the upstream sequence of target site(s) and

its DYW domain catalyzes C-to-U editing. Editosomes, including the non-PPR proteins involved in editing, have not yet been identified in *Physcomitrium*. In *Arabidopsis*, DYW- and E+-type PPR proteins are responsible for editing-site recognition. Most E+ type proteins function *in trans* with a short, atypical DYW2 protein for editing. Examples of an RNA editing event, the PPR and the non-PPR protein components required for editing at *rps14-C2* (pt), *nad5-C598* (mt), *ndhD-C878*(pt) sites, are illustrated. pt and mt in parentheses indicate plastid and mitochondrial editing sites, respectively. Members of non-PPR families (MORF/RIP, ORRM and OZ) are partially redundant. The editosome requires multiple copies of non-PPR factors. Most plastid and mitochondrial editosomes usually contain multiple non-PPR proteins in *Arabidopsis* but not in *Physcomitrium*. Models of *Arabidopsis* editosomes were modified from Andreés-Colaás et al. [150] and Sun et al. [148]. In the early land plants (mosses), the non-PPR editing factors that were identified in *Arabidopsis* were not encoded in their nuclear genomes. Unlike the complex editosome of seed plants, RNA editing may occur in a simpler editing complex, composed of a single DYW-type PPR editing protein and a few other unidentified non-PPR editing factors, at least in mosses.

9. Additional Function of the DYW Domain

Although most DYW-type proteins are responsible for RNA editing, some are involved in RNA splicing [53] or RNA cleavage [21]. The DYW-type PpPPR_43 protein is required for the group II intron splicing of the mitochondrial *cox1* transcript [53] (Figure 3). Its DYW domain is distinct from the other nine DYW domains of *Physcomitrium* proteins. *Arabidopsis* DYW-type CRR2 is required for the intergenic RNA cleavage of plastid *rps7-ndhB* pre-mRNA [21]. The DYW domain of CRR2 was indispensable for the cleavage of the target RNA *in vivo* [154]. The DYW domain of OTP85 (At2g02980) possessed endoribonuclease activity *in vitro* [155]. Its DYW domain contains the cytochrome *c* family heme-binding site signature (CxxCH), which overlaps with the active site of cytidine deaminase. The mutation of this signature to GxxGH resulted in a significant reduction in RNA cleavage activity. This suggests that the CxxCH motif is required for endoribonuclease activity of the DYW domain. The DYW domains of PpPPR_56, 71 and 77 proteins also showed endoribonuclease activity *in vitro* [116]. Thus, *Physcomitrium* DYW-type proteins are involved in RNA editing but may also function in certain RNA processing events in organelles. This possibility needs to be further investigated.

10. Conclusions and Future Perspective

As described above, the size and constitution of the PPR protein family are largely different between *Physcomitrium* and flowering plants. Some PPR proteins such as PpPPR_66/AtPPR66L and PpPPR_63, 67,104/AtPRORPs show the same or similar function, but some proteins including pTAC2 and its homolog, PpPPR_64, have a diverse function in *Physcomitrium* and *Arabidopsis*. Mitochondrial PpPPR_9, 31 and 43 are PPR splicing factors that are specific to *Physcomitrium* because their target introns are present in this moss but are not in *Arabidopsis*. There are only 13 RNA editing sites in *Physcomitrium*, but there are over 500 in *Arabidopsis*. At least three types of PPR proteins (E+, DYW and P type) are involved in RNA editing in *Arabidopsis* while only the DYW-type proteins are required for editing in *Physcomitrium*. Of the three editing events that are conserved in both plants, editing at the mitochondrial *nad5-C598* site requires DYW-type PpPPR_79 in *Physcomitrium* but it needs the tripartite CWM1 (E+ type), NUWA (P type) and DYW2 (atypical DYW type) in *Arabidopsis*. Several PPR proteins are found in the plastid nucleoids [47,62] or mitochondrial ribosomes of *Arabidopsis* [156]. In contrast, at present, no PPR proteins in nucleoids or ribosomes have been identified in *Physcomitrium*. The PPR protein, GUN1, is known to be required for chloroplast-to-nucleus retrograde signaling [70]. Although GUN1 is an ancient protein that evolved within the streptophyte algal ancestors of land plants, but it has no role in chloroplast retrograde signaling in the streptophyte alga, *Coleochate orbicularis*, or in the liverwort, *M. polymorpha* [157]. Its role in chloroplast retrograde signaling probably evolved more recently. Thus, the evolutionarily conserved PPR

proteins are not always functional orthologs and their function may have been expanded and diversified during plant evolution.

The proplastids of seed plants differentiate to various types of plastids including etioplasts, chloroplasts and chromoplasts. However, such plastid differentiation is not observed in *Physcomitrium*, and the moss gametophytes always contain chloroplasts that develop even while they are in the dark [158,159]. Therefore, it is considered that plastid gene expression is differentially regulated in a plastid-type specific manner in seed plants, although this is not likely in *Physcomitrium*. This may imply that the post-transcriptional regulation in plastids is more complex in seed plants than it is in mosses. Alternatively, seed plants may require more PPR proteins to accomplish the post-transcriptional regulation of gene expression specific to each plastid type. In contrast, mosses, including *Physcomitrium*, might retain a minimum set of PPR proteins that are required for the post-transcriptional regulation of plant organellar gene expression. To understand the basal and molecular mechanism of post-transcriptional regulation in plastid and mitochondrial gene expression, further identification of all the target RNA molecules recognized by the *Physcomitrium* PPR proteins and a characterization of their functions needs to be urgently performed. This will provide clues to identify the primary (primordial) function of PPR proteins in land plant lineages.

Funding: JSPS KAKENHI Grant Numbers 17K08195, 20K05957 to MS.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: I thank all members of the Sugita lab for contributing to functional analysis of *Physcomitrium* PPR proteins.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Gray, M.W. The evolutionary origins of organelles. *Trends Genet.* **1989**, *5*, 294–299. [[CrossRef](#)]
2. Herrmann, R.G.; Westhoff, P.; Link, G. Biogenesis of plastids in higher plants. In *Cell Organelles. Plant Gene Research*; Herrmann, R.G., Ed.; Springer: Vienna, Austria, 1992; Volume 6, pp. 275–349. [[CrossRef](#)]
3. Binder, S.; Brennicke, A. Gene expression in plant mitochondria: Transcriptional and post-transcriptional control. *Phil. Trans. R. Soc. Lond. B* **2003**, *358*, 181–189. [[CrossRef](#)] [[PubMed](#)]
4. Barkan, A. Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic mRNAs. *EMBO J.* **1988**, *7*, 2637–2644. [[CrossRef](#)] [[PubMed](#)]
5. Sugita, M.; Sugiura, M. Regulation of gene expression in chloroplasts of higher plants. *Plant Mol. Biol.* **1996**, *32*, 315–326. [[CrossRef](#)]
6. Stern, D.B.; Goldschmidt-Clermont, M.; Hanson, M.R. Chloroplast RNA metabolism. *Ann. Rev. Plant Biol.* **2010**, *61*, 125–155. [[CrossRef](#)]
7. Hammani, K.; Giegé, P. RNA metabolism in plant mitochondria. *Trends Plant Sci.* **2014**, *19*, 380–389. [[CrossRef](#)]
8. Schmitz-Linneweber, C.; Small, I. Pentatricopeptide repeat proteins: A socket set for organelle gene expression. *Trends Plant Sci.* **2008**, *13*, 663–670. [[CrossRef](#)]
9. Gutmann, B.; Gobert, A.; Giegé, P. Mitochondrial genome evolution and the emergence of PPR proteins. In *Advances in Botanical Research 63, Mitochondrial Genome Evolution. Edition: Mitochondria, RNA Binding Proteins, Post-Transcriptional Regulation*; Drouard, L., Ed.; Elsevier: Amsterdam, The Netherlands, 2012; Chapter 10; pp. 253–313. [[CrossRef](#)]
10. Barkan, A.; Small, I. Pentatricopeptide repeat proteins in plants. *Annu. Rev. Plant Biol.* **2014**, *65*, 415–442. [[CrossRef](#)]
11. Small, I.D.; Peeters, N. The PPR-motif—A TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* **2000**, *25*, 46–47. [[CrossRef](#)]
12. Aubourg, S.; Boudet, N.; Kreis, M.; Lecharny, A. In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. *Plant Mol. Biol.* **2000**, *42*, 603–613. [[CrossRef](#)]
13. Groves, M.R.; Barford, D. Topological characteristics of helical repeat protein. *Curr. Opin. Struct. Biol.* **1999**, *9*, 383–389. [[CrossRef](#)]
14. Lurin, C.; Andrés, C.; Aubourg, S.; Bellaoui, M.; Bitton, F.; Bruyère, C.; Caboche, M.; Debast, C.; Gualberto, J.; Hoffmann, B.; et al. Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* **2004**, *16*, 2089–2103. [[CrossRef](#)]

15. Cheng, S.; Gutmann, B.; Zhong, X.; Ye, Y.; Fisher, M.F.; Bai, F.; Castleden, I.; Song, Y.; Song, B.; Huang, J.; et al. Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. *Plant J.* **2016**, *85*, 532–547. [[CrossRef](#)]
16. Colcombet, J.; Lopez-Obando, M.; Heurtevin, L.; Bernard, C.; Martin, K.; Berthomé, R.; Lurin, C. Systematic study of subcellular localization of Arabidopsis PPR proteins confirms a massive targeting to organelles. *RNA Biol.* **2013**, *10*, 1557–1575. [[CrossRef](#)]
17. Barkan, A.; Walker, M.; Nolasco, M.; Johnson, D. A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. *EMBO J.* **1994**, *13*, 3170–3181. [[CrossRef](#)]
18. Fisk, D.G.; Walker, M.B.; Barkan, A. Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO J.* **1999**, *18*, 2621–2630. [[CrossRef](#)]
19. Bentolila, S.; Alfonso, A.A.; Hanson, M.R. A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10887–10892. [[CrossRef](#)]
20. Meierhoff, K.; Felder, S.; Nakamura, T.; Bechtold, N.; Schuster, G. HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast *psbB-psbT-psbH-petB-petD* RNAs. *Plant Cell* **2003**, *15*, 1480–1495. [[CrossRef](#)]
21. Hashimoto, M.; Endo, T.; Peltier, G.; Tasaka, M.; Shikanai, T. A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in *Arabidopsis*. *Plant J.* **2003**, *36*, 541–549. [[CrossRef](#)]
22. Kazama, T.; Toriyama, K. A pentatricopeptide repeat-containing gene that promotes the processing of aberrant *atp6* RNA of cytoplasmic male-sterile rice. *FEBS Lett.* **2003**, *544*, 1873–3468. [[CrossRef](#)]
23. Williams, P.M.; Barkan, A. A chloroplast-localized PPR protein required for plastid ribosome accumulation. *Plant J.* **2003**, *36*, 675–686. [[CrossRef](#)] [[PubMed](#)]
24. Yamazaki, H.; Tasaka, M.; Shikanai, T. PPR motif of the nucleus-encoded factor, PGR3, function in the selective and distinct steps of chloroplast gene expression in *Arabidopsis*. *Plant J.* **2004**, *38*, 152–163. [[CrossRef](#)] [[PubMed](#)]
25. Hattori, M.; Hasebe, M.; Sugita, M. Identification and characterization of cDNAs encoding pentatricopeptide repeat proteins in the basal land plant, the moss *Physcomitrella patens*. *Gene* **2004**, *343*, 305–311. [[CrossRef](#)] [[PubMed](#)]
26. Cove, D. The moss *Physcomitrella patens*. *Annu. Rev. Genet.* **2005**, *39*, 339–358. [[CrossRef](#)]
27. Cove, D.; Bezanilla, M.; Harries, P.; Quatrano, R. Mosses as model systems for the study of metabolism and development. *Annu. Rev. Plant Biol.* **2006**, *57*, 497–520. [[CrossRef](#)]
28. Sugiura, C.; Kobayashi, Y.; Aoki, S.; Sugita, C.; Sugita, M. Complete chloroplast DNA sequence of the moss *Physcomitrella patens*: Evidence for the loss and relocation of *rpoA* from the chloroplast to the nucleus. *Nucleic Acids Res.* **2003**, *31*, 5324–5331. [[CrossRef](#)]
29. Terasawa, K.; Odahara, M.; Kabeya, Y.; Kikugawa, T.; Sekine, Y.; Fugiwara, M.; Sato, N. The mitochondrial genome of the moss *Physcomitrella patens* sheds new light on mitochondrial evolution in land plants. *Mol. Biol. Evol.* **2007**, *24*, 699–709. [[CrossRef](#)]
30. Rensing, S.A.; Lang, D.; Zimmer, A.D.; Terry, A.; Salamov, A.; Shapiro, H.; Nishiyama, T.; Perroud, P.-F.; Lindquist, E.A.; Kamisugi, Y.; et al. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **2008**, *319*, 64–69. [[CrossRef](#)]
31. Schaefer, D.G.; Zryd, J.-P. Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J.* **1997**, *11*, 1195–1206. [[CrossRef](#)]
32. Sugiura, C.; Sugita, M. Plastid transformation reveals that moss tRNA^{Arg}-CCG is not essential for plastid function. *Plant J.* **2004**, *40*, 314–321. [[CrossRef](#)]
33. Nishiyama, T.; Hiwatashi, Y.; Sakakibara, K.; Kato, M.; Hasebe, M. Tagged mutagenesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis. *DNA Res.* **2000**, *7*, 9–17. [[CrossRef](#)] [[PubMed](#)]
34. Bowman, J.L.; Kohchi, T.; Yamato, K.T.; Jenkins, J.; Shu, S.; Ishizaki, K.; Yamaoka, S.; Nishihama, R.; Nakamura, Y.; Berger, F.; et al. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* **2017**, *171*, 287–304. [[CrossRef](#)] [[PubMed](#)]
35. Nishiyama, T.; Sakayama, H.; de Vries, J.; Buschmann, H.; Saint-Marcoux, D.; Ullrich, K.K.; Haas, F.B.; Vanderstraeten, L.; Becker, D.; Lang, D.; et al. The *Chara* genome: Secondary complexity and implications for plant terrestrialization. *Cell* **2018**, *174*, 448–464.e24. [[CrossRef](#)]
36. O’Toole, N.; Hattori, M.; Andres, C.; Iida, K.; Lurin, C.; Schmitz-Linneweber, C.; Sugita, M.; Small, I. On the expansion of the pentatricopeptide repeat gene family in plants. *Mol. Biol. Evol.* **2008**, *25*, 1120–1128. [[CrossRef](#)]
37. Schmitz-Linneweber, C.; Williams-Carrier, R.E.; Williams-Voelker, P.M.; Kroeger, T.S.; Vichas, A.; Barkan, A. A pentatricopeptide repeat protein facilitates the *trans*-splicing of the maize chloroplast *rps12* pre-mRNA. *Plant Cell* **2006**, *18*, 2650–2663. [[CrossRef](#)]
38. Tahini, L.; Ferrari, R.; Lehninger, M.-K.; Mizzotti, C.; Moratti, F.; Resentini, F.; Colombo, M.; Costa, A.; Masiero, S.; Pesaresi, P. *Trans*-splicing of plastid *rps12* transcripts, mediated by *AtPPR4*, is essential for embryo patterning in *Arabidopsis thaliana*. *Planta* **2018**, *248*, 257–265. [[CrossRef](#)]
39. Goto, S.; Kawaguchi, Y.; Sugita, C.; Ichinose, M.; Sugita, M. P-class pentatricopeptide repeat protein PTSF1 is required for splicing of the plastid pre-tRNA^{Ile} in *Physcomitrella patens*. *Plant J.* **2016**, *86*, 493–503. [[CrossRef](#)]
40. de Longevialle, A.F.; Hendrickson, L.; Taylor, N.L.; Delannoy, E.; Lurin, C.; Badger, M.; Millar, A.H.; Small, I. The pentatricopeptide repeat gene *OTP51* with two LAGLIDADG motifs is required for the *cis*-splicing of plastid *ycf3* intron 2 in *Arabidopsis thaliana*. *Plant J.* **2008**, *56*, 157–168. [[CrossRef](#)]
41. Ichinose, M.; Ishimaru, A.; Nakajima, K.; Kawaguchi, Y.; Sugita, M. Two novel PLS-class pentatricopeptide repeat proteins are involved in the group II intron splicing of mitochondrial transcripts in the moss *Physcomitrella patens*. *Plant Cell Physiol.* **2020**, *61*, 1687–1698. [[CrossRef](#)]

42. Jin, H.; Fu, M.; Duan, Z.; Duan, S.; Li, M.; Dong, X.; Liu, B.; Feng, D.; Wang, J.; Peng, L.; et al. Low photosynthetic efficiency 1 is required for light-regulated photosystem II biogenesis in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E6075–E6084. [[CrossRef](#)]
43. Beick, S.; Schmitz-Linneweber, C.; Williams-Carrier, R.; Jensen, B.; Barkan, A. The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. *Mol. Cell. Biol.* **2008**, *28*, 5337–5347. [[CrossRef](#)] [[PubMed](#)]
44. Johnson, X.; Wostrikoff, K.; Finazzi, G.; Kuras, R.; Schwarz, C.; Bujaldon, S.; Nickelsen, J.; Stern, D.B.; Wollman, F.-A.; Vallon, O. MRL1, a conserved pentatricopeptide repeat protein, is required for stabilization of *rbcl* mRNA in *Chlamydomonas* and *Arabidopsis*. *Plant Cell* **2010**, *22*, 234–248. [[CrossRef](#)] [[PubMed](#)]
45. Ebihara, T.; Matsuda, Y.; Sugita, C.; Ichinose, M.; Yamamoto, H.; Shikanai, T.; Sugita, M. The P-class pentatricopeptide repeat protein PpPPR_21 is needed for accumulation of the *psbI-ycf12* dicistronic mRNA in *Physcomitrella* chloroplasts. *Plant J.* **2019**, *97*, 1120–1131. [[CrossRef](#)]
46. Lee, K.; Park, S.J.; Han, J.H.; Jeon, Y.; Pai, H.-S.; Kang, H. A chloroplast-targeted pentatricopeptide repeat protein PPR287 is crucial for chloroplast function and *Arabidopsis* development. *BMC Plant Biol.* **2019**, *19*, 244. [[CrossRef](#)]
47. Majeran, W.; Friso, G.; Asakura, Y.; Qu, X.; Hung, M.; Ponnala, L.; Watkins, K.P.; Barkan, A.; van Wijk, K.J. Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: A new conceptual framework for nucleoid functions. *Plant Physiol.* **2012**, *158*, 156–189. [[CrossRef](#)]
48. Suzuki, R.; Sugita, C.; Aoki, S.; Sugita, M. *Physcomitrium patens* pentatricopeptide repeat protein PpPPR_32 is involved in the accumulation of *psaC* mRNA encoding the iron sulfur protein of photosystem I. *Genes Cells* **2022**, *27*, 293–304. [[CrossRef](#)]
49. Hattori, M.; Miyake, H.; Sugita, M. A pentatricopeptide repeat protein is required for RNA processing of *clpP* pre-mRNA in moss chloroplasts. *J. Biol. Chem.* **2007**, *282*, 10773–10782. [[CrossRef](#)]
50. Hattori, M.; Sugita, M. A moss pentatricopeptide repeat protein binds to the 3' end of plastid *clpP* pre-mRNA and assists with the mRNA maturation. *FEBS J.* **2009**, *276*, 5860–5869. [[CrossRef](#)]
51. Wang, X.; Yang, Z.; Zhang, Y.; Zhou, W.; Zhang, A.; Lu, C. Pentatricopeptide repeat protein PHOTOSYSTEM I BIOGENESIS FACTOR2 is required for splicing of *ycf3*. *J. Integr. Plant Biol.* **2020**, *62*, 1741–1761. [[CrossRef](#)]
52. Chi, W.; Ma, J.; Zhang, D.; Guo, J.; Chen, F.; Lu, C.; Zhang, L. The pentatricopeptide repeat protein DELAYED GREENING1 is involved in the regulation of early chloroplast development and chloroplast gene expression in *Arabidopsis*. *Plant Physiol.* **2008**, *147*, 573–584. [[CrossRef](#)]
53. Ichinose, M.; Tasaki, E.; Sugita, C.; Sugita, M. A PPR-DYW protein is required for splicing of a group II intron of *cox1* pre-mRNA in *Physcomitrella patens*. *Plant J.* **2012**, *70*, 271–278. [[CrossRef](#)] [[PubMed](#)]
54. Ichinose, M.; Uchida, M.; Sugita, M. Identification of a pentatricopeptide repeat RNA editing factor in *Physcomitrella patens* chloroplasts. *FEBS Lett.* **2014**, *588*, 4060–4064, Erratum in *FEBS Lett.* **2015**, *589*, 1171. [[CrossRef](#)]
55. Zhang, J.; Xiao, J.; Li, Y.; Su, B.; Xu, H.; Shan, X.; Song, C.; Xie, J.; Li, R. PDM3, a pentatricopeptide repeat-containing protein, affects chloroplast development. *J. Exp. Bot.* **2017**, *68*, 5615–5627. [[CrossRef](#)] [[PubMed](#)]
56. Du, L.; Zhang, J.; Qu, S.; Zhao, Y.; Su, B.; Lv, X.; Li, R.; Wan, Y.; Xiao, J. The pentatricopeptide repeat protein pigment-defective mutant2 is involved in the regulation of chloroplast development and chloroplast gene expression in *Arabidopsis*. *Plant Cell Physiol.* **2017**, *58*, 747–759. [[CrossRef](#)] [[PubMed](#)]
57. Khrouchtchova, A.; Monde, R.A.; Barkan, A. A short PPR protein required for the splicing of specific group II introns in angiosperm chloroplasts. *RNA* **2012**, *18*, 1197–1209. [[CrossRef](#)]
58. Ohtani, S.; Ichinose, M.; Tasaki, E.; Aoki, Y.; Komura, Y.; Sugita, M. Targeted gene disruption identifies three PPR-DYW proteins involved in RNA editing for five editing sites of the moss mitochondrial transcripts. *Plant Cell Physiol.* **2010**, *51*, 1942–1949. [[CrossRef](#)]
59. Senkler, J.; Senkler, M.; Eubel, H.; Hildebrandt, T.; Lengwenus, C.; Schertl, P.; Schwarzländer, M.; Wagner, S.; Wittig, I.; Braun, H.-P. The mitochondrial complexome of *Arabidopsis thaliana*. *Plant J.* **2017**, *89*, 1079–1092. [[CrossRef](#)]
60. Sugita, C.; Komura, Y.; Tanaka, K.; Kometani, K.; Satoh, H.; Sugita, M. Molecular characterization of three PRORP proteins in the moss *Physcomitrella patens*: Nuclear PRORP protein is not essential for moss viability. *PLoS ONE* **2014**, *9*, e108962-10. [[CrossRef](#)]
61. Gobert, A.; Gutmann, B.; Taschner, A.; Gößringer, M.; Holzmann, J.; Hartmann, R.K.; Rossmanith, W.; Giegé, P. A single *Arabidopsis* organellar protein has RNase P activity. *Nat. Struct. Mol. Biol.* **2010**, *17*, 740–744. [[CrossRef](#)]
62. Pfalz, J.; Liere, K.; Kandlbinder, A.; Dietz, K.J.; Oelmüller, R. pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell* **2006**, *18*, 176–197. [[CrossRef](#)]
63. Takahashi, A.; Sugita, C.; Ichinose, M.; Sugita, M. Moss PPR-SMR protein PpPPR_64 influences the expression of a *psaA-psaB-rps14* gene cluster and processing of the 23S–4.5S rRNA precursor in chloroplasts. *Plant Mol. Biol.* **2021**, *107*, 417–429. [[CrossRef](#)] [[PubMed](#)]
64. Ichinose, M.; Sugita, C.; Yagi, Y.; Nakamura, T.; Sugita, M. Two DYW subclass PPR proteins are involved in RNA editing of *ccmFc* and *atp9* transcripts in the moss *Physcomitrella patens*: First complete set of PPR editing factors in plant mitochondria. *Plant Cell Physiol.* **2013**, *54*, 1907–1916. [[CrossRef](#)] [[PubMed](#)]
65. Schallenberg-Rüdinger, M.; Kindgren, P.; Zehrmann, A.; Small, I.; Knoop, V. A DYW-protein knockout in *Physcomitrella* affects two closely spaced mitochondrial editing sites and causes a severe developmental phenotype. *Plant J.* **2013**, *76*, 420–432. [[CrossRef](#)]

66. Ito, A.; Sugita, C.; Ichinose, M.; Kato, Y.; Yamamoto, H.; Shikanai, T.; Sugita, M. An evolutionarily conserved P-subfamily pentatricopeptide repeat protein is required to splice the plastid *ndhA* transcript in the moss *Physcomitrella patens* and *Arabidopsis thaliana*. *Plant J.* **2018**, *94*, 638–648. [[CrossRef](#)] [[PubMed](#)]
67. Zhang, H.D.; Cui, Y.L.; Huang, C.; Yin, Q.-Q.; Qin, X.-M.; Xu, T.; He, X.-F.; Zhang, Y.; Li, Z.-R.; Yang, Z.-N. PPR protein PDM1/SEL1 is involved in RNA editing and splicing of plastid genes in *Arabidopsis thaliana*. *Photosynth. Res.* **2015**, *126*, 311–321. [[CrossRef](#)]
68. Kushwaha, H.R.; Singh, A.K.; Sopory, S.K.; Singla-Pareek, S.L.; Pareek, A. Genome wide expression analysis of CBS domain containing proteins in *Arabidopsis thaliana* (L.) Heynh and *Oryza sativa* L. reveals their developmental and stress regulation. *BMC Genom.* **2009**, *10*, 200. [[CrossRef](#)]
69. Tasaki, E.; Hattori, M.; Sugita, M. The moss pentatricopeptide repeat protein with a DYW domain is responsible for RNA editing of mitochondrial *ccmF*c transcript. *Plant J.* **2010**, *62*, 560–570. [[CrossRef](#)]
70. Koussevitzky, S.; Nott, A.; Mockler, T.C.; Hong, F.; Sachetto-Martins, G.; Surpin, M.; Lim, J.; Mittler, R.; Chory, J. Signals from chloroplasts converge to regulate nuclear gene expression. *Science* **2007**, *316*, 715–719. [[CrossRef](#)]
71. Uchida, M.; Ohtani, S.; Ichinose, M.; Sugita, C.; Sugita, M. The PPR-DYW proteins are required for RNA editing of *rps14*, *cox1* and *nad5* transcripts in *Physcomitrella patens* mitochondria. *FEBS Lett.* **2011**, *585*, 2367–2371. [[CrossRef](#)] [[PubMed](#)]
72. Rüdinger, M.; Szövényi, P.; Rensing, S.A.; Knoop, V. Assigning DYW-type PPR proteins to RNA editing sites in the funariid mosses *Physcomitrella patens* and *Funaria hygrometrica*. *Plant J.* **2011**, *67*, 370–380. [[CrossRef](#)]
73. Aryamanesh, N.; Ruwe, H.; Sanglard, L.V.P.; Eshraghi, L.; Bussell, J.D.; Howell, K.A.; Small, I. Colas des Francs-Small, C. The pentatricopeptide repeat protein EMB2654 is essential for *trans*-splicing of a chloroplast small ribosomal subunit transcript. *Plant Physiol.* **2017**, *173*, 1164–1176. [[CrossRef](#)] [[PubMed](#)]
74. Lee, K.; Park, S.J.; Colas des Francs-Small, C.; Whitby, M.; Small, I.; Kang, H. The coordinated action of PPR4 and EMB2654 on each intron half mediates *trans*-splicing of *rps12* transcripts in plant chloroplasts. *Plant J.* **2019**, *100*, 1193–1207. [[CrossRef](#)] [[PubMed](#)]
75. Schmitz-Linneweber, C.; Williams-Carrier, R.; Barkan, A. 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* **2005**, *17*, 2791–2804. [[CrossRef](#)] [[PubMed](#)]
76. Ferrari, R.; Tadini, L.; Moratti, F.; Lehniger, M.-K.; Costa, A.; Rossi, F.; Colombo, M.; Masiero, S.; Schmitz-Linneweber, C.; Pesaresi, P. CRP1 Protein: (dis)similarities between *Arabidopsis thaliana* and *Zea mays*. *Front. Plant Sci.* **2017**, *8*, 163. [[CrossRef](#)]
77. Zhang, L.; Zhou, W.; Che, L.; Rochaix, J.D.; Lu, C.; Li, W.; Peng, L. PPR protein BFA2 is essential for the accumulation of the *atpH/F* transcript in chloroplasts. *Front. Plant Sci.* **2019**, *10*, 446. [[CrossRef](#)]
78. Zhelyazkova, P.; Hammani, K.; Rojas, M.; Voelker, R.; Vargas-Suárez, M.; Börner, T.; Barkan, A. Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. *Nucleic Acids Res.* **2012**, *40*, 3092–3105. [[CrossRef](#)]
79. Yagi, Y.; Ishizaki, Y.; Nakahira, Y.; Tozawa, Y.; Shiina, T. Eukaryotic-type plastid nucleoid protein pTAC3 is essential for transcription by the bacterial-type plastid RNA polymerase. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 7541–7546. [[CrossRef](#)]
80. Pfalz, J.; Bayramtar, O.A.; Prikryl, J.; Barkan, A. Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. *EMBO J.* **2009**, *28*, 2042–2052. [[CrossRef](#)]
81. Barkan, A.; Rojas, M.; Fujii, S.; Yap, A.; Chong, Y.S.; Bond, C.S.; Small, I. A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet.* **2012**, *8*, e1002910. [[CrossRef](#)]
82. Yagi, Y.; Hayashi, S.; Kobayashi, K.; Hirayama, T.; Nakamura, T. Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PLoS ONE* **2013**, *8*, e57286. [[CrossRef](#)]
83. Takenaka, M.; Zehrmann, A.; Brennicke, A.; Graichen, K. Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. *PLoS ONE* **2013**, *8*, e65343. [[CrossRef](#)]
84. Yin, P.; Li, Q.; Yan, C.; Liu, Y.; Liu, J.; Yu, F.; Wang, Z.; Long, J.; He, J.; Wang, H.-W.; et al. Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* **2013**, *504*, 168–171. [[CrossRef](#)] [[PubMed](#)]
85. Ke, J.; Chen, R.-Z.; Ban, T.; Zhou, X.E.; Gu, X.; Tan, M.H.E.; Chen, C.; Kang, Y.; Brunzelle, J.B.; Zhu, J.-K.; et al. Structural basis for RNA recognition by a dimeric PPR-protein complex. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1377–1382. [[CrossRef](#)] [[PubMed](#)]
86. Colas des Francs-Small, C.; Vincis Pereira Sanglard, L.; Small, I. Targeted cleavage of *nad6* mRNA induced by a modified pentatricopeptide repeat protein in plant mitochondria. *Commun. Biol.* **2018**, *1*, 166. [[CrossRef](#)] [[PubMed](#)]
87. Rojas, M.; Yu, Q.; Williams-Carrier, R.; Maliga, P.; Barkan, A. Engineered PPR proteins as inducible switches to activate the expression of chloroplast transgenes. *Nat. Plants* **2019**, *5*, 505–511. [[CrossRef](#)]
88. Yu, Q.; Barkan, A.; Maliga, P. Engineered RNA-binding protein for transgene activation in non-green plastids. *Nat. Plants* **2019**, *5*, 486–490. [[CrossRef](#)] [[PubMed](#)]
89. Künstner, P.; Guardiola, A.; Takahashi, Y.; Rochaix, J.D. A mutant strain of *Chlamydomonas reinhardtii* lacking the chloroplast photosystem II *psbI* gene grows photoautotrophically. *J. Biol. Chem.* **1995**, *270*, 9651–9654. [[CrossRef](#)]
90. Schwenkert, S.; Umate, P.; Dal Bosco, C.; Volz, S.; Mlčochová, L.; Zoryan, M.; Eichacker, L.A.; Ohad, I.; Herrmann, R.G.; Meurer, J. PsbI affects the stability, function, and phosphorylation patterns of photosystem II assemblies in tobacco. *J. Biol. Chem.* **2006**, *281*, 34227–34238. [[CrossRef](#)]

91. del Campo, E.M.; Sabater, B.; Martin, M. Post-transcriptional control of chloroplast gene expression: Accumulation of stable *psaC* mRNA is due to downstream RNA cleavages in the *ndhD* gene. *J. Biol. Chem.* **2002**, *277*, 36457–36464. [[CrossRef](#)]
92. Li, X.; Luo, W.; Zhou, W.; Yin, X.; Wang, X.; Li, X.; Jiang, C.; Zhang, Q.; Kang, X.; Zhang, A.; et al. CAF proteins help SOT1 regulate the stability of chloroplast *ndhA* transcripts. *Int. J. Mol. Sci.* **2021**, *22*, 12639. [[CrossRef](#)]
93. Hirose, T.; Sugiura, M. Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast *ndhD* mRNA: A possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. *EMBO J.* **1997**, *16*, 6804–6811. [[CrossRef](#)] [[PubMed](#)]
94. Burrows, P.A.; Sazanov, L.A.; Svab, Z.; Maliga, P.; Nixon, P.J. Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid *ndh* genes. *EMBO J.* **1998**, *17*, 868–876. [[CrossRef](#)] [[PubMed](#)]
95. Takahashi, Y.; Goldschmidt-Clermont, M.; Soen, S.Y.; Franzén, L.G.; Rochaix, J.D. Directed chloroplast transformation in *Chlamydomonas reinhardtii*: Insertional inactivation of the *psaC* gene encoding the iron sulfur protein destabilizes photosystem I. *EMBO J.* **1991**, *10*, 2033–2040. [[CrossRef](#)] [[PubMed](#)]
96. Schöttler, M.A.; Albus, C.A.; Bock, R. Photosystem I: Its biogenesis and function in higher plants. *J. Plant Physiol.* **2011**, *168*, 1452–1461. [[CrossRef](#)]
97. Watkins, K.P.; Kroeger, T.S.; Cooke, A.M.; Williams-Carrier, R.E.; Friso, G.; Belcher, S.E.; van Wijk, K.J.; Barkan, A. A ribonuclease III domain protein functions in group II intron splicing in maize chloroplasts. *Plant Cell* **2007**, *19*, 2606–2623. [[CrossRef](#)]
98. Kroeger, T.S.; Watkins, K.P.; Friso, G.; van Wijk, K.J.; Barkan, A. A plant-specific RNA-binding domain revealed through analysis of chloroplast group II intron splicing. *Proc. Natl. Acad. Sci. USA* **2006**, *106*, 4537–4542. [[CrossRef](#)]
99. Asakura, Y.; Galarneau, E.; Watkins, K.P.; Barkan, A.; van Wijk, K.J. Chloroplast RH3 DEAD box RNA helicases in maize and Arabidopsis function in splicing of specific group II introns and affect chloroplast ribosome biogenesis. *Plant Physiol.* **2012**, *159*, 961–974. [[CrossRef](#)]
100. Hammani, K.; Barkan, A. An mTERF domain protein functions in group II intron splicing in maize chloroplasts. *Nucleic Acids Res.* **2014**, *42*, 5033–5042. [[CrossRef](#)]
101. Zoschke, R.; Nakamura, M.; Liere, K.; Sugiura, M.; Börner, T.; Schmitz-Linneweber, C. An organellar maturase associates with multiple group II introns. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 3245–3250. [[CrossRef](#)]
102. Schmitz-Linneweber, C.; Lampe, M.-K.; Sultan, L.D.; Ostersetzer-Biran, O. Organellar maturases: A window into the evolution of the spliceosome. *Biochim. Biophys. Acta* **2015**, *1847*, 798–808. [[CrossRef](#)]
103. Jenkins, B.; Kulhanek, D.; Barkan, A. Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. *Plant Cell* **1997**, *9*, 283–296. [[CrossRef](#)] [[PubMed](#)]
104. Ostheimer, G.J.; Williams-Carrier, R.; Belcher, S.; Osborne, E.; Gierke, J.; Barkan, A. Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. *EMBO J.* **2003**, *22*, 3919–3929. [[CrossRef](#)] [[PubMed](#)]
105. Asakura, Y.; Barkan, A. A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts. *Plant Cell* **2007**, *19*, 3864–3875. [[CrossRef](#)] [[PubMed](#)]
106. Frank, D.N.; Pace, N.R. Ribonuclease P: Unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem.* **1998**, *67*, 153–180. [[CrossRef](#)]
107. Gutmann, B.; Gobert, A.; Giegé, P. PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. *Genes Dev.* **2012**, *26*, 1022–1027. [[CrossRef](#)]
108. Anantharaman, V.; Aravind, L. The NYN domains. Novel predicted RNases with a PIN domain-like fold. *RNA Biol.* **2006**, *3*, 18–27. [[CrossRef](#)]
109. Moreira, D.; Philippe, H. Smr: A bacterial and eukaryotic homologue of the C-terminal region of the MutS2 family. *Trends Biochem. Sci.* **1999**, *24*, 298–300. [[CrossRef](#)]
110. Liu, S.; Melonek, J.; Boykin, L.M.; Small, I.; Howell, K.A. PPR-SMRs. Ancient proteins with enigmatic functions. *RNA Biol.* **2013**, *10*, 1501–1510. [[CrossRef](#)]
111. Liu, X.; Yu, F.; Rodermeil, S. An Arabidopsis pentatricopeptide repeat protein, SUPPRESSOR OF VARIATION7, is required for FtsH-mediated chloroplast biogenesis. *Plant Physiol.* **2010**, *154*, 1588–1601. [[CrossRef](#)]
112. Wu, W.; Liu, S.; Ruwe, H.; Zhang, D.; Melonek, J.; Zhu, Y.; Hu, X.; Gusewski, S.; Yin, P.; Small, I.; et al. SOT1, a pentatricopeptide repeat protein with a small MutS-related domain, is required for correct processing of plastid 23S–4.5S rRNA precursors in *Arabidopsis thaliana*. *Plant J.* **2016**, *85*, 607–621. [[CrossRef](#)]
113. Zoschke, R.; Watkins, K.P.; Miranda, R.G.; Barkan, A. The PPR-SMR protein PPR 53 enhances the stability and translation of specific chloroplast RNAs in maize. *Plant J.* **2016**, *85*, 594–606. [[CrossRef](#)] [[PubMed](#)]
114. Zhou, W.; Lu, Q.; Li, Q.; Wang, L.; Ding, S.; Zhang, A.; Wen, X.; Zhang, L.; Lu, C. PPR-SMR protein SOT1 has RNA endonuclease activity. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E1554–E1563. [[CrossRef](#)] [[PubMed](#)]
115. Zhang, Y.; Lu, C. The enigmatic roles of PPR-SMR proteins in plants. *Adv. Sci.* **2019**, *6*, 1900361. [[CrossRef](#)] [[PubMed](#)]
116. Sugita, M.; Ichinose, M.; Ide, M.; Sugita, C. Architecture of the PPR gene family in the moss *Physcomitrella patens*. *RNA Biol.* **2013**, *10*, 1439–1445. [[CrossRef](#)]
117. Wang, D.; Liu, H.; Zhai, G.; Wang, L.; Shao, J.; Tao, Y. *OspTAC2* encodes a pentatricopeptide repeat protein and regulates rice chloroplast development. *J. Genet. Genom.* **2016**, *43*, 601–608. [[CrossRef](#)]

118. Pyo, Y.J.; Kwon, K.C.; Kim, A.; Cho, M.H. Seedling Lethal1, a pentatricopeptide repeat protein lacking an E/E+ or DYW domain in Arabidopsis, is involved in plastid gene expression and early chloroplast development. *Plant Physiol.* **2013**, *163*, 1844–1858. [CrossRef]
119. Chen, L.; Huang, L.; Dai, L.; Gao, Y.; Zou, W.; Lu, X.; Wang, C.; Zhang, G.; Ren, D.; Hu, J.; et al. PALE-GREEN LEAF12 encodes a novel pentatricopeptide repeat protein required for chloroplast development and 16S rRNA processing in rice. *Plant Cell Physiol.* **2019**, *60*, 587–598. [CrossRef]
120. Knoop, V.; Rüdinger, M. DYW-type PPR proteins in a heterolobosean protist: Plant RNA editing factors involved in an ancient horizontal gene transfer? *FEBS Lett.* **2010**, *584*, 4287–4291. [CrossRef]
121. Takenaka, M.; Zehrmann, A.; Verbitskiy, D.; Härtel, B.; Brennicke, A. RNA editing in plants and its evolution. *Annu. Rev. Genet.* **2013**, *47*, 335–352. [CrossRef]
122. Ichinose, M.; Sugita, M. RNA editing and its molecular mechanism in plant organelles. *Genes* **2017**, *8*, 5. [CrossRef]
123. Okuda, K.; Nakamura, T.; Sugita, M.; Shimizu, T.; Shikanai, T. A pentatricopeptide repeat protein is a site-recognition factor in the plastid RNA editing. *J. Biol. Chem.* **2006**, *281*, 37661–37667. [CrossRef] [PubMed]
124. Small, I.D.; Schallenberg-Rüdinger, M.; Takenaka, M.; Mireau, H.; Osteretzer-Biran, O. Plant organellar RNA editing: What 30 years of research has revealed. *Plant J.* **2020**, *101*, 1040–1056. [CrossRef] [PubMed]
125. Plant RNA Editing-Prediction & Analysis Computer Tool (PREPACT). Available online: <http://www.prepact.de/prepact-main.php> (accessed on 1 July 2022).
126. Miyata, Y.; Sugiura, C.; Kobayashi, Y.; Hagiwara, M.; Sugita, M. Chloroplast ribosomal S14 protein transcript is edited to create a translation initiation codon in the moss *Physcomitrella patens*. *Biochim. Biophys. Acta* **2002**, *1576*, 346–349. [CrossRef]
127. Rüdinger, M.; Funk, H.T.; Rensing, S.A.; Maier, U.G.; Knoop, V. RNA editing: Only eleven sites are present in the *Physcomitrella patens* mitochondrial transcriptome and a universal nomenclature proposal. *Mol. Genet. Genom.* **2009**, *281*, 473–481. [CrossRef]
128. Hayes, M.L.; Giang, K.; Mulligan, R.M. Molecular evolution of pentatricopeptide repeat genes reveals truncation in species lacking an editing target and structural domains under distinct selective pressures. *BMC Evol. Biol.* **2012**, *12*, 66. [CrossRef] [PubMed]
129. Miyata, Y.; Sugita, M. Tissue- and stage-specific RNA editing of *rps14* transcripts in moss (*Physcomitrella patens*) chloroplasts. *J. Plant Physiol.* **2004**, *161*, 113–115. [CrossRef]
130. Salone, V.; Rüdinger, M.; Polsakiewicz, M.; Hoffmann, B.; Groth-Malonek, M.; Szurek, B.; Small, I.; Knoop, V.; Lurin, C. A hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Lett.* **2007**, *581*, 4132–4138. [CrossRef]
131. Rüdinger, M.; Polsakiewicz, M.; Knoop, V. Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat proteins in jungermanniid but not in marchantiid liverworts. *Mol. Biol. Evol.* **2008**, *25*, 1405–1414. [CrossRef] [PubMed]
132. Oldenkott, B.; Yang, Y.; Lesch, E.; Knoop, V.; Schallenberg-Rüdinger, M. Plant-type pentatricopeptide repeat proteins with a DYW domain drive C-to-U RNA editing in *Escherichia coli*. *Commun. Biol.* **2019**, *2*, 85. [CrossRef]
133. Hayes, M.L.; Santibanez, P.I. A plant pentatricopeptide repeat protein with a DYW-deaminase domain is sufficient for catalyzing C-to-U RNA editing in vitro. *J. Biol. Chem.* **2020**, *295*, 3497–3505. [CrossRef]
134. Takenaka, M.; Takenaka, S.; Barthel, T.; Frink, B.; Haag, S.; Verbitskiy, D.; Oldenkott, B.; Schallenberg-Rüdinger, M.; Feiler, C.G.; Weiss, M.S.; et al. DYW domain structures imply an unusual regulation principle in plant organellar RNA editing catalysis. *Nat. Catal.* **2021**, *4*, 510–522. [CrossRef] [PubMed]
135. Schuster, G.; Gruissem, W. Chloroplast mRNA 3' end processing requires a nuclear-encoded RNA-binding protein. *EMBO J.* **1991**, *10*, 493–502. [CrossRef] [PubMed]
136. Nakamura, T.; Ohta, M.; Sugiura, M.; Sugita, M. Chloroplast ribonucleoproteins function as a stabilizing factor of ribosome free mRNAs in the stroma. *J. Biol. Chem.* **2001**, *276*, 147–152. [CrossRef] [PubMed]
137. Kupsch, C.; Ruwe, H.; Gusewski, S.; Tillich, M.; Small, I.; Schmitz-Linneweber, C. Arabidopsis chloroplast RNA binding proteins CP31A and CP29A associate with large transcript pools and confer cold stress tolerance by influencing multiple chloroplast RNA processing steps. *Plant Cell* **2012**, *24*, 4266–4280. [CrossRef] [PubMed]
138. Teubner, M.; Fuß, J.; Kühn, K.; Krause, K.; Schmitz-Linneweber, C. The RNA recognition motif protein CP33A is a global ligand of chloroplast mRNAs and is essential for plastid biogenesis and plant development. *Plant J.* **2017**, *89*, 472–485. [CrossRef] [PubMed]
139. Hirose, T.; Sugiura, M. Involvement of a site-specific trans-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: Development of a chloroplast in vitro RNA editing system. *EMBO J.* **2001**, *20*, 1144–1152. [CrossRef]
140. Tillich, M.; Hardel, S.L.; Kupsch, C.; Armbruster, U.; Delannoy, E.; Gualberto, M.; Lehwark, P.; Leister, D.; Small, I.D.; Schmitz-Linneweber, C. Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6002–6007. [CrossRef] [PubMed]
141. Mueller, S.J.; Lang, D.; Hoernstein, S.N.W.; Lang, E.G.E.; Schuessele, C.; Schmidt, A.; Fluck, M.; Leisibach, D.; Niegl, C.; Zimmer, A.D.; et al. Quantitative analysis of the mitochondrial and plastid proteomes of the moss *Physcomitrella patens* reveals protein macrocompartmentation and microcompartmentation. *Plant Physiol.* **2014**, *164*, 2081–2095. [CrossRef]
142. Uchiyama, H.; Ichinose, M.; Sugita, M. Chloroplast ribonucleoprotein-like proteins of the moss *Physcomitrella patens* are not involved in RNA stability and RNA editing. *Photosynthetica* **2018**, *56*, 62–66. [CrossRef]
143. Sun, T.; Germain, A.; Giloteaux, L.; Hammani, K.; Barkan, A.; Hanson, M.R.; Bentolila, S. An RNA recognition motif-containing protein is required for plastid RNA editing in *Arabidopsis* and maize. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E1169–E1178. [CrossRef]

144. Takenaka, M.; Zehrmann, A.; Verbitskiy, D.; Kugelman, M.; Härtel, B.; Brennicke, A. Multiple organellar RNA editing factor (MORF) family proteins are required for RNA editing in mitochondria and plastids of plants. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5104–5109. [[CrossRef](#)] [[PubMed](#)]
145. Bentolila, S.; Heller, W.P.; Sun, T.; Babina, A.M.; Friso, G.; van Wijk, K.J.; Hanson, M. RIP1, a member of an *Arabidopsis* protein family, interacts with the protein RARE1 and broadly affects RNA editing. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E1453–E1461. [[CrossRef](#)] [[PubMed](#)]
146. Zehrmann, A.; Härtel, B.; Glass, F.; Bayer-Császár, E.; Obata, T.; Meyer, E.; Brennicke, A.; Takenaka, M. Selective homo- and heteromer interactions between the multiple organellar RNA editing factor (MORF) proteins in *Arabidopsis thaliana*. *J. Biol. Chem.* **2015**, *290*, 6445–6456. [[CrossRef](#)] [[PubMed](#)]
147. Zhang, F.; Tang, W.; Hedtke, B.; Zhong, L.; Liu, L.; Peng, L.; Lu, C.; Grimm, B.; Lin, R. Tetrapyrrole biosynthetic enzyme protoporphyrinogen IX oxidase 1 is required for plastid RNA editing. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 2023–2028. [[CrossRef](#)] [[PubMed](#)]
148. Sun, T.; Shi, X.; Friso, G.; van Wijk, K.; Bentolila, S.; Hanson, M.R. A zinc finger motif-containing protein is essential for chloroplast RNA editing. *PLoS Genet.* **2015**, *11*, e1005028. [[CrossRef](#)]
149. Guillaumot, D.; Lopez-Obando, M.; Baudry, K.; Avon, A.; Rigai, G.; Falcon De Longevialle, A.; Broche, B.; Takenaka, M.; Berthomé, R.; De Jaeger, G.; et al. Two interacting PPR proteins are major *Arabidopsis* editing factors in plastid and mitochondria. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8877–8882. [[CrossRef](#)]
150. Andreés-Colaás, N.; Zhu, Q.; Takenaka, M.; De Rybel, B.; Weijers, D.; Van Der Straeten, D. Multiple PPR protein interactions are involved in the RNA editing system in *Arabidopsis* mitochondria and plastids. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8883–8888. [[CrossRef](#)]
151. García-Andrade, J.; Ramiírez, V.; Loópez, A.; Vera, P. Mediated plastid RNA editing in plant immunity. *PLoS Pathog.* **2013**, *9*, e1003713. [[CrossRef](#)]
152. Huang, C.; Yu, Q.-B.; Li, Z.-R.; Ye, L.-S.; Xu, L.; Yang, Z.-N. Porphobilinogen deaminase HEMC interacts with the PPR-protein AtECB2 for chloroplast RNA editing. *Plant J.* **2017**, *92*, 546–556. [[CrossRef](#)]
153. Gutmann, B.; Royan, S.; Small, I. Protein complexes implicated in RNA editing in plant organelles. *Mol. Plant* **2017**, *10*, 1255–1257. [[CrossRef](#)]
154. Okuda, K.; Chateigner-Boutin, A.L.; Nakamura, T.; Delannoy, E.; Sugita, M.; Myouga, F.; Motohashi, R.; Shinozaki, K.; Small, I.; Shikanai, T. Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in *Arabidopsis* chloroplasts. *Plant Cell* **2009**, *21*, 146–156. [[CrossRef](#)] [[PubMed](#)]
155. Nakamura, T.; Sugita, M. A conserved DYW domain of the pentatricopeptide repeat protein possesses a novel endoribonuclease activity. *FEBS Lett.* **2008**, *582*, 4163–4168. [[CrossRef](#)] [[PubMed](#)]
156. Waltz, F.; Nguyen, T.T.; Arrivé, M.; Bochler, A.; Chicher, J.; Hammann, P.; Kuhn, L.; Quadrado, M.; Mireau, H.; Hashem, Y.; et al. Small is big in *Arabidopsis* mitochondrial ribosome. *Nat. Plants* **2019**, *5*, 106–117. [[CrossRef](#)] [[PubMed](#)]
157. Honkanen, S.; Small, I. The GENOME UNCOUPLED1 protein has an ancient, highly conserved role but not in retrograde signaling. *New Phytol.* **2022**, *in press*. [[CrossRef](#)] [[PubMed](#)]
158. Duckett, J.G.; Renzaglia, K.S. Ultrastructure and development of plastids in bryophytes. *Adv. Bryol.* **1988**, *3*, 33–93.
159. Sugita, M.; Aoki, S. Chloroplasts. In *The Moss Physcomitrella patens. Annual Plant Reviews*; Knight, C.D., Perroud, P.-F., Cove, D.J., Eds.; Wiley-Blackwell: Hoboken, NJ, USA, 2009; Chapter 8; Volume 36, pp. 182–210. ISBN 978-1-4051-8189-1.