Two interacting PPR proteins are major Arabidopsis editing factors in plastid and mitochondria

Damien Guillaumot[a,b,1], Mauricio Lopez-Obando[a,b,1,2], Kevin Baudry[a,b], Alexandra Avon[a,b], Guillem Rigaill[a,b,c], Andéol Falcon de Longevialle[a,b,3], Benjamin Broche[a,b], Mizuki Takenaka[d], Richard Berthomé[a,b,4], Geert De Jaeger[e,f], Etienne Delannoy[a,b], and Claire Lurin[a,b,5]

[a]Institute of Plant Sciences Paris-Saclay (IPS2), CNRS, Institut National de la Recherche Agronomique (INRA), Université Paris-Sud, Université Evry, Université Paris-Saclay, 91405 Orsay, France; [b]Institute of Plant Sciences Paris-Saclay (IPS2), Paris Diderot, Sorbonne Paris-Cité, 91405 Orsay, France; [c]Laboratoire de Mathématiques et Modélisation (LaMME), Université Evry, 91000 Evry, France; [d]Molekulare Botanik, Universität Ulm, 89069 Ulm, Germany; [e]Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium; and [f]VIB Center for Plant Systems Biology, 9052 Ghent, Belgium

Edited by Robert Haselkorn, University of Chicago, Chicago, IL, and approved July 3, 2017 (received for review April 7, 2017)

RNA editing is converting hundreds of cytosines into uridines during organelle gene expression of land plants. The pentatricopeptide repeat (PPR) proteins are at the core of this posttranscriptional RNA modification. Even if a PPR protein defines the editing site, a DYW domain of the same or another PPR protein is believed to catalyze the deamination. To give insight into the organelle RNA editosome, we performed tandem affinity purification of the plastidial CHLOROPLAST BIOGENESIS 19 (CLB19) PPR editing factor. Two PPR proteins, dually targeted to mitochondria and chloroplasts, were identified as potential partners of CLB19. These two proteins, a P-type PPR and a member of a small PPR-DYW subfamily, were shown to interact in yeast. Insertional mutations resulted in embryo lethality that could be rescued by embryo-specific complementation. A transcriptome analysis of these complemented plants showed major editing defects in both organelles with a very high PPR type specificity, indicating that the two proteins are core members of E+-type PPR editosomes.

RNA editing | organelles | pentatricopeptide repeat

In vascular plant organelle RNAs, hundreds of specific cytidines are converted into uridines by the so-called RNA editing mechanism (C to U editing). This phenomenon remained very enigmatic for a long time, raising numerous questions about its purpose, its evolution, and the molecular mechanism behind its very high specificity. Even if editing finality is still a matter of debate, many components of plant editosomes and the molecular elements required for editing specificity have been described (1, 2).

The editable cytidine is identified by a pentatricopeptide repeat (PPR) protein through the recognition of 20–25 bases upstream of the cytidine (1). However, the 5′ cis-elements, defining RNA editing sites, are not conserved between sites. Each editing site is targeted by a specific PPR protein. For example, in *Arabidopsis thaliana*, a total of 56 PPR proteins were shown so far to be each required for the editing of one to eight specific sites (Table S1). The PPR domain is a degenerated polypeptide showing a conserved structural conformation able to bind RNA molecules when it is repeated in tandem (3–5). A code for RNA recognition by PPR proteins was proposed (6–10). In this code, the nucleotide recognition is achieved by the combination of three amino acids of each PPR motif.

The nature of the PPR domains within proteins is used to divide the PPR family into two subfamilies, the PPR-P and the PPR-PLS. This last subfamily is subdivided in subgroups according to their E1, E2, E+, and DYW C-terminal domains (11, 12). Most members of the P-type PPR subfamily have been implicated in RNA metabolism such as 5′ or 3′ transcript stabilization and processing, splicing, and translation (5), whereas most editing PPR proteins belong to the PLS subfamily (1). Although a function in selecting editing sites is well defined for their PPR domains, the functions of the E1, E2, E+, and DYW domains remain unclear and controversial. Molecular and phylogenetic evidences suggest that the DYW domain is required for the editing activity (13, 14). Despite the lack of definitive biochemical evidence, it has been hypothesized that it could contain the RNA editing enzymatic activity required for the deamination of cytidines into uridines (13–16). However, some editing PPRs do not carry any DYW domain and end with either an E1, E2, or E+ domain (1). Moreover, the DYW domain could be deleted in some PPR-DYW proteins without affecting their function in editing (17). To reconcile the different pieces of evidence, it has been proposed that the cytidine deaminase activity could be provided either in cis by a PPR-DYW specificity factor or in trans when a PPR-E factor is required for the site recognition. This was shown, for example, for the editing of the chloroplastic *ndhD*-1 site, where the target site is recognized by CRR4, a PPR-E specificity factor, whereas a DYW domain is provided by DYW1, a small protein containing only a DYW domain (18).

Besides PPR proteins, numerous additional proteins were shown to be required for the same editing events, suggesting the existence of high molecular mass editosome protein complexes (2). In particular, three classes of essential non-PPR components of the editosomes were shown to be involved in C to U RNA editing. These proteins are members of small families and are suspected to have partially redundant functions as general factors involved

Significance

Breaking the "central dogma" of molecular biology, RNA editing is a specific posttranscriptional modification of RNA sequences. In seed plant organelle editosomes, each editable cytidine is identified by a specific pentatricopeptide repeat (PPR) protein. Some of these sequence-specific proteins contain an additional C-terminal "DYW" domain, which is supposed to carry the catalytic activity for editing. However, many PPR editing factors lack this domain. In this article, we show that a subfamily of about 60 Arabidopsis proteins might all require two additional PPR proteins for the editing of their sites. One of them, DYW2, is a specific cofactor containing a DYW domain, supporting the hypothesis that this domain might bring the cytidine deaminase activity to these editosomes.

Author contributions: D.G., M.L.-O., A.F.d.L., R.B., E.D., and C.L. designed research; D.G., M.L.-O., K.B., A.A., A.F.d.L., B.B., M.T., G.D.J., and E.D. performed research; D.G., M.L.-O., G.R., M.T., G.D.J., E.D., and C.L. analyzed data; and M.L.-O., E.D., and C.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE100298).

[1]D.G. and M.L.-O. contributed equally to this work.

[2]Present address: Department of Plant Biology, Swedish University of Agricultural Sciences, The Linnean Centre of Plant Biology in Uppsala, SE-75007 Uppsala, Sweden.

[3]Present address: Laboratoire Plant Process Innovation, GENOPOLE, 91000 Evry, France.

[4]Present address: Laboratoire des Interactions Plantes-Microorganismes (LIPM), Université de Toulouse, INRA, CNRS, 31326 Castanet-Tolosan, France.

[5]To whom correspondence should be addressed. Email: claire.lurin@inra.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1705780114/-/DCSupplemental.

Table 1. Proteins purified by TAP using CLB19 as bait in *Arabidopsis* cell suspension culture PSB-L

AGI*	Name†	Prot. mass, KDa	Loc.‡	No. identified in four TAPs§
AT1G05750	CLB19	56.4	C (26)	4
AT2G15690	DYW2	66.3	M/C (42)	4
AT3G49240	NUWA	71.7	M/C (34)	4

**Arabidopsis* genome initiative annotation identifier in TAIR database version 9.
†CLB19, ChLoroplast Biogenesis 19.
‡Loc., subcellular localization of proteins.
§See Dataset S1 for mass spectrometry analysis details.

in the editing of organelle transcripts (2). In *Arabidopsis*, nine Multiple Organellar RNA editing Factors (MORF/RIPs) were described as required for many editing sites of plant organelles (19–21). Members of the ORRM family and the CP31 protein, containing RNA Recognition Motifs (RRMs), were also found to influence RNA editing in plant organelles (22, 23). More recently, OZ proteins were found to copurify with components of the editosomes and also be required for organellar editing (24).

Although extensive studies of plant editosomes have already identified many factors, further studies are needed to discover new components as well as their relations in the protein network. Here, we implemented a tandem affinity purification (TAP) approach to gain insight into the composition of a chloroplast editing complex. We use the known chloroplast editing factor CHLOROPLAST BIOGENESIS 19 (CLB19) required for *rpoA* and *clpP* editing (25) as bait for purification. Two unknown PPR proteins, dually targeted to mitochondria and chloroplasts and required for *Arabidopsis* embryo development, were identified in the CLB19 editing complex. A transcriptome analysis of the mutants showed major editing defects in both organelles with a very high PPR-type specificity indicating that the two proteins are core members of E+-type PPR editosomes.

Results

Exploring the CLB19 Chloroplast Editing Complex. To improve our knowledge of the in vivo composition of an RNA editing complex of land plant chloroplasts, a TAP approach was performed using the previously characterized chloroplast editing factor CLB19 as bait. CLB19 was fused to a G protein and a streptavidin-binding peptide (GS) tag at its C terminus (26, 27) and was expressed under the control of the cauliflower mosaic virus (CaMV) 35S promoter (Fig. S1*A*). In *clb19-1* mutant plants, the CLB19-TAP tag protein was able to complement the macroscopic phenotype of the mutant (Fig. S1 *B* and *C*), indicating that the fusion protein was functionally similar to the wild-type protein. After production in *Arabidopsis* cell suspension culture PSB-L, two proteins were copurified with CLB19-GS in four independent experiments (Table 1 and Dataset S1). The fourth sample was subjected to RNase treatment before purification without any modification of the proteins identified in the complex (Dataset S1). Both identified proteins, AT3G49240 and AT2G15690, are members of the PPR family. According to the PPR classification, AT3G49240 belongs to the P-type PPR subfamily. This protein was recently shown to be encoded by a maternal imprinted gene named *NUWA* (28). AT2G15690 belongs to the PPR-DYW subfamily and was named DYW2 (detailed in *DYW2 and NUWA Are Two Distant PPR Proteins*).

To identify proteins interacting with CLB19, NUWA, and DYW2, we screened these three PPR proteins against a library of more than 12,000 *Arabidopsis* proteins using an improved high-throughput binary interactome mapping pipeline based on yeast two-hybrid (Y2H). Among the interactions involving DYW2, we identified a direct link between NUWA and DYW2 (Fig. S2). In contrast, we did not identify any interactor of the CLB19 protein in this screen.

DYW2 and NUWA Are Two Distant PPR Proteins. DYW2 is an atypical PPR-DYW protein containing five predicted PPR domains and a C-terminal DYW domain separated by an amino acid sequence that do not clearly correspond to an E domain (11) (Fig. 1*A*). This unusual architecture of a PPR protein carrying a DYW domain without any regular E1 and E2 domains is shared by only five other proteins in the *A. thaliana* genome, among which is the DYW1 chloroplast editing factor (18). The other members of this small subfamily (Fig. S3), called here after the DYW1-like subfamily, are two mitochondrial editing factors, MEF8 and MEF8S (29), and two uncharacterized proteins, AT2G34370 (DYW3) and AT1G29710 (DYW4).

In silico prediction of PPR domains using the PPR Gene Database (11) and TPRpred (30) websites showed that NUWA harbors up to 12 PPR domains covering most of its amino acid sequence. As reported previously (28), a coiled-coil domain is predicted in the *N*-terminal region of the NUWA protein, whereas a 106-amino acid sequence without any conserved domain is present upstream of the PPR motifs (Fig. 1*B*). Its closest homolog in the *Arabidopsis* genome is GRP23, a nuclear and mitochondrial PPR (31, 32). GRP23 shares 34% amino acid identity with NUWA but does not carry any coiled coil domain (31).

NUWA and DYW2 Are Dually Targeted to Chloroplast and Mitochondria. Several independent experiments previously showed that *A. thaliana* NUWA and DYW2 proteins, or some of their orthologs, are dually targeted to mitochondria and plastids. Whereas NUWA was recently published as localized in mitochondria (28), AT3G49240 (NUWA) was identified in several proteomic data either in *A. thaliana* mitochondria (33, 34) or chloroplastic samples (35, 36). In accordance with these results, we and Andrés-Colás et al. (32, 37) observed NUWA presequence and full-length fusions to fluorescent proteins in mitochondria and chloroplast (Fig. S4). The two maize orthologs of DYW2 (GRMZM2G073551 and GRMZM2G017821) were identified in plastid nucleoids (38), whereas its rice ortholog was described in mitochondria (39) samples. In *A. thaliana*, AT2G15690 (DYW2) was observed in both mitochondria and plastids when fused to a GFP protein (Fig. S4) (37, 40). These dual subcellular localizations were further confirmed by reverse

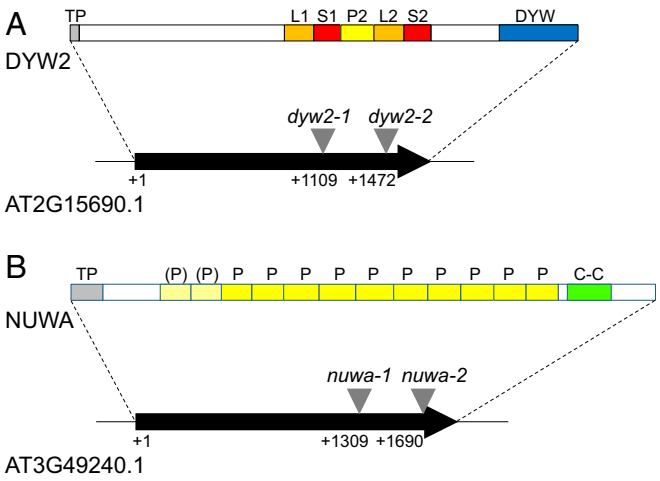

Fig. 1. DYW2 and NUWA are members of the PPR family. (*A*) DYW2. (*B*) NUWA. Schematic structures of the loci and the proteins. Predicted targeting peptide (TP, gray box), S (red), L (orange), P (yellow), DYW (blue), and coiled-coil region (C-C, green) domains are labeled on the protein sequence. The targeting peptide was predicted using the TargetP software at www.cbs.dtu.dk/services/TargetP/ for plant organisms using no cutoffs. S, L, P, and DYW domains were located according to the PPR Gene Database (11) and using the TPRpred software at https://toolkit.tuebingen.mpg.de/#/tools/tprpred [(P), light yellow]. Sequence-verified locations of the T-DNA insertions used in this study are indicated (+1 is the transcription start).

genetic analyses and the identification of molecular phenotypes in both mitochondria and plastids of *dyw2* and *nuwa* mutants (see *DYW2 and NUWA Proteins Are Functionally Linked and Involved in Editing of Chloroplast and Mitochondria Transcripts*).

NUWA and DYW2 Proteins Are Required for Embryo Development. A reverse genetics approach was used to identify the molecular functions of DYW2 and NUWA proteins. Two T-DNA insertions in each gene were selected from the T-DNA Express database (41) and were named *dyw2-1* (GK_332A07), *dyw2-2* (FLAG_435F11), *nuwa-1* (SALK_069042), and *nuwa-2* (SAIL_784_A11). The position of each T-DNA was confirmed by sequencing and is indicated in Fig. 1. *nuwa-1* and *nuwa-2* were previously characterized as two embryo-defective alleles of the *EMB1796* locus during the seedgenes project (42). Similarly, another T-DNA insertion mutant, *nuwa*, was recently shown to be affected in early embryogenesis and endosperm development (28). Accordingly, no homozygous seedling for *nuwa-1* and *nuwa-2,* but also *dyw2-1* and *dyw2-2*, insertions was found from large screens of heterozygous plant progenies, whereas aborted embryos were observed when opening siliques of heterozygous plants (Fig. 2*A*). Pollen viability was assayed by Alexander staining of mature anthers from heterozygous plants. Results indicated that all pollen grains were viable in both heterozygous mutants and that mutations did not affect male gametophyte viability (Fig. 2*A*). Finally, genetic complementation assays between, on one hand, *dyw2-1* and *dyw2-2* lines and, on the other hand, *nuwa-1* and *nuwa-2* lines confirmed that *dyw2-1* and *dyw2-2* and *nuwa-1* and *nuwa-2* are allelic mutations responsible for the observed embryo lethal phenotype (Table S2).

Complementation of *dyw2* and *nuwa* in Embryos and Seeds. To bypass the embryo lethality of the mutants, we complemented them by expressing NUWA and DYW2 wild-type proteins under the control of the embryo-specific *ABI3* promoter (43). After seedling development, the *ABI3* promoter is expected to be no longer active, leading to its absence of expression in seedlings and at the adult stage. This strategy allowed the development of homozygous *dyw2* and *nuwa* mutant embryos in siliques of heterozygous plants and the germination of homozygous seedlings in their progeny (Fig. 2*B*). The absence of expression of *NUWA* and *DYW2* transcripts in adult plants was confirmed by qRT-PCR (Fig. 2*C*) and subsequently when analyzing RNA-sequencing (RNA-seq) data (see dyw2-1 *and* nuwa-2 *Complete Transcriptome Analysis*). Whereas the *nuwa* mutants were almost indistinguishable from WT, the *dyw2* mutants were small pale green plants producing sterile flowers (Fig. 2 *B* and *D*).

DYW2 and NUWA Proteins Are Functionally Linked and Involved in Editing of Chloroplast and Mitochondria Transcripts. As DYW2 and NUWA are PPR proteins interacting with the editing PPR protein CLB19, we tested their involvement in editing of organelle RNA by total RNA-seq analysis of the rescued *dyw2-1* and *nuwa-2* mutants. The organellar editing quantification identified 392 and 223 differentially edited sites in *dyw2-1* and *nuwa-2*, respectively (Fig. 3 *A* and *B* and Dataset S2 A and B). The differentially edited sites were either mitochondrial or chloroplastic and included previously unidentified editing sites (109 in *dyw2-1* and 44 in *nuwa-2*). Surprisingly, one site (position 20299 of the plastid genome) was edited only in *dyw2-1*. Targeted Sanger sequencing of

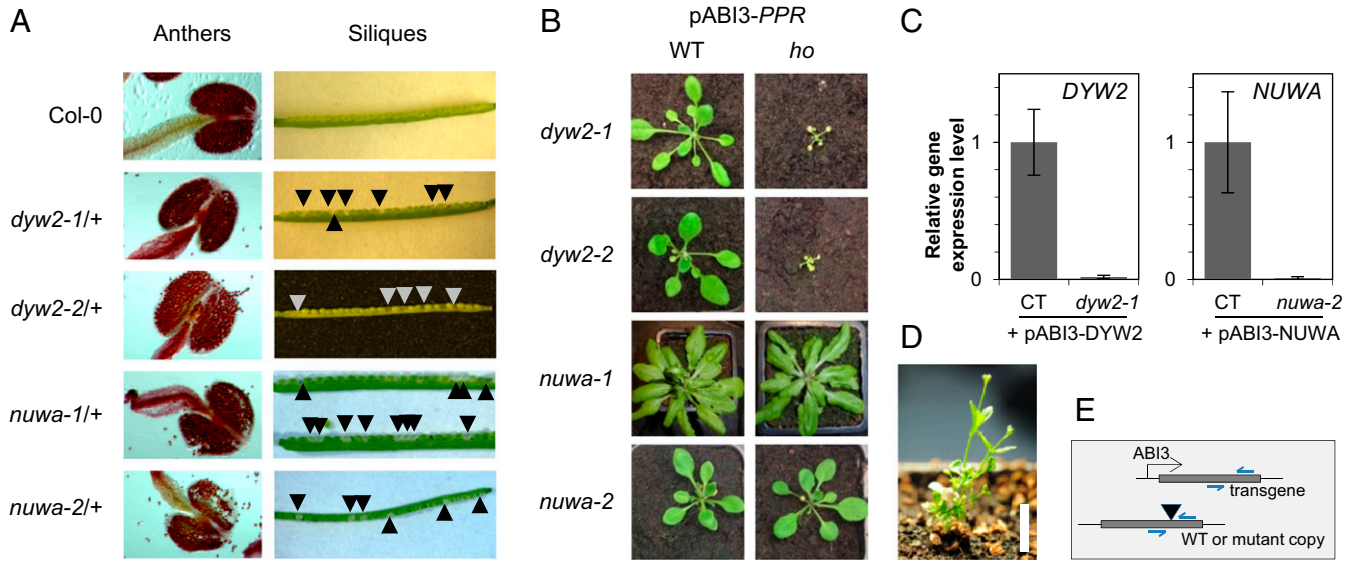

Fig. 2. Characterization of *dyw2* and *nuwa* mutants. (*A*) Phenotype of pollen and embryo observed on heterozygous *dyw2* and *nuwa* plants. (*Left*) The viability of pollen was assayed by Alexander staining of wild-type and heterozygous plant anthers. (*Right*) Siliques resulting from wild-type or heterozygote self-pollination were opened 10 d after pollination. Abnormal seeds (arrows) accounted for ∼25% of the total observed. (*B*) Macroscopic phenotype of 1-mo-old homozygous mutant plants obtained after complementation with an embryo-specific construct. Heterozygous plants for *dyw2-1* and *dyw2-2* mutations were transformed with full-length DYW2 under the control of the embryo-specific p*ABI3* promoter, whereas heterozygous plants for *nuwa-1* and *nuwa-2* mutations were transformed with full-length NUWA under the control of the p*ABI3* promoter. Progeny seedlings of heterozygous T1 plants were genotyped to identify wild-type (*Left*), heterozygous (not shown), and homozygous (*Right*) sibling plants carrying the p*ABI3* embryo-specific construct. After germination on MS + hygromycin media, 10-d-old seedlings were transferred onto soil in a growth chamber with long-day conditions. (*C*) *DYW2* and *NUWA* gene expression in homozygous mutant plants obtained after complementation with an embryo-specific construct. Gene expression of 1-mo-old plants grown in long-day conditions was measured by qRT-PCR using total RNA extracted from leaves of four biological replicates of *dyw2-1* and *nuwa-2* homozygous plants expressing the corresponding p*ABI3* construct and of control sibling plants. For each biological replicate, the mean expression level of three technical qRT-PCR replicates was normalized with the mean of actin2-8 expression, used as reference gene. Controls (CTs) refer to siblings of homozygous mutant plants coming from the same self-progeny, wild type or heterozygous for the mutation, and carrying the corresponding p*ABI3* transgene. (*D*) Adult phenotype of *dyw2* mutant expressing pABI3-DYW2. A homozygous *dyw2-2* adult plant was observed after 7 wk of culture in soil in greenhouse with short day condition (white bar, 1 cm). Flower buds were produced but did not further develop in flowers and siliques. (*E*) Schematic representation of the primers used for qRT-PCR in Fig. 2*C*. Primers are surrounding positions of T-DNA insertion. Thus, in homozygous plants expressing pABI3 construct, the expression level reflects the expression of the pABI3 construct, whereas, in CT plants, the expression level corresponds to the expression of both endogenous and transgenic genes.

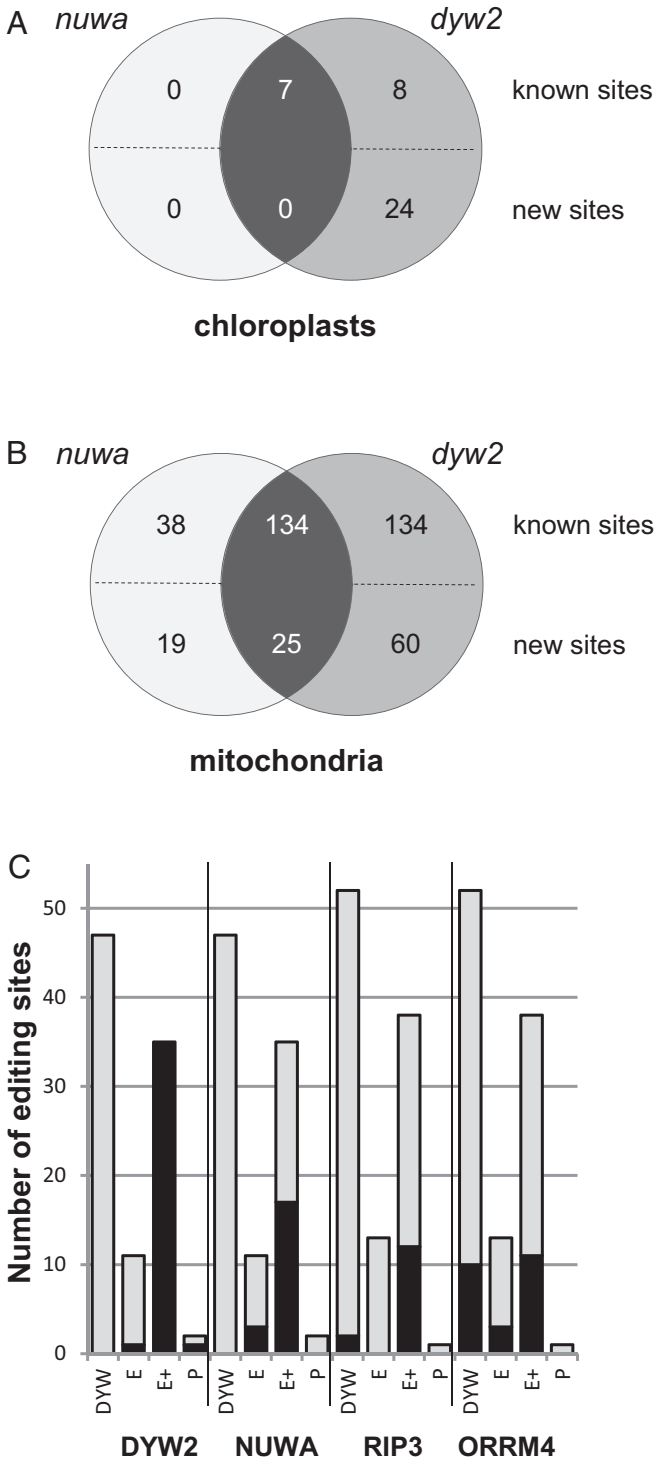

Fig. 3. Editing activity of DYW2 and NUWA. The detailed results are provided in Dataset S2 A–D and Dataset S3. (*A*) Plastidial editing sites affected in *dyw2* and *nuwa*. (*B*) Mitochondrial editing sites affected in *dyw2* and *nuwa*. Venn diagrams summarizing the number of differentially edited sites in *dyw2-1* and *nuwa-2*. The known sites correspond to the sites identified in Bentolila et al. (44), Sun et al. (24), Shi et al. (60), and Shi et al. (45). The new sites are the editing sites identified in this study. (*C*) DYW2 and NUWA-dependent sites are targeted by PPR-E+ proteins The PPR specificity of DYW2, NUWA, RIP3, and ORRM4 was estimated by counting the dependent (black) and independent (gray) editing sites associated with PPR of the various subfamilies [DYW, E, E+, and pure (P)]. An editing site is considered as depending on a particular protein if it is differentially edited between CT and mutant (P value < 5% after Bonferroni correction) and its editing extent is decreased by 10% or more in the mutant. It is independent otherwise. Editing sites associated with known PPRs are listed in Table S1 with their corresponding primary references. Values for RIP3 and ORRM4 were obtained by applying our statistical protocol to the raw data from Bentolila et al. (44) and Shi et al. (45), respectively. The total number of sites differs from one study to the other because of missing data from some editing sites in each study.

several RT-PCR products from the corresponding second allele (rescued *dyw2-2* and *nuwa-1*) mutants showed similar results (Dataset S3), indicating that editing defects were genetically linked to mutations in either *DYW2* or *NUWA* locus. Interestingly, *dyw2-1* and *nuwa-2* shared 166 differentially edited sites, a number significantly higher to what would be randomly expected ($P = 0$), and these common sites included the *clpP* and *rpoA* editing sites associated with CLB19. This result strongly suggests that CLB19, DYW2, and NUWA are editing partners supporting the physical interactions previously observed.

DYW2 and NUWA Are Required for Editing by PPR-E+. As DYW2 and NUWA are working together with CLB19 (this work) and SLO2 (37), two PPR-E+s, we explored their association with PPR-E+ in general. Using the same criteria as Bentolila et al. (44), we considered only the editing sites depending on DYW2 or NUWA. There was a strong positive bias for sites associated with PPR-E+ proteins. Out of 47 analyzed sites associated with PPR-DYW proteins, none are depending on DYW2 or NUWA. Conversely, the 35 known PPR-E+ sites analyzed in this study are depending on DYW2 and 17 of them are depending on NUWA (Fig. 3*C*). Applying the same statistical analysis for RIP3 (44) (Fig. 3*C* and Dataset S2C) and ORRM4 (45) (Fig. 3*C* and Dataset S2D), two editing factors controlling numerous chloroplastic and mitochondrial sites, showed no such specificity. These results and the work of Andrés-Colás et al. (37) support an extension of the function of DYW2 and NUWA to all PPR-E+ proteins.

***dyw2-1* and *nuwa-2* Complete Transcriptome Analysis.** The total RNA-seq approach allowed further complete and parallel quantitative analyses of nuclear, mitochondrial, and plastidial transcriptomes. As PPR proteins of the pure subfamily such as NUWA have rarely been involved in editing (46) and to confirm that DYW2 and NUWA function primarily in RNA editing, total RNA-seq data were used to quantify organelle transcript splicing, processing, and accumulation in *nuwa-2* and *dyw2-1*. The organelle transcriptome of *dyw2-1* was highly impacted with 182 differentially expressed genes out of 239 and 21 differentially spliced introns out of 37 (Dataset S2 E–G). Noteworthily, the plastid gene expression profile of *dyw2-1* was similar to the expression profile of *clb19* as described by Chateigner-Boutin et al. (25), suggesting that the perturbations in the *dyw2-1* transcriptome were the consequences of the numerous editing defects, especially in *rpoA*, which encodes a subunit of the plastid encoded RNA polymerase. On the other hand, the organelle transcriptome of *nuwa-2* showed limited perturbations with only 28 differentially expressed genes and 6 differentially spliced introns, including only one that was partially impaired (Dataset S2 E, F, and H). As most of these perturbations were also found in *dyw2-1* and no strong processing defect likely to explain the editing defects of *nuwa-2* was detected (Dataset S2I), these results strongly suggest that both DYW2 and NUWA are genuine editing factors.

The analysis of the nuclear transcripts confirmed that no functional RNA of *DYW2* or *NUWA* was detected in the corresponding mutants. Indeed, although reads were mapping to the genes, the mutants showed no read overlapping the T-DNA insertion sites as opposed to the controls, indicating that despite normal counts, no full-length RNA was produced in these mutants. The nuclear transcriptome analysis also showed that 12,485 genes were differentially expressed in *dyw2-1* versus only 1,097 in *nuwa-2*, in agreement with their macroscopic phenotype (Dataset S2J). Interestingly, the analysis of the nuclear transcriptome with MapMan

showed that the PPR gene family was significantly affected in both mutants ($P = 0$; Dataset S2K), suggesting that the organelle gene expression was impacted. In particular, *NUWA* was induced four times in *dyw2-1*, and *DYW2* was induced twice in *nuwa-2*. Although major, these nuclear transcriptome modifications very probably reflected secondary effects of severe organelle dysfunctions as expected in mitochondria and plastid mutants because of organelle-nuclear signaling (47).

Discussion

Different RNA editing complexes have been described in several organisms, some of them having high molecular mass quaternary structures such as the 20S editosome of *Trypanosoma brucei* (48), for example. In contrast, it was first suggested that plant editing complexes could simply constitute one or two (PPR) proteins (18, 49, 50), similarly to the initial model of C-to-U mammal editosome where a specificity factor, ACF, binds the RNA sequence and recruits APOBEC-1, the enzyme catalyzing the reaction (51). The composition of plant RNA editosomes recently appeared to be more convoluted and heterogeneous with numerous additional proteins whose functions are still poorly understood (2). In these complexes, one or two PPR proteins of the PLS subfamily are considered to be key factors providing both the specificity and probably the enzymatic activity. Here, we show that two PPR proteins, DYW2 and NUWA, are physically and functionally part of the E+ editosomes and required for the editing activity of probably all PPR-E+ proteins.

Our results predicted that at least three different PPR proteins are at the core of each E+ editosome: a PPR-E+ specific of the target site and two common PPR proteins, DYW2 and NUWA. Whereas, in such a complex, the function of the PPR-E+ is well known as the specificity factor binding the target RNA, the molecular functions of the two other PPR proteins remain unclear. Unexpectedly for PPR proteins, reverse genetics analyses indicated that they could have numerous potential binding sites without any sequence similarity. This suggests that unlike most PPR proteins, DYW2 and NUWA may not bind to RNA or bind RNA with low sequence specificity. As proposed in the companion paper from Andrés-Colás et al. (37) and supported by our results, an interesting hypothesis is that NUWA may bridge and stabilize the interaction between PPR-E+ and DYW2 proteins. As proposed for the DYW1 protein, which brings a DYW domain to the CRR4 protein (18), it is probable that DYW2 brings the cytidine deaminase activity to the E+ editosomes. Thus, the core of any plant organelle editosome would be organized with a PPR protein targeting the editing site and a DYW domain bringing the enzymatic activity. This domain is provided in cis by the PPR specificity factor when it belongs to the PPR-DYW subfamily or could be brought in trans by a member of the DYW1-like clade when the specificity factor is a PPR-PLS, PPR-E, or PPR-E+ protein.

In addition to the severe editing defects observed in E+ sites, we also showed that *dyw2* and *nuwa* mutants had numerous minor negative as well as positive defects in non-E+ editing sites. These results suggested that DYW2 and NUWA editosomes compete with other editosomes for unknown editing factors, supporting Sun et al.'s (2) recent review, which proposed that editosomes result from complex assembly equilibria of numerous editing factors. However, we tested the functional overlap of DYW2 and NUWA with RIP/MORFs, ORRMs, and OZ1 by comparing the lists of impacted editing sites, but none of these comparisons showed significant overlaps. Surprisingly, whereas the bait protein used in our approach, CLB19, is required for editing of two plastidial sites, which are also targets of general factors such as MORF2, MORF9, and ORRM1, we were not able to purify them together with CLB19, NUWA, and DYW2. One possible explanation of the absence of these proteins could be due to weak and/or transient interactions of these factors within the complex. Indeed, when screening the Y2H library with CLB19, NUWA, and DYW2, we were not able to identify any of these editing factors.

Besides these targeted questions regarding the molecular functions of NUWA and DYW2 PPR proteins, an intriguing observation of our study is the requirement of both of them during embryo development, whereas they are dispensable for further plant growth. Mutant studies showed that a large number of nuclear mutations impairing embryo development are associated with proteins targeted to organelles (42). Interestingly, most of these proteins are involved in the regulation of organelle gene expression, from editing to translation via splicing and processing on different RNA targets. The general consensus is that the embryo lethal phenotype of mitochondrial *emb* mutations is associated with a lack of energy production (52, 53), whereas the understanding of the essential role of plastid function in plant embryogenesis is still very limited. NUWA and DYW2 are required for numerous editing events in both mitochondria and plastids. The impact of these defects on embryo development is not surprising. However, further studies will be needed to understand why, in contrast, these editing defects are not lethal at the adult stage, especially in *NUWA* mutants that do not show any macroscopic phenotype.

Materials and Methods

Plant Material and Phenotype Characterization of T-DNA Insertion Lines. The T-DNA insertional mutants, GK_332A07 (*dwy2-1*), FLAG_435F11 (*dyw2-2*), SALK_069042 (*nuwa-1*), and SAIL_784_A11 (*nuwa-2*), were obtained from the ABRC stock center. *NUWA* and *DYW2* wild-type ORFs were cloned under the control of the embryo-specific p*ABI3* promoter using the pH7WG-ABI3 vector from Aryamanesh et al. (43). Detailed information on the plant methods used in this study is included in *SI Materials and Methods*. All sequences of primers used in this study are available in Table S3.

Protein Interaction Methods. The detailed methods of TAP and Y2H screening are given in *SI Materials and Methods*. In brief, constructions and *Arabidopsis* transformation were carried out as previously described (54), protein complex purification was done as described in Van Leene et al. (55), and peptide isolation and analysis were performed according to Van Leene et al. (56). Protein identification details are listed in Dataset S1.

RNA-Seq Analysis. The detailed methods are given in *SI Materials and Methods*. In brief, the RNA-seq analysis was performed following the recommendations of Rigaill et al. (57). The organelle transcriptome was studied after mapping of the reads with STAR (v020201) (58) using in-house scripts adapted from the ChloroSeq package (59).

ACKNOWLEDGMENTS. M.L.-O. and K.B.'s research was supported by French PhD fellowships from "Ministère de la Recherche et de l'Enseignement Supérieur." D.G.'s research was supported by a Labex Saclay Plant Sciences-SPS postdoctoral fellowship. M.T. is a Heisenberg fellow, and his work is supported by grants from the DFG. The IPS2 benefits from the support of the Labex Saclay Plant Sciences-SPS (Grant ANR-10-LABX-0040-SPS).

1. Ichinose M, Sugita M (2016) RNA editing and its molecular mechanism in plant organelles. *Genes (Basel)* 8:5.
2. Sun T, Bentolila S, Hanson MR (2016) The unexpected diversity of plant organelle RNA editosomes. *Trends Plant Sci* 21:962–973.
3. Small ID, Peeters N (2000) The PPR motif - A TPR-related motif prevalent in plant organellar proteins. *Trends Biochem Sci* 25:46–47.
4. Delannoy E, Stanley WA, Bond CS, Small ID (2007) Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. *Biochem Soc Trans* 35:1643–1647.
5. Barkan A, Small I (2014) Pentatricopeptide repeat proteins in plants. *Annu Rev Plant Biol* 65:415–442.
6. Ke J, et al. (2013) Structural basis for RNA recognition by a dimeric PPR-protein complex. *Nat Struct Mol Biol* 20:1377–1382.
7. Barkan A, et al. (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet* 8:e1002910.
8. Takenaka M, Zehrmann A, Brennicke A, Graichen K (2013) Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. *PLoS One* 8: e65343.
9. Yagi Y, Hayashi S, Kobayashi K, Hirayama T, Nakamura T (2013) Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PLoS One* 8:e57286.
10. Yin P, et al. (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* 504:168–171.
11. Cheng S, et al. (2016) Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. *Plant J* 85: 532–547.

12. Lurin C, et al. (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16:2089–2103.
13. Salone V, et al. (2007) A hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Lett* 581:4132–4138.
14. Boussardon C, et al. (2014) The cytidine deaminase signature HxE(x)n CxxC of DYW1 binds zinc and is necessary for RNA editing of ndhD-1. *New Phytol* 203:1090–1095.
15. Rüdinger M, Volkmar U, Lenz H, Groth-Malonek M, Knoop V (2012) Nuclear DYW-type PPR gene families diversify with increasing RNA editing frequencies in liverwort and moss mitochondria. *J Mol Evol* 74:37–51.
16. Wagoner JA, Sun T, Lin L, Hanson MR (2015) Cytidine deaminase motifs within the DYW domain of two pentatricopeptide repeat-containing proteins are required for site-specific chloroplast RNA editing. *J Biol Chem* 290:2957–2968.
17. Okuda K, et al. (2009) Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in Arabidopsis chloroplasts. *Plant Cell* 21:146–156.
18. Boussardon C, et al. (2012) Two interacting proteins are necessary for the editing of the NdhD-1 site in Arabidopsis plastids. *Plant Cell* 24:3684–3694.
19. Takenaka M, et al. (2012) Multiple organellar RNA editing factor (MORF) family proteins are required for RNA editing in mitochondria and plastids of plants. *Proc Natl Acad Sci USA* 109:5104–5109.
20. Bentolila S, et al. (2012) RIP1, a member of an Arabidopsis protein family, interacts with the protein RARE1 and broadly affects RNA editing. *Proc Natl Acad Sci USA* 109: E1453–E1461.
21. Bayer-Császár E, et al. (2017) The conserved domain in MORF proteins has distinct affinities to the PPR and E elements in PPR RNA editing factors. *Biochim Biophys Acta* 1860:813–828.
22. Kupsch C, et al. (2012) 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* 24:4266–4280.
23. Sun T, et al. (2013) An RNA recognition motif-containing protein is required for plastid RNA editing in Arabidopsis and maize. *Proc Natl Acad Sci USA* 110:E1169–E1178.
24. Sun T, et al. (2015) A zinc finger motif-containing protein is essential for chloroplast RNA editing. *PLoS Genet* 11:e1005028.
25. Chateigner-Boutin AL, et al. (2008) CLB19, a pentatricopeptide repeat protein required for editing of rpoA and clpP chloroplast transcripts. *Plant J* 56:590–602.
26. Bürckstümmer T, et al. (2006) An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nat Methods* 3:1013–1019.
27. Van Leene J, Witters E, Inzé D, De Jaeger G (2008) Boosting tandem affinity purification of plant protein complexes. *Trends Plant Sci* 13:517–520.
28. He S, et al. (2017) A novel imprinted gene NUWA controls mitochondrial function in early seed development in Arabidopsis. *PLoS Genet* 13:e1006553.
29. Verbitskiy D, Zehrmann A, Härtel B, Brennicke A, Takenaka M (2012) Two related RNA-editing proteins target the same sites in mitochondria of Arabidopsis thaliana. *J Biol Chem* 287:38064–38072.
30. Karpenahalli MR, Lupas AN, Söding J (2007) TPRpred: A tool for prediction of TPR-, PPR- and SEL1-like repeats from protein sequences. *BMC Bioinformatics* 8:2.
31. Ding YH, Liu NY, Tang ZS, Liu J, Yang WC (2006) Arabidopsis GLUTAMINE-RICH PROTEIN23 is essential for early embryogenesis and encodes a novel nuclear PPR motif protein that interacts with RNA polymerase II subunit III. *Plant Cell* 18:815–830.
32. Colcombet J, et al. (2013) Systematic study of subcellular localization of Arabidopsis PPR proteins confirms a massive targeting to organelles. *RNA Biol* 10:1557–1575.
33. Ito J, Heazlewood JL, Millar AH (2006) Analysis of the soluble ATP-binding proteome of plant mitochondria identifies new proteins and nucleotide triphosphate interactions within the matrix. *J Proteome Res* 5:3459–3469.
34. Klodmann J, Senkler M, Rode C, Braun HP (2011) Defining the protein complex proteome of plant mitochondria. *Plant Physiol* 157:587–598.
35. Olinares PD, Ponnala L, van Wijk KJ (2010) Megadalton complexes in the chloroplast stroma of Arabidopsis thaliana characterized by size exclusion chromatography, mass spectrometry, and hierarchical clustering. *Mol Cell Proteomics* 9:1594–1615.
36. Zybailov B, et al. (2008) Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS One* 3:e1994.
37. Andrés-Colás N, et al. (2017) A tripartite PPR protein interaction is involved in the SLO2 editosome in plant mitochondria. *Proc Natl Acad Sci USA* 114:8883–8888.
38. Majeran W, et al. (2012) Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: A new conceptual framework for nucleoid functions. *Plant Physiol* 158:156–189.
39. Huang S, et al. (2009) Experimental analysis of the rice mitochondrial proteome, its biogenesis, and heterogeneity. *Plant Physiol* 149:719–734.
40. Tanz SK, et al. (2013) SUBA3: A database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. *Nucleic Acids Res* 41:D1185–D1191.
41. Alonso JM, et al. (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* 301:653–657.
42. Bryant N, Lloyd J, Sweeney C, Myouga F, Meinke D (2011) Identification of nuclear genes encoding chloroplast-localized proteins required for embryo development in Arabidopsis. *Plant Physiol* 155:1678–1689.
43. Aryamanesh N, et al. (2017) The pentatricopeptide repeat protein EMB2654 is essential for trans-splicing of a chloroplast small ribosomal subunit transcript. *Plant Physiol* 173:1164–1176.
44. Bentolila S, Oh J, Hanson MR, Bukowski R (2013) Comprehensive high-resolution analysis of the role of an Arabidopsis gene family in RNA editing. *PLoS Genet* 9: e1003584.
45. Shi X, Germain A, Hanson MR, Bentolila S (2016) RNA recognition motif-containing protein ORRM4 broadly affects mitochondrial RNA editing and impacts plant development and flowering. *Plant Physiol* 170:294–309.
46. Leu K-C, Hsieh M-H, Wang H-J, Hsieh H-L, Jauh G-Y (2016) Distinct role of Arabidopsis mitochondrial P-type pentatricopeptide repeat protein-modulating editing protein, PPME, in nad1 RNA editing. *RNA Biol* 13:593–604.
47. Kleine T, Leister D (2016) Retrograde signaling: Organelles go networking. *Biochim Biophys Acta* 1857:1313–1325.
48. Göringer HU (2012) 'Gestalt,' composition and function of the Trypanosoma brucei editosome. *Annu Rev Microbiol* 66:65–82.
49. Shikanai T (2006) RNA editing in plant organelles: Machinery, physiological function and evolution. *Cell Mol Life Sci* 63:698–708.
50. Chateigner-Boutin AL, Small I (2010) Plant RNA editing. *RNA Biol* 7:213–219.
51. Blanc V, Davidson NO (2010) APOBEC-1-mediated RNA editing. *Wiley Interdiscip Rev Syst Biol Med* 2:594–602.
52. Xiu Z, et al. (2016) EMPTY PERICARP16 is required for mitochondrial nad2 intron 4 cis-splicing, complex I assembly and seed development in maize. *Plant J* 85:507–519.
53. Zmudjak M, et al. (2013) mCSF1, a nucleus-encoded CRM protein required for the processing of many mitochondrial introns, is involved in the biogenesis of respiratory complexes I and IV in Arabidopsis. *New Phytol* 199:379–394.
54. Van Leene J, et al. (2007) A tandem affinity purification-based technology platform to study the cell cycle interactome in *Arabidopsis thaliana*. *Mol Cell Proteomics* 6: 1226–1238.
55. Van Leene J, et al. (2011) Isolation of transcription factor complexes from Arabidopsis cell suspension cultures by tandem affinity purification. *Methods Mol Biol* 754:195–218.
56. Van Leene J, et al. (2010) Targeted interactomics reveals a complex core cell cycle machinery in Arabidopsis thaliana. *Mol Syst Biol* 6:397.
57. Rigaill G, et al. (October 11, 2016) Synthetic data sets for the identification of key ingredients for RNA-seq differential analysis. *Brief Bioinform*, 10.1093/bib/bbw092.
58. Dobin A, et al. (2013) STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21.
59. Castandet B, Hotto AM, Strickler SR, Stern DB (2016) ChloroSeq, an optimized chloroplast RNA-seq bioinformatic pipeline, reveals remodeling of the organellar transcriptome under heat stress. *G3 (Bethesda)* 6:2817–2827.
60. Shi X, Hanson MR, Bentolila S (2015) Two RNA recognition motif-containing proteins are plant mitochondrial editing factors. *Nucleic Acids Res* 43:3814–3825.
61. Alexander MP (1969) Differential staining of aborted and nonaborted pollen. *Stain Technol* 44:117–122.
62. Dreze M, et al. (2010) High-quality binary interactome mapping. *Methods Enzymol* 470:281–315.
63. James P, Halladay J, Craig EA (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* 144:1425–1436.
64. Clough SJ, Bent AF (1998) Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* 16:735–743.
65. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.
66. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359.
67. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–140.
68. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 100:9440–9445.
69. Thimm O, et al. (2004) MAPMAN: A user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J* 37:914–939.
70. Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency by Yoav Benjamini 1 and Daniel Yekutieli 2. *Ann Stat* 29: 1165–1188.
71. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550.
72. Dunn OJ (1961) Multiple comparisons among means. *J Am Stat Assoc* 56:52–64.
73. Auger IE, Lawrence CE (1989) Algorithms for the optimal identification of segment neighborhoods. *Bull Math Biol* 51:39–54.
74. Cleynen A, Koskas M, Lebarbier E, Rigaill G, Robin S (2014) Segmentor3IsBack: An R package for the fast and exact segmentation of Seq-data. *Algorithms Mol Biol* 9:6.
75. Baudry JP, Maugis C, Michel B (2012) Slope heuristics: Overview and implementation. *Stat Comput* 22:455–470.
76. Brehme N, Bayer-Császár E, Glass F, Takenaka M (2015) The DYW subgroup PPR protein MEF35 targets RNA editing sites in the mitochondrial rpl16, nad4 and cob mRNAs in Arabidopsis thaliana. *PLoS One* 10:e0140680.
77. Bentolila S, Babina AM, Germain A, Hanson MR (2013) Quantitative trait locus mapping identifies REME2, a PPR-DYW protein required for editing of specific C targets in Arabidopsis mitochondria. *RNA Biol* 10:1520–1525.
78. Hu Z, et al. (2016) Mitochondrial defects confer tolerance against cellulose deficiency. *Plant Cell* 28:2276–2290.
79. Dahan J, et al. (2014) Disruption of the CYTOCHROME C OXIDASE DEFICIENT1 gene leads to cytochrome c oxidase depletion and reorchestrated respiratory metabolism in Arabidopsis. *Plant Physiol* 166:1788–1802.
80. Xie T, et al. (2016) Growing Slowly 1 locus encodes a PLS-type PPR protein required for RNA editing and plant development in Arabidopsis. *J Exp Bot* 67:5687–5698.
81. Härtel B, Zehrmann A, Verbitskiy D, Takenaka M (2013) The longest mitochondrial RNA editing PPR protein MEF12 in Arabidopsis thaliana requires the full-length E domain. *RNA Biol* 10:1543–1548.
82. Weißenberger S, Soll J, Carrie C (2017) The PPR protein SLOW GROWTH 4 is involved in editing of nad4 and affects the splicing of nad2 intron 1. *Plant Mol Biol* 93:355–368.
83. Arenas-M A, Takenaka M, Moreno S, Gómez I, Jordana X (2013) Contiguous RNA editing sites in the mitochondrial nad1 transcript of Arabidopsis thaliana are recognized by different proteins. *FEBS Lett* 587:887–891.