

## Supplementary Information

### **A synthetic RNA editing factor edits its target site in chloroplasts and bacteria**

Santana Royan<sup>1</sup>, Bernard Gutmann<sup>2</sup>, Catherine Colas des Francs-Small<sup>2</sup>, Suvi Honkanen<sup>2,3</sup>, Jason Schmidberger<sup>1</sup>, Ashley Soet<sup>1</sup>, Yueming Kelly Sun<sup>2</sup>, Lilian Vincis Pereira Sanglard<sup>2</sup>, Charles S. Bond<sup>1</sup>, Ian Small<sup>2</sup>

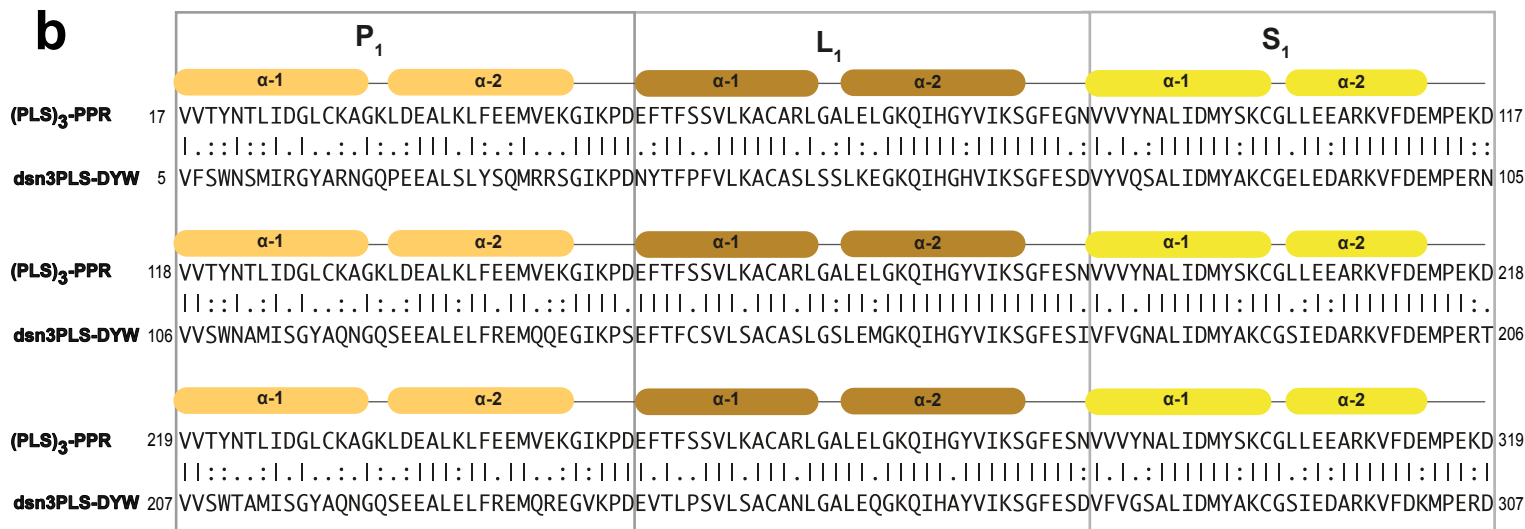
<sup>1</sup>School of Molecular Sciences, The University of Western Australia, Crawley 6009, Australia

<sup>2</sup>Australian Research Council Centre of Excellence in Plant Energy Biology, School of Molecular Sciences, The University of Western Australia, Crawley 6009, Australia

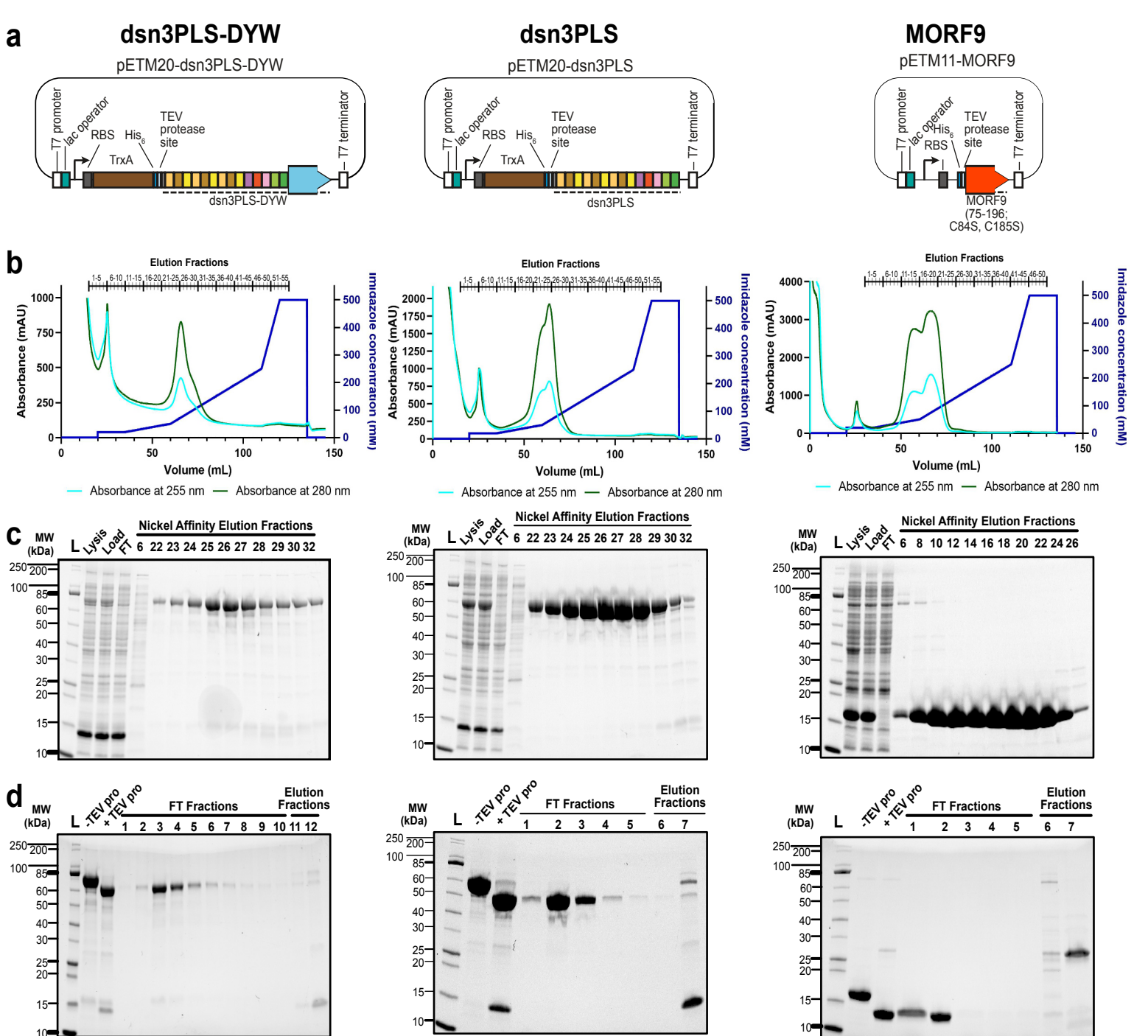
<sup>3</sup>Synthetic Biology Future Science Platform, CSIRO, Australia

**a****dsn3PLS-DYW**

Cap	1	MGNS	4				
P1	5	VFSWNSMIRGYARNGQPEEALSLYSQMRRSGIKPD	40	P2	308	VVSWNAMISGYAMHGHGKEALELFEEMQQSGIKPS	342
L1	41	NYTFPFVLKACASLSSLKEGKQIHGHVVIKSGFESD	74	L2	343	HVTFTGVLSACSHAGLVDEGRQYFNSMKKDYGIEPR	378
S1	75	VYVQSALIDMYAKCGELEDARKVFDEMPERN	105	S2	379	VEHYGCMVDLLGRAGRLDEAYEFIESMPIEPN	410
P1	106	VVSWNAMISGYAQNGQSEEALELFREMQQEGIKPS	140	E1	411	AVVWGALLGACRIHGNVELGERAAEKLFELEPES	444
L1	141	EFTFCSVLSACASLGSLEMKGQIHGYVIKSGFESI	175	E2	445	SGNYVLLSNIYASAGRWDDVAKVRKMMKERGIKK	478
S1	176	VFVGNALIDMYAKCGSIEDARKVFDEMPET	206	479	EPGCSWIEVKNVHEFVAGDRSHPQSEEIYAKLE	512	
P1	207	VVSWTAMISGYAQNGQSEEALELFREMQRGVKPD	241	513	ELSEKMKEAGYVPDTSFVLHDVEEEEKEQMLSYH	546	
L1	242	EVTLPVSVLSACANLGALEQKGKQIHAYVIKSGFESD	276	DYW	547	SEKLAIAFGLISTPPGTPRIVKNLRVCGDCHTA	580
S1	277	VFVGSALIDMYAKCGSIEDARKVFDKMPERD	307	581	IKFISKIVGREIIVRDSNRFHFKDGCSCGDYW	614	

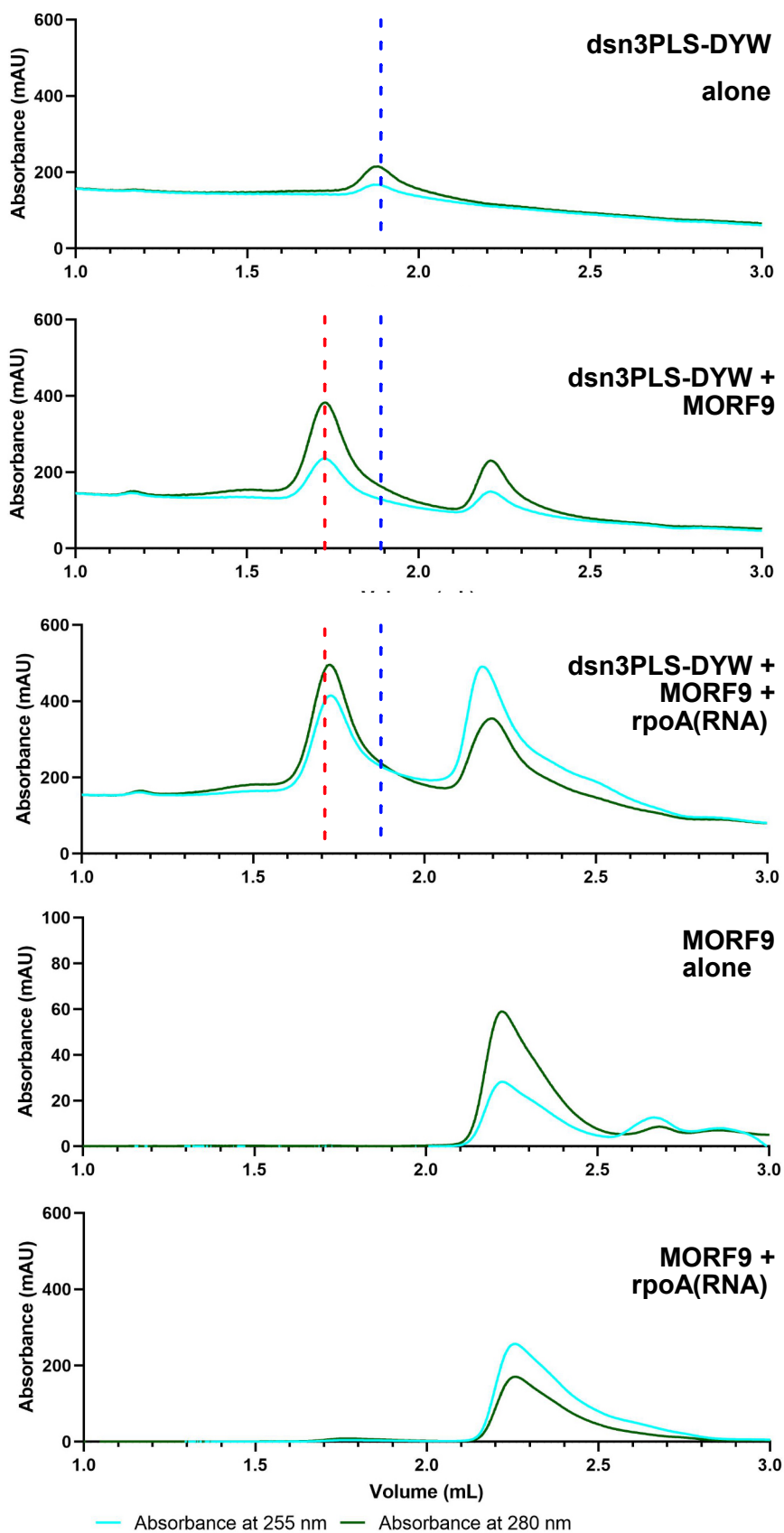
**b****Supplementary Figure S1. Comparison of the sequences of dsn3PLS-DYW and (PLS)<sub>3</sub>-PPR.**

**a:** Sequence and motif structure of dsn3PLS-DYW. The catalytic E70 residue within the DYW domain is underlined in red. **b:** Alignment of the (P1L1S1)<sub>3</sub> regions of dsn3PLS-DYW and (PLS)<sub>3</sub>-PPR<sup>1</sup>. Overall, throughout this region the two synthetic proteins are only 64% identical.



**Supplementary Figure S2. Purification of dsn3PLS-DYW, dsn3PLS and MORF9.**

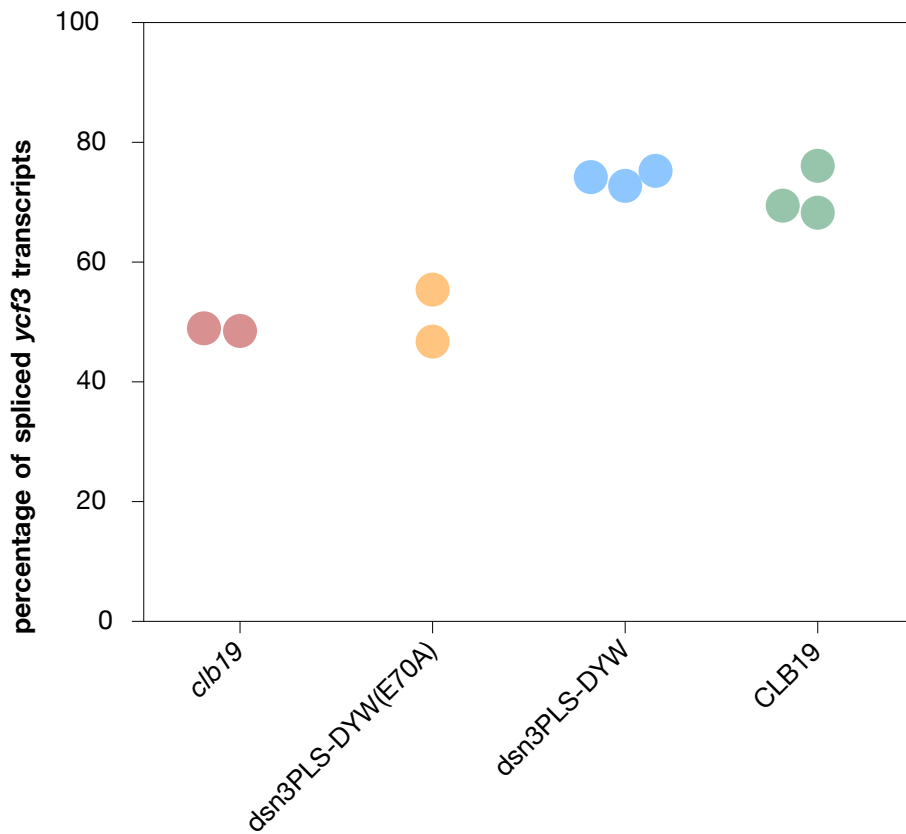
**a:** Vector map of pETM20-dsn3PLS-DYW (left), pETM20-dsn3PLS (middle), and pETM11-MORF9 (right).  
**b:** Elution profiles from nickel affinity chromatography using imidazole (dark blue trace) as a competitor. Protein content was measured by absorbance at 255 nm (cyan trace) and 280 nm (dark green trace).  
**c:** Protein profiles of selected elution fractions. L, molecular weight ladder; Lysis, unpurified bacterial lysate; Load, fraction loaded on HisTrap column; FT, flow-through from column; 6-32, specific elution fractions numbered as in **b**.  
**d:** Reverse nickel-affinity purification of dsn3PLS-DYW, dsn3PLS and MORF9 following TEV cleavage of the N-terminal Ni-binding His-tag. L, molecular weight ladder; -TEV, control with no TEV added; +TEV, TEV cleavage products; FT fractions, flow-through fractions containing the desired unbound product; Elution fractions, fractions eluted from the column showing the bound, cleaved His-tag. Expected molecular weights for dsn3PLS-DYW, dsn3PLS, and MORF9 with their purification tags: 82.88 kDa, 68.39 kDa, and 17.22 kDa respectively. Expected molecular weights for dsn3PLS-DYW, dsn3PLS, and MORF9 without their purification tags: 68.64, 54.14, and 14.22 kDa respectively.



**Supplementary Figure S3. Analytical size-exclusion chromatography of complexes between dsn3PLS-DYW, MORF9 and their RNA target.**

Fractions leaving the column were analysed for protein and RNA content by absorbance at 280 nm and 255 nm respectively. The peak containing dsn3PLS-DYW (vertical blue dashed line) was shifted to a larger molecular weight complex (vertical red dashed line) by the addition of MORF9. The addition of the rpoA-78691 target oligonucleotide raised the absorbance at 255 nm of this peak, indicating the incorporation of the RNA into the high molecular weight complex. The samples shown contain 4.88 nmol of dsn3PLS-DYW, with 19.92 nmol of MORF9, and 10 nmol RNA oligonucleotide. This equates to a 4:1 ratio of MORF9 to dsn3PLS-DYW and a 2:1 ratio of RNA oligonucleotide to dsn3PLS-DYW.





**Supplementary Figure S5. Splicing of *ycf3* intron 2 is unaffected by editing of *ycf3-43350*.** The extent of *ycf3* intron 2 splicing was estimated from RNA-seq data by counting reads crossing the donor (exon-intron) junction, the acceptor (intron-exon) junction and the spliced (exon-exon) junction. Percentage splicing was calculated as  $\text{spliced} / (\text{spliced} + (\text{donor} + \text{acceptor}) / 2)$ . The percentage splicing in plants expressing dsn3PLS-DYW (no editing of *ycf3-43350*) and CLB19 (~20% editing of *ycf3-43350*) is not statistically distinguishable (p-value 0.33, t-test, n=3 biological replicates).

dsn3PLS-DYW	CLB19	position	functional annotation	effect of editing
23/34518	<b>31384/36063</b>	<b>69942</b>	<i>clpP1</i> coding sequence	<b>CAU</b> (His) -> <b>UAU</b> (Tyr)
<b>471/1071</b>	<b>1279/1575</b>	<b>78691</b>	<i>rpoA</i> coding sequence	<b>UCU</b> (Ser) -> <b>UUU</b> (Phe)
2/3622	<b>829/4255</b>	43350	<i>ycf3</i> intron 2	
0/3810	<b>64/3940</b>	59002	intergenic ( <i>accD-psaI</i> )	
<b>113/7898</b>	0/9704	114860	<i>ccsA</i> coding sequence	<b>CUC</b> (Leu) -> <b>UUC</b> (Phe)
<b>160/11446</b>	<b>99/12332</b>	118393	<i>ndhG</i> coding sequence	<b>UCU</b> (Ser) -> <b>UUU</b> (Phe)
<b>340/28513</b>	6/33229	116203	<i>ndhD</i> coding sequence	<b>UCU</b> (Ser) -> <b>UUU</b> (Phe)
1/1647	<b>19/1709</b>	9538	3' <i>atpA</i>	
0/2842	<b>37/3501</b>	111658	<i>ndhF</i> coding sequence	<b>CUC</b> (Leu) -> <b>CUU</b> (Leu)
5/19019	<b>139/22668</b>	14647	<i>atpI</i> coding sequence	<b>CUU</b> (Leu) -> <b>UUU</b> (Phe)
<b>76/13655</b>	1/16473	49646	<i>ndhK</i> coding sequence	<b>CCU</b> (Pro) -> <b>UCU</b> (Ser)
<b>149/26854</b>	1/30538	97693	<i>rps7</i> coding sequence	<b>CAU</b> (His) -> <b>UAU</b> (Tyr)
4/24888	<b>136/29717</b>	13561	5' <i>atpH</i>	
<b>49/9909</b>	1/11722	119998	<i>ndhA</i> coding sequence	<b>CUU</b> (Leu) -> <b>UUU</b> (Phe)
1/10196	<b>45/11786</b>	97277	3' <i>rps7</i>	
6/63955	<b>345/92791</b>	38871	<i>psaB</i> coding sequence	<b>CCC</b> (Pro) -> <b>UCC</b> (Ser)
0/36770	<b>122/39249</b>	66041	<i>petG</i> coding sequence	<b>CCU</b> (Pro) -> <b>CUU</b> (Leu)
<b>150/45178</b>	7/46938	69992	<i>clpP1</i> coding sequence	<b>UCC</b> (Ser) -> <b>UUC</b> (Phe)
<b>105/58902</b>	11/79032	39020	<i>psaB</i> coding sequence	<b>UCA</b> (Ser) -> <b>UUA</b> (Leu)
20/105478	<b>233/107140</b>	13258	3' <i>atpH</i>	
<b>69/102074</b>	25/114170	73588	<i>psbB</i> coding sequence	<b>CUC</b> (Leu) -> <b>CUU</b> (Leu)

**Supplementary Table S1. Editing events detected in this study and their predicted functional impact.**

The primary targets of CLB19 are indicated first, in bold. The other sites are presumed 'off-target' events. The values in the first two columns are editing events/read counts. Ratios in bold were statistically significant in comparison to the equivalent values from the negative control, i.e. *clb19* plants expressing dsn3PLS-DYW (E70A). Where editing leads to a predicted amino acid change in the encoded protein, this is indicated in the final column, with the altered nucleotide in bold.

Oligonucleotide name	Sequence (5' - 3')	Modification	Purpose
<i>rpoA_L</i>	UAUUACACGUGCAAAAUCUG	None	Analytical size-exclusion chromatography
<i>rpoA_REMSA</i>	/5Cy5/AUGUAUUACACGUGCAAAAUCUGAGA	5'-Cy5 labelled	REMSA experiments
<i>clpP1_REMSA</i>	/5Cy5/CAGCAACAGAAGCCCAAGCUCAUGGA	5'-Cy5 labelled	REMSA experiments

**Supplementary Table S2. RNA oligonucleotides used in this study.**





<i>clpP1_ES</i> .REV	TGAACCGCTACAAGATCAAC	Amplifying <i>A. thaliana clpP1</i> cDNA
<i>clpP1_PS</i> .FOR	GTATAGCATTCCCTCACGCTAG	Sequencing <i>A. thaliana clpP1</i> editing site
MORF2_73.FOR	GGTGGTCCATGGCGACGGAGATGGCTCCTTTG TTCCGGGATCCGATTATG	Cloning MORF2 73-193, introducing C82S mutation
MORF2_193.REV	GGTGGTCTCGAGTCACTCAACCCGCCTCTGCC TCTCTGGTGATCGTTGGACTATC	Cloning MORF2 73-193

**Supplementary Table S3. DNA primers used in this study.**

## Supplementary References

1. Yan, J. *et al.* MORF9 increases the RNA-binding activity of PLS-type pentatricopeptide repeat protein in plastid RNA editing. *Nat Plants* **3**, 17037 (2017).