THE *ORGANELLE ZINC FINGER* FAMILY OF PROTEINS CONTRIBUTES TO RNA PROCESSING IN PLANT ORGANELLES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Andrew Brendan Gipson

August 2021

THE *ORGANELLE ZINC FINGER* FAMILY OF PROTEINS CONTRIBUTES TO RNA PROCESSING IN PLANT ORGANELLES

Andrew Brendan Gipson, Ph.D.

Cornell University 2021

Our lab previously discovered a small family of organelle-localized proteins in plants characterized by the presence of zinc finger domains, the Organelle Zinc finger (OZ) family. Two members of this family, OZ1 and OZ2, are necessary for successful processing and maturation of RNA produced from the circular genomes of plant chloroplasts and mitochondria. We discovered that OZ2 is a splicing factor, and I performed experiments showing that OZ2 binds to a set of both general and specific splicing factors.

In the chloroplast, OZ1 is required for many RNA editing events, a process by which specific cytosines are enzymatically converted to uracils as a correction mechanism for evolutionarily conserved missense mutations in the organelle genomes. RNA editing is carried out by a large multi-protein complex called the "editosome." As well as OZ1, editosomes contain members of the PPR protein family, the RIP/MORF family, and the ORRM family. OZ1 is an 82-kDa protein with distinct domains, including a family-specific N-terminal region, a pair of zinc finger domains, and a unique C-terminal region. To elucidate the functions of these domains, I generated truncations of OZ1 for use in protein-protein interaction assays that identified the C-terminal region of OZ1 as the primary interactor with PPR proteins, the editosome factors required for site-specificity and enzymatic editing. Expression of

these OZ1 constructs *in vivo* showed that the zinc finger domains were required to restore editing in *oz1* knockout plants. Mutation of key structural residues in the zinc finger domains showed that they are necessary for editing and required for interaction with ORRM1, a general editing factor with an RNA-binding domain.

This work establishes the OZ protein family as a novel group of factors necessary for RNA metabolism in plant organelles.

BIOGRAPHICAL SKETCH

Andrew B. Gipson was born in Wadsworth, Ohio in 1991.

His love of nature was nurtured by his parents, who always took him to the zoo, aquarium, or wildlife preserve near whatever place they were visiting on family vacations, even if they were a bit too tired to hike around a swamp.

After a brief flirtation with linguistics in high school, he majored in Molecular Biology at Kenyon College from 2009 to 2013. He dabbled in different areas of biology but settled on molecular biology, especially as it pertains to plants, as his undergraduate pursuit. A summer internship in Dr. Maureen Hanson's lab in 2011 gave him the chance to work with motor proteins and fluorescence microscopy in Arabidopsis. For his undergraduate research thesis under Dr. Kerry Rouhier at Kenyon, he studied the effect of a mutation in the short-chain fatty acid catabolic pathway of Arabidopsis on the accumulation of lipid reserves in seeds. He received a Bachelor of Arts with High Honors in 2013.

To experience biomedical science in a large research institution setting, Andrew took a position as researcher and lab manager for the start-up cancer biology lab of Dr. Cheuk Leung in the Department of Pharmacology at the University of Minnesota, Minneapolis campus.

After applying to many biomedical translational-focused graduate programs, Andrew found himself drawn back to Cornell University in 2015. He returned to Dr. Maureen Hanson's lab to study the recently discovered *Organelle Zinc Finger* family of proteins and their involvement in RNA processing.

During his time at Cornell, Andrew served as a teaching assistant for the MBG department for ten semesters, teaching courses on introductory cell biology, molecular biology, structural biology, and a lab in biochemistry. He also worked as a writing consultant for the English Language Support Office, assisting international graduate students and postdocs with academic writing projects.

Outside of academia, he pursues photography, using modern digital equipment for birding and rediscovering old film shooting and analog printing techniques for artistic and travel photography. He also writes adventures for horror roleplaying games and enjoys roping his friends into playing new board games.

This work is dedicated to my father, Thomas B. Gipson, who always supported my interests and was proud of everything I achieved.

ACKNOWLEDGEMENTS

First and foremost, thank to Maureen Hanson for taking me on as an advisee. She has been supportive of my scientific career since I worked as a BTI REU student in her lab in the summer of 2011. I have always appreciated her even-keel and have sung her praises to other students whenever they ask about my lab experiences.

Thanks to Stéphane Bentolila for patience, and for clarifying and articulating results when I could not. When people speak of science standing on the shoulders of giants, Stéphane is the giant I'm standing on. Sorry about the shoulders!

As my other committee members, Ailong Ke and Klaas van Wijk have been great sounding boards and sources of advice over the past years. Their feedback was a crucial part of reorienting my project after the A-exam, and I would probably still be wallowing with proteins in the cold room if not for them.

I also have to thank the other members of the Hanson lab. Ludovic Giloteaux's expertise in sequence alignment and phylogenetic trees was key for insights gained into the RanBP2 zinc finger proteins in plants. Thanks to Carl Franconi for keeping the gears of the lab well-oiled, and to Arnaud Germain for a sharp eye during lab meetings. I want to thank Kevin Hines, Alex Mandarano, Adam O'Neal, Jessica Maya, Myat Lin, and Vishal Chaudhari for commiserating and chatting about experiments or life in general. Alex Kehl was an excellent summer student who contributed the majority of the work towards discovering the localization of the OZ family proteins.

Thanks go to Tim Huffaker, Jim Blankenship, and Stephen Jesch, all of whom were role models for teaching and were wonderful professors during my long tenure as a teaching assistant.

Ed Partlow has been a great friend, fellow scholar, and enthusiastic gaming partner since the first time we stepped onto the cliffs above Cayuga. I thank Chris Furman of the E. Alani lab for advice on yeast culturing and transformation.

Hye-Rim Hong has been my patient and supportive girlfriend for the last two years of my PhD, and without her "chop-chop," I wouldn't be graduating for at least another year

Of course, I've been born on the wind of support from my family. I'm glad Dad got to see me in cap and gown, but I wish he could have been here to see me as Dr. Gipson.

This work was funded by a grant from the National Science Foundation (MCB-1615393 to S.B. and M.R.H.). Confocal microscopy was performed at the Cornell BRC Imaging Facility using equipment funded by a grant from the NIH (S10RR025502).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Arabidopsis RanBP2-type zinc finger proteins related to chloroplast RNA editing factor OZ1*

ABSTRACT

OZ1, an RNA editing factor that controls the editing of 14 cytidine targets in Arabidopsis chloroplasts, contains two RanBP2-type zinc finger (Znf) domains. The RanBP2 Znf is a C4-type member of the broader zinc finger family with unique functions and an unusually diverse distribution in plants. The domain can mediate interactions with proteins or RNA and appears in protein types such as proteases, RNA editing factors, and chromatin modifiers; however, few characterized Arabidopsis proteins containing RanBP2 Znfs have been studied specifically with the domain in mind. In humans, RanBP2 Znf-containing proteins are involved in RNA splicing, transport, and transcription initiation. We present a phylogenetic overview of Arabidopsis RanBP2 Znf proteins and the functional niches that these proteins occupy in plants. OZ1 and its four-member family represent a branch of this family with major impact on the RNA biology of chloroplasts and mitochondria in Arabidopsis.

We discuss what is known about other plant proteins carrying the RanBP2 Znf domain and point out how phylogenetic information can provide clues to functions of uncharacterized Znf proteins.

---

* This work was originally published as "Arabidopsis RanBP2-type zinc finger proteins related to chloroplast RNA editing factor OZ1" by Andrew B. Gipson, Ludovic Giloteaux, Maureen R. Hanson, and Stephane Bentolila. *Plants*. 2020; 9 (3): 307.

INTRODUCTION

The zinc finger domain is most well-known as a DNA-binding domain present in numerous transcription factors, but it is in fact a superfamily with several sub-families characterized by various structural and functional differences. One zinc finger (Znf) family that breaks with the classical DNA-binding function is the RanBP2 family. The namesake of the RanBP2 zinc finger family is Ran Binding Protein 2, a human protein whose cluster of eight zinc fingers participate in protein binding with the nuclear export factor exportin-1 [1]. There are other proteins associated with the nuclear pore whose RanBP2-Znf domains participate in protein-protein interactions, such as Nup153 [2]. Particularly relevant to this review is ZRANB2, a human protein that is known to be a component of the nuclear spliceosome and has influence over alternative splicing in several transcripts related to cell migration (e.g., SPATA13) and chromatin remodeling (e.g., SMARCC2), among other functions [3,4]. The RanBP2 Znf domains of ZRANB2 have a high affinity for the RNA sequence GGU [5], which is the core sequence in the 5' splice site of the majority of spliced transcripts in both humans [6] and plants [7]. ZRANB2 also binds to and stabilizes a long non-coding RNA, SNHG20, which is involved in reducing the formation of nutrient-obtaining structures in glioma tumors [8]. ZRANB2 is not alone in this subclass of RNA-binding RanBP2 proteins; seven human proteins, including ZRANB2, share highly similar RanBP2-Znf domains, and almost all are involved in RNA splicing, transport, or transcription initiation [9].

Our group became interested in the RanBP2 family of Znfs after our discovery of OZ1, an essential Arabidopsis chloroplast RNA-editing factor that has two tandem

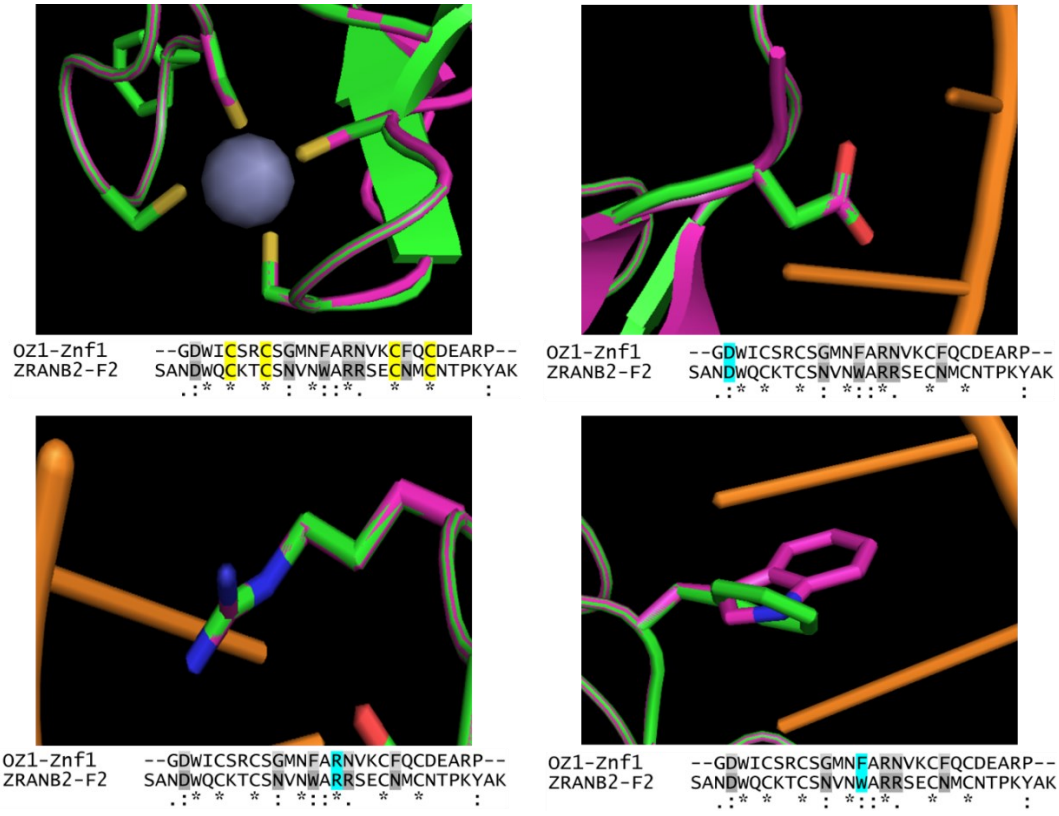
RanBP2-type Znfs [10]. The sequences of the OZ1 Znfs are remarkably similar to those of ZRANB2 (**Figure 1.1a**). The consensus sequence of the RanBP2 Znf is R/K-X-G-D-W-X-C-X(2,4)-C-X(3)-N-X(6)-C-X(2)-C-X(3)-R/K (30 residues); the four Cys residues distinguish this Znf from some other Znf families and are required for coordination of a Zn^{2+} necessary for the structural fold of the domain. Predictive structure modeling of an OZ1 Znf, using the solved structure of the second Znf of ZRANB2 as a starting point and comparison, shows several residues of OZ1 in the same position as the RNA-binding residues of ZRANB2 (**Figure 1.1b**), suggesting that this organellar RNA editing factor may have similar RNA-binding properties. As will be discussed later, there is already evidence that other Arabidopsis proteins with the RanBP2 Znf domain perform RNA binding. In the remainder of this paper, we will discuss a number of other RanBP2 Znf proteins found in Arabidopsis, comparing their sequences and localization along with uncharacterized RanBP2 Znf proteins, and then highlight the need to further study this family in plants.

Figure 1.1. Model comparison of ZRANB2-Znf2 crystal structure with predicted structure of OZ1-Znf1. (**a**) Sequence alignment of the zinc finger domains of human ZRANB2 ("ZRANB2-F1 and ZRANB2-F2") and of OZ1 ("OZ1-Znf1 and OZ1-Znf2"). Asterisks = conserved RanBP2 Znf residues; double dots = chemically similar residues; gray shading = putative RNA-binding residues. (**b**) Overlaid models of ZRANB2-F2 w/ RNA-target crystal structure and OZ1-Znf1 predicted structure comparing ZRANB2-F2 RNA-binding residues with OZ1-Znf1 residues in the same position. Blue highlighting in sequence alignment corresponds to the residue in the image. Green = OZ1 Znf-1; purple = ZRANB2; orange = RNA from ZRANB2 crystal structure; gray sphere = Zn^{2+} ion. Upper left: Zn^{2+}-coordinating Cys residues; upper right: Glu residue of ZRANB2-F2 hydrogen bonds with guanine through water molecule; bottom left: Arg residue of ZRANB2-F2 hydrogen bonds with guanine at two positions; bottom right: Trp residue of ZRANB2-F2 base-stacks between adjacent guanines. PDB for ZRANB2 structure: 3g9y. OZ1 Znf1 modeled with Phyre2 [11] using ZRANB2-F2 structure as a base (Phyre2 homology confidence = 99.2%; percent ID = 36).

(a)

```
OZ1-Znf1        --GDWIC--SRCSGMNFARNVKCFQCDEARP--    27
OZ1-Znf2        --SEWEC--PQCDFYNYGRNVACLRCDCKRP--    27
ZRANB2-F1       SDGDWICPDKKCGNVNFARRTSCNRCGREKTTE    33
ZRANB2-F2       SANDWQC--KTCSNVNWARRSECNMCNTPKYAK    31
                  .:* *    *.  *:.*.  *  *.  :
```

(b)



SUBCELLULAR LOCATION OF RANBP2 ZINC FINGERS IN ARABIDOPSIS

Arabidopsis proteins possessing RanBP2 Znfs (InterPro ID: IPR001876) were identified by querying the Interpro protein database (version 77.0) [12], and redundant entries were identified with CD-HIT [13], narrowing down the selection to 26 proteins (**Table 1.1**). In considering what proportion of these proteins may have organellar localization, a mix of experimental and algorithmic prediction results from the SUBA4 protein localization database indicate that 3.9% of Arabidopsis proteins are mitochondrial, and 4.4% are plastidial, for a total of 8.3% of proteins being in one of those organelles [14]. Analysis of the proteins in this study show that the majority of Arabidopsis RanBP2 Znf proteins are nuclear-localized (~58%), with a portion being predicted or demonstrated to be in plastids and/or mitochondria (23%), largely due to the OZ family (**Table 1.1**).

Comparison of the results from TargetP 2.0 [15] and Predotar [16] localization algorithms against the SUBA4 database, which compiles information from TargetP and Predotar along with other algorithms and empirical evidence, shows that localization prediction is not always accurate (**Table 1.1**). OZ1, for instance is rather weakly predicted to be plastid-localized by TargetP and strongly predicted to be non-organellar by Predotar, but all localization experiments have shown it to be exclusively in the plastid [10]. RBL14 is an example of a protein with controversial localization; the primary Arabidopsis algorithm TargetP predicts mitochondria, and SUBA4 strongly suggests plasma membrane localization based on a large-scale fluorescent labeling experiment [17], but the image data from the study does not

exclude mitochondrial localization in addition to plasma membrane localization. In any case, empirical studies focused on the protein in question are required for confident statements about its location.

Table 1. RanBP2 Znf proteins in Arabidopsis and representative model plants

Name	UniProt ID	TAIR ID	# of RanBP2 Znf Domains	Znf Function[1]	SUBA Localization[2]	TargetP[2]	Predotar (v. 1.04, 2016)[2]	Function[1]
OZ1	Q8S9K3	At5g17790	2	N/A	Plastid (1)	Chloroplast (0.4837)	Non-Organellar (0.78)	RNA editing
OZ2	Q9C7M2	At1g55040	2	N/A	Plastid (1)	Mitochondrial (0.8581)	Plastid (0.41), Mitochondria (0.38)	N/A
OZ3	F4I6V3	At1g70650	3	N/A	Cytosol (0.579)	Mitochondrial (0.7018)	Mitochondria (0.71)	N/A
OZ4	Q9LP67	At1g48570	4	N/A	Plastid (0.993)	Mitochondrial (0.8274)	Non-Organellar (0.75)	N/A
RHOMBOID-like protein 14 (RBL14)	Q8RXW0	At3g17611	1	*Protein interaction*	Plasma membrane (1)	Mitochondria (0.9658)	Non-Organellar (0.72)	Serine protease
SKI7	F4KI84	At5g10630	1	*Protein interaction*	Cytosol (0.998)	Non-Organellar (0.9999)	Non-Organellar (0.99)	RNA degradation
SKI7-variant	Q9LXB6	At5g10630	1	*Protein interaction*	Cytosol (1)	Non-Organellar (0.9913)	Non-Organellar (0.97)	*RNA degradation*
HBS1	A0A1P8BFF4	At5g10630	1	*Protein interaction*	Cytosol (0.998)	Non-Organellar (0.9984)	Non-Organellar (0.99)	Stalled ribosome rescue
TBP-associated factor 15 (TAF15)	Q9AST1	At1g50300	2	*RNA binding*	Nucleus (1)	Non-Organellar (0.9998)	Non-Organellar (0.99)	Component of the general transcription factor TFIID
TBP-associated factor 15B (TAF15b)	Q94KD0	At5g58470	1	*RNA binding*	Nucleus (1)	Non-Organellar (0.9999)	Non-Organellar (0.99)	PolII inhibition
SUPPRESSOR OF ABI3-5 (SUA)	F4JCU0	At3g54230	1	*RNA binding*	Nucleus (1)	Non-Organellar (1)	Non-Organellar (0.99)	Splicing Factor
Histone Deacetylase 15 (HDA15)	Q8GXJ1	At3g18520	1	Protein interaction	Nucleus (1)	Non-Organellar (0.9977)	Non-Organellar (0.99)	Chromatin modification
ARI8	Q8W468	At1g65430	1	*Protein interaction*	Nucleus (1)	Non-Organellar (1)	Non-Organellar (0.99)	Protein turnover/ubiquitination
ARI13	Q9FFN9	At5g63750	1	*Protein interaction*	Nucleus (1)	Non-Organellar (0.9997)	Non-Organellar (0.99)	*Protein turnover/ubiquitination*

Name	UniProt ID	TAIR ID	# of RanBP2 Znf Domains	Znf Function[1]	SUBA Localization[2]	TargetP[2]	Predotar (v. 1.04, 2016)[2]	Function[1]
ARI14	Q9FFP1	At5g63730	1	*Protein interaction*	Nucleus (0.996)	Non-Organellar (1)	Non-Organellar (0.99)	*Protein turnover/ubiquitination*
ARI15	Q84RQ8	At5g63760	1	*Protein interaction*	Nucleus (0.997)	Non-Organellar (0.9967)	Non-Organellar (0.99)	*Protein turnover/ubiquitination*
ARI16	Q9C5A4	At5g08730	1	*Protein interaction*	Nucleus (0.999)	Non-Organellar (0.9999)	Non-Organellar (0.99)	*Protein turnover/ubiquitination*
Stress-associated RNA-binding Protein 1 (SRP1)	Q8S8K1	At2g17975	3	RNA binding	Nucleus (0.98)	Non-Organellar (0.9979)	Non-Organellar (0.99)	RNA turnover; binding of 3'-UTR of ABI2
-	Q9SA95	At1g11800	1	N/A	Mitochondria (0.999)	Mitochondria (0.7178)	Mitochondria (0.55)	N/A
-	Q8GWD1	At5g25490	3	N/A	Nucleus (0.54)	Non-Organellar (0.9895)	Non-Organellar (0.99)	N/A
-	F4JM55	At4g28990	1	N/A	Nucleus (1)	Non-Organellar (1)	Non-Organellar (0.99)	N/A
-	Q8GZ43	At1g67325	3	N/A	Nucleus (1)	Non-Organellar (0.9915)	Non-Organellar (0.98)	N/A
-	A0A1P8ARI2	At1g55915	2	N/A	Nucleus (0.999)	Non-Organellar (0.9174)	Non-Organellar (0.86)	N/A
-	O64715	At2g02620	2	N/A	Cytosol (0.542)	Non-Organellar (0.9994)	Non-Organellar (0.99)	N/A
-	Q6ID73	At2g26695	3	N/A	Plasma membrane (0.545)	Non-Organellar (0.7971)/Signal peptide (0.2023)	Non-Organellar (0.98)	N/A
-	Q9LW11	At3g15680	3	N/A	Nucleus (0.879)	Non-Organellar (0.9787)	Non-Organellar (0.99)	N/A
Pp_OZ	A0A2K1J4G7	PHYPA_022275	3	N/A	N/A	Mitochondria (0.4898)/ Chloroplast (0.4733)	Plastid (0.88)	N/A
Sm_OZ	D8SXV1	SELMODRAFT_447225	2	N/A	N/A	Non-Organellar (0.6019)/ Mitochondria (0.3215)	Non-Organellar (0.74)	N/A
Zm_OZ1	A0A1D6N8W5	GRMZM2G312244	2	N/A	N/A	Mitochondria (0.4279)/ Chloroplast (0.3028)	ER (0.76), Plastid (0.41)	*RNA editing*

[1]Italics indicate hypothetical function; non-italics indicate experimentally confirmed function. [2]Number in parentheses indicates the "confidence score" on a scale of 0 to 1; scores are determined independently within each database or algorithm.

PHYLOGENETIC ANALYSIS OF ARABIDOPSIS PROTEINS CARRYING RANBP2 ZINC FINGERS

Full-length protein sequences for phylogenetic analysis were obtained from UniProt [18] and verified against sequences in TAIR [19]. MUSCLE alignment [20] was used to align protein sequences, and phylogenetic relationships were drawn with the Maximum Likelihood method as described in Whelan and Goldman [21]. Using the Znf sequence from OZ1 as a BLAST query, RanBP2 Znf proteins from model plants in other clades (*Zea mays*, the moss *Physcomitrella patens*, and the lycophyte *Selaginella moellendorffii*) were identified and included in the phylogenetic analysis to act as reference points for evolutionary speculation. The most noticeable feature of the phylogenetic tree is the grouping of OZ proteins into a single clade (**Figure 1.2a**). In spite of the evolutionary distance between OZ proteins in Arabidopsis, *P. patens*, and *S. moellendorffii*, their grouping may imply similar functions of these proteins, possibly in organellar RNA processing like OZ1 [10]. Another notable clade is the Ariadne (ARI) family, E3-type ubiquitin ligases with unique C-terminal RanBP2 Znfs. This protein family is discussed in detail in later, but it is worth mentioning here because this group is like the OZ group in that all members have at least been named and commented upon in the literature, if not individually characterized [22,23]. Conversely, the remaining nodes on the tree in **Figure 1.2** include at least one wholly uncharacterized protein. Each of these RanBP2 Znf proteins can be seen as a potential new splicing factor, chromatin modifier, protease, or RNA processing factor, based on their association with RanBP2 Znf proteins whose functions have been established.

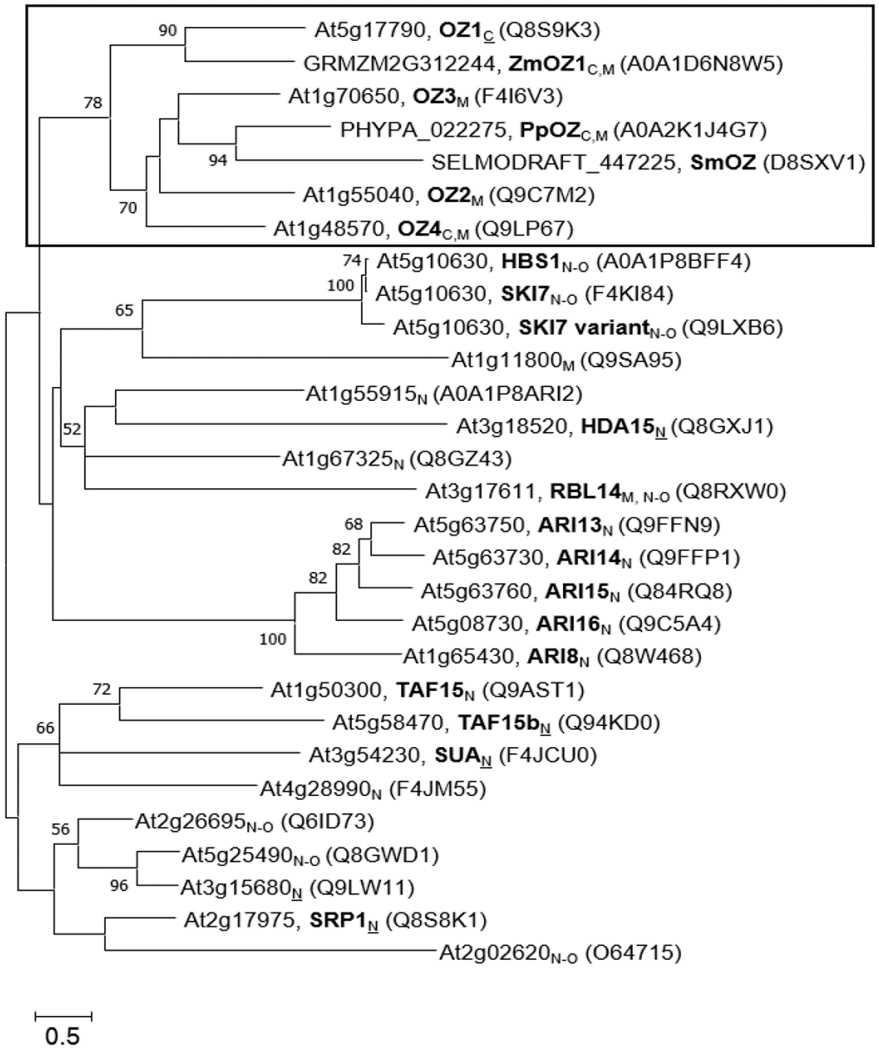


Figure 1.2. Phylogenetic tree of full-length proteins containing RanBP2 zinc finger domains in Arabidopsis. RanBP2 Znf protein sequences from *Zea mays* (ZmOZ1), *Physcomitrella patens* (PpOZ), and *Selaginella moellendorffii* (SmOZ) are included for evolutionary context. Entries list the TAIR accession number (or the species-specific accession number), the protein name in bold, and the UniProt ID in parentheses. Subscripts indicate localization: C = chloroplast, M = mitochondria, N = nucleus, N-O = non-organellar; non-underlined letter indicates predicted localization, underlined letter indicates experimentally confirmed localization. The evolutionary relationships were examined by sequence alignment using MUSCLE [20] and analysis was conducted in MEGA X [24]. A phylogenetic tree was inferred by using the Maximum Likelihood method and the Whelan and Goldman model [21]. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Percentages of 500 bootstrap resampling that supported the branching orders in each analysis are presented above or near the relevant nodes and are shown for branches with more than 50% bootstrap support. This analysis involved 29 amino acid sequences with a total of 1288 positions considered in the final dataset.

The nature of protein domains as tools within different protein "toolboxes" led us to analyze the RanBP2 Znf domains themselves and how they interrelate, both for insights on their function as well as how they are related evolutionarily. The RanBP2 Znf has a unique signature (see consensus sequence in the Introduction) and is well-conserved across its different iterations, but examination of their number and sequence differences can give clues as to their function within the full-length protein. Sequences of individual RanBP2 Znf domains were extracted from UniProt domain designations, then the alignment and tree were assembled as described above. The Znf domains of the human splicing factor ZRANB2 were included as a reference point because their function and structure have been exhaustively determined and can provide insight into how the neighboring Znfs from Arabidopsis may function. Looking at this domain-based tree, we see again that the OZ family Znfs group together (**Figure 1.3**, black box). All OZ proteins have at least two RanBP2 Znf domains, and interestingly we see that the Znfs in the most C-terminal position group together within the OZ clade (**Figure 1.3**, gray box). Preliminary data indicate that the Znf in this position has the most impact on OZ1 editing function [25], so the conservation we observe may be indicative of selection on Znfs in that position related to their functional relevance.

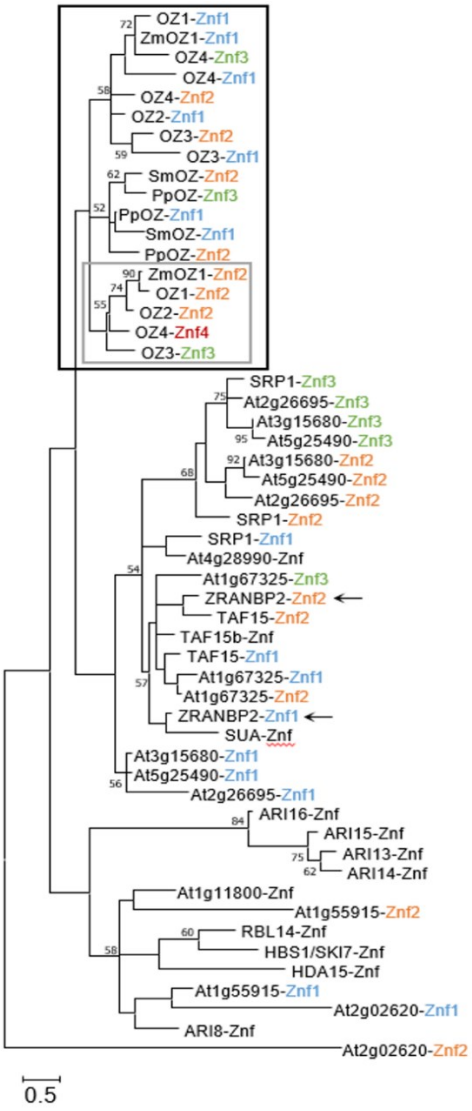


Figure 1.3. Phylogenetic tree of Arabidopsis RanBP2 Zinc Finger domains. ZRANB2 (marked with arrows) is a human RanBP2 Znf protein included for comparison. Black box: OZ clade; gray box: clade of C-terminal OZ Znf domains. Text color based on the order of the zinc finger domain in the protein starting from the N-terminus: blue = first Znf, orange = second Znf, green = third Znf, red = fourth Znf; Znfs from proteins with only one such domain are colored black. RanBP2 Znf domain sequences were aligned using MUSCLE [20], and evolutionary analysis was conducted in MEGA X [24]. The tree was inferred by using the Maximum Likelihood method and the Whelan and Goldman model [21]. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Percentages of 500 bootstrap resampling that supported the branching orders in each analysis are presented above or near the relevant nodes and are shown for branches with more than 50% bootstrap support. This analysis involved 53 amino acid sequences with a total of 38 positions considered in the final dataset.

Looking at other groups, we see that the divergence of the RanBP2 Znf domain does not always correspond to what is observed in the full-length protein phylogeny. For instance, the Znf of the ubiquitin ligase ARI8 is not in a clade with the other ARI proteins, reflecting its genetic distance from those proteins in spite of the similarity between the full-length ARI proteins (see "Ariadne Family" section). The second Znf of At2g02620 is particularly unusual, as it has lost both the well-conserved asparagine and the third putatively Zn^{2+}-coordinating cysteine residue (**Figure 1.4**). As in the full-length phylogeny, looking at the closest characterized relative to a Znf from an uncharacterized protein in this tree can provide hypotheses as to their function. Znfs from the uncharacterized protein At1g67325 all group with known RNA-binding Znfs from ZRANB2 as well as putatively RNA-binding Znfs from TAF15, TAF15b, and SUA, strongly suggesting that At1g67325 could be an RNA-binding protein itself. We will explore more of these possibilities in the section titled, "Uncharacterized RanBP2 Znf proteins."

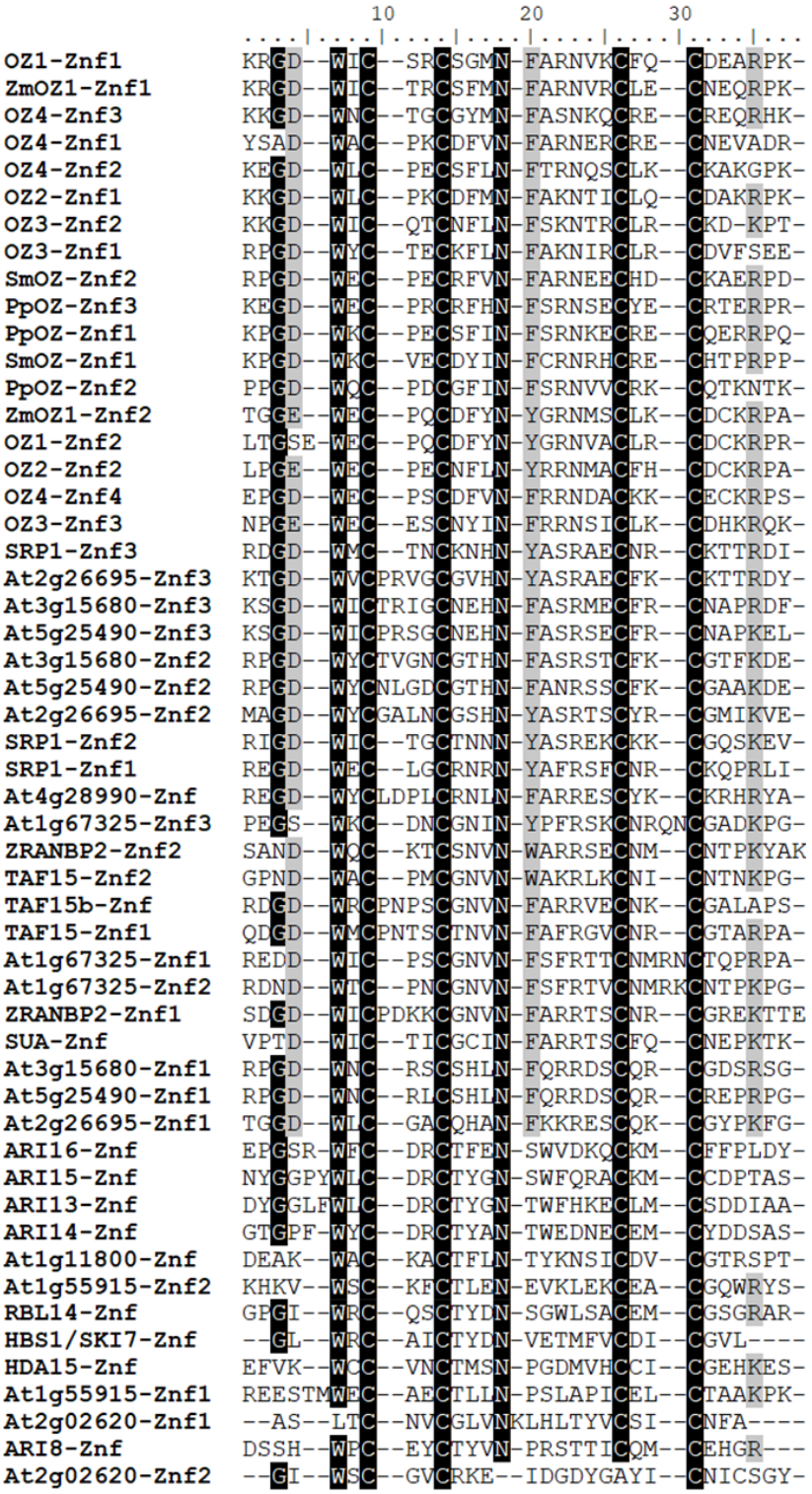


Figure 1.4. Alignment of RanBP2 Znf domains in Arabidopsis. MUSCLE alignment that was used to generate tree in Figure 1.3. Shading applied to positions with either 70% identical amino acids (black) or similar amino acids (gray).

RANBP2 PROTEINS WITH KNOWN FUNCTIONS

OZ1 and Family

The Organellar Zinc Finger (OZ) family was named as a result of the identification of OZ1 as a chloroplast RNA editing factor [10]. Although OZ1 had been described before as a variegated mutant of unknown function named VAR3, its importance for chloroplast RNA editing was not detected at that time [26]. In the *oz1-1* insertional mutant used in Sun *et al*. [10], the absence of OZ1 leads to delayed chloroplast development and decreased germination rates (**Figure 1.5**), but not variegation. The *var3* mutant is a *Ds* mutant in Landsberg *erecta*, while the *oz1-1* line is a T-DNA mutant in Columbia background; the reason for the discrepancy in phenotype is not known. The small protein family that includes OZ1, which has two Znfs, is also comprised of OZ2 (two Znfs), OZ3 (three Znfs) and OZ4 (four Znfs). These proteins have Znf domains well-conserved among them as well as a heretofore uncharacterized and uncatalogued N-terminal set of motifs unique to the family. At least two of the OZ proteins have been identified to be crucial for RNA processing in their respective organelles, and the remaining two are observed to be located in plastids and/or mitochondria as well [10,25].

Figure 1.5. WT Arabidopsis: 10 days old (top left) and 2 months old (top right); *oz1-1* homozygous mutant: 11 days old (bottom left) and 2 months old (bottom right).

It is not yet known whether the OZ1 Znf domains are capable of binding ssRNA like the ZRANB2 Znf domains. Recent experiments, however, indicate that the Znf domains of OZ1 are necessary for its function in chloroplast RNA editing. Full-length OZ1 is known to bind to other chloroplast RNA editing factors; indeed, the protein was originally discovered due to its interaction with the editing factor ORRM1 [10]. Experiments in Chapter 2 show that the Znfs, specifically Znf2 of OZ1, are necessary for the interaction with ORRM1, and this interaction may be linked to its function in RNA editing.

RHOMBOID-like Protein 14 (RBL14)

Rhomboids are membrane-anchored serine proteases chiefly involved in the proteolytic maturation, activation, or inhibition of protein targets [27]. RBL14 (named RBL10 in some studies) is considered to be in the "secretase-B" class, characterized by having six trans-membrane domain helices. The closest well-characterized homolog of RBL14 is the human protein RHBDL4 (alternatively named RHBDD1, UniProt acc. Q8TEB9), which is involved in proteolytic maturation of ER proteins, among other functions [28]. Based on the sequence similarity to ER-localized RHBDL4, Lemberg and Freeman [29] state that secretase-B rhomboids are not mitochondrial, contrary to prediction by algorithms (**Table 1.1**). However, AtRBL14 has only 44.3% sequence similarity and only 31.2% sequence identity with RHBDL4 [29]. Importantly, RHBDL4 does NOT have a Znf domain, but the rice ortholog of AtRBL14 does [30], so either the RanBP2 Znf became incorporated into this protein after the plant ancestor split from the animal/fungal ancestor, or the Znf was lost in the other lineage. Based on a loose prediction of RBL14's structure [29], the C-terminally positioned Znf of RBL14 could serve to interact with soluble domains of the protein targets unique to RBL14. There is precedent for Arabidopsis homologs of characterized RBL proteins in other taxa to have different substrate specificity; another Arabidopsis RBL, RBL12, does not process yeast targets of the corresponding yeast rhomboid, unlike the corresponding human rhomboid [31], and AtRBL1 could not cleave substrates of its *Drosophila* homolog RHO-1 [32]. This implies that the substrates of AtRBL14 may differ significantly from those of the homologous yeast, *Drosophila,* and human proteins.

Using Genevestigator, Knopf and Adam [27] determined that RBL14 (named RBL10 in their review) is expressed consistently across plant tissues and across development, but they also found that RBL14 may be upregulated in response to heat shock. An attempt to identify the location of the protein through GFP labeling was inconclusive [31]. The published images of the transfected protoplasts labeled with MitoTracker Orange do not rule out mitochondrial localization. The contribution of the Znf to the putative function of RBL14 has not been assessed, but as noted earlier, it might facilitate interaction with proteolytic targets.

Histone Deacetylase 15 (HDA15)

HDA15 was first found by querying the Arabidopsis genome sequence with histone deacetylase protein sequences from yeast, *Drosophila*, human, maize, and mouse [33]. It is considered to be part of the Class II group of histone deacetylases [33]. HDA15 is expressed highly in the stems throughout the life of the plant, and it may undergo partial export from the nucleus under dark conditions, reentering the nucleus when the protoplasts are returned to light [34].

HDA 15 is involved in the repression of chlorophyll synthesis in the dark that is mediated by PIF3, a transcription factor [35]. PIF3 binds to genes related to photosynthetic activity and represses them through the histone deacetylase activity of its binding partner, HDA15. PIF3 recruits HDA15 to the promoter regions of target genes by binding to the G-box sequence. HDA15 is one of the few proteins we reviewed for which there is direct evidence for the function of the RanBP2 Znf: a Znf-containing truncation of HDA15 binds to PIF3 *in vitro*. HDA15 binds to the

transcriptional repressor PIF3 in the absence of light, and HDA15 is confirmed to be nuclear-localized [35]. The location of HDA15 did not change in response to light treatment in the Liu *et al.* study [35], so the putative cytoplasmic shuttling proposed by Alinsug *et al.* [34] may not be related to its function in repression of genes involved in photosynthesis. HDA15 levels are not changed by light; it is the phosphorylation and degradation of PIF3 that reduces HDA15 deacetylation of histones at the affected genes. This represents a classic case of transcriptional repression via recruitment of histone modifiers [35]. HDA15 has also been recently shown to be a direct repressor of plant thermal responsive genes at normal temperature [36].

Other transcription factors interact with the RanBP2 Znf domain of HDA15 to utilize it as a transcriptional repressor. For example, PIF1 opposes seed germination in the absence of light by repressing gibberellic acid-synthesizing genes [37], and MYB96 recruits HDA15 to inhibit abscisic acid (ABA) signaling genes [38]. Nuclear Factor-YC proteins (NF-YCs) are transcription factors that function in a group with other factors to induce chromatin modifications of genes controlling germination and light responses, and they also interact with HDA15 [39]. Unfortunately, these latter interaction studies did not use truncations nor domain mutations of either protein, so we cannot draw any conclusions about which domains are required for interactions with NF-YC transcription factors. The most recent study of HDA15 found yet another protein interaction with HFR1, a transcription factor that works with HDA15 to repress the warm-temperature response [36]. That study detected protein-protein interactions between HFR1 and a Znf-containing truncation of HDA15 using yeast

two-hybrid assays. Another truncation containing the HDA15 Znf did not interact with HFR1; that construct removed the first two residues of the Znf.

TATA-binding Protein-associated Factor 15 and 15b (TAF15 and TAF15b)

Using a consensus sequence of TAF proteins built from humans, *Drosophila*, and yeast, Lago *et al.* [40] found, among other Arabidopsis TAF proteins, two that matched with human TAF15: TAF15 and TAF15b. Human TAF15 is most similar to Arabidopsis TAF15b, not to Arabidopsis TAF15. TAF15 was found to be a component of the general transcription factor TFIID. TAF15b was shown to be located in both the nucleus and cytosolic p-bodies [41]. TAF15b has an RGG-rich region in the C-terminus that may direct it to the p-bodies, based on similar domains confirmed to localize human TAF15 to p-bodies [42]. TAF15b may control the stability or RNA processing of SNC1, a Toll-like receptor involved in plant immunity, in the p-bodies [41].

TAF15b suppresses flowering under conditions of vernalization [43]. TAF15b binds to Pol II in various states of phosphorylation, as shown by co-immunoprecipitation. TAF15b represses transcription elongation of Flowering Locus C (FLC), as determined by analysis of the phosphorylation state of Pol II in wild-type (WT) vs. *taf15b* mutants. Bertolotti *et al*. [44] found that the TAF15b-related human protein TAF15 (named $hTAF_{II}68$ in that study) binds both RNA and single-stranded DNA; however, their tested truncations did not separate the RanBP2 Znf from the neighboring RNA Recognition Motif (RRM), so those binding activities cannot be assigned to one domain over the other.

Hsp70 Subfamily B Suppressor (HBS1)/Superkiller Protein 7 (SKI7)

The gene products of At5g10630 have splice isoforms predicted to have drastically different functions. The shorter splice form, named HBS1, may recognize and release stalled ribosomes, while the longer splice form, SKI7, likely incorporates into an RNA exosome [45]. These names and putative functions are based on their similarity to human and yeast homologs. The third predicted protein (UniProt ID: Q9LXB6; **Table 1.1**) is likely the product of *SKI7* regulation via alternative splicing, producing an RNA targeted for nonsense-mediated decay. The Znf domain is present in all isoforms, as are a GTP-ase domain, Patch 4-like domain, and negatively charged N-terminal region [45]. The main feature separating SKI7 from HBS1 is the presence of a 66-amino acid region called the SKI7-like motif, which is known to be the domain that tethers SKI7 to the RNA exosome [46].

A recent study showed that HBS1 contributes to degradation of 5' cleavage products from RNA silencing as well as releasing RNA from stalled ribosomes [47]. Aside from these "non-stop" RNA decay processes, carried out on RNAs lacking stop codons, HBS1 is also necessary for "no-go" decay, where transcripts are degraded when the ribosome is stalled [48]. It does this by entering the ribosome A-site along with two other proteins, Pelota2 and SKI2, and exposing the RNA to endonucleolytic cleavage. SKI7 associates with a completely different set of proteins to participate in RNA degradation; it is not an endonuclease itself, but at least one other protein in the SKI-exosome complex does have endonucleolytic activity [49]. The presence of a RanBP2 Znf in both isoforms could be an example of a protein-binding function for this Znf domain; the evidence from the reviewed studies point to the Znf being used to

bind protein partners in these RNA-processing complexes. Its similarity to the Znf of HDA15 (**Figure 1.3**), which binds to a variety of different protein partners, further supports this possibility.

Suppressor of ABI3-5 (SUA)

SUA is a splicing factor containing two RRM domains, a RanBP2 Znf, an octamer repeat region, and a glycine-rich domain [50]. SUA prevents the splicing of an intron in ABI3, a transcription factor in the ABA signaling pathway [51], which is responsible for seed maturation; if spliced, this intron produces an ABI3 truncation in WT plants that does not function. *abi3-5* mutant plants are insensitive to ABA and produce green seeds with low viability, but these mutants produce a functional form of ABI3 when that intron is spliced in *abi3-5 sua-1* double mutants. Expression of SUA:GFP under the control of the SUA promoter in *abi3-5 sua-1* double knockout plants caused them to revert to the green seed phenotype of the *abi3-5* mutant. SUA is concentrated in the nucleus and binds with the spliceosomal factor U2AF65, suggesting that SUA is directly integrated with the spliceosome [50].

SNC4, a receptor-like kinase involved in plant immunity, retains an intron in *sua* mutants. SUA is also needed for splicing of CERK1, another receptor-like kinase, and lack of CERK1 splicing in *sua* mutants degrades their ability to resist pathogens [52]. Both the Sugliani *et al.* [50] and Zhang *et al.* [52] studies implicate SUA as a splicing factor. It is possible that the Znf in SUA binds to either the 5’ or 3’ splice site in a manner similar to ZRANB2.

Stress Associated RNA-binding Protein 1 (SRP1)

First discovered for its homology to a stress response RNA-binding protein in rice, Stress Associated RNA-binding Protein 1 (SRP1) has been recently described as a post-transcriptional regulatory protein containing three RanBP2 Znf domains [53]. Its specific target is the 3' UTR of *ABI2*, a transcript coding for a phosphatase involved in regulation of the ABA signaling pathway [51]. The expression of known ABA signaling genes was altered in an opposite manner in *srp1* knockouts vs. SRP1 overexpression plants (e.g., ABI2 expression increased in *srp1* mutants vs. WT but was decreased in SRP1 overexpression plants). Interestingly, ABI2 appears to be negatively regulated by SRP1 binding. It was shown to bind to 3'-UTR RNAs *in vitro*, potentially through the AUUUA sequences. This is the first Arabidopsis RanBP2 Znf-containing protein demonstrated to have RNA binding activity likely mediated by its Znf domain, as there are no other known motifs in SRP1.

Ariadne Family

The Ariadne (ARI) family is a group of putative E3-type ubiquitin ligases [22], the final enzymatic actor in a chain of reactions that ubiquitinate protein targets, marking them for degradation [54]. ARI proteins are typified by a pair of RING finger-type Znf domains, an unusual Znf typical of E3 ubiquitin ligases that binds two Zn^{2+} atoms with a 3-Cys, 1-His, 4-Cys motif [55]. The ARI proteins mentioned in this study (**Table 1.1**) are distinguished from the other Arabidopsis ARI proteins by the presence of a single RanBP2 Znf at the extreme C-terminal end. *ARI13*, *ARI14*, *ARI15*, and *ARI16* are close together on chromosome 5 of Arabidopsis and share similar gene

architectures, which led Mladek *et al.* [22] to hypothesize that the cluster resulted from a series of gene duplication events; the subsequent diversification has led to notable functional differences. For example, ARI14 is regulated by a native siRNA, *KOKOPELLI* (*KPL*), in sperm cells to potentially counteract the ubiquitin-mediated protein degradation effected by ARI13 [56]. Due to ARI14's mutations in the N-terminal RING finger domain, Ron *et al.* [56] speculated that it cannot act as an E3 ubiquitin ligase but can still bind to the ubiquitination complex in sperm cells, thus acting as a negative regulator of that activity.

The gene encoding ARI8 is located on a different chromosome and has an architecture much different from the ARI13/14/15/16 group. ARI8 has been demonstrated to perform ubiquitination in conjunction with a number of E2-type conjugating enzymes, as determined by *in vitro* assay [23]; these findings are the basis for considering the other ARI proteins to be ubiquitin ligases. Despite the evolutionary distance of ARI8 from the ARI13/14/15/16 group, it also has a C-terminal RanBP2 Znf; however, the Znf in ARI8 does differ in sequence from the other ARI Znfs (**Figure 1.3**, **Figure 1.4**). The presence of the C-terminal RanBP2 Znf was not noted in the original description of the family [22], and no subsequent study has explored its function [23,56,57]. The close relationship of ARI RanBP2 Znf domains with that of HDA15, a protein which interacts with other proteins (**Figure 1.3**), suggests that they should be investigated further to determine whether they may mediate protein-protein interactions.

UNCHARACTERIZED RANBP2 ZNF PROTEINS

After having surveyed the literature on Arabidopsis RanBP2 Znf proteins, we can now discuss the uncharacterized proteins and hypothesize as to their function by comparing with closely related characterized proteins and looking at the other domains present in the protein using NCBI Conserved Domain Search [58]. We can generally see that, in cases where the RanBP2 Znf is RNA-binding, it has a specific sequence to which it binds (e.g., ZRANB2 and SRP1). In contrast, several proteins containing RanBP2 Znfs apparently interact with multiple protein partners (e.g., HDA15 and RBL14). At1g55915 is on a node with HDA15 in the full-length tree (**Figure 1.2**) and possesses a WLM domain, which is a putative metalloprotease domain. If it does indeed function as a protease, perhaps the two Znf domains of At1g55915 conduct multiple protein interactions to facilitate proteolysis of multiple protein targets. At1g11800 is worth noting because it is predicted to be mitochondrially-localized. It contains a TDP2 domain, putatively a phosphodiesterase domain. This protein groups with HBS1/SKI7 (**Figure 1.2**), and so At1g11800 could have either ribosome-rescuing and/or RNA degradation duties in the mitochondrion. At4g28990 is a simple protein with just one Znf and no other detectable domains. It groups with TAF15 and TAF15b, but the Znf groups with SRP1-Znf1 (**Figure 1.3**). These associations suggest that At4g28990 may bind to RNA, but experimentation will be needed to learn the function of this protein and others. At1g67325, mentioned in the section, "SUBCELLULAR LOCATION OF RANBP2 ZINC FINGERS IN ARABIDOPSIS," may have three RNA-binding Znfs (**Figure 1.3**), but the full-length protein loosely groups with HDA15 and RBL14, one of which is confirmed to have a protein-binding

Znf. At2g02620 is a sister protein to SRP1 in the full-length phylogeny, but neither of its Znfs cluster with those of SRP1 (**Figure 1.3**); indeed, its first Znf is missing the signature tryptophan, and the second zinc finger has lost one of the four signature cysteines (**Figure 1.4**). This may be a protein for which the Znf is no longer necessary for its function. At5g25490, At2g26695, and At3g15680 are all small proteins with three RanBP2 Znfs. They are in the same clade as SRP1 (**Figure 1.2**) and may be additional RNA-binding proteins affecting the turnover and/or stability of transcripts. This group is the least studied of all the clades mentioned here and could represent a unique family of RNA-regulating factors.

CONCLUSIONS AND FUTURE DIRECTIONS

Determining the function of these RanBP2 Znf domains in each protein can be a simple matter of producing truncated constructs of the protein that include and exclude the Znfs, or contain mutated Znfs, and testing those constructs for protein interaction and function in knockout mutants *in vivo*, as was done with HDA15. Expressing the domain itself to perform protein and RNA interaction experiments could help identify the function of Znf-containing proteins. Many of the proteins in this review are members of larger families, but they are distinguished from those families by their possession of a RanBP2 Znf. How does this domain contribute to the function of the protein in which it is located? During evolution, why has this domain become located in such a variety of proteins? What functional differences arise from having one versus three or four Znfs? These are questions that can only be answered by directly studying this versatile protein domain in multiple proteins.

REFERENCES

1. Singh, B. B., Patel, H. H., Roepman, R., Schick, D. & Ferreira, P. A. The zinc finger cluster domain of RanBP2 is a specific docking site for the nuclear export factor Exportin-1. *J. Biol. Chem.* **274**, 37370–37378 (1999).

2. Higa, M. M., Alam, S. L., Sundquist, W. I. & Ullman, K. S. Molecular characterization of the Ran-binding zinc finger domain of Nup153. *J. Biol. Chem.* **282**, 17090–17100 (2007).

3. Adams, D. J., van der Weyden, L., Mayeda, A., Stamm, S., Morris, B. J. & Rasko, J. E. J. ZNF265—a novel spliceosomal protein able to induce alternative splicing. *J. Cell Biol.* **154**, 25–32 (2001).

4. Yang, Y. H. J., Markus, M. A., Mangs, A. H., Raitskin, O., Sperling, R. & Morris, B. J. ZRANB2 localizes to supraspliceosomes and influences the alternative splicing of multiple genes in the transcriptome. *Mol. Biol. Rep.* **40**, 5381–5395 (2013).

5. Loughlin, F. E. *et al.* The zinc fingers of the SR-like protein ZRANB2 are single-stranded RNA-binding domains that recognize 5' splice site-like sequences. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 5581–6 (2009).

6. Roca, X., Krainer, A. R. & Eperon, I. C. Pick one, but be quick: 5' splice sites and the problems of too many choices. *Genes Dev.* **27**, 129–144 (2013).

7. Brown, J. W. S., Smith, P. & Simpson, C. G. Arabidopsis consensus intron sequences. *Plant Mol. Biol.* **32**, 531–535 (1996).

8. Li, X. *et al.* ZRANB2/SNHG20/FOXK1 Axis regulates vasculogenic mimicry formation in glioma. *J. Exp. Clin. Cancer Res.* **38**, 68 (2019).

9. Nguyen, C. D., Mansfield, R. E., Leung, W., Vaz, P. M., Loughlin, F. E., Grant, R. P. & Mackay, J. P. Characterization of a family of RanBP2-type zinc fingers that can recognize single-stranded RNA. *J. Mol. Biol.* **407**, 273–283 (2011).

10. 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.* **11**, 1–23 (2015).

11. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).

12. Mitchell, A. L. *et al.* InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **47**, D351–D360 (2019).

13. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**, 680–682 (2010).

14. Hooper, C. M., Castleden, I. R., Tanz, S. K., Aryamanesh, N. & Millar, A. H. SUBA4: the interactive data analysis centre for Arabidopsis subcellular protein locations. *Nucleic Acids Res.* **45**, D1064–D1074 (2017).

15. Millar, A. H., Sweetlove, L. J., Giegé, P. & Leaver, C. J. Analysis of the Arabidopsis mitochondrial proteome. *Plant Physiol.* **127**, 1711–1727 (2001).

16. Small, I., Peeters, N., Legeai, F. & Lurin, C. Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* **4**, 1581–90 (2004).

17. Inzé, A., Vanderauwera, S., Hoeberichts, F. A., Vandorpe, M., Van Gaever, T. & Van Breusegem, F. A subcellular localization compendium of hydrogen peroxide-induced proteins. *Plant. Cell Environ.* **35**, 308–20 (2012).

18. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).

19. Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. & Huala, E. The Arabidopsis Information Resource: Making and Mining the 'Gold Standard' Annotated Reference Plant Genome. doi:10.1002/dvg.22877.

20. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

21. Whelan, S. & Goldman, N. A General Empirical Model of Protein Evolution

Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. *Mol. Biol. Evol.* **18**, 691–699 (2001).

22. Mladek, C., Guger, K. & Hauser, M. T. Identification and characterization of the ARIADNE gene family in Arabidopsis. A group of putative E3 ligases. *Plant Physiol.* **131**, 27–40 (2003).

23. Kraft, E., Stone, S. L., Ma, L., Su, N., Gao, Y., Lau, O., Deng, X. & Callis, J. Genome Analysis and Functional Characterization of the E2 and RING-Type E3 Ligase Ubiquitination Enzymes of Arabidopsis. *Plant Physiol.* **139**, 1597–1611 (2005).

24. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).

25. Gipson, A., Bentolila, S., Kehl, A. & Hanson, M. R. *Unpublished*.

26. Naested, H. *et al.* Arabidopsis VARIEGATED 3 encodes a chloroplast-targeted, zinc-finger protein required for chloroplast and palisade cell development. *J. Cell Sci.* **117**, 4807–18 (2004).

27. Knopf, R. R. & Adam, Z. Rhomboid proteases in plants - still in square one? *Physiol. Plant.* **145**, 41–51 (2012).

28. Wunderle, L., Knopf, J. D., Kühnle, N., Morlé, A., Hehn, B., Adrain, C., Strisovsky, K., Freeman, M. & Lemberg, M. K. Rhomboid intramembrane protease RHBDL4 triggers ER-export and non-canonical secretion of membrane-anchored TGFα. *Sci. Rep.* **6**, 27342 (2016).

29. Lemberg, M. K. & Freeman, M. Functional and evolutionary implications of enhanced genomic analysis of rhomboid intramembrane proteases. *Genome Res.* **17**, 1634–1646 (2007).

30. Tripathi, L. P. & Sowdhamini, R. Cross genome comparisons of serine proteases in Arabidopsis and rice. *BMC Genomics* **7**, 200 (2006).

31. Kmiec-Wisniewska, B., Krumpe, K., Urantowka, A., Sakamoto, W., Pratje, E.

& Janska, H. Plant mitochondrial rhomboid, AtRBL12, has different substrate specificity from its yeast counterpart. *Plant Mol. Biol.* **68**, 159–171 (2008).

32. Kanaoka, M. M., Urban, S., Freeman, M. & Okada, K. An Arabidopsis Rhomboid homolog is an intramembrane protease in plants. *FEBS Lett.* **579**, 5723–8 (2005).

33. Pandey, R., Müller, A., Napoli, C. A., Selinger, D. A., Pikaard, C. S., Richards, E. J., Bender, J., Mount, D. W. & Jorgensen, R. A. Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Res.* **30**, 5036–5055 (2002).

34. Alinsug, M. V, Chen, F. F., Luo, M., Tai, R., Jiang, L. & Wu, K. Subcellular Localization of Class II HDAs in Arabidopsis thaliana: Nucleocytoplasmic Shuttling of HDA15 Is Driven by Light. *PLoS One* **7**, e30846 (2012).

35. Liu, X. *et al.* PHYTOCHROME INTERACTING FACTOR3 Associates with the Histone Deacetylase HDA15 in Repression of Chlorophyll Biosynthesis and Photosynthesis in Etiolated Arabidopsis Seedlings. *Plant Cell* **25**, 1258–1273 (2013).

36. Shen, Y., Lei, T., Cui, X., Liu, X., Zhou, S., Zheng, Y., Guérard, F., Issakidis-Bourguet, E. & Zhou, D. Arabidopsis histone deacetylase HDA15 directly represses plant response to elevated ambient temperature. *Plant J.* **100**, 991–1006 (2019).

37. Gu, D., Chen, C.-Y., Zhao, M., Zhao, L., Duan, X., Duan, J., Wu, K. & Liu, X. Identification of HDA15-PIF1 as a key repression module directing the transcriptional network of seed germination in the dark. *Nucleic Acids Res.* **45**, 7137–7150 (2017).

38. Lee, H. G. & Seo, P. J. MYB96 recruits the HDA15 protein to suppress negative regulators of ABA signaling in Arabidopsis. *Nat. Commun.* **10**, 1713 (2019).

39. Tang, Y., Liu, X., Liu, X., Li, Y., Wu, K. & Hou, X. Arabidopsis NF-YCs Mediate the Light-Controlled Hypocotyl Elongation via Modulating Histone Acetylation. *Mol. Plant* **10**, 260–273 (2017).

40. Lago, C., Clerici, E., Mizzi, L., Colombo, L. & Kater, M. M. TBP-associated factors in Arabidopsis. *Gene* **342**, 231–241 (2004).

41. Dong, O. X. *et al.* Arabidopsis TAF15b Localizes to RNA Processing Bodies and Contributes to snc1 -Mediated Autoimmunity. *Mol. Plant-Microbe Interact.* **29**, 247–257 (2016).

42. Marko, M., Vlassis, A., Guialis, A. & Leichter, M. Domains involved in TAF15 subcellular localisation: Dependence on cell type and ongoing transcription. *Gene* **506**, 331–338 (2012).

43. Eom, H., Park, S. J., Kim, M. K., Kim, H., Kang, H. & Lee, I. TAF15b, involved in the autonomous pathway for flowering, represses transcription of FLOWERING LOCUS C. *Plant J.* **93**, 79–91 (2018).

44. Bertolotti, A., Lutz, Y., Heard, D. J., Chambon, P. & Tora, L. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J.* **15**, 5022–5031 (1996).

45. Brunkard, J. O. & Baker, B. A Two-Headed Monster to Avert Disaster: HBS1/SKI7 Is Alternatively Spliced to Build Eukaryotic RNA Surveillance Complexes. *Front. Plant Sci.* **9**, 1–17 (2018).

46. Kalisiak, K., Kuliński, T. M., Tomecki, R., Cysewski, D., Pietras, Z., Chlebowski, A., Kowalska, K. & Dziembowski, A. A short splicing isoform of HBS1L links the cytoplasmic exosome and SKI complexes in humans. *Nucleic Acids Res.* **45**, 2068–2080 (2016).

47. Szádeczky-Kardoss, I., Csorba, T., Auber, A., Schamberger, A., Nyikó, T., Taller, J., Orbán, T. I., Burgyán, J. & Silhavy, D. The nonstop decay and the RNA silencing systems operate cooperatively in plants. *Nucleic Acids Res.* **46**, 4632–4648 (2018).

48. Szádeczky-Kardoss, I., Gál, L., Auber, A., Taller, J. & Silhavy, D. The No-go decay system degrades plant mRNAs that contain a long A-stretch in the coding region. *Plant Sci.* **275**, 19–27 (2018).

49. Lange, H. *et al.* RST1 and RIPR connect the cytosolic RNA exosome to the Ski

complex in Arabidopsis. *Nat. Commun.* **10**, 3871 (2019).

50. Sugliani, M., Brambilla, V., Clerkx, E. J. M., Koornneef, M. & Soppe, W. J. J. The Conserved Splicing Factor SUA Controls Alternative Splicing of the Developmental Regulator ABI3 in Arabidopsis. *Plant Cell* **22**, 1936–1946 (2010).

51. Finkelstein, R. R., Gampala, S. S. L. & Rock, C. D. Abscisic acid signaling in seeds and seedlings. *Plant Cell* **14 Suppl**, S15–S45 (2002).

52. Zhang, Z., Liu, Y., Ding, P., Li, Y., Kong, Q. & Zhang, Y. Splicing of Receptor-Like Kinase-Encoding SNC4 and CERK1 is Regulated by Two Conserved Splicing Factors that Are Required for Plant Immunity. *Mol. Plant* **7**, 1766–1775 (2014).

53. Xu, J. *et al.* A Novel RNA-Binding Protein Involves ABA Signaling by Post-transcriptionally Repressing ABI2. *Front. Plant Sci.* **8**, 1–12 (2017).

54. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479 (1998).

55. Joazeiro, C. A. P. & Weissman, A. M. RING Finger Proteins. *Cell* **102**, 549–552 (2000).

56. Ron, M., Alandete Saez, M., Eshed Williams, L., Fletcher, J. C. & McCormick, S. Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in Arabidopsis. *Genes Dev.* **24**, 1010–1021 (2010).

57. Marín, I. & Ferrús, A. Comparative Genomics of the RBR Family, Including the Parkinson's Disease–Related Gene Parkin and the Genes of the Ariadne Subfamily. *Mol. Biol. Evol.* **19**, 2039–2050 (2002).

58. Marchler-Bauer, A. *et al.* CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* **45**, D200–D203 (2017).

CHAPTER 2

The RanBP2 zinc finger domains of chloroplast RNA editing factor OZ1 are required for protein-protein interactions and conversion of C to U

ABSTRACT

In the chloroplast, OZ1 is a RanBP2-type zinc finger protein required for many RNA editing events, a process by which specific cytosines are enzymatically converted to uracils as a correction mechanism for missense mutations in the organelle genomes. RNA editing is carried out by a large multi-protein complex called the "editosome" that contains members of the PPR protein family, the RIP/MORF family, and the ORRM family in addition to OZ1. OZ1 is an 82-kDa protein with distinct domains, including a pair of zinc finger domains and a unique C-terminal region. To elucidate the functions of these domains, I have generated truncations of OZ1 for use in protein-protein interaction assays that identified the C-terminal region of OZ1 as well as the zinc finger domains as the primary interactor with PPR proteins, the editosome factors required for site-specificity and enzymatic editing. Expression of these OZ1 constructs *in vivo* showed that the zinc finger domains were required to restore editing in *oz1* knockout plants. Mutation of key structural residues in the zinc finger domains showed that they are necessary for editing and required for interaction with ORRM1, a general editing factor with an RNA-binding domain.

INTRODUCTION

RNA editing in plants compensates for missense mutations in organellar genes by converting select cytidines in mRNA to uridines [1–3]. To date in *Arabidopsis thaliana*, 43 sites in the chloroplast and over 600 sites in the mitochondria have been

identified as targets for editing [4,5]. These editing events are important for proper plant growth. Many cases of mutant Arabidopsis plants that lack editing at key sites are characterized by defects linked to incomplete development of the chloroplasts or mitochondria, such as lack of pigmentation, slow growth in the absence of an external sugar supply, or embryo lethality [6–9].

The "editosome" complex is composed of nuclear-encoded proteins whose diversity and subfunctions are not fully characterized. Evidence from genetic assays and protein interaction studies revealed a core set of four protein families: the numerous pentatricopeptide repeat (PPR) proteins [10], the RNA editing factor interacting proteins (RIPs), also known as MORFs [11,12], the organelle RNA-recognition motif (ORRM) protein family [13], and the organelle zinc finger family (OZ) [14,15]. The particular protein members of the families within an editosome differ from one edited site to the next. The identification of these editing factors and their interactions with each other and the RNA substrate have led to a model for the editosome, but the individual contributions of these factors to the editing process is still being uncovered.

Plant RNA editing factors exhibit a variety of functionally distinct domains, such as the RNA-binding, editing-essential RRM domain vs. the protein-binding Gly-rich regions of ORRM proteins [7,9,16], and the RNA-binding PPR tract, protein-binding E-domains, and nucleotide deaminase DYW domain of editing PPR proteins [17–22]. OZ1 also contains a number of distinct domains: a group of N-terminal motifs shared by all four members of the OZ family, two zinc finger (Znf) domains, and a C-terminal region specific to OZ1 that contains three spaced motifs with strong

conservation across many species [14]. Many Znf proteins perform various roles in plant physiology, with some of these roles being as yet unidentified [15]. The particular subfamily of Znf in the OZ family is the RanBP2 family, named for the nuclear pore factor that binds to the Ran nuclear localization factor [23]. Previous analysis [14] showed that the RanBP2 Znfs in OZ1 are similar to those of ZRANB2, a human splicing factor known to bind to single-stranded RNA via its Znf domains [24].

As an 82-kDa protein with several predicted domains, I hypothesized that OZ1 would have similar specialized domain functions as other editing factors. Based on the three-part domain structure, I prepared truncated forms of OZ1 for protein-protein interaction testing in yeast two-hybrid (Y2H) and bimolecular fluorescence complementation (BiFC) assays with known binding partners from the other editosome protein families. These experiments have identified the C-terminal domain of OZ1 and the zinc finger domains as the primary interactors with PPR proteins, whereas the Znf domains alone are necessary for interaction with ORRM1. Using mutational analysis of the Znf domains of the editing factor OZ1, I have demonstrated the domains' importance to the contribution of OZ1 to RNA editing. The ability of OZ1 truncations consisting of either the Znf domains alone or Znfs with a C-terminal domain to rescue RNA editing defects in the knockout mutants further establishes that the Znfs of OZ1 are necessary for its role in editing at target sites.

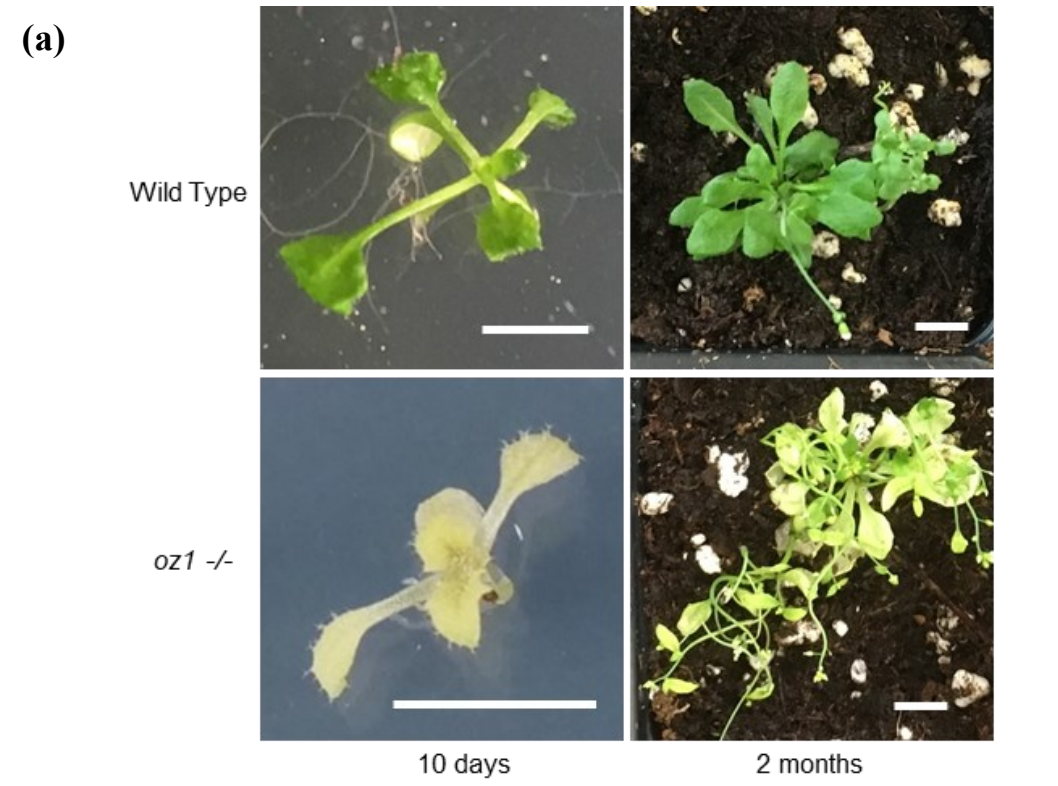
RESULTS

Domain boundary analysis of OZ1 predicts several topologically independent regions

oz1 knockout mutants exhibit reduced seed viability and early chlorosis (**Figure 2.1a**), although they eventually develop functional chloroplasts that allow them to grow on soil autotrophically [14].

To determine possible secondary structure and delineate domains in preparation for cloning truncation constructs, I performed domain boundary analysis on the OZ1 protein sequence using a suite of online tools for structure prediction [25]. My analysis determined cutoff points for three putative functional regions: the N-terminal domain, shared between all four OZ family proteins and spanning OZ1 residues 34–273; the pair of RanBP2 zinc fingers spanning residues 274–337; and the C-terminal stretch, unique to OZ1 and spanning residues 358–stop (**Figure 2.1b**).

Figure 2.1. OZ1 is necessary for normal plant development and consists of multiple domains. (**a**) WT Col and *oz1* homozygous mutant Arabidopsis plants at 10 days (left two images; scale bars = 5 mm) were grown in Magenta boxes, and at two months (right two images; scale bars = 1 cm) they were transferred to soil. (**b**) Map of OZ1 domains as identified in Sun *et al*. [14] and truncations used in this study. Numbers above cartoons label the first and last amino acid residues of each construct. All constructs were cloned with the first 33 amino acids removed (the estimated size of the transit peptide) and then fused with the first 65 amino acids of RecA as a replacement chloroplast transit peptide at the N-terminus. Red = 19 residue-long motif, blue = 60 residue-long motif, purple = 16 residue-long motif, teal = RanBP2 zinc fingers, yellow = 47 residue-long motif.


(a)
Wild Type
oz1 -/-
10 days
2 months
(b)
Full-length
34
758
N-terminus+Znf
34
337
N-terminus
34
273
Znf
274
337
Znf+C-terminus
274
758
C-terminus
358
758

The original analysis of the OZ family protein sequences using MEME [14] showed the presence of four unique sequence motifs. The N-terminal region of every OZ protein (**Figure 2.1b**) contains a 19 residue-long motif (red), a 60 residue-long motif (blue), and a 16 residue-long motif (purple), all with variation in the length of linkers between them. Based on the domain boundary analysis of OZ1, the 19 residue-long motif is predicted to contain an alpha helix; the 60 residue-long motif contains three alpha helices along with a short, disordered region; and the 16 residue-long motif contains another short alpha helix. The fourth motif (**Figure 2.1b**) identified is unique to the long C-terminal region of OZ1, which contains three of these 47 residue-long motifs. Domain boundary analysis was inconsistent with secondary structure prediction of the three C-terminal motifs, predicting two alpha helices in the most C-terminal motif but largely disordered regions in the first two such motifs.

The C-terminal domain of OZ1 mediates protein-protein interactions with other editing factors except for ORRM1

Previous work identified OZ1 as a protein binding partner with the chloroplast RNA editing factor ORRM1 along with chloroplast editing PPR proteins CRR28 and OTP82 [14]. To evaluate which domain(s) of OZ1 participate in these interactions, I prepared truncated constructs of OZ1 based on our domain boundary analysis (**Figure 2.1b**) for expression in a GAL4-based yeast two-hybrid assay. Along with CRR28 and OTP82, I included RARE1 and QED1 (named in older literature as "OTP81") in the assays because these PPR proteins share target editing sites with those of OZ1 [14,17,26,27]. Finally, I assayed for interaction with DYW2, an unusual member of

the PPR family in having a very short tract of just five PPR domains upstream of the E and DYW domain [28]; DYW2 associates with so-called E-type PPR proteins lacking a DYW domain of their own and provides editing catalysis *in trans* to those editosomes [29,30]. Testing these constructs for interaction with RARE1, CRR28, OTP82, QED1, and DYW2 showed that the C-terminal domain construct alone was able to establish the interactions that occur between those editing factors and mature-length OZ1 (**Table 2.1**; **Figure 2.2**).

Table 2.1. Summary of Y2H and BiFC interaction assay results between truncations of OZ1 and known RNA-editing factor binding partners.

	RARE1		OTP82		CRR28		QED1		DYW2		ORRM1	
OZ1_full-length	Yes†	Yes†	Yes	Yes	Yes	Yes	Yes†	Yes	Yes	Yes	Yes	Yes
OZ1_N-terminus	No	Yes	No	Yes	No	Yes	No	No	No	Yes	No	No
OZ1_N-terminus+Znf	No	No	Yes	Yes	Yes	Yes	No	Yes†	Yes	Yes	Yes	Yes
OZ1_Znf	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes†	Yes	Yes
OZ1_Znf+C-terminus	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
OZ1_C-terminus	Yes†	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No
	Y2H	BiFC	Y2H	BiFC	Y2H	BiFC	Y2H	BiFC	Y2H	BiFC	Y2H	BiFC

†Weaker than average interactions, based on growth density in Y2H or fluorescent puncta abundance in BiFC.

These interactions also occurred with both the N-terminal domain plus the zinc finger domains (N-terminus+Znf) or the zinc finger domains alone (Znf), albeit with a weaker strength (Fig. 2a). I also observed an interaction between ORRM1 and both the N-terminus+Znf truncation of OZ1 as well as the Znf alone (**Figure 2.2**). This interaction was the only one that was not sustained by the C-terminal domain of OZ1 (**Figure 2.2**). The N-terminal domain of OZ1 did not show interaction with any of the

editing factors tested (**Figure 2.2**). Furthermore, it seems that this N-terminal domain has a repressive effect on the binding ability of the zinc finger (Znf) domain. The Znf alone is able to interact with all the editing factors tested, including QED1, while N-terminus+Znf does not show an interaction with QED1. Moreover, the Znf alone showed interaction with DYW2 and ORRM1 at the lowest concentration tested, which is not observed in the N-terminus+Znf truncation of OZ1.

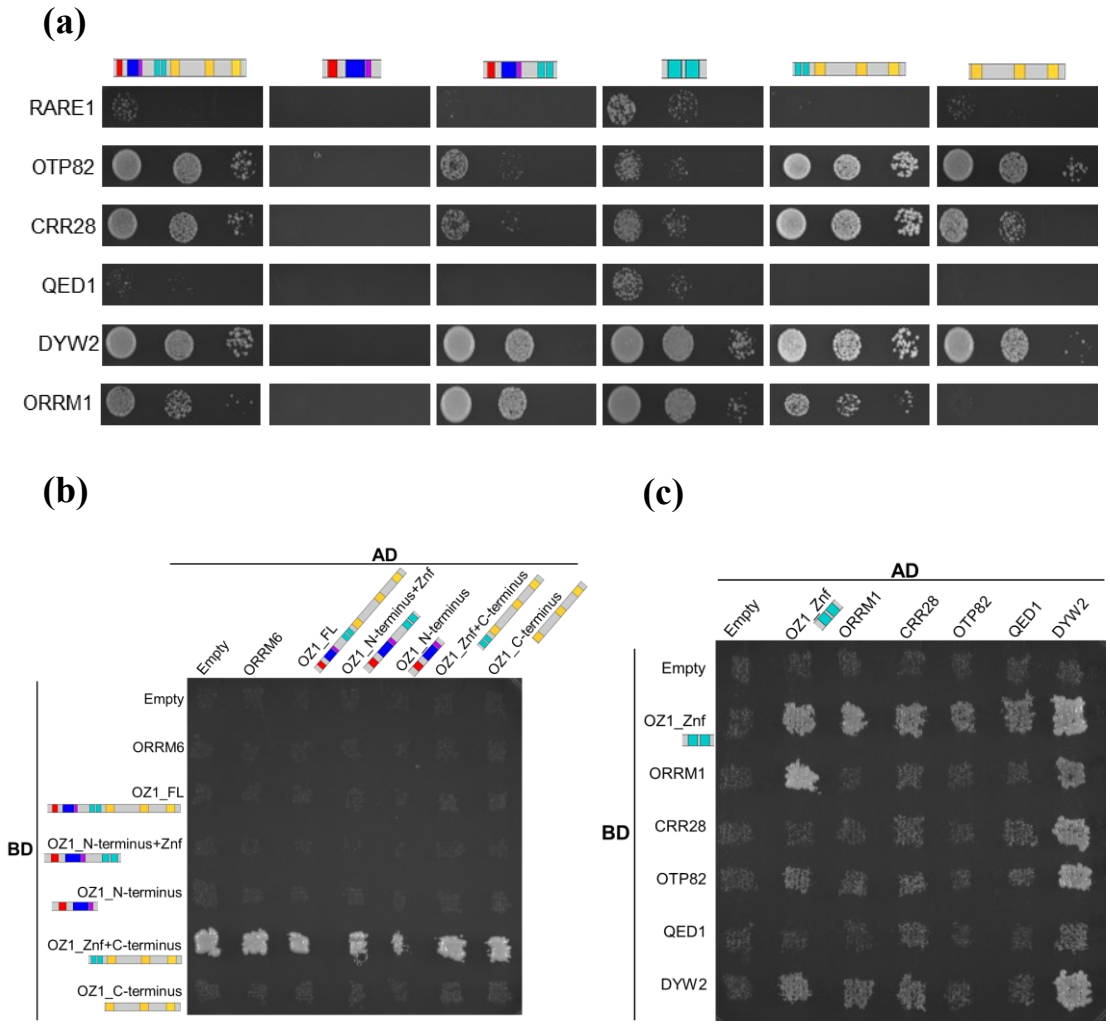


Figure 2.2. Yeast two-hybrid shows that the Znfs of OZ1 bind to other editing factors, and the C-terminal domain of OZ1 mediates interaction with most other editing factors except ORRM1. (**a**) Yeast two-hybrid tests between mature-length OZ1 or OZ1 truncations (OZ1_FL, OZ1_N-terminus, OZ1_N-terminus+Znf, OZ1_Znf, OZ1_Znf+C-terminus, and OZ1_C-terminus) and PPR proteins (RARE1, CRR28, OTP82, QED1), DYW2, or ORRM1. (**b**, **c**) Negative control yeast two-hybrid assays. Plates were prepared by streaking lines of individual yeast with proteins of interest either in the pGADT7 (AD) prey plasmid or the pGBKT7 (BD) bait plasmid onto selective media (-Leu for pGADT7, -Trp for pGBKT7), stamping onto YPAD media in a matrix pattern to produce diploid yeast at the intersections, then stamping onto diploid selective media (-Leu -Trp), and finally stamping onto assay media (-Leu -Trp -His -Ade). (**b**) Matrix Y2H of OZ1 truncations and ORRM6. No construct grows when paired with an empty vector except for BD-OZ1_Znf+C-terminus, which is auto-activating and thus not used in panel a; (**c**) matrix Y2H of OZ1_Znf and other editing factors used in this study.

I confirmed these yeast two-hybrid results through bimolecular fluorescence complementation (BiFC) in *N. benthamiana* leaves (**Figure 2.3**; **Table 2.1**). In several cases, BiFC detected interactions that had not been observed in the yeast two-hybrid assays, such as the interaction of QED1 and the OZ1 C-terminus, and most notably the interaction of the N-terminal domain of OZ1 with RARE1, OTP82, CRR28, and DYW2 (**Figure 2.3g–i, k**; **Table 2.1**). In conclusion, when compiling the results from both the Y2H and the BiFC assays, the Znf domain alone is able to recapitulate all the interactions observed for the full length OZ1. This observation is particularly relevant for ORRM1, for which the presence of the Znf is both sufficient and necessary for the interaction to occur (**Table 2.1**).

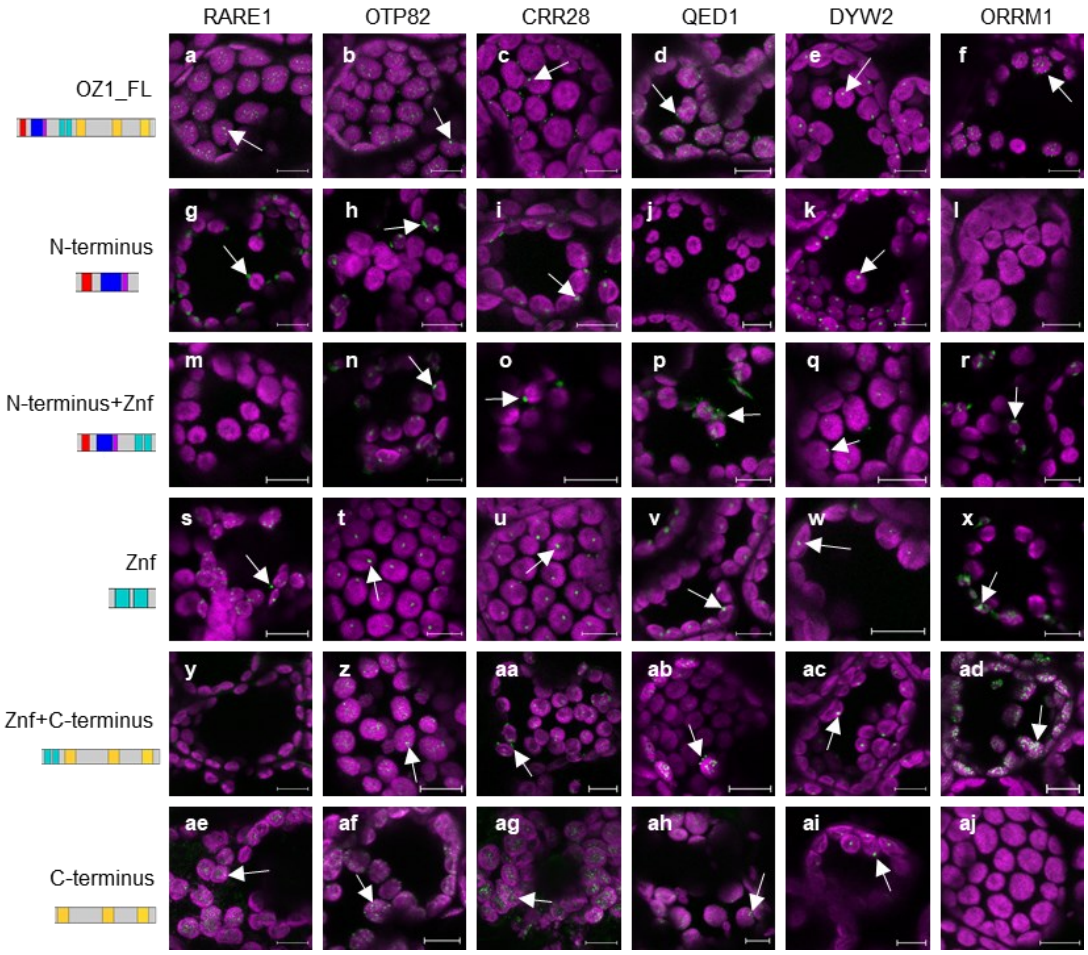


Figure 2.3. Bimolecular fluorescence complementation of OZ1 truncations with other editing factors shows that all domains interact with PPR proteins, but ORRM1 only interacts with truncations containing the Znf domains. (**a-aj**) Confocal images of interactions between OZ1 truncations and RARE1, OTP82, CRR28, QED1, DYW2, and ORRM1. Arrows indicate fluorescent puncta indicative of a positive interaction. Magenta = chlorophyll autofluorescence; green = YFP. Scale bars = 10 µm.

The zinc finger domains of OZ1 are essential for chloroplast RNA editing

After observing *in vitro* that the Znf domains of OZ1 are responsible for all of the protein-protein interactions, I produced stable transformants of *oz1* homozygous plants expressing OZ1 truncations under control of the 35S promoter and targeted to the

chloroplast with the RecA transit peptide (the first 65 amino acids of RecA). Seedlings and mature plants expressing the full-length OZ1 construct resembled WT Arabidopsis, but plants expressing truncation constructs exhibited varying degrees of chlorosis and delayed growth (**Figure 2.4a**). The plants expressing the N-terminal part of OZ1 or the N-terminus+Znf showed severe chlorosis like the *oz1* mutant plant. On the other hand, the transgenic plants expressing constructs with the Znf alone, Znf+C-terminus, or the C-terminal part of OZ1 exhibited some greening (**Figure 2.4a**). I wanted to find out whether the restoration of greening could be coupled with an ability of the constructs to restore plastid editing extent. After chloroplast RNA extraction and quantitative RT-PCR experiments, I assayed several chloroplast sites for their editing extent in the various OZ1 truncation transgenic plants. The *ndhB*-C1255 editing site exhibited severe reduction of its editing extent in the *oz1* mutant compared to WT, and expression of full-length OZ1 in the *oz1* mutant rescued editing at that site (**Figure 2.4b**). Expression of the OZ1 N-terminus, N-terminus+Znf, and C-terminus did not rescue *ndhB*-C1255 editing, but the Znf+C-terminus and Znf did rescue editing, albeit not to the extent of the full-length construct in the case of the OZ1_Znf construct (**Figure 2.4b**). Other sites repeated this pattern, where the Znf and the Znf+C-terminus truncation are able to rescue editing (**Figure Apx3.1**). The ability of these constructs to rescue editing is also apparently concentration-dependent, as biological replicates that have lower expression levels of either the Znf or Znf+C-terminus construct have correspondingly lower editing extents (**Figure 2.5a**; **Figure Apx3.1**). However, most other chloroplast editing sites, such as *ndhB*-C467, remain at the *oz1* knockout editing level in the OZ1_Znf truncation-expressing transgenic plants

but are rescued in the Znf+C-terminus transgenic plants (**Figure 2.4b**; **Figure Apx3.1**). More importantly, the inability of certain truncation constructs to restore editing extent in the *oz1* mutant, such as the N-terminus or N-terminus+Znf, cannot be attributed to a defective level of expression; both of these constructs are expressed at a higher or similar level relative to the OZ1_Znf constructs in all the transgenic plants tested (**Figure 2.5a**). In conclusion, there is a good correspondence between the morphological aspect (level of greening) of the transgenic plants expressing the different truncation constructs and the restoration of plastid editing sustained by these constructs.

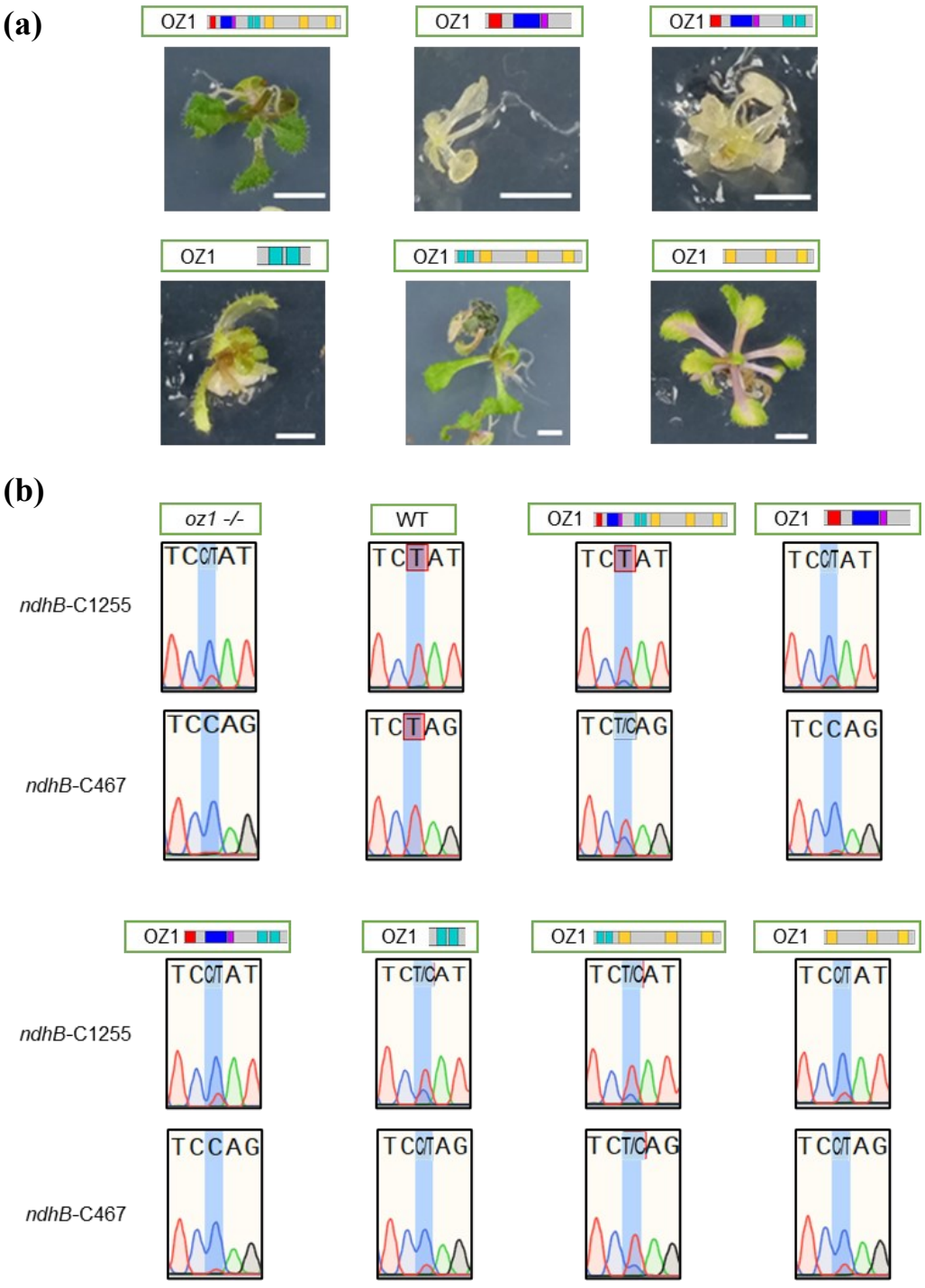
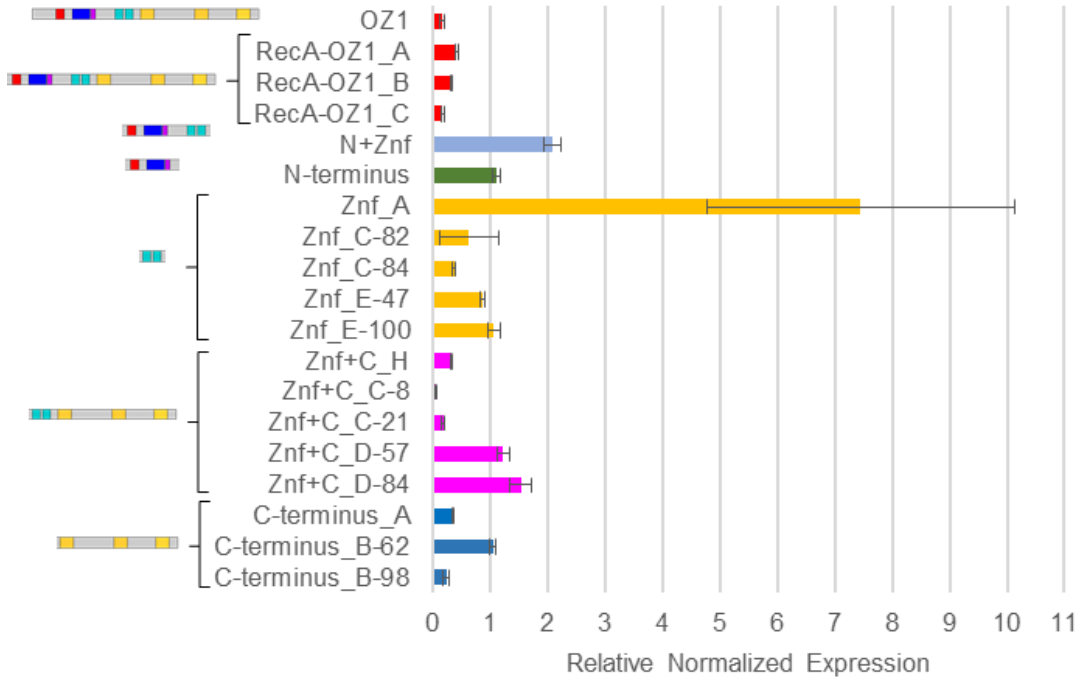


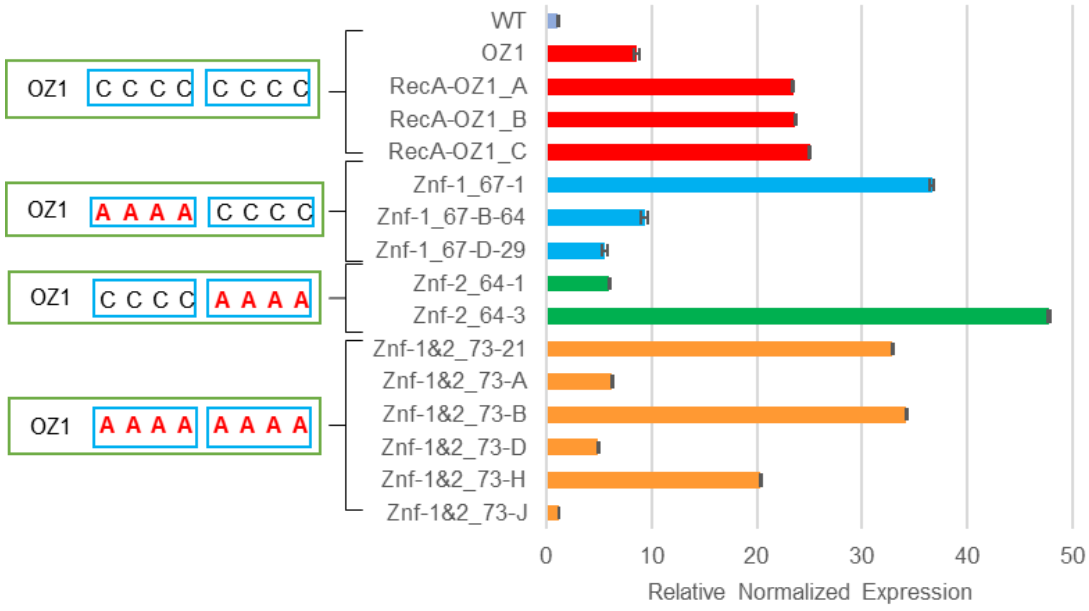
Figure 2.4. The OZ1 Znf domains alone can partially rescue certain editing sites. (**a**) One-month old *oz1* homozygous seedlings expressing OZ1 truncation constructs. Scale bars = 5 mm. (**b**) Sanger sequencing traces of editing sites *ndhB*-C1255 and *ndhB*-C467 in *oz1* mutants, WT-sequence OZ1 in *oz1* mutants, and *oz1* mutants expressing OZ1 truncation constructs. Chosen traces are representative of ≥3 different plants analyzed per construct.

Figure 2.5. qPCR of plants expressing OZ1 constructs. (**a**) qPCR of OZ1 truncation expression constructs. All plants are *oz1* homozygous background. "OZ1" construct refers to the full-length OZ1 sequence, while "RecA-OZ1" refers to constructs of the RecA chloroplast transit peptide fused to the N-terminus of the OZ1 sequence minus the first 100 nucleotides. Primers used: pBI121_qPCR_F and pBI121_qPCR_R for OZ1 constructs, SAND_qPCR_F and SAND_qPCR_R for the SAND reference gene. Error bars = SEM, corrected for expression differences relative to reference gene. (**b**) qPCR quantification of OZ1 Znf mutants, obtained as above except using OZ1q_1464-1681_F and OZ1q_1464-1681_R for OZ1 constructs and native OZ1 in the WT sample.

(a)

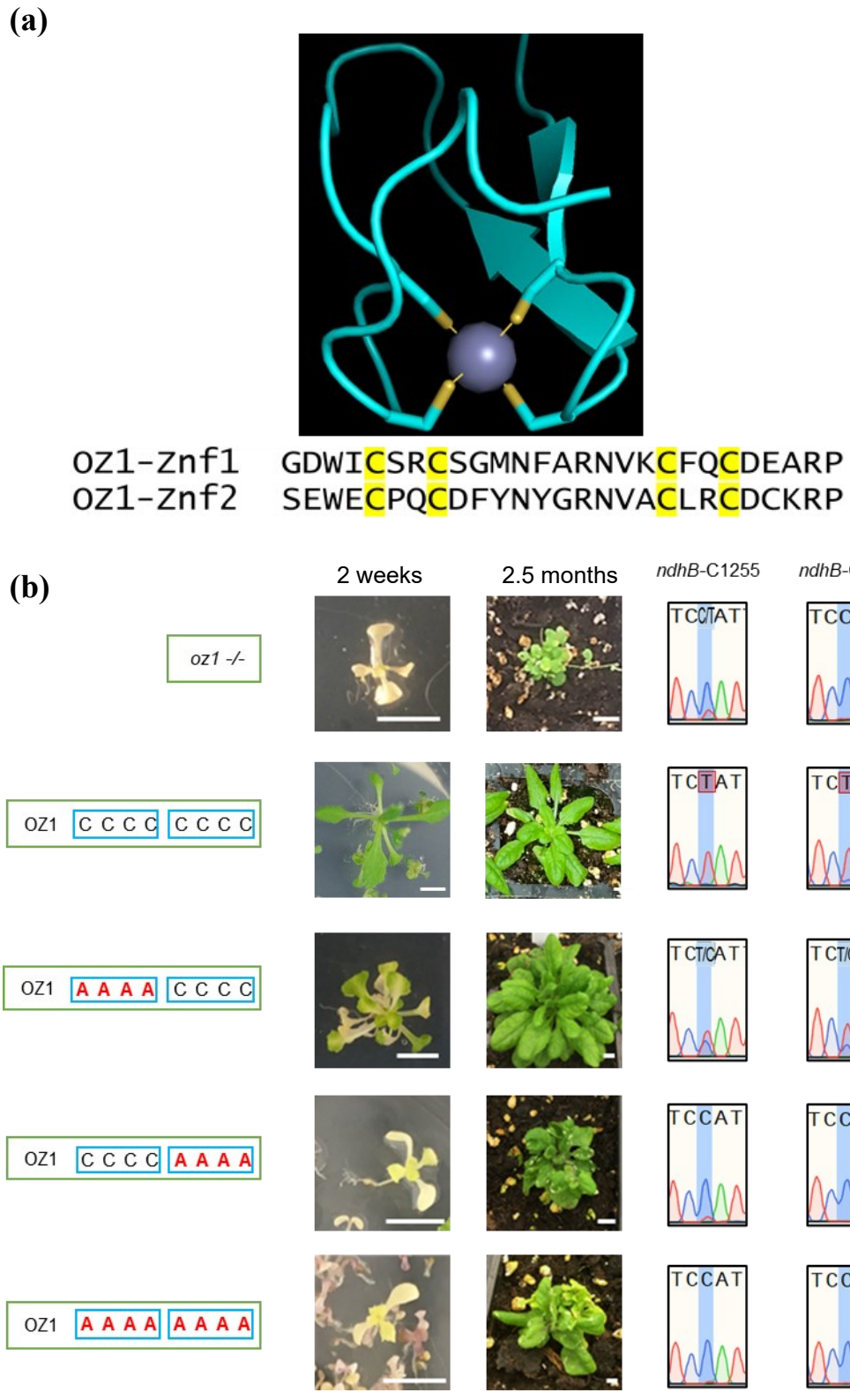

OZ1
RecA-OZ1_A
RecA-OZ1_B
RecA-OZ1_C
N+Znf
N-terminus
Znf_A
Znf_C-82
Znf_C-84
Znf_E-47
Znf_E-100
Znf+C_H
Znf+C_C-8
Znf+C_C-21
Znf+C_D-57
Znf+C_D-84
C-terminus_A
C-terminus_B-62
C-terminus_B-98
0 1 2 3 4 5 6 7 8 9 10 11
Relative Normalized Expression

(b)


OZ1 C C C C C C C C
OZ1 A A A A C C C C
OZ1 C C C C A A A A
OZ1 A A A A A A A A
WT
OZ1
RecA-OZ1_A
RecA-OZ1_B
RecA-OZ1_C
Znf-1_67-1
Znf-1_67-B-64
Znf-1_67-D-29
Znf-2_64-1
Znf-2_64-3
Znf-1&2_73-21
Znf-1&2_73-A
Znf-1&2_73-B
Znf-1&2_73-D
Znf-1&2_73-H
Znf-1&2_73-J
0 10 20 30 40 50
Relative Normalized Expression

To verify that the Znf domains of OZ1 are essential domains for RNA editing, I mutated the four conserved cysteines in each Znf domain, as those cysteines are predicted to be essential for their structural folding (**Figure 2.6a**) [15]. I transformed *oz1* plants with full-length OZ1 carrying the zinc finger cysteine-to-alanine mutations and observed more robust growth and greening in the Znf-1 plants (where the four C have been changed to A in the first Znf domain) vs. the Znf-2 (where the four C have been changed to A in the second Znf domain) or Znf-1&2 mutants (where all the C have been changed to A in both Znf domains; **Figure 2.6b**).

Figure 2.6. The zinc finger domains of OZ1 are essential for chloroplast RNA editing. (**a**) Predicted structure of the second OZ1 zinc finger domain based on the crystal structure of ZRANB2 (3g9y). Gray sphere = coordinated Zn^{2+} ion; yellow = zinc-coordinating cysteine residues; teal = rest of Znf structure, showing predicted beta sheet secondary structure. Underneath, alignment of the sequences of the first and seconds Znfs of OZ1. Yellow highlights indicate the putative zinc-coordinating cysteine residues. (**b**) Zinc finger mutant seedlings and representative editing traces. *oz1* homozygous seedlings and mature plants expressing OZ1 zinc finger mutant constructs. Scale bars = 5 mm. Sanger sequencing traces of the *ndhB*-C1255 and *ndhB*-C467 editing site in *oz1* mutants, *oz1* mutants expressing WT-sequence OZ1, and *oz1* mutants expressing cysteine-to-alanine zinc finger mutant constructs.


(a)
OZ1-Znf1 GDWICSRCSGMNFARNVKCFQCDEARP
OZ1-Znf2 SEWECPQCDFYNYGRNVACLRCDCKRP
(b)
2 weeks
2.5 months
ndhB-C1255
ndhB-C467
oz1 -/-
OZ1 C C C C C C C C
OZ1 A A A A C C C C
OZ1 C C C C A A A A
OZ1 A A A A A A A A
TCCTAT
TCCAG
TCTAT
TCTAG
TCT/CAT
TCT/CAG
TCCAT
TCCAG
TCCAT
TCCAG

As performed with the plants carrying truncation constructs, I wanted to assess if the morphological aspect of the transgenic plants could be explained by the level of plastid editing caused by the different constructs. RNA editing analysis revealed no editing rescue in *oz1* homozygous background plants expressing OZ1 constructs with the cysteine-to-alanine mutations in both Znfs or just Znf2 (**Figure 2.6b**; **Figure Apx3.2**), regardless of how highly the construct was expressed (**Figure 2.5b**), demonstrating that the Znf domains are essential for the editing function of OZ1. For instance, the transgenic plant 64-3 expressing the Znf-2 construct had the highest level of expression of the transgene among all plants tested, but it still failed to show any restoration of editing (**Figure 2.5b**; **Figure Apx3.2**). Furthermore, editing is largely restored by the OZ1_Znf-1 mutant construct (**Figure 2.6b**; **Figure Apx3.2**). As observed in plants expressing OZ1 truncations capable of rescue, plants expressing OZ1_Znf-1 more highly experience a greater extent of editing rescue (**Figure 2.5b**; **Figure Apx3.2**). Altogether, these results indicate that Znf2 alone is the editing-essential domain.

Zinc finger mutation does not affect PPR-OZ1 interactions but does perturb ORRM1-OZ1 binding

In order to understand the role of the second Znf domain in the editing function of OZ1, the question arose whether this domain might be important for interaction with other editing factors. I therefore evaluated the possibility that the putative structural disruption caused by mutating zinc-coordinating cysteines into alanines could be interfering with the ability of OZ1 to bind with other editing factors and incorporate

into the editosome. Y2H with PPR proteins and either the Znf-1 mutation or Znf-1&2 mutant of OZ1 showed that mutation of the Znfs does not interrupt binding between OZ1 and PPR proteins (**Figure 2.7a**). However, the interaction with ORRM1 was abolished when both zinc fingers were mutated, though mutation of Znf1 did not interrupt ORRM1/OZ1 binding in Y2H.

The requirement for Znf2 was also observed when interactions were tested by BiFC (**Figure 2.7b**), where PPR proteins interact with OZ1 regardless of OZ1 mutation, while ORRM1 does not interact with OZ1 protein that has the cysteine-to-alanine mutations in Znf2.

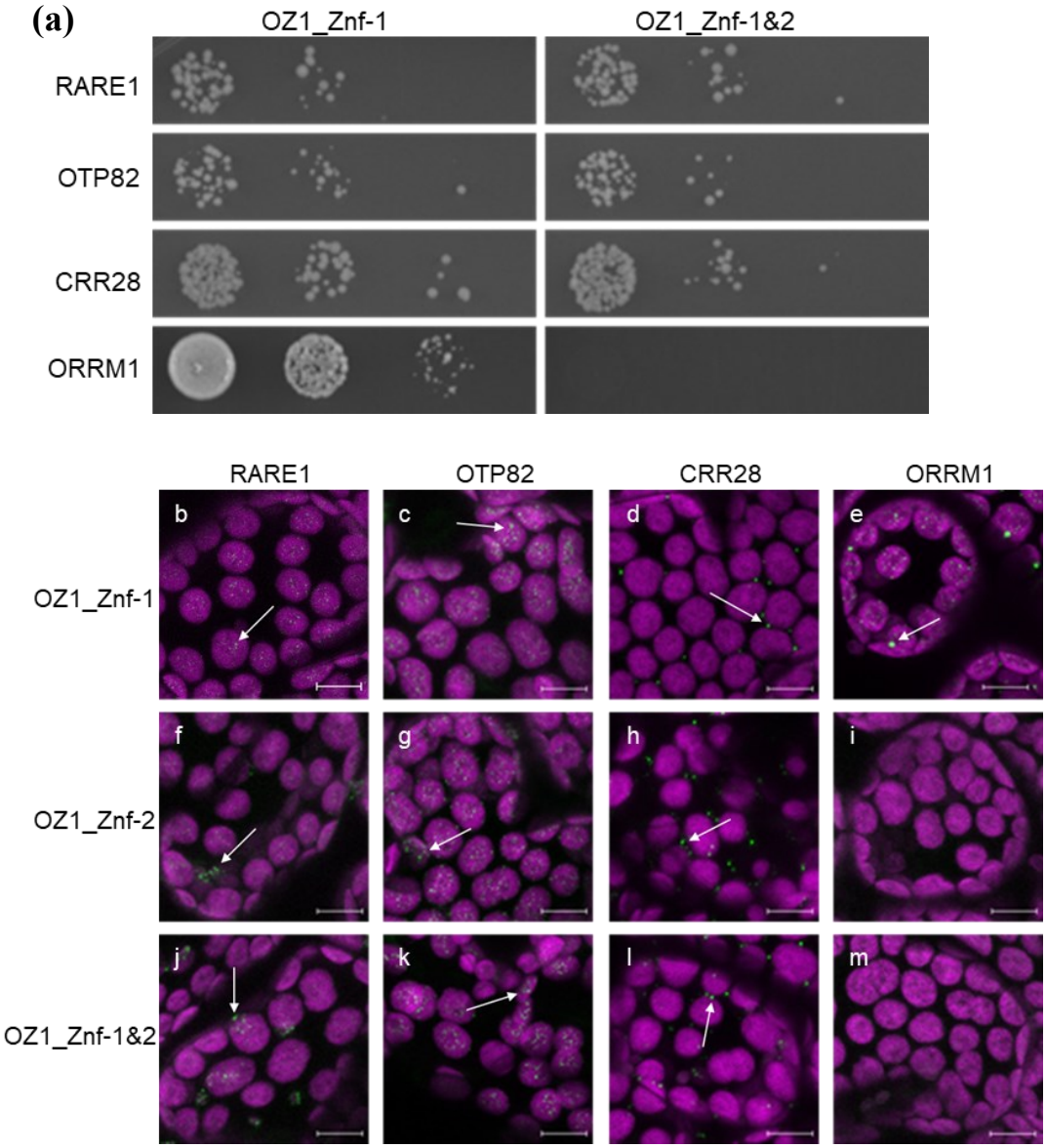


Figure 2.7. Mutation of the OZ1 Znf domains perturbs interaction with ORRM1 but not PPR proteins. (**a**) Yeast two-hybrid assays between OZ1 Znf mutants and editing factors RARE1, OTP82, CRR28 (PPR proteins), or ORRM1. (**b-m**) Bimolecular fluorescence complementation confocal images of interactions between OZ1 Znf mutants and RARE1, OTP82, CRR28, and ORRM1. Magenta = chlorophyll autofluorescence; green = YFP. Scale bars = 10 μm.

DISCUSSION

In this work I tried to delineate functional domains in the plastid editing factor OZ1. Three modular domains, the N-terminus, a central zinc finger pair, and a C-terminus were tested individually or in combination for their binding ability to other editing factors. In addition, these domains were assayed for their ability to restore a wild-type morphology and/or editing extent in an *oz1* mutant background. The goal was to determine whether there was a connection between these different biological processes.

In every instance where the full-length OZ1 protein interacted with a PPR protein, I found that the C-terminal portion of the protein recapitulated protein-protein binding (**Table 2.1**), leading to the conclusion that the C-terminal domain of OZ1 is necessary for those interactions. Yeast two-hybrid assays of OZ1 with truncations of the PPR proteins RARE1, OTP82, and CRR28 showed that the E-plus-DYW domain of each protein interacted with OZ1, suggesting that this may be the point of contact for the OZ1 C-terminal domain (**Table Apx3.1**). The motifs in the C-terminal region of OZ1, shown as yellow in **Figure 2.1b** and in the sequence displayed in Sun *et al.* [14], are unique to OZ1. Future structural elucidation of this editing factor will likely deepen the understanding of OZ1's incorporation into the editosome.

Surprisingly, the OZ1_Znf truncation was also able to interact not only with PPR proteins (**Figure 2.2**; **Figure 2.3**; **Table 2.1**), but also with ORRM1, a major component of the plastid editosome, as defined by the complete loss of editing for 12 sites in the *orrm1* mutant [31]. I found that both the C-terminal domain and the zinc finger domains of OZ1 contribute to protein-protein interactions, but the physical

nature of interactions mediated by the C-terminal domain is likely distinct from those mediated by the zinc fingers. There is precedent for RanBP2 Znfs acting as protein-binding domains; indeed, the namesake protein of the family is Ran-binding protein 2, which resides in the nuclear pore complex of human cells and binds a nuclear export factor through its zinc finger domains [23]. The rigidly structured zinc finger domains of OZ1 may contribute to ORRM1 interactions, but OZ1's ability to accumulate in a sufficiently high concentration to be associated with putative editosomes in spite of its low native expression level [32] could be explained by the presence of low-complexity domains in its C-terminal region that were predicted by domain boundary analysis. Much recent research has characterized the ubiquity and necessity of so-called phase-separated droplets, membraneless organelles created by the loose interactions between proteins with intrinsically disordered domains, as well as those between such proteins and RNAs [33]. These droplets are defined by the ability to fuse with other droplets, distortion of shape when bumping into physical barriers (e.g., lipid membranes), rapid recomposition, and a roughly spherical shape due to surface tension. These features separate them from protein aggregates, which are insoluble and irreversible.

Many critical molecular functions occur in the context of phase-separated droplets, such as splicing, transcription, DNA repair, and stress response [34–37]. Editosomes, and thus OZ1, are likely localized in the chloroplast nucleoids, membraneless structures in the stroma that contain the chromosomes and myriad associated proteins and RNAs involved in transcription and RNA processing [38]. Indeed, the maize ortholog of ORRM1 (GRMZM2G044422_P03) was identified in nucleoid-enriched proteomes of chloroplast maize leaves [39]. The dynamic nature of the nucleoid and

its separation into functionally distinct layers in spite of the absence of membranes is suggestive of phase-separated droplets [40,41]. The association of OZ1 into this environment could be dependent on its low-complexity C-terminal domain, but it could also be dependent on the presence of the Znfs, which mediate the interaction with ORRM1. ORRM1, like OZ1, is predicted to carry a low-complexity disordered peptide sequence directly upstream of its RRM domain, which might be responsible for its localization in the nucleoids. The ratio of protein to RNA in a given droplet has a large effect on its composition and behavior [42]; if OZ1 is associated with editosome-containing phase-separated droplets, then sensitivity to component concentration could explain the variation seen in fluorescent puncta morphology in our BiFC experiments (**Figure 2.3**; **Figure 2.7b-m**).

The BiFC assay revealed more interactions than the Y2H assay, an observation that was already encountered during our study on the mitochondrial splicing factor OZ2 (At1g55040) [43]. As proposed in this previous work, the Y2H assay, which is a heterologous system, might not allow the proper folding of some of the proteins tested, preventing their interactions. In addition, the removal of the transit peptide from the factors assayed in the Y2H assay is based on prediction software and, as such, might be inaccurate in removing too much or too little of the protein, impairing its proper folding.

Complementation experiments demonstrated that the presence of the Znf domains is necessary but not sufficient to restore editing in the *oz1* mutant background. The N-terminus+Znf construct was unable to complement the editing defect in the mutant, while the Znf alone and the Znf+C-terminus was able to significantly increase the

plastid editing extent (**Figure 2.4b**). The failure of the N-terminus+Znf construct to restore editing is somewhat puzzling but might be explained by the repressive effect of the N-terminus on the binding of Znf that was observed in the Y2H assay (**Figure 2.2**). For the constructs exhibiting any ability to rescue editing (e.g., OZ1_Znf-1, OZ1_Znf+C-terminus), the expression level of the construct in a given plant as measured by qRT-PCR (**Figure 2.5a**) correlates with the extent of editing rescue in that plant (**Figure Apx3.1**, **Figure Apx3.2**). Furthermore, relatively low expression levels of the Znf+C-terminus construct are able to rescue editing to a much greater extent than even high expression of Znf only; this may indicate that the presence of the C-terminus assists with the association of the truncation into the editosome more effectively than the Znf alone can do. This consideration is also relevant to the lack of restoration observed with the N-terminus+Znf construct.

In the light of the interaction and complementation experiments, it is tempting to link the ability of the Znf domain to bind to ORRM1, a major component of the plastid editosome, and the capacity of OZ1 to fulfill its editing function. This model is strongly supported by the mutation analysis we performed. The inability of ORRM1 to interact with the OZ1_Znf-2 and OZ1_Znf-1&2 mutations is thus likely impairing the incorporation of the mutated OZ1 into a functional editosome, which could be the reason for the lack of editing rescue in plants expressing this construct. Furthermore, through the mutant analysis, I was able to implicate the second zinc finger domain as being essential for both the binding of OZ1 to ORRM1 and for the editing function of OZ1; I confirmed that the loss of editing function in OZ1_Znf-2 and OZ1_Znf-1&2 could not be attributed to the misfolding and degradation of those proteins, as YFP

labeling demonstrated that OZ1_Znf-1&2 still accumulates in the chloroplast (**Figure Apx3.3**).

The structure of the RanBP2 Znf domain was solved with the first domain of the ZRANB2, a human splicing factor [44]. NMR spectra in both the presence and absence of zinc demonstrated that this domain was a genuine zinc-binding domain. Previously, the Znf domains of OZ2 were observed to be able to bind to two zinc atoms by performing mass spectrometry analysis under native and denaturing conditions [43]. Given the high similarity between the Znf domains of OZ1 and OZ2 (**Figure Apx3.4**), and the severe impact of mutating the cysteines predicted to coordinate the zinc atom, it seems highly likely that the Znf domains of OZ1 also bind zinc.

Previous work aligning the RNA sequences surrounding editing sites focused on PPR binding [17,45,46], which involves sequences ranging from 15-30 nucleotides upstream of the editing site [10]. The two RanBP2-type Znf domains found in ZRANB2 are able to bind ssRNA with high affinity and specificity [24], making it likely that the Znf domains found in OZ1 are also able to bind to the RNA target. However, if OZ1 interacts with the RNA, it would only be with very short sequences, no more than three to six nucleotides long, based on the length of known RanBP2 Znf binding sites [24]. De Franco *et al.*'s [47] extensive study of RNA-binding specificity of RanBP2 Znf domains from multiple lineages, including the Znf from the Arabidopsis protein Suppressor of ABI3-5, suggests that the aspartate in position 4 (D4) and arginine 20 (R20) of the RanBP2 Znf is critical for RNA binding to a GGN ssRNA sequence. Conservation of this residue in many of the OZ family zinc fingers

[14,15] suggest that they may indeed be RNA-binding domains that also recognize a GGN sequence. However, some of these domains deviate from D4 with a glutamate instead, and OZ1-Znf2 deviates significantly with a serine inserted at position 4 (**Figure Apx3.4**) [15]. These differences raise the possibility that there is a different preferred sequence for OZ1-Znf2 that may not quite resemble the sequence preferred by the human RanBP2 Znf protein ZRANB2 [24]. The possibility of OZ1 binding to RNA through its Znf domains adds another level of complexity that was not explored in this work, and future studies may further refine how OZ1 is integrated in the editosome.

I have demonstrated that OZ1 interacts with PPR proteins through the C-terminal domain, which is likely to be necessary for efficient incorporation into the editosome, while interaction with ORRM1 through its second Znf is absolutely required for RNA editing at sites targeted by OZ1. A recent study by our group revealed OZ2, the closest relative of OZ1 in the Arabidopsis proteome, to be a mitochondrial splicing factor [43]. Because the C-terminal region of OZ1 that mediates PPR interaction is unique to OZ1, it is possible that the other two OZ proteins may function, like OZ2, in organelle RNA metabolism other than editing.

MATERIALS AND METHODS

Plant materials and growth conditions

Arabidopsis Col-3 plants were used for WT controls. *oz1* T-DNA insertional mutants (SAIL_358_H03) were purchased from the Arabidopsis Biological Resource Center (https://abrc.osu.edu/). Because of the homozygous mutant's weak germination and chlorotic phenotype, seeds were sown on 0.5x MS with Gamborg's Vitamins

(Caisson Labs, Smithfield, UT) + 1% sucrose agar plates and grown at 25 °C with a 16-hour day length. After 5–8 weeks, seedlings with greening leaves were transferred to LM-111 soil and grown in short-day (10-hr) conditions for RNA analysis or long-day (16-hr) conditions for seed harvesting and floral dipping.

Domain boundary analysis of OZ1

In designing truncation constructs of OZ1 to test the function of different domains, I performed domain boundary analysis [25] to identify the best cutoff points for each construct that would still preserve predicted domain folding. In short, PSIPRED, FoldIndex [48], and GlobPlot 2 [49] were used to analyze the OZ1 amino acid sequence for secondary structure prediction and disordered region identification, providing a basis for determining construct boundaries.

Yeast two-hybrid

The mature coding sequences (lacking the first 33 amino acids, the predicted length of the plastid transit peptide [14]) of OZ1, OZ1 truncations (containing the last eight residues of the RecA chloroplast transit peptide at the N-terminus as an artifact of cloning), OZ1 zinc finger mutants, and editing factors RARE1, OTP82, CRR28, QED1, DYW2, CLB19, and ORRM1 were amplified from Col-3 Arabidopsis cDNA using Phusion polymerase (Thermo Scientific, Waltham, MA) and the Y2H primer pairs in **Table 2.2**. 3'-A overhangs were added with Taq DNA Polymerase (QIAGEN, Germantown, MD) by incubating at 37 °C for 10 minutes. After purification, the amplicons were TA cloned into pCR8/GW/TOPO (Invitrogen, Waltham, MA) to use

in Gateway cloning reactions with the yeast two-hybrid destination vectors pGADT7GW and pGBKT7GW [50]. Empty pGADT7GW and pGBKT7GW vectors were used as negative controls in yeast two-hybrid assays. Yeast mating strains PJ69-4a and PJ69-4α were individually transformed with pGADT7GW and pGBKT7GW plasmids, respectively. Single transformants were mated to produce diploid double-transformant yeast on YPAD agar plates. Diploid yeast were grown in -Leu and -Trp media overnight (Takara Bio USA, Mountain View, CA), then 10 µL of each culture was spotted onto -Leu, -Trp, -His, and -Ade media plates (Takara Bio USA, Mountain View, CA) after being diluted with water to OD_{600} 0.5, 0.05, or 0.005. Survival/growth plates were imaged after three days of incubation at 30°C.

Bimolecular fluorescence complementation

The full-length coding sequences (including N-terminal plastid transit peptides) of OZ1, OZ1 zinc finger mutants, and editing factors RARE1, OTP82, CRR28, QED1, DYW2, and ORRM1 were amplified using BiFC primer pairs in **Table 2.2**. For experiments with OZ1 truncations, all OZ1 constructs were cloned without the first 33 amino acids (the predicted transit peptide) and in its place was a RecA chloroplast targeting sequence, corresponding to the first 65 residues of that plastid-localized protein [51,52]. PCR products were first cloned into pCR8/GW/TOPO and then pEXSG-nYFP and pEXSG-cYFP BiFC destination vectors via Gateway cloning as above.

Agroinfiltration of *N. benthamiana* with pEXSG plasmid-carrying *Agrobacteria* was performed as in Sparkes *et al.* [53], summarized as follows. *A. tumefaciens*

GV3101::pMP90RK were transformed with pEXSG plasmids via electroporation using 1 μg plasmid. Electroporation was conducted with the following parameters: capacitance 25 μF, voltage 2.0 kV, resistance 200 Ω, pulse length approximately 5 msec. Agrobacteria were selected on LB agar plates containing kanamycin (50 μg/mL) and carbenicillin (25 μg/mL). 5-mL cultures of individual transformed Agrobacteria were incubated for two days at 28°C and resuspended in a solution of 50 mM 2-(N-morpholine)-ethanesulphonic acid (pH 5.6), 2 mM Na_3PO_4, 1.4 mM glucose, and 100 mM acetosyringone. Infiltration samples were made by mixing 0.3 OD_{600} of each pEXSG plasmid-carrying bacteria and 0.3 OD of P19-carrying bacteria to a final OD of 0.9. Leaves of 4- to 6-week-old *N. benthamiana* plants grown in long-day conditions were injected with the *Agrobacteria* solution. 2–3 days post-infiltration, 2-mm squares were cut from the infiltrated leaf area and imaged using a Zeiss Axio Observer LSM 710 microscope and C-Apochromat 40x/1.20 W Korr M27 objective. Chlorophyll was excited at 514 nm, and the emission filter range was set to 617–735 nm; YFP was excited at 514 nm, emission at 519–602 nm.

Plant transformation

OZ1 truncation constructs with the RecA targeting sequence (as described above) were cloned into pBI121. OZ1 zinc finger point mutants with the native transit peptide were synthesized with the Q5 Site-Directed Mutagenesis kit (NEB, Ipswich, MA), TA cloned into pCR8, then cloned into a Gateway-compatible pBI121 vector as above. *A. tumefaciens* GV3101 were transformed with pBI121 expression plasmids as above, selecting with kanamycin (50 μg/mL) and gentamicin (25 μg/mL).

oz1 heterozygous plants underwent floral dipping using standard protocols [54]. After 24 hours of recovery in the dark, floral-dipped plants were returned to long-day conditions and grown for 4–6 weeks before drying and harvesting seeds. Seeds were sterilized and stratified before plating for selection on 0.5x MS with Gamborg's Vitamins + 1% sucrose agar plates containing kanamycin (50 µg/mL), BASTA, and cefotaxime (250 µg/mL). Surviving seedlings were transferred to soil 5–6 weeks after sowing.

RNA analysis

Leaves were taken from one- to five-month-old Arabidopsis, and RNA was extracted using Trizol and the PureLink RNA Mini Kit (ThermoFisher, Waltham, MA). RNA was treated with TURBO DNAse (Invitrogen, Waltham, MA) and quantified with the Qubit II (ThermoFisher, Waltham, MA). cDNA was amplified using Superscript III (Invitrogen, Waltham, MA) and pooled chloroplast transcript primers (**Table 2.2**). Gene-specific amplicons were then amplified with the corresponding primer pairs (**Table 2.2**) and Sanger sequenced.

Quantitative RT-PCR (qPCR) was used to evaluate the relative expression levels of OZ1 constructs in transformed *oz1* plants. Total RNA was extracted and treated as above, and cDNA was amplified as above with random hexamer primers. qPCR was performed using the iTaq Universal SYBR Green Supermix (Bio-Rad, Hercules, CA) and primers in **Table 2.2** in a CFX Connect reaction module (Bio-Rad, Hercules, CA). ΔCq was calculated from three technical replicates per sample by comparing Cq

values of the OZ1 expression construct target to those of the *SAND* (At3g28390) reference gene.

Table 2.2. Primers used in OZ1 study.

Oligo Name	Sequence	Use
OZ1_100F	CGTTTCCACCGCCGTGCGTTT	Y2H
OZ1_R	TCATTTATCTCCTTTACCAGTGGGATC	
Rare1_+100F	GGATCCATGTCGAGCACTTCTTCTCCGTCT	
Rare1_+2259R	TCACCAGTAATCGTTGCAAGAACA	
OTP82-292F	AACCTGTTGATTTGGAACACGATGTTT	
OTP82_R	CTACCAGTAGTCATTGCAGGAACAAACA	
rCRR28-FW	TCCACCGCCGGTAACCAT	
CRR28_R	CTACCAGTAGTCTAAACAAGAGCA	
QED1_127_F	CTCCGACAACTAAAGCAAAC	
QED1_R	TCACCAGAAATCGTTACAGGAAC	
DYW2_97_F	TCTCAATTCCACTTCTCCG	
DYW2_R	CTACCAGTAATCCCCGCAAG	
CLB19_163_F	GCAGCGAAGGAATTCTCC	
CLB19_R	TCAAGCATTGAGGAGATCACC	
ORRM1_163_F	GGAAGTGGGGGTTCAGGATCCTATATAGTCGACTCTTCTGCAATTTCCGCACCG	
ORRM1_R	CTAGAGCCCGAAACTTGGTTG	
ORRM1_F	ATGGAAGCTCTTATTGCTTCC	BiFC
ORRM1_R_WO	GAGCCCGAAACTTGGTTGACTTCT	
RARE1_F	ATGACGATTCTCACTGTACAGTCTTC	
RARE1_nostop_R	CCAGTAATCGTTGCAAGAACA	
OTP82_F	ATGATGCTCTCGTGTTCTCCTCTC	
OTP82_nostop_R	CCAGTAGTCATTGCAGGAACAAACA	
CRR28_F	ATGGTTGTTCGTTCAATTATC	
CRR28_nostop_R	CCAGTAGTCTAAACAAGAGC	
QED1_F	ATGGCTATCTTCTCCACAGC	
QED1_nostop_R	CCAGAAATCGTTACAGGAAC	
DYW2_F	ATGTCTTCTCTAATGGCCATTC	
DYW2_nostop_R	CCAGTAATCCCCGCAAG	
CLB19_F	ATGGGTCTCCTTCCCGTCG	
CLB19_nostop_R	AGCATTGAGGAGATCACCAGC	
OZ1-819_nostop_R	AACGTTATTCTGAAAAGACCTG	
OZ1-1011_nostop_R	GGGTCGCTTGCAGTC	

CtpOZ1-100_F	CACAAAATCAGTTCTGAATTCGATCGTTTCCACCGCCG	Y2H/BiFC/stable transformation
CtpOZ1-820_F	CACAAAATCAGTTCTGAATTCGATGAGATGAAACGTGGTGACTGG	
CtpOZ1-1072_F	CACAAAATCAGTTCTGAATTCGATAGACTGGTCGAAAACGAAAAGA	
OZ1-819_R	tcaAACGTTATTCTGAAAAGACCTG	Y2H/stable transformation
OZ1-1011_R	tcaGGGTCGCTTGCAGTC	
recA-Ctp_F	ATGGATTCACAGCTAGTCTTGTCTC	BiFC/stable transformation
recA-Ctp_R	ATCGAATTCAGAACTGATTTTGTG	
Q5_OZ1_Znf1C14A_F	gcgagaaatgtcaaagccttccaggctGATGAAGCAAGACCAAAG	Znf mutagenesis
Q5_OZ1_Znf1C14A_R	aaagttcatgccactggccctcgaggcAATCCAGTCACCACGTTTC	
Q5_OZ1_Znf2C14A_F	gggaggaatgtagctgccttaagagctGACTGCAAGCGACCCAGG	
Q5_OZ1_Znf2C14A_R	gtaattatagaaatcagcttgaggagcCTCCCATTCACTACCAGTAAGC	
ndhB-F2	TTTTATGTGGTGCTAACGATTTAA	Plastid transcript cDNA sequencing
ndhB-F3	GCATGTACAGAATGAAAATTTCATTCT	
ndhB-R2	AATCGCAATAATCGGGTTCATT	
ndhD-F	AACAACTCGAAGTATGGGTC	
ndhD-R	GGCAATGCAAGGGAAGCC	
rpoA-F2	CTCGGACACTACAGTGGAAGTGTG	
rpoA-R2	CTGGGAGGCAATTCTAATTGGTC	
clpP-F2	TATTGGCGTTCCAAAAGTACCTT	
clpP-R2	GAACCGCTACAAGATCAACAATTC	
ndhG-F2	TTTGCCTGGACCAATACATG	
ndhG-R2	AGCCACAGAAATTGCACCTAT	
accD-F2	ATGGAAAAATCGTGGTTCAATTTTATG	
accD-R2	GTTTGTCTAGTCTAATTTGAACTTCCC	
accD-seq-F	GTGCCTGAAGGTTCACAAGC	
pBI121_qPCR_F	AATTCGACCCAGCTTTCT	qPCR
pBI121_qPCR_R	TTGTCATCGTCATCCTTGTA	
OZ1q_1464-1681_F	GTTTGCCAAGAAACCTAAGGAG	
OZ1q_1464-1681_R	AGTCTGAAACATTGTGCAGTTC	
SAND_qPCR_F	AACTCTATGCAGCATTTGATCCACT	
SAND_qPCR_R	TGATTGCATATCTTTATCGCCATC	

REFERENCES

1. Small, I. D., Schallenberg-Rüdinger, M., Takenaka, M., Mireau, H. & Ostersetzer-Biran, O. Plant organellar RNA editing: What 30 years of research has revealed. *Plant J.* **101**, 1040–1056 (2020).

2. Smith, H. C., Gott, J. M. & Hanson, M. R. A guide to RNA editing. *RNA* **3**, 1105–23 (1997).

3. Stern, D. B., Goldschmidt-Clermont, M. & Hanson, M. R. Chloroplast RNA metabolism. *Annu. Rev. Plant Biol.* **61**, 125–55 (2010).

4. Bentolila, S., Oh, J., Hanson, M. R. & Bukowski, R. Comprehensive high-resolution analysis of the role of an Arabidopsis gene family in RNA editing. *PLoS Genet.* **9**, e1003584 (2013).

5. Ruwe, H., Castandet, B., Schmitz-Linneweber, C. & Stern, D. B. Arabidopsis chloroplast quantitative editotype. *FEBS Lett.* **587**, 1429–1433 (2013).

6. Wang, X., An, Y., Xu, P. & Xiao, J. Functioning of PPR proteins in organelle RNA metabolism and chloroplast biogenesis. *Front. Plant Sci.* **12**, (2021).

7. Shi, X., Hanson, M. R. & Bentolila, S. Functional diversity of Arabidopsis organelle-localized RNA-recognition motif-containing proteins. *Wiley Interdiscip. Rev. RNA* **1**, e1420 (2017).

8. Hackett, J. B., Shi, X., Kobylarz, A. T., Lucas, M. K., Wessendorf, R. L., Hines, K. M., Bentolila, S., Hanson, M. R. & Lu, Y. An organelle RNA recognition motif protein is required for photosystem II subunit psbF transcript editing. *Plant Physiol.* **173**, 2278–2293 (2017).

9. Shi, X., Germain, A., Hanson, M. R. & Bentolila, S. RNA recognition motif-containing protein ORRM4 broadly affects mitochondrial RNA editing and impacts plant development and flowering. *Plant Physiol.* **170**, 294–309 (2016).

10. Barkan, A. & Small, I. Pentatricopeptide repeat proteins in plants. *Annu. Rev. Plant Biol.* **65**, 415–442 (2014).

11. Sun, T., Bentolila, S. & Hanson, M. R. The unexpected diversity of plant organelle RNA editosomes. *Trends Plant Sci.* **21**, 962–973 (2016).

12. Takenaka, M., Zehrmann, A., Verbitskiy, D., Kugelmann, M., Hartel, 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.* **109**, 5104–5109 (2012).

13. Shi, X., Bentolila, S. & Hanson, M. R. Organelle RNA recognition motif-containing (ORRM) proteins are plastid and mitochondrial editing factors in Arabidopsis. *Plant Signal. Behav.* **11**, e1167299 (2016).

14. 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.* **11**, 1–23 (2015).

15. Gipson, A. B., Giloteaux, L., Hanson, M. R. & Bentolila, S. Arabidopsis RanBP2-type zinc finger proteins related to chloroplast RNA editing factor OZ1. *Plants* **9**, 307 (2020).

16. Shi, X., Hanson, M. R. & Bentolila, S. Two RNA recognition motif-containing proteins are plant mitochondrial editing factors. *Nucleic Acids Res.* **43**, 3814–3825 (2015).

17. Wagoner, J. A., Sun, T., Lin, L. & Hanson, M. R. 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 (2015).

18. Boussardon, C., Avon, A., Kindgren, P., Bond, C. S., Challenor, M., Lurin, C. & Small, I. 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 (2014).

19. Boussardon, C., Salone, V., Avon, A., Berthomé, R., Hammani, K., Okuda, K., Shikanai, T., Small, I. & Lurin, C. Two interacting proteins are necessary for the editing of the ndhD-1 site in Arabidopsis plastids. *Plant Cell* **24**, 3684–94 (2012).

20. Okuda, K., Nakamura, T., Sugita, M., Shimizu, T. & Shikanai, T. A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. *J. Biol. Chem.* **281**, 37661–7 (2006).

21. Bayer-Császár, E., Haag, S., Jörg, A., Glass, F., Härtel, B., Obata, T., Meyer, E. H., Brennicke, A. & Takenaka, M. The conserved domain in MORF proteins has distinct affinities to the PPR and E elements in PPR RNA editing factors. *Biochim. Biophys. Acta Gene Regul. Mech.* **1860**, 813–828 (2017).

22. 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.* **295**, 3497–3505 (2020).

23. Singh, B. B., Patel, H. H., Roepman, R., Schick, D. & Ferreira, P. A. The zinc finger cluster domain of RanBP2 is a specific docking site for the nuclear export factor Exportin-1. *J. Biol. Chem.* **274**, 37370–37378 (1999).

24. Loughlin, F. E. *et al.* The zinc fingers of the SR-like protein ZRANB2 are single-stranded RNA-binding domains that recognize 5' splice site-like sequences. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 5581–6 (2009).

25. Cooper, C. D. O. & Marsden, B. D. N- and C-terminal truncations to enhance protein solubility and crystallization: Predicting protein domain boundaries with bioinformatics tools. in *Heterologous Gene Expression in E. coli: Methods and Protocols* (ed. Burgess-Brown, N. A.) vol. 1586 11–31 (Springer New York, 2017).

26. Robbins, J. C., Heller, W. P. & Hanson, M. R. A comparative genomics approach identifies a PPR-DYW protein that is essential for C-to-U editing of the Arabidopsis chloroplast accD transcript. *RNA* **15**, 1142–1153 (2009).

27. Hammani, K., Okuda, K., Tanz, S. K., Chateigner-Boutin, A.-L., Shikanai, T. & Small, I. A study of new Arabidopsis chloroplast RNA editing mutants reveals general features of editing factors and their target sites. *Plant Cell* **21**, 3686–99 (2009).

28. Andrés-Colás, 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. U. S. A.* **114**, 8883–8888 (2017).

29. Brehme, N., Glass, F., Jörg, A. & Takenaka, M. MEF46 and MEF47 are novel specificity factors for RNA editing sites in mitochondrial nad transcripts. *Mitochondrion* **53**, 121–127 (2020).

30. Malbert, B. *et al.* The analysis of the editing defects in the dyw2 mutant provides new clues for the prediction of RNA targets of Arabidopsis E+-class PPR proteins. *Plants* **9**, 280 (2020).

31. 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.* **110**, E1169–E1178 (2013).

32. Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W. & Zimmermann, P. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Boinformatics* 420747 (2008) doi:10.1155/2008/420747.

33. Cuevas-Velazquez, C. L. & Dinneny, J. R. Organization out of disorder: liquid–liquid phase separation in plants. *Curr. Opin. Plant Biol.* **45**, 68–74 (2018).

34. Guo, Y. E. *et al.* Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543–548 (2019).

35. Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V. & Cisse, I. I. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science (80-.).* **361**, 412–415 (2018).

36. Singatulina, A. S., Hamon, L., Sukhanova, M. V., Desforges, B., Joshi, V., Bouhss, A., Lavrik, O. I. & Pastré, D. PARP-1 activation directs FUS to DNA damage sites to form PARG-reversible compartments enriched in damaged DNA. *Cell Rep.* **27**, 1809-1821.e5 (2019).

37. Leung, A. K. L., Vyas, S., Rood, J. E., Bhutkar, A., Sharp, P. A. & Chang, P. Poly(ADP-ribose) regulates stress responses and microRNA activity in the cytoplasm. *Mol. Cell* **42**, 489–499 (2011).

38. Germain, A., Hotto, A. M., Barkan, A. & Stern, D. B. RNA processing and decay in plastids. *Wiley Interdiscip. Rev. RNA* **4**, 295–316 (2013).

39. Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, 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.* **158**, 156–89 (2012).

40. Krupinska, K., Melonek, J. & Krause, K. New insights into plastid nucleoid structure and functionality. *Planta* **237**, 653–664 (2013).

41. Powikrowska, M., Oetke, S., Jensen, P. E. & Krupinska, K. Dynamic composition, shaping and organization of plastid nucleoids. *Front. Plant Sci.* **5**, 3060–73 (2014).

42. Rhine, K., Vidaurre, V. & Myong, S. RNA droplets. *Annu. Rev. Biophys.* **49**, 247–265 (2020).

43. Bentolila, S., Gipson, A. B., Kehl, A. J., Hamm, L. N., Hayes, M. L., Mulligan, R. M. & Hanson, M. R. A RanBP2-type zinc finger protein functions in intron splicing in Arabidopsis mitochondria and is involved in the biogenesis of respiratory complex I. *Nucleic Acids Res.* **49**, 3490–3506 (2021).

44. Plambeck, C. A., Kwan, A. H. Y., Adams, D. J., Westman, B. J., Van der Weyden, L., Medcalf, R. L., Morris, B. J. & Mackay, J. P. The structure of the zinc finger domain from human splicing factor ZNF265 fold. *J. Biol. Chem.* **278**, 22805–22811 (2003).

45. 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.* **8**, e1002910 (2012).

46. Ruwe, H., Gutmann, B., Schmitz-Linneweber, C., Small, I. & Kindgren, P. The E domain of CRR2 participates in sequence-specific recognition of RNA in plastids. *New Phytol.* **222**, 218–229 (2019).

47. De Franco, S. *et al.* Exploring the suitability of RanBP2-type zinc fingers for RNA-binding protein design. *Sci. Rep.* **9**, 2484 (2019).

48. Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., Silman, I. & Sussman, J. L. FoldIndex©: a simple tool to predict whether a given protein sequence is intrinsically unfolded.

Bioinformatics **21**, 3435–3438 (2005).

49. Linding, R. GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic Acids Res.* **31**, 3701–3708 (2003).

50. Lu, Q. *et al.* Arabidopsis homolog of the yeast TREX-2 mRNA export complex: components and anchoring nucleoporin. *Plant J.* **61**, 259–270 (2010).

51. Lin, M. T., Occhialini, A., Andralojc, P. J., Devonshire, J., Hines, K. M., Parry, M. A. J. & Hanson, M. R. β-Carboxysomal proteins assemble into highly organized structures in Nicotiana chloroplasts. *Plant J.* **79**, 1–12 (2014).

52. Köhler, R. H., Cao, J., Zipfel, W. R., Webb, W. W. & Hanson, M. R. Exchange of protein molecules through connections between higher plant plastids. *Science.* **276**, 2039–42 (1997).

53. Sparkes, I. A., Runions, J., Kearns, A. & Hawes, C. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.* **1**, 2019–25 (2006).

54. Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W. & Chua, N.-H. Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. *Nat. Protoc.* **1**, 641–646 (2006).

APPENDIX 1

OZ2 interacts with mitochondrial splicing factors*

As discussed in Bentolia S. *et al.* [1], OZ2 is a mitochondrial splicing factor essential for embryo viability. Rescue of the *oz2* mutant plants from embryo lethality was done by expression of wild type *OZ2* under the ABSCISIC ACID-INSENSITIVE3 (ABI3) promoter, which is active during the early stages of seed development and germination but quickly shuts down afterwards, allowing us to analyze the morphological and molecular phenotype of the *oz2* mutant. These plants exhibited splicing defects of several *nad* transcripts involved in complex I formation and *rps3*. I verified the mitochondrial localization of OZ2 through transformation of Arabidopsis and *Nicotiana benthamiana* protoplasts (**Figure Apx1.1**).

* Data in this appendix was included in "A RanBP2-type zinc finger protein functions in intron splicing in Arabidopsis mitochondria and is involved in the biogenesis of respiratory complex I" by Stephane Bentolila, Andrew B. Gipson, Alexander J. Kehl, Lauren N. Hamm, Michael L. Hayes, R. Michael Mulligan, and Maureen R. Hanson**.** *Nucleic Acids Res.* © The Author(s) 2021; 49 (6):3490-3506. The localization experiments in this appendix were performed by A. Kehl under my supervision. I performed all other experiments shown here.

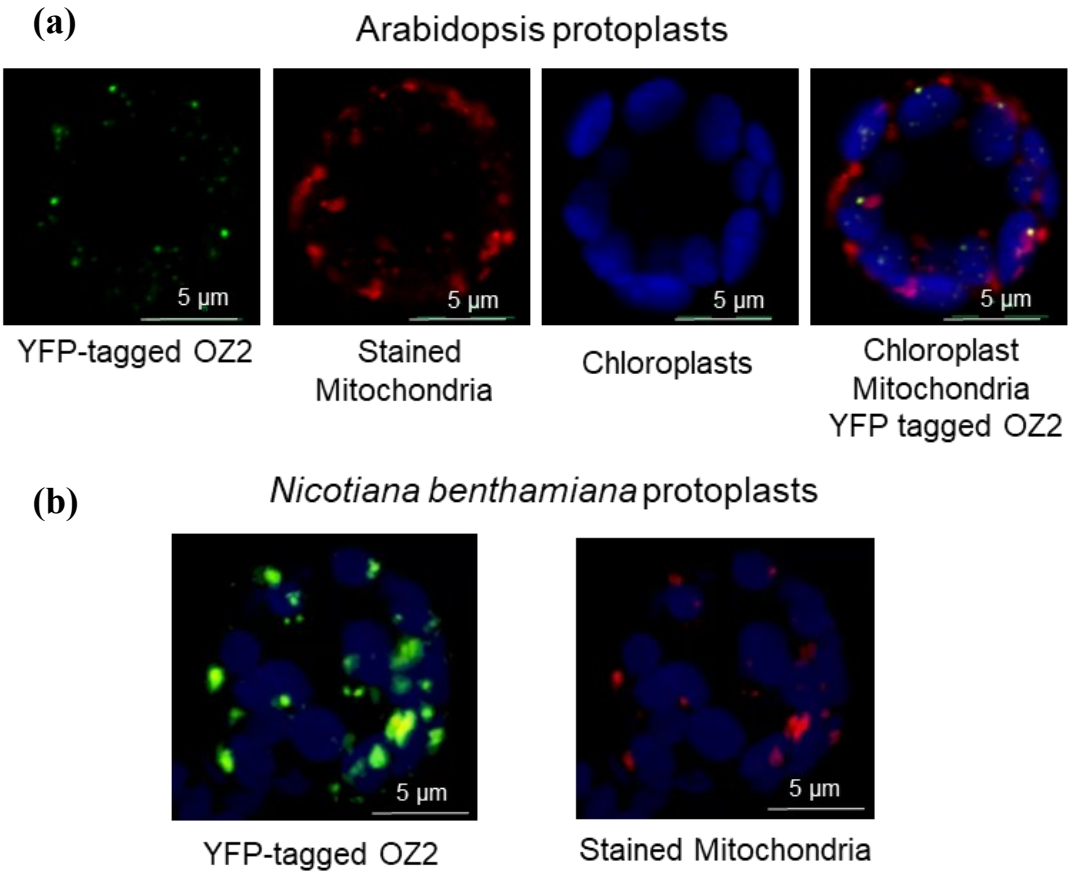


Figure Apx1.1. OZ2 is targeted to mitochondria. (**a**) Transfection of Arabidopsis protoplasts with OZ2 fused to YFP. From left panel to right panel: OZ2 location appears as green dots, mitochondria are stained in red with MitoTracker, chloroplasts autofluorescence is artificially colored blue, overlay of these three panels demonstrate the location of OZ2 in mitochondria as yellow (green + red) dots. (**b**) Transfection of Nicotiana benthamiana protoplasts with OZ2 fused to YFP. Some of the green spots (OZ2-YFP) can be seen to co-localize with the red spots (mitochondria).

Based on previous knowledge of which mitochondrial splicing factors target which splice sites, I conducted protein-protein interaction assays between OZ2 and mitochondrial splicing factors that share target splicing sites with OZ2 (**Table Apx1.1**), along with some that do not share splice sites with OZ2 as possible negative controls (BIR6, MISF74, and WTF9) to demonstrate that OZ2 associates with these factors.

Table Apx1.1. Summary of protein-protein interactions between OZ2 and mitochondrial splice factors. Y2H = yeast two-hybrid; BiFC = bimolecular fluorescence complementation.

Splicing factor	Target introns	OZ2 interaction (Y2H)	OZ2 interaction (BiFC)
ABO5	nad2-i3*	Yes	Yes
BIR6	nad7-i1	Yes	Yes
MAT2	cox2, nad1-i2, nad7-i2*	No	Yes
mCSF1	rps3*, nad2-i3*, nad5-i1*, -i2*, -i3*, nad7-i2* and others	No	Yes
MISF26	nad2-i3*	Yes	Yes
MISF74	nad1-i4, nad2-i4	No	No
mTERF15	nad2-i3*	No	No
MTL1	nad7-i2*	No	Yes
OTP439	nad5-i2*	No	No
PMH2	nad5-i1*, -i2*, -i3*, rps3* and others	Yes	Yes
RUG3	nad2-i2, nad2-i3*	No	No
SLO3	nad7-i2*	No	No
WTF9	rpl2, ccmFc	No	No

*Splice sites targeted by OZ2.

Some of the mitochondrial splicing factors sharing common targets with OZ2 are general factors like MAT2, mCSF1 and PMH2, which affect the splicing of several sites, while other are specific and control one to two splicing events, e.g. ABO5 and RUG3, respectively. The Y2H assays showed fewer OZ2/splice factor interactions than the BiFC assay, 4 versus 7, respectively (**Table Apx1.1**, **Figure Apx1.2**). All the interactions detected by the Y2H assay were confirmed by the BiFC assay (**Table Apx1.1**, **Figure Apx1.2**). Among the four positive interactions detected by both assays, three involved splicing factors sharing targets with OZ2, one general factor

(PMH2), and two specific factors (ABO5 and MISF26). One consistent interaction identified by both Y2H and BiFC assays involves BIR6, a protein that does not share a target splicing site with OZ2. The other two negative controls, MISF74 and WTF9 did not show any interaction with OZ2. The BiFC assay revealed three additional interactions not detectable by the Y2H assay, two with the general splicing factors MAT2 and mCSF1 and one with the specific factor MTL1, all of which share common splicing targets with OZ2 (**Table Apx1.1**, **Figure Apx1.2**). Overall, my interaction data are consistent with the specificity of the OZ2 splicing function being mediated in some instances by the interaction of OZ2 with other mitochondrial splicing factors.

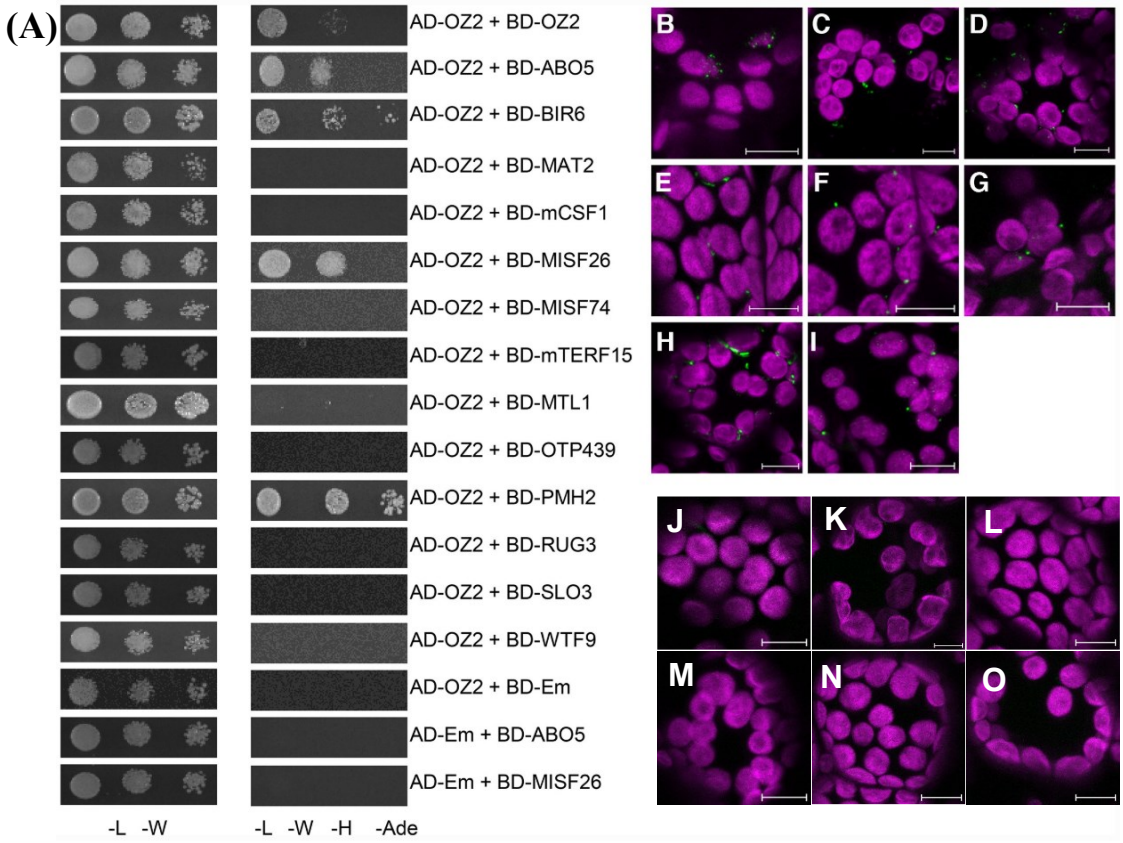


Figure Apx1.2. OZ2 interacts with several site-specific and general mitochondrial splicing factors *in vitro* and *in vivo*. (**A**) Yeast two-hybrid assay for OZ2/mitochondrial splice factors demonstrate interactions with certain factors that share splicing targets with OZ2, as well as OZ2 dimerization. Haploid yeast were transformed with constructs expressing OZ2 or a mitochondrial splicing factor with N-terminal fusions of parts of the yeast GAL4 transcription factor. Transformed yeast were mated to produce diploids expressing two fusion proteins and selected on synthetic dropout media lacking leucine and tryptophan (-L-W). Interactions were tested on plates lacking leucine, tryptophan, histidine, and adenine (-L-W-H-Ade). Yeast dilution spots contained 10^6, 10^5, and 10^4 cells per mL. Em = empty vector. (**B–O**) Bimolecular fluorescence complementation (BiFC) was performed in *N. benthamiana* using transient nYFP/cYFP fusion protein expression constructs. Green = YFP fluorescence; magenta = chlorophyll autofluorescence; scale bars = 10 µm. (**B**) ABO5 + OZ2. (**C**) BIR6 + OZ2. (**D**) MAT2 + OZ2. (**E**) mCSF1 + OZ2. (**F**) MISF26 + OZ2. (**G**) MTL1 + OZ2. (**H** PMH2 + OZ2. (**I**) OZ2 + OZ2. The interaction of OZ2 with itself serves as a positive control for this assay. (**J–O**) Negative results from BiFC. (**J**) MISF74 + OZ2. (**K**) mTERF15 + OZ2. (**L**) OTP439 + OZ2. (**M**) RUG3 + OZ2. (**N**) SLO3 + OZ2. (**O**) WTF9 + OZ2.

MATERIALS AND METHODS

OZ2 localization

The cDNA clone of OZ2 (At1g55040) used in this study was reverse-transcribed by SuperScript® III Reverse Transcriptase (Invitrogen) with the OZ2-R1 primer (**Table Apx1.2**) from RNA extracted from wild-type Col Arabidopsis using PureLink® RNA Mini Kit (Invitrogen). *OZ2* was amplified from *OZ2* cDNA with Phusion polymerase (Thermo Scientific) and the primers OZ2-F1 and OZ2-nostop-R1 using standard protocols. 3'-A overhangs were added with Taq (QIAGEN) by incubating at 37 °C for 10 minutes. After purification, the amplicons were cloned into pCR8/GW/TOPO (Invitrogen) to use in a Gateway cloning reaction with a modified pEXSG vector [2] containing an EYFP C-terminal tag using LR Clonase II (Invitrogen) to produce pEXSG-OZ2-YFP.

Arabidopsis Col-0 plants and *Nicotiana benthamiana* were grown on soil in long day (16 hr) conditions for 3–5 weeks for Arabidopsis and 5–6 weeks for *N. benthamiana*. pEXSG-OZ2-YFP was transfected into Arabidopsis and *N. benthamiana* protoplasts using the method outlined in [3], using 3.0×10^5 cells per transformation.

Protoplast mitochondria were stained with MitoTracker™ Orange CM-H_2TMRos (500 nM; ThermoScientific), using DMSO as the solvent and W5 buffer. Protoplasts were incubated in the dark for 45 minutes and then were pelleted (1000 xg, 5 min) and resuspended in W5 buffer (500 µL). Protoplasts were imaged using a Zeiss Axio Observer LSM 710 microscope and C-Apochromat 40x/1.20 W Korr M27 objective.

Yeast two-hybrid assay

The mature coding sequences (without predicted N-terminal mitochondrial transit peptide) of OZ2 and mitochondrial splicing factors ABO5, BIR6, MAT2, mCSF1, MISF26, MISF74, MTL1, OTP439, PMH2, RUG3, SLO3, and WTF9 were amplified using Y2H primer pairs listed in **Table Apx1.2**. PCR products were first cloned into pCR8/GW/TOPO and then pGADT7GW and pGBKT7GW yeast two-hybrid destination vectors via Gateway cloning. Empty pGADT7GW and pGBKT7GW vectors were used as negative controls in yeast two-hybrid assays. Yeast mating strains PJ69-4a and PJ69-4α were individually transformed with pGADT7GW and pGBKT7GW plasmids, respectively. Single transformants were mated to produce diploid double-transformant yeast on YPAD agar plates. Yeast harboring testing pairs were grown in leucine- and tryptophan-deficient media overnight; then, 10 μL of each culture was spotted onto leucine-, tryptophan-, histidine-, adenine-deficient media plates after being diluted with water to OD600 0.5, 0.05, and 0.005. Survival/growth plates were imaged after three days of incubation at 30°C.

Bimolecular fluorescence complementation

The full-length coding sequences (including N-terminal mitochondrial transit peptides but without the stop codon) of OZ2 and mitochondrial splicing factors ABO5, BIR6, MAT2, mCSF1, MISF26, MISF74, MTL1, OTP439, PMH2, RUG3, SLO3, and WTF9 were amplified using BiFC primer pairs listed in **Table Apx1.2**. PCR products were first cloned into pCR8/GW/TOPO and then pEXSG-nYFP and pEXSG-cYFP BiFC destination vectors via Gateway cloning. *A. tumefaciens* GV3101::pMP90RK

cells were transformed via electroporation using 1 µg plasmid. Electroporation was conducted with the following parameters: capacitance 25 µF, voltage 2.0 kV, resistance 200 Ω, pulse length ~5 msec. Agrobacteria were selected on LB agar plates containing kanamycin (50 µg/mL), gentamicin (25 µg/mL), and carbenicillin (50 µg/mL).

5-mL cultures of individual transformed Agrobacteria were incubated for two days at 28°C and resuspended in a solution of 2-(N-morpholine)-ethanesulphonic acid (pH 5.6), 10 mM $MgCl_2$, and 150 mM acetosyringone. Infiltration samples were made by mixing bacterial cultures carrying pEXSG-nYFP, pEXSG-cYFP, and P19 at equal OD_{600} to a final OD of 0.9. Leaves of 4- to 6-week-old *N. benthamiana* plants grown in long-days conditions were agroinfiltrated as described in [4]. 2–3 days post-infiltration, 5-mm squares were cut from the infiltrated leaf area and imaged using fluorescence microscopy as described above.

Table Apx1.2. Primers used in OZ2 study.

Name	Sequence	Purpose
OZ2-F1	ATGGCTGCTTCAATCTCTCTTCTTCTC	Cloning of OZ2 CDS
OZ2-R1	CTATCTCTCGATAACTCTTCGACTGTTTCTG	
OZ2-nostop-R1	TCTCTCGATAACTCTTCGACTGTTTCTG	OZ2 BiFC
OZ2-Y2H-F	AGCAACACTGAATCAACCCATGAA	Primers used to amplify genes whose products were tested in Y2H
ABO5-Y2H-F	GCCACAAAGTATGTCGCCAAAG	
ABO5-Y2H-R	CTACAAAGGGCTGACAACCCA	
BIR6-Y2H-F	TCTTCGAAACCAGATTCTATGCTT	
BIR6-Y2H-R	TCAAGCTGCAGCACCAAAG	
MAT2-Y2H-F	TGGCTAAAACCTTCTTCTACATATACC	
MAT2-Y2H-R	TTACATGCGTGCAATGCGAAGC	
mCSF1-Y2H-F	GCCTCCGAAAATCCTGAC	
mCSF1-Y2H-R	CTAGGTTGTCTCGTCAGGAG	
MISF26-Y2H-F	GCATTGACGATTCATATCCTTGTCAAAGC	
MISF26-Y2H-R	TCATCTGACTGAAATCATCTCATTATAAACTCTATC	
MISF74-Y2H-F	GAAGCTCGAAAACCGATTGTCTCG	

MISF74-Y2H-R	CTATAAATGTCTCCTCCTCTCTCTAGCA	
mTERF15-Y2H-F	TCCTCACGAATTCTCACACCTATCAA	
mTERF15-Y2H-R	CTAAGCAAGTGACTCTATAAAGGAC	
MTL1-Y2H-F	TCTTTGCAAAGAATCTGCTACTACG	
MTL1-Y2H-R	TCAAAAAGCTGCATTATAAAACCTTCG	
OTP439-Y2H-F	AATCTCAATGTGAATCATCTCCTC	
OTP439-Y2H-R	CTATATGGTAAACTGATCATGGG	
PMH2-Y2H-F	GCTGGATTTGCGATCTCTG	
PMH2-Y2H-R	TCAGTAAGATCTTTTCCCATCATTTG	
RUG3-Y2H-F	ACAAGCCCAGACATCGACTCCG	
RUG3-Y2H-R	TTAAGGTGATCTTGAGACTAAACACAGAGC	
SLO3-Y2H-F	TCTTCATCTTTATCATCTACATCGA	
SLO3-Y2H-R	CTACATTAGCTGTATTTCAGGAGGAG	
WTF9-Y2H-F	CACTTCTTAAGGAAGTTTCCTTCAATCTTTG	
WTF9-Y2H-R	TTAGCCTTCAAAATCCAAATCCAAATCTTTATC	
ABO5-BiFC-F	ATGAAGCTTCTCCGCCGC	Primers used to amplify genes whose products were tested in BiFC
ABO5-BiFC-R	CAAAGGGCTGACAACCCAC	
BIR6-BiFC-F	ATGTACAGATCAATGGCAATCCTG	
BIR6-BiFC-R	AGCTGCAGCACCAAAGAGTTC	
MAT2-BiFC-F	ATGCGTAGAAGCTTCTCTGTTTTGGG	
MAT2-BiFC-R	CATGCGTGCAATGCGAAGC	
mCSF1-BiFC-F	ATGTTCTTGATTCGTCTCTCCCG	
mCSF1-BiFC-R	GGTTGTCTCGTCAGGAGAATCTTG	
MISF26-BiFC-F	ATGGCGTCAGCTTTGCG	
MISF26-BiFC-R	TCTGACTGAAATCATCTCATTATAAACTCTATCAGC	
MISF74-BiFC-F	ATGATTCGCCGGCCGATC	
MISF74-BiFC-R	TAAATGTCTCCTCCTCTCTCTAGCATTTTTACG	
mTERF15-BiFC-F	ATGGCCTCAAAACTCAAAACCTTCA	
mTERF15-BiFC-R	AGCAAGTGACTCTATAAAGGACTTCATGT	
MTL1-BiFC-F	ATGGTTATGCTAGCGAGATCAAAGCTAG	
MTL1-BiFC-R	AAAAGCTGCATTATAAAACCTTCGTCTAT	
OTP439-BiFC-F	ATGTACTTGAAAGAAGGATGTTCG	
OTP439-BiFC-R	TATGGTAAACTGATCATGGGCTCT	
PMH2-BiFC-F	ATGATCACTACAGTGCTACGACGA	
PMH2-BiFC-R	GTAAGATCTTTTCCCATCATTTGA	
RUG3-BiFC-F	ATGGCAGCGTTAAGCCACC	
RUG3-BiFC-R	AGGTGATCTTGAGACTAAACACAGAGC	
SLO3-BiFC-F	ATGCTTCAGAAGATCTCCTCCGAT	
SLO3-BiFC-R	CATTAGCTGTATTTCAGGAGGAGG	
WTF9-BiFC-F	ATGCTCTCTATTCGCCGCCA	
WTF9-BiFC-R	GCCTTCAAAATCCAAATCCAAATCTTTATCTAC	

REFERENCES

1. Bentolila, S., Gipson, A. B., Kehl, A. J., Hamm, L. N., Hayes, M. L., Mulligan, R. M. & Hanson, M. R. A RanBP2-type zinc finger protein functions in intron splicing in Arabidopsis mitochondria and is involved in the biogenesis of respiratory complex I. *Nucleic Acids Res.* **49**, 3490–3506 (2021).

2. Jakoby, M. J., Weinl, C., Pusch, S., Kuijt, S. J. H., Merkle, T., Dissmeyer, N. & Schnittger, A. Analysis of the subcellular localization, function, and proteolytic control of the Arabidopsis cyclin-dependent kinase inhibitor ICK1/KRP1. *Plant Physiol.* **141**, 1293–305 (2006).

3. Yoo, S.-D., Cho, Y.-H. & Sheen, J. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* **2**, 1565–1572 (2007).

4. Lin, M. T., Occhialini, A., Andralojc, P. J., Devonshire, J., Hines, K. M., Parry, M. A. J. & Hanson, M. R. β-Carboxysomal proteins assemble into highly organized structures in Nicotiana chloroplasts. *Plant J.* **79**, 1–12 (2014).

APPENDIX 2

OZ3 and OZ4 have unknown molecular functions*

The OZ family consists of four members in total: OZ1 and OZ2 have been discussed previously, but the other two family members, OZ3 (At1g70650) and OZ4 (At1g48570), have as yet unknown functions. As first described in Sun *et al.* [1], these two proteins share their N-terminal and RanBP2-type zinc fingers with OZ1 and OZ2, although OZ3 has three Znfs and OZ4 has four (**Figure Apx2.1**).

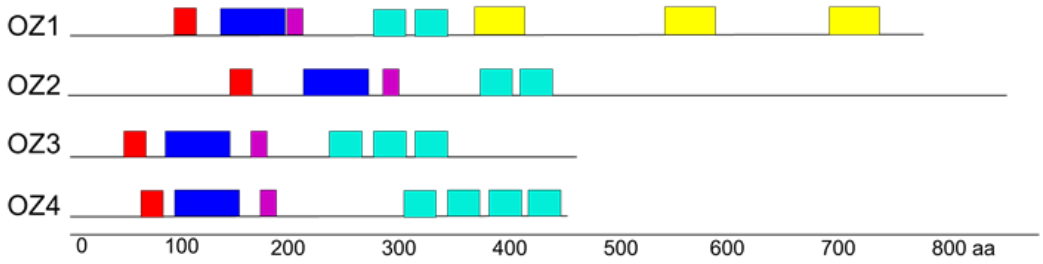


Figure Apx2.1. Linear schematic of OZ family proteins and their domains. Red = 19 residue-long motif, blue = 60 residue-long motif, purple = 16 residue-long motif, teal = RanBP2 zinc fingers, yellow = 47 residue-long motif.

Previous localization studies of the OZ family were done not with full-length protein, but only with the first 100 amino acids of each protein fused to a C-terminal YFP [2]. The localization experiments were repeated with full-length OZ3 and OZ4, transiently expressing in *N. benthamiana* protoplasts. Contrary to localization predictions based on TargetP analysis [1] and even the prior truncated protein

* Localization experiments were performed by Alexander J. Kehl under my supervision. I performed all other experiments in this appendix.

experiments, OZ3 was found to be a mitochondrial protein, and OZ4 was found to be dual-localized to the mitochondria and chloroplasts (**Figure Apx2.2**).

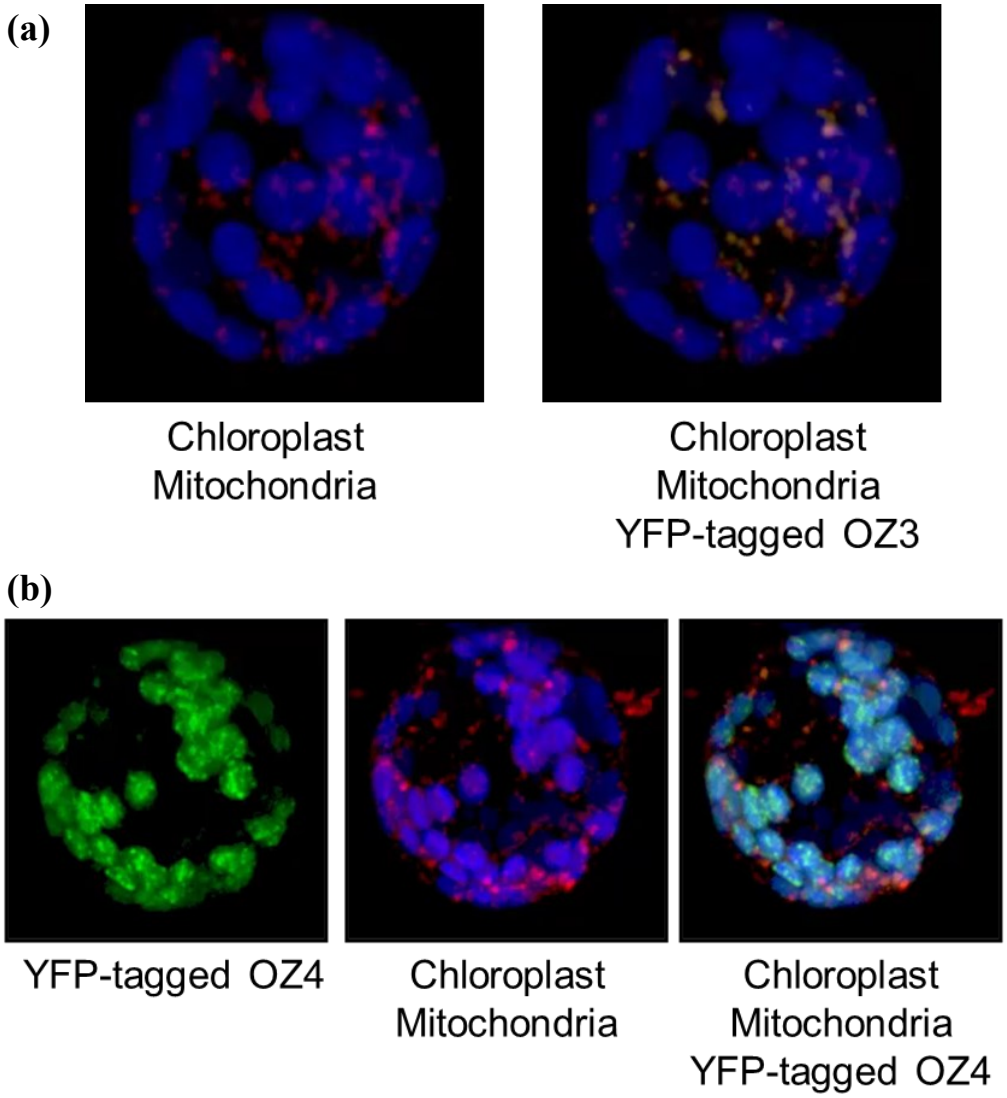


Figure Apx2.2. OZ3 is localized to the mitochondria, and OZ4 is dual-localized. (**a**) Transiently transformed *N. benthamiana* protoplasts expressing OZ3 fused to YFP. Blue = chloroplast autofluorescence, red = MitoTracker Orange, yellow-orange = overlap of YFP signal (green) and red MitoTracker signal. (**b**) Transiently transformed *N. benthamiana* protoplasts expressing OZ4 fused to YFP. Blue = chloroplast autofluorescence, red = MitoTracker Orange, green = YFP-tagged OZ4. YFP overlaps both with chloroplasts (green + blue = cyan) and mitochondria (green + red = yellow-orange), indicating dual localization.

T-DNA insertional mutants for OZ3 (SALK_089645) and OZ4 (SAIL_533_G03) were obtained to analyze the mutant morphology and molecular phenotype; based on the phenotypes of *oz1* and *oz2*, I expected to see significant effects of *oz3* and *oz4* mutations. Surprisingly, pre-flowering rosette mutant plants grown in short-day conditions did not exhibit any notable defects (**Figure Apx2.3a)**. Once reaching the flowering stage, *oz3* mutants displayed some defects in flower development (**Figure Apx2.3b**) but were otherwise normal, while *oz4* plants exhibited no aberrancy. Bulk Sanger sequencing analysis of chloroplast *ndhB*, *ndhD*, and *rpoA* cDNA from *oz4* mutant plants did not reveal any defects in RNA editing (**Figure Apx2.4**), although a more comprehensive analysis is warranted.

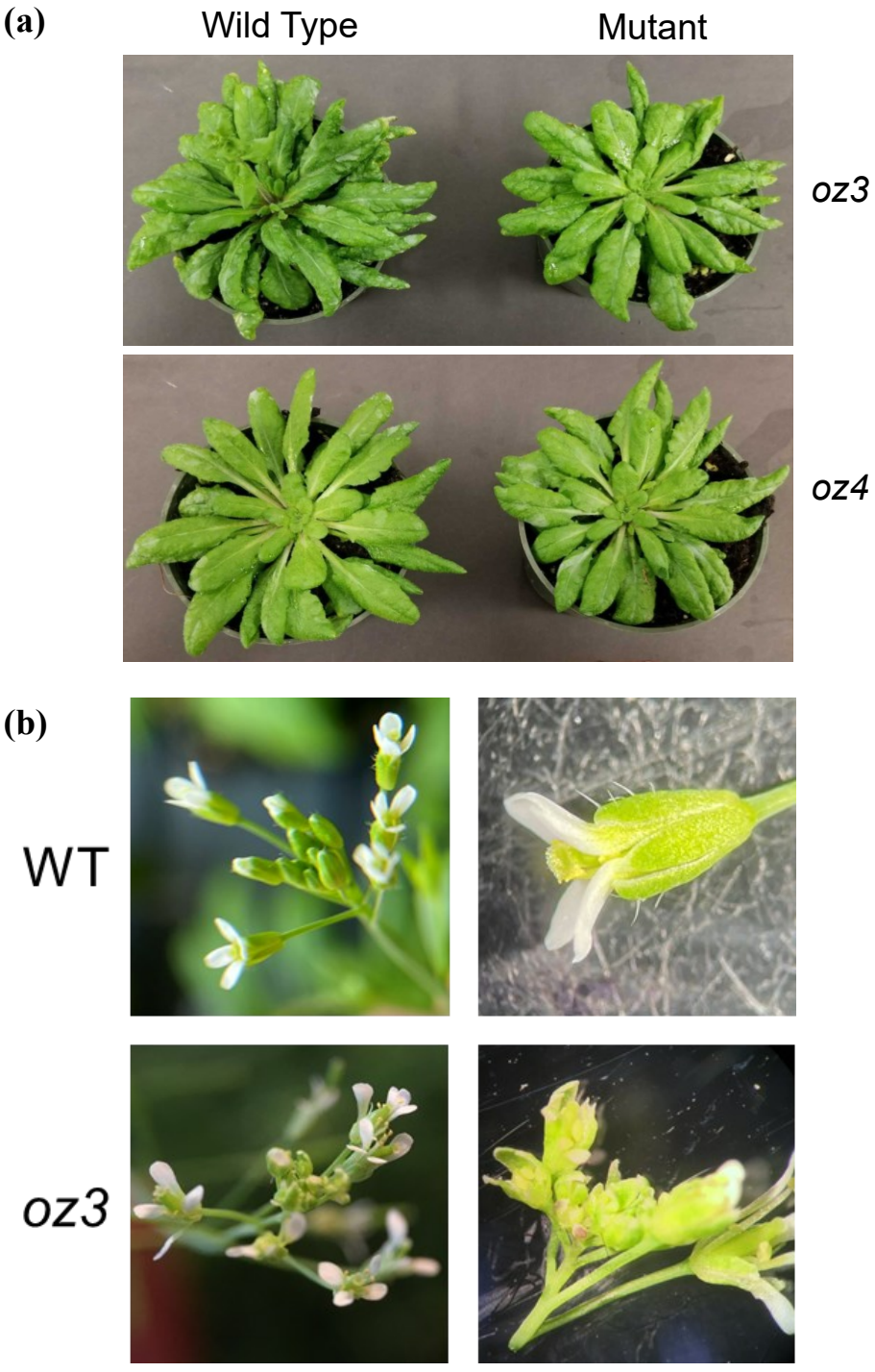


Figure Apx2.3. *oz3* and *oz4* mutant plants differ little from WT Arabidopsis. (**a**) 1-month-old Col WT, *oz3*, and *oz4* plants grown in short-day (8 hr) conditions (photo credit: M. R. Hanson). (**b**) Magnified images of flowers from WT and *oz3* plants.

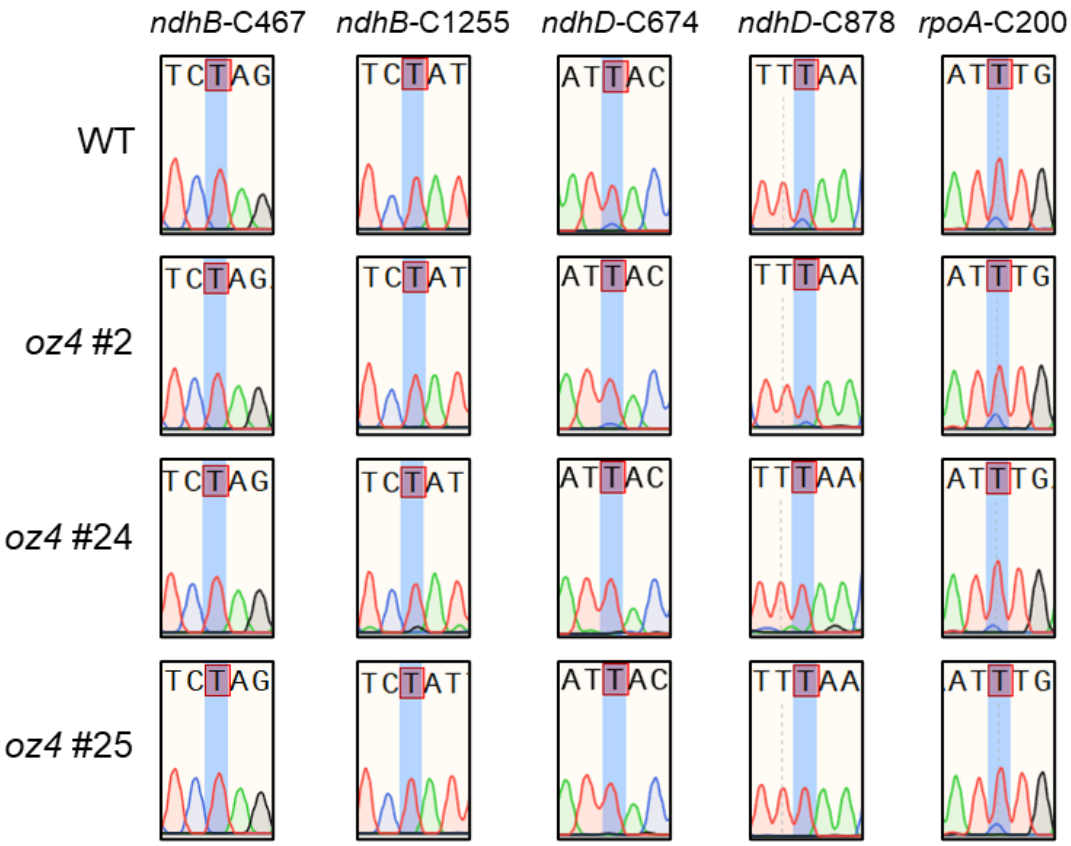


Figure Apx2.4. *oz4* plants do not have chloroplast RNA editing defects. Sanger sequencing traces of editing sites in WT and three *oz4* plants.

A somewhat accidental discovery in very old (4-month-old) *oz1* plants (**Figure Apx2.5a)** showed that editing was rescued at several sites in these old *oz1* plants, although not all editing sites analyzed were rescued (**Figure Apx2.5b**). Inspection of expression levels of OZ family proteins over the life cycle of Arabidopsis via Genevestigator [3] shows that expression of *OZ4* is increased during senescence (**Figure Apx2.6**, rightmost stage of development). Although it is speculation, this editing rescue in old *oz1* Arabidopsis could be due to the more highly expressed OZ4 taking over editing for the absent OZ1 protein.

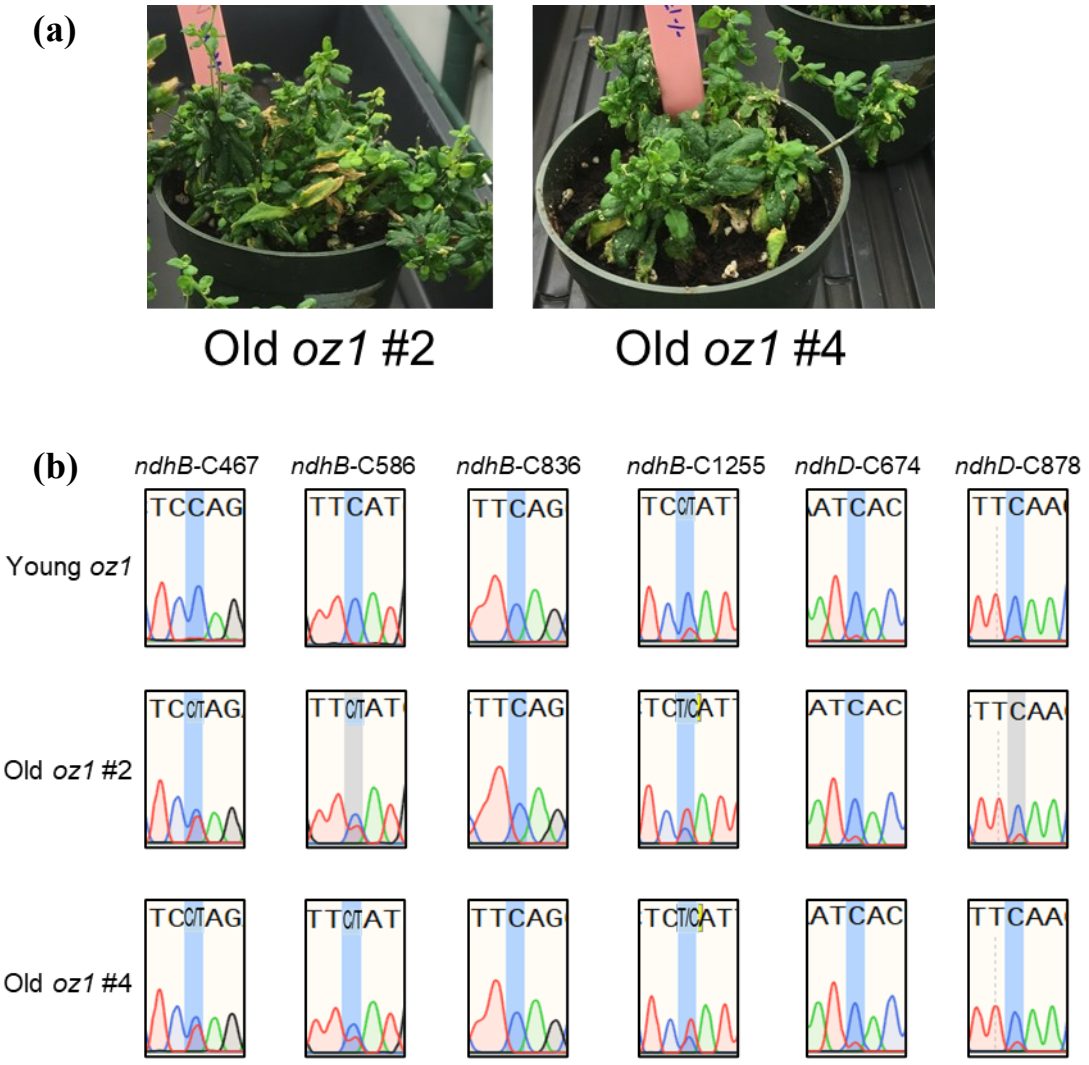


Figure Apx2.5. Senescent *oz1* plants exhibit RNA editing rescue at several chloroplast editing sites. (**a**) Four-month-old *oz1* homozygous plants grown in short-day (8 hr) conditions. (**b**) Sanger sequencing traces from young (3 weeks old) and two old (4 months old) *oz1* plants.

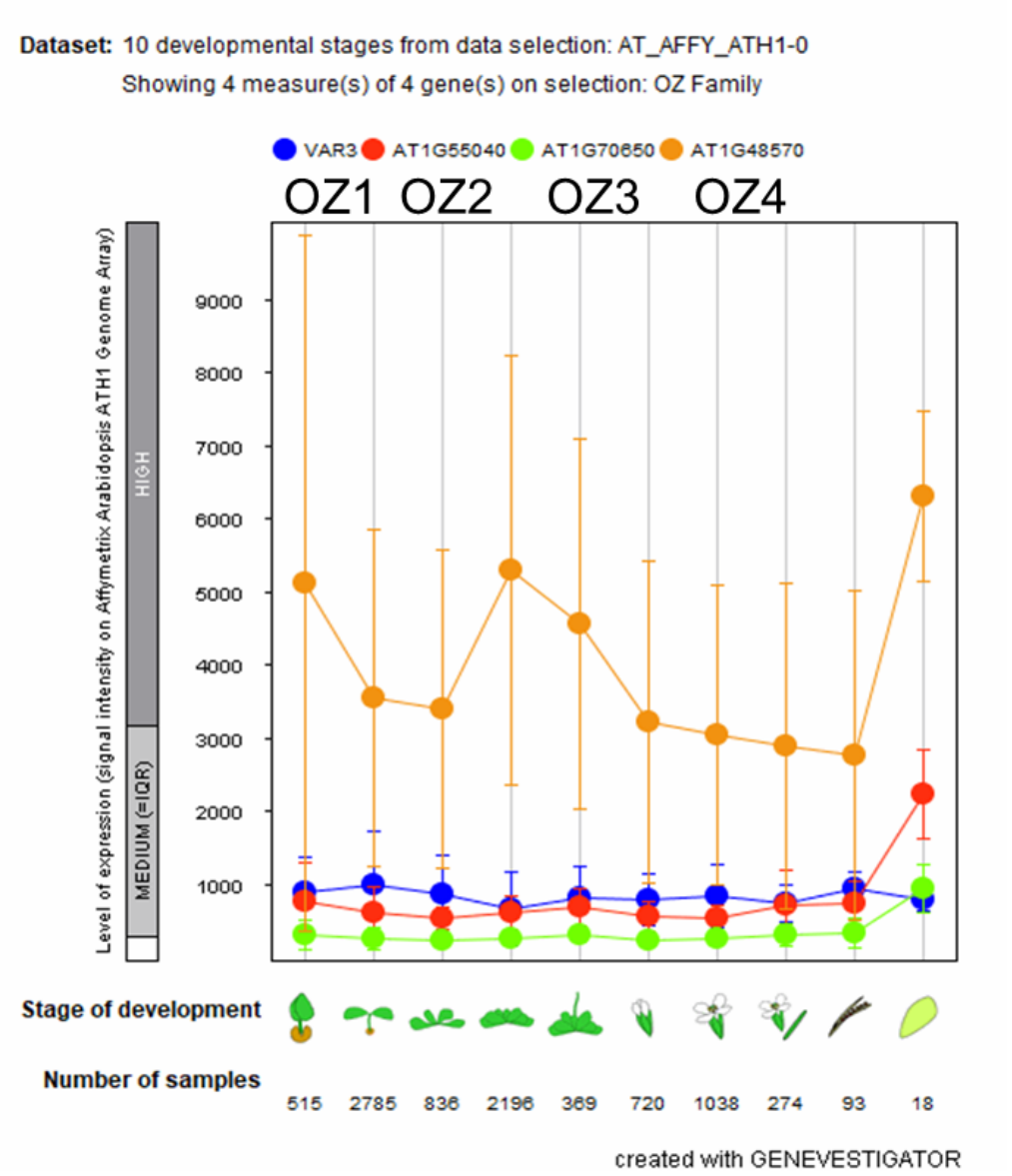


Figure Apx2.6. Expression levels of OZ family proteins in different stages of development as output by Genevestigator, based on microarray data. Blue = OZ1, red = OZ2, green = OZ3, orange = OZ4.

To pursue the possibility of OZ4 being a redundant editing factor in the chloroplast, I conducted yeast two-hybrid (Y2H) studies with OZ4 and other chloroplast editing factors which have already been tested with OZ1, including RARE1, OTP82, CRR28, and DYW2 (**Figure Apx2.7**). Except for RARE1, OZ4 does

in fact interact with these editing factors. A cross-reference with the editing sites rescued in old *oz1* plants does not reveal a discernable pattern with the Y2H data; *ndhB*-C467, which is targeted by CRR28 (binds to OZ4), is rescued, while *ndhD*-C878, another editing site targeted by CRR28, is barely rescued if at all in the old *oz1* plants (**Figure Apx2.5b**). Conversely, *ndhB*-C836, which is targeted by OTP82 (binds to OZ4), remains unedited in the old *oz1* plants (**Figure Apx2.5b**). This warrants further development to characterize the editing of these old *oz1* plants in more transcripts.

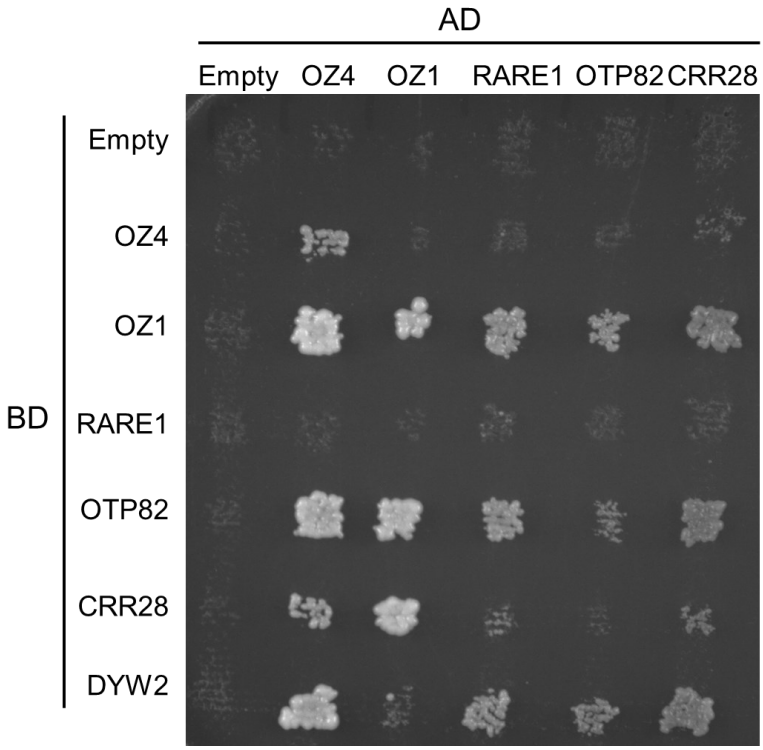


Figure Apx2.7. Yeast two-hybrid shows that OZ4 interacts with chloroplast RNA editing factors. Plates were prepared by streaking lines of individual yeast with proteins of interest either in the pGADT7 (AD) prey plasmid or the pGBKT7 (BD) bait plasmid onto selective media (-Leu for pGADT7, -Trp for pGBKT7), stamping onto YPAD media in a matrix pattern to produce diploid yeast at the intersections, then stamping onto diploid selective media (-Leu -Trp), and finally stamping onto assay media (-Leu -Trp -His -Ade).

I have crossed *oz1* and *oz4* plants to produce double knockout offspring with the intent of allowing them to reach 4 months of age and analyzing their editing; if OZ4 is indeed "filling in" for OZ1 in senescent Arabidopsis, I would expect these double mutants to not display editing rescue in the later stages. I initially found double homozygous mutants in F2 offspring of the cross (**Figure Apx2.8**), but none of these mutants survived after transfer to soil. This is perhaps a result of the double mutation, and in the future, growth in Magenta boxes may prevent early death.

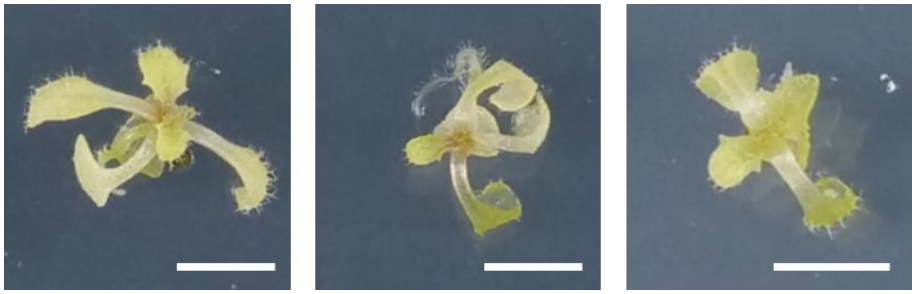


Figure Apx2.8. *oz1/oz4* double mutant seedlings (2 weeks old). Scale bars = 5 mm.

Considering that OZ3 and OZ4 both occur in the mitochondria, they may share a function in that organelle, such that single mutants of either gene would not reveal a phenotype. I crossed *oz3* and *oz4* to obtain double mutants and test for either morphological or molecular defects. This yielded two double mutant plants (**Figure Apx2.9**); however, sequencing of mitochondrial cDNA for transcripts *ccmB*, *ccmC*, *nad1*, *nad2*, *nad4*, *nad5*, *nad7*, and *rps3* showed no sign of splicing or editing defects. As of yet, the precise molecular function of OZ3 and OZ4 is unknown, and we cannot say if they are joined with OZ1 and OZ2 in being major actors in plant organelle RNA processing.

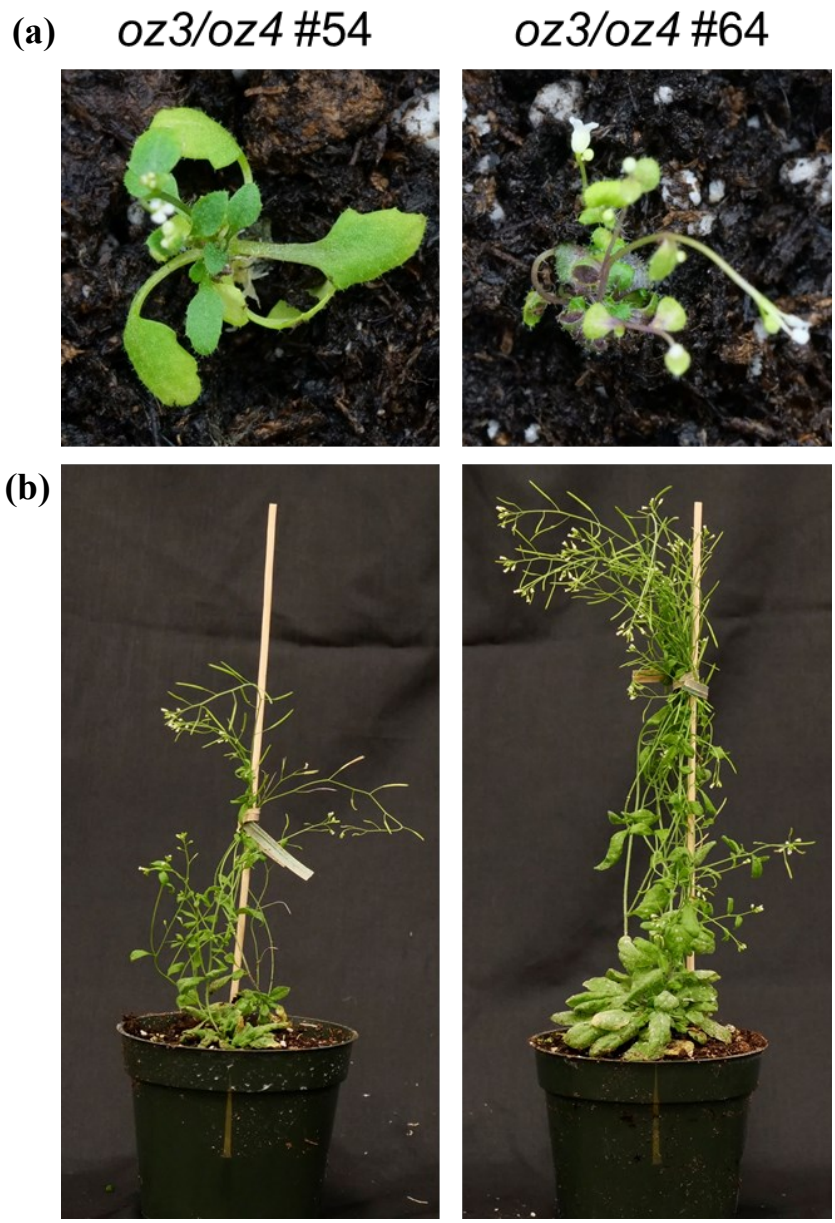


Figure Apx2.9. *oz3/oz4* double mutant plants. (**a**) 1-month-old seedlings grown on MS media and transferred to soil. (**b**) 2.5-month-old *oz3/oz4* plants.

MATERIALS AND METHODS

OZ2 localization

The cDNA clones of OZ3 (At1g70650) and OZ4 (At1g48570) used in this study were reverse-transcribed by SuperScript® III Reverse Transcriptase (Invitrogen) with oligo-dT primers from RNA extracted from wild-type Col Arabidopsis using PureLink® RNA Mini Kit (Invitrogen). *OZ3* and *OZ4* were amplified from the cDNA with Phusion polymerase (Thermo Scientific) and the primers listed in **Table Apx2.1** using standard protocols. 3'-A overhangs were added with Taq (QIAGEN) by incubating at 37 °C for 10 minutes. After purification, the amplicons were cloned into pCR8/GW/TOPO (Invitrogen) to use in a Gateway cloning reaction with a modified pEXSG vector [4] containing an EYFP C-terminal tag using LR Clonase II (Invitrogen) to produce pEXSG-OZ3-YFP and pEXSG-OZ4-YFP.

Nicotiana benthamiana were grown on soil in long day (16 hr) conditions for 5–6 weeks. pEXSG-OZ3-YFP and pEXSG-OZ4-YFP were transfected into *N. benthamiana* protoplasts using the method outlined in [5], using 3.0×10^5 cells per transformation.

Protoplast mitochondria were stained with MitoTracker™ Orange CM-H_2TMRos (500 nM; ThermoScientific), using DMSO as the solvent and W5 buffer. Protoplasts were incubated in the dark for 45 minutes and then were pelleted (1000 xg, 5 min) and resuspended in W5 buffer (500 µL). Protoplasts were imaged using a Zeiss Axio Observer LSM 710 microscope and C-Apochromat 40x/1.20 W Korr M27 objective.

Yeast two-hybrid assay

The mature coding sequences (without predicted N-terminal plastid/mitochondrial transit peptide) of OZ4 was amplified using Y2H primer pairs listed in **Table Apx2.1**. PCR products were first cloned into pCR8/GW/TOPO and then pGADT7GW and pGBKT7GW Y2H destination vectors via Gateway cloning. RARE1, OTP82, CRR28, and DYW2 Y2H plasmids and yeast strains were prepared in Chapter 2. Empty pGADT7GW and pGBKT7GW vectors were used as negative controls. Yeast mating strains PJ69-4a and PJ69-4α were individually transformed with pGADT7GW and pGBKT7GW plasmids, respectively. Single transformants were mated to produce diploid double-transformant yeast on YPAD agar plates. Yeast harboring testing pairs were stamped onto leucine- and tryptophan-deficient plates, grown, and then stamped onto leucine-, tryptophan-, histidine-, adenine-deficient media plates. Survival/growth plates were imaged after three days of incubation at 30°C.

RNA analysis

Leaves were taken from Arabidopsis plants, and RNA was extracted using Trizol and the PureLink RNA Mini Kit (ThermoFisher, Waltham, MA). RNA was treated with TURBO DNAse (Invitrogen, Waltham, MA) and quantified with the Qubit II (ThermoFisher, Waltham, MA). cDNA was amplified using Superscript III (Invitrogen, Waltham, MA) and either pooled plastid transcript primers or pooled mitochondrial transcript primers (**Table Apx2.1**). Gene-specific amplicons were then amplified with the corresponding primer pairs (**Table Apx2.1**) and Sanger sequenced.

Table Apx2.1. Primers used in OZ3, OZ4 study.

Name	Sequence	Purpose
OZ3_F	ATGTTGAGATTCTTCAAGACCGACC	Localization
OZ3_nostop_R	AGTCGATCTATCTTGTTTTGAAACAAACC	
OZ4_F	ATGTCATCTTCCAGAATATTCCTAGTCGG	
OZ4_nostop_R	CACTAAGGCGGGTCGTTTCC	
OZ4_139_F	GCAACCACCGTGGATTCG	Y2H
OZ4_withstop_R	TTACACTAAGGCGGGTCGTTTCC	
ndhB-F2	TTTTATGTGGTGCTAACGATTTAA	Plastid transcript cDNA sequencing
ndhB-F3	GCATGTACAGAATGAAAATTTCATTCT	
ndhB-R2	AATCGCAATAATCGGGTTCATT	
ndhD-F	AACAACTCGAAGTATGGGTC	
ndhD-R	GGCAATGCAAGGGAAGCC	
rpoA-F2	CTCGGACACTACAGTGGAAGTGTG	
rpoA-R2	CTGGGAGGCAATTCTAATTGGTC	
ccmB-F2	CAGCCTTGAAGTGAATGAATT	Mitochondrial transcript cDNA sequencing
ccmB-R2	TTAATCTTGTAAACTAATCGAGACC	
ccmC-F2	CTACGCGCAAATTCTCATTGG	
ccmC-R2	GAGCGAGTGAACTAGGTTTTGGTA	
nad1-F2	CCAGCTGAAATACTTGGAATAAT	
nad1-R1	AAAGGTGACTAAAAGACCAGAAAC	
nad2-F4	GACCGTAACGTAAGTGACTCAGTG	
nad2-R4	ACGGCCTACCCTTTCTTTGAA	
nad4-F4	ATGTTAGAACATTTCTGTGAATG	
nad4-R1	TTTGCCATGTTGCACTAAGTTACT	
nad5-F9	ATCAGAAGGAAGCGCTATAATGACCA	
nad5-R8	TAAAAACTACTCACTATCAAAATGAAAG	
nad7-F3	ATGACGACTAGGAAAAGGCAAATC	
nad7-R2	ATCCACCTCTCCAAACACAATA	
rps3-F1	AATCCGATTTCGGTAAGACTT	
rps3-R1	CGTTTCGGATATAGCACGTC	

REFERENCES

1. 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.* **11**, 1–23 (2015).

2. Shi, X. A family of trans-acting RNA editing factors in plant organelles. (Cornell University, 2016).

3. Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W. & Zimmermann, P. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Boinformatics* 420747 (2008) doi:10.1155/2008/420747.

4. Jakoby, M. J., Weinl, C., Pusch, S., Kuijt, S. J. H., Merkle, T., Dissmeyer, N. & Schnittger, A. Analysis of the subcellular localization, function, and proteolytic control of the Arabidopsis cyclin-dependent kinase inhibitor ICK1/KRP1. *Plant Physiol.* **141**, 1293–305 (2006).

5. Yoo, S.-D., Cho, Y.-H. & Sheen, J. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* **2**, 1565–1572 (2007).

APPENDIX 3

Supplemental OZ1 RNA editing, sequence alignment, and protein interaction data

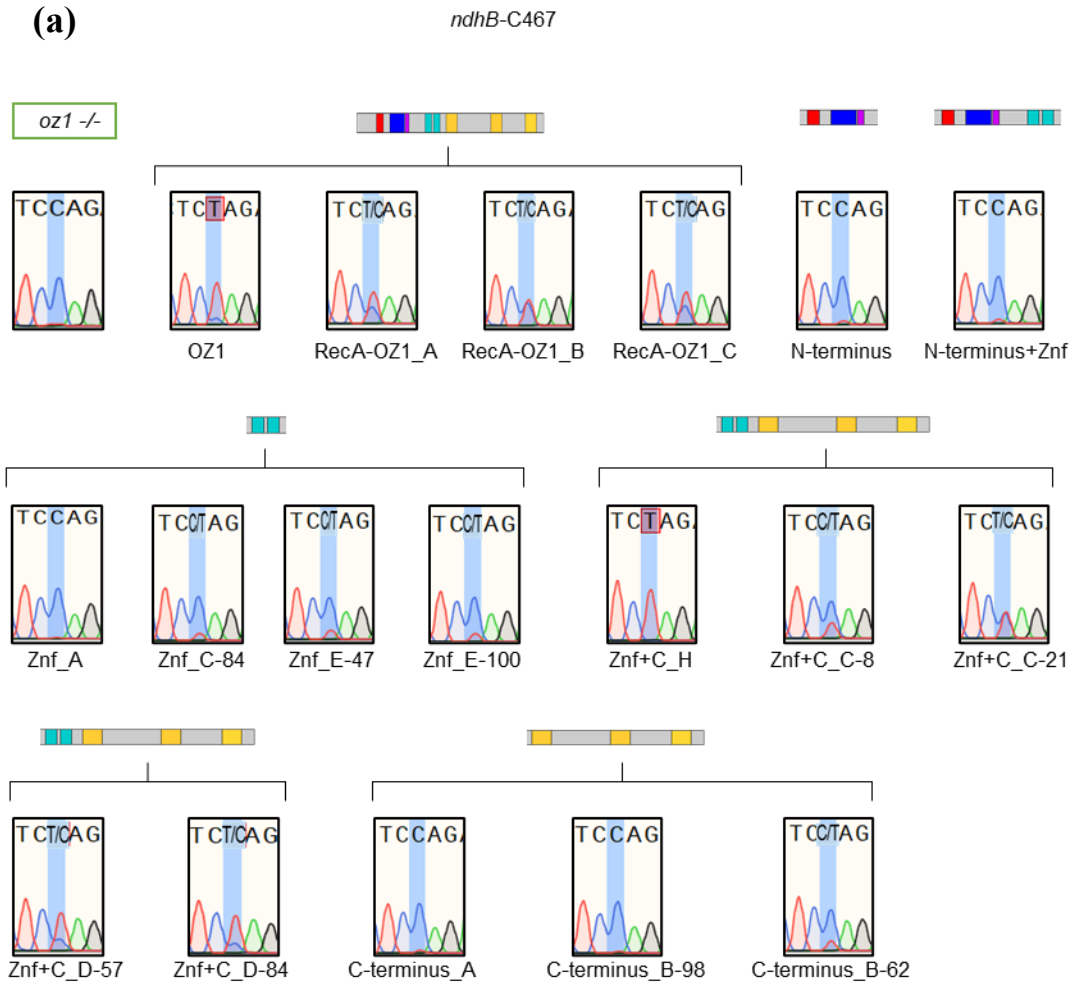


Figure Apx3.1. Sanger sequencing traces of RNA editing sites from OZ1 truncation construct-expressing plants and homozygous *oz1*. "OZ1" corresponds to an *oz1* plant expressing full-length OZ1, including the native transit peptide; "RecA-OZ1" refers to *oz1* plants expressing a construct of OZ1 with the first 33 amino acids replaced with 65 amino acids of the RecA transit peptide. Information about editing extent percentages in *oz1* and WT plants, and targeting PPR protein, are from [1]. (**a**) *ndhB*-C467, *oz1* editing = 0%, WT = 84%, PPR = CRR28; (**b**) *ndhB*-C836, *oz1* = 0%, WT = 95%, PPR = OTP82; (**c**) *ndhB*-C872, *oz1* = 0%, WT = 90%, PPR = QED1; (**d**) *ndhB*-C1255, *oz1* = 0%, WT = 99%, PPR = CREF7; (**e**) ndhD-C674, *oz1* = 16%, WT = 91%, PPR = OTP85; (**f**) ndhD-C878, *oz1* = 5%, WT = 85%, PPR = CRR28; (**g**) *ndhG*-C50, *oz1* = 0%, WT = 84%, PPR = OTP82; (**h**) *rpoA*-C200, *oz1* = 0%, WT = 71%, PPR = CLB19/DYW2; (**i**) *accD*-C1568, *oz1* = 2%, WT = 77%, PPR = QED1; (**j**) *clpP*-C559, *oz1* = 0%, WT = 61%, PPR = CLB19/DYW2.

Figure Apx3.1 cont.

(b)

Figure Apx3.1 cont.

(c)

Figure Apx3.1 cont.

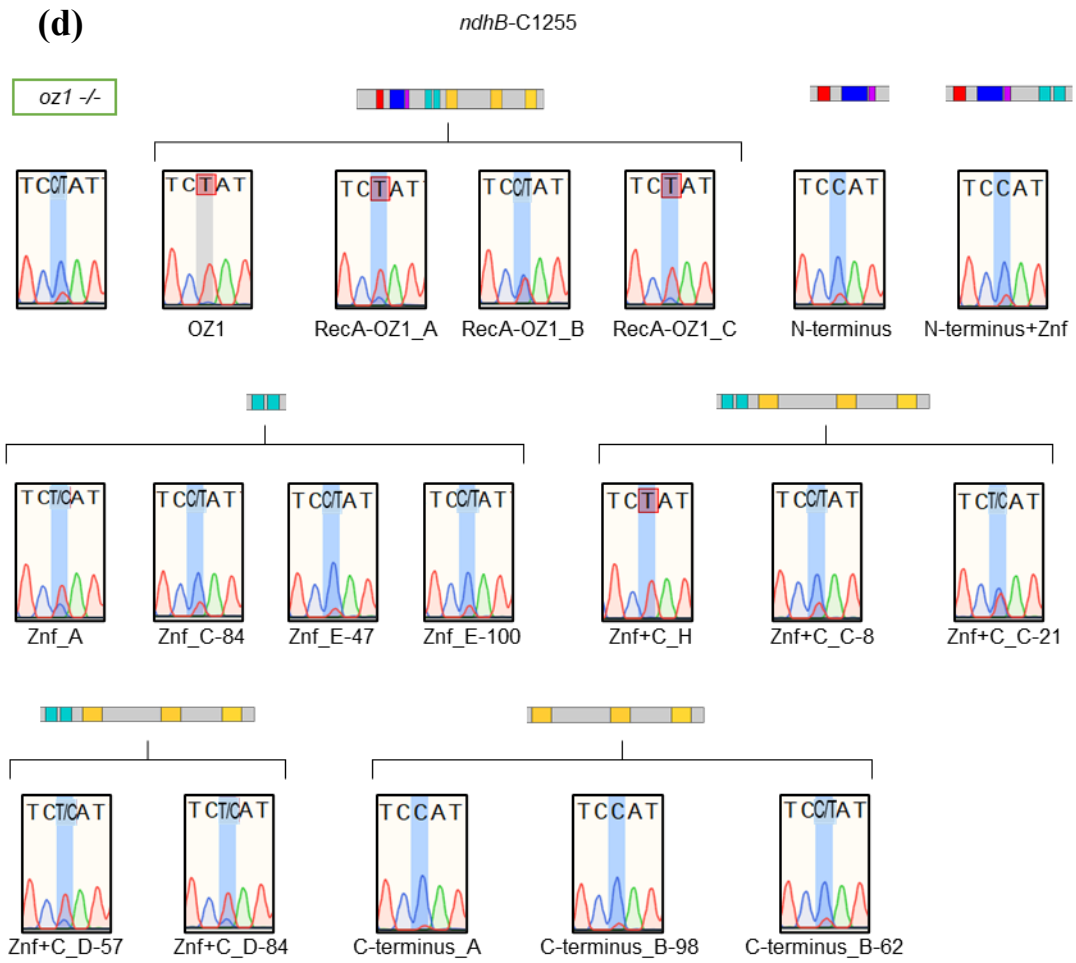


Figure Apx3.1 cont.

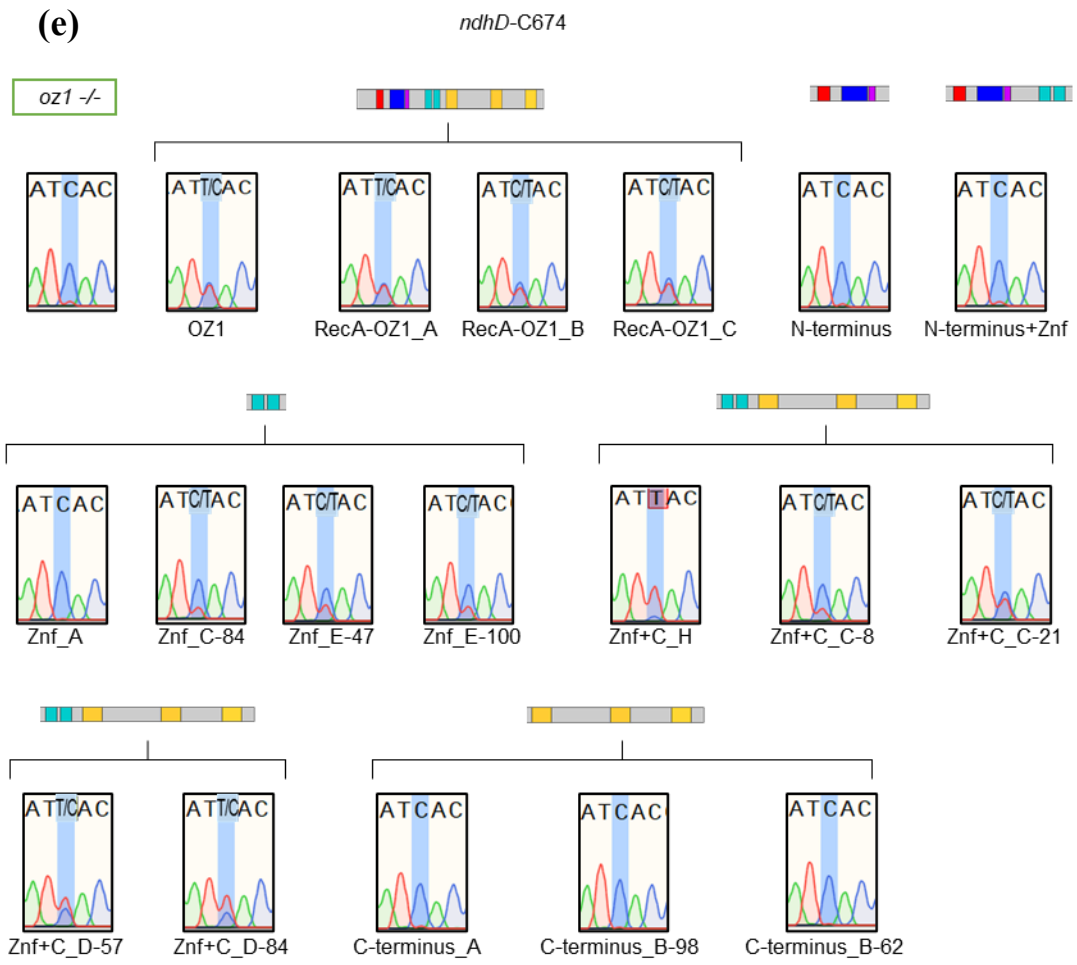


Figure Apx3.1 cont.

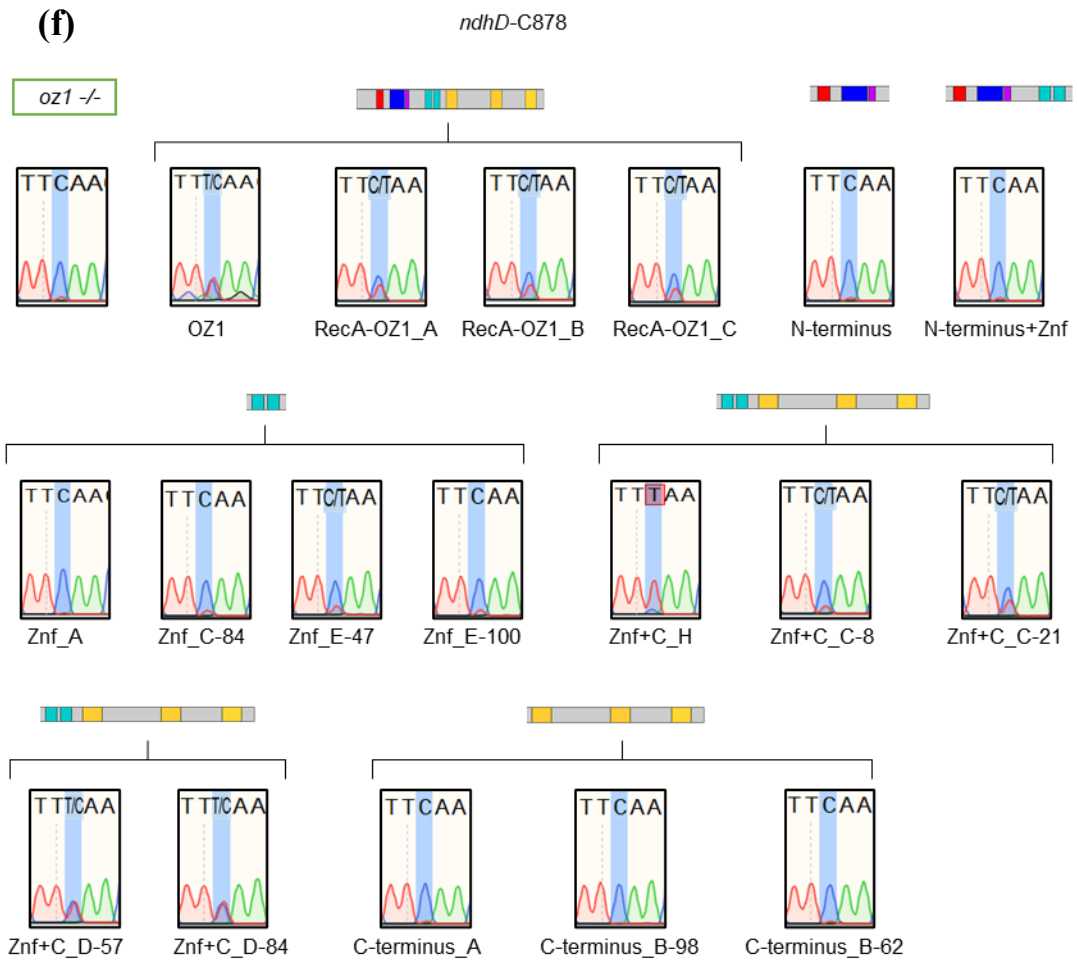


Figure Apx3.1 cont.

Figure Apx3.1 cont.

Figure Apx3.1 cont.

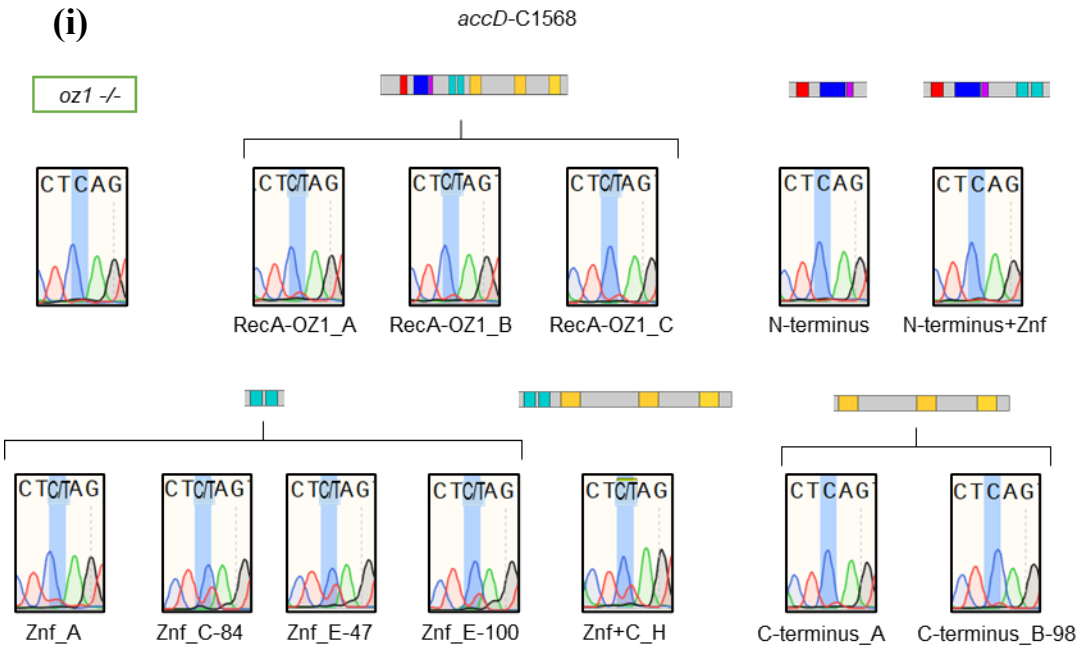


Figure Apx3.1 cont.

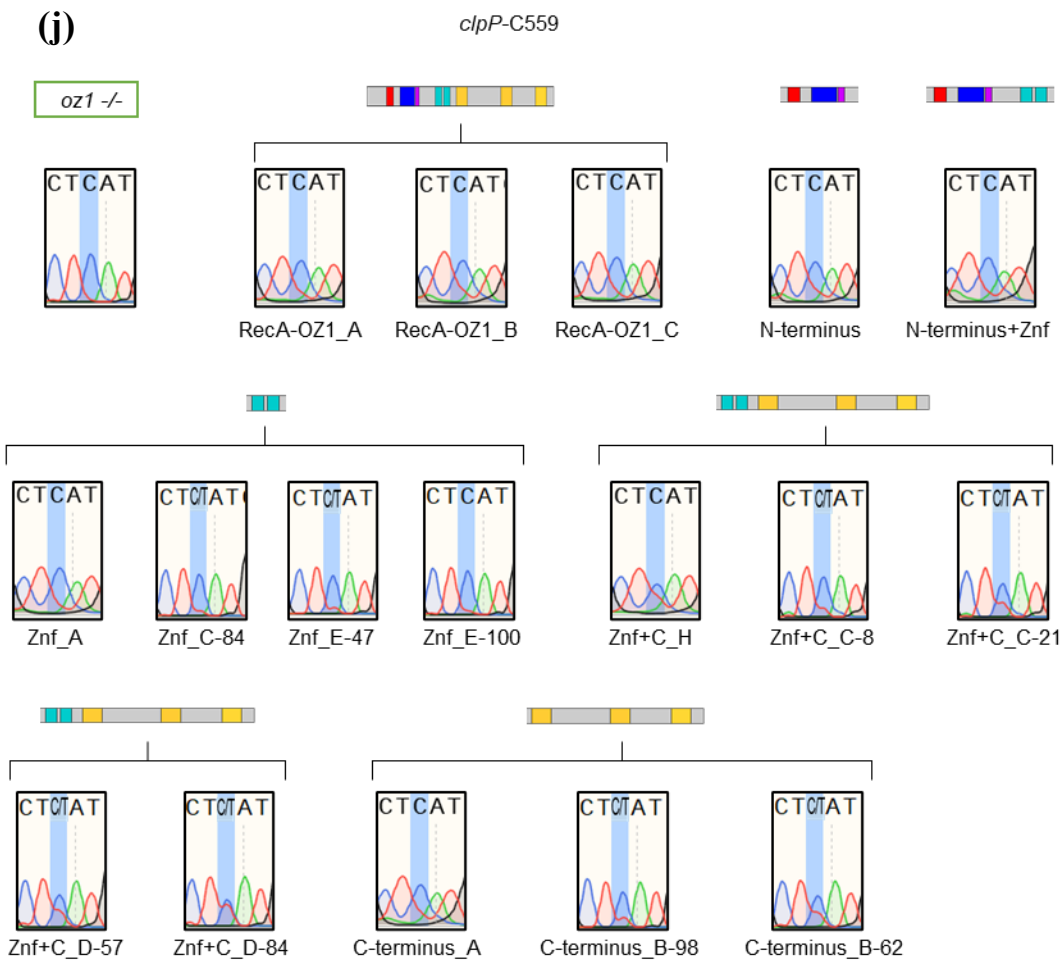


Figure Apx3.2. Sanger sequencing traces of RNA editing sites from OZ1 zinc finger mutant construct-expressing plants, homozygous *oz1*, and WT Arabidopsis. "OZ1" corresponds to an *oz1* plant expressing full-length OZ1, including the native transit peptide; "RecA-OZ1" refers to *oz1* plants expressing a construct of OZ1 with the first 33 amino acids replaced with 65 amino acids of the RecA transit peptide. Information about editing extent percentages in *oz1* and WT plants, and targeting PPR protein, are from [1]. (**a**) *ndhB*-C467, *oz1* editing = 0%, WT = 84%, PPR = CRR28; (**b**) *ndhB*-C836, *oz1* = 0%, WT = 95%, PPR = OTP82; (**c**) *ndhB*-C872, *oz1* = 0%, WT = 90%, PPR = QED1; (**d**) *ndhB*-C1255, *oz1* = 0%, WT = 99%, PPR = CREF7; (**e**) *ndhD*-C674, *oz1* = 16%, WT = 91%, PPR = OTP85; (**f**) *ndhD*-C878, *oz1* = 5%, WT = 85%, PPR = CRR28; (**g**) *ndhG*-C50, *oz1* = 0%, WT = 84%, PPR = OTP82; (**h**) *rpoA*-C200, *oz1* = 0%, WT = 71%, PPR = CLB19/DYW2; (**i**) *accD*-C1568, *oz1* = 2%, WT = 77%, PPR = QED1; (**j**) *clpP*-C559, *oz1* = 0%, WT = 61%, PPR = CLB19/DYW2.

Figure Apx3.2 cont.

Figure Apx3.2 cont.

Figure Apx3.2 cont.

Figure Apx3.2 cont.

Figure Apx3.2 cont.

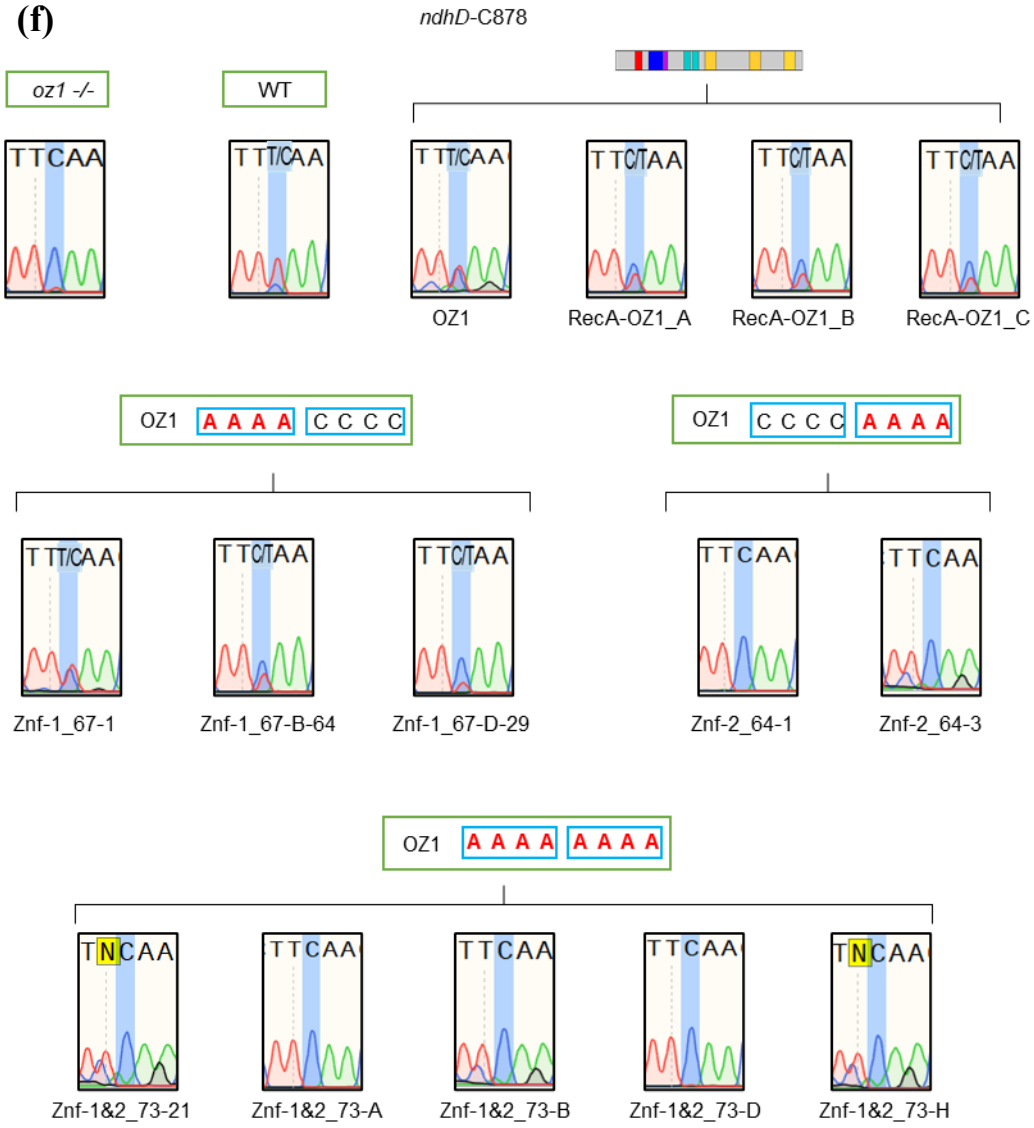


Figure Apx3.2 cont.

Figure Apx3.2 cont.

Figure Apx3.2 cont.

(i)

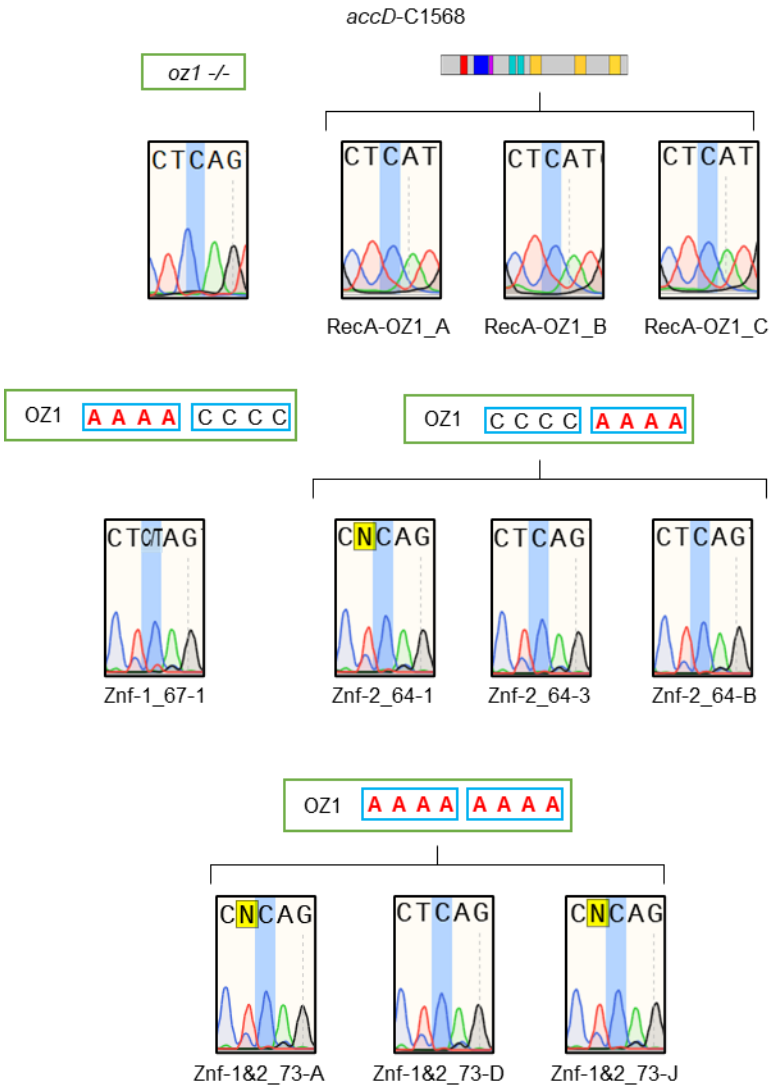
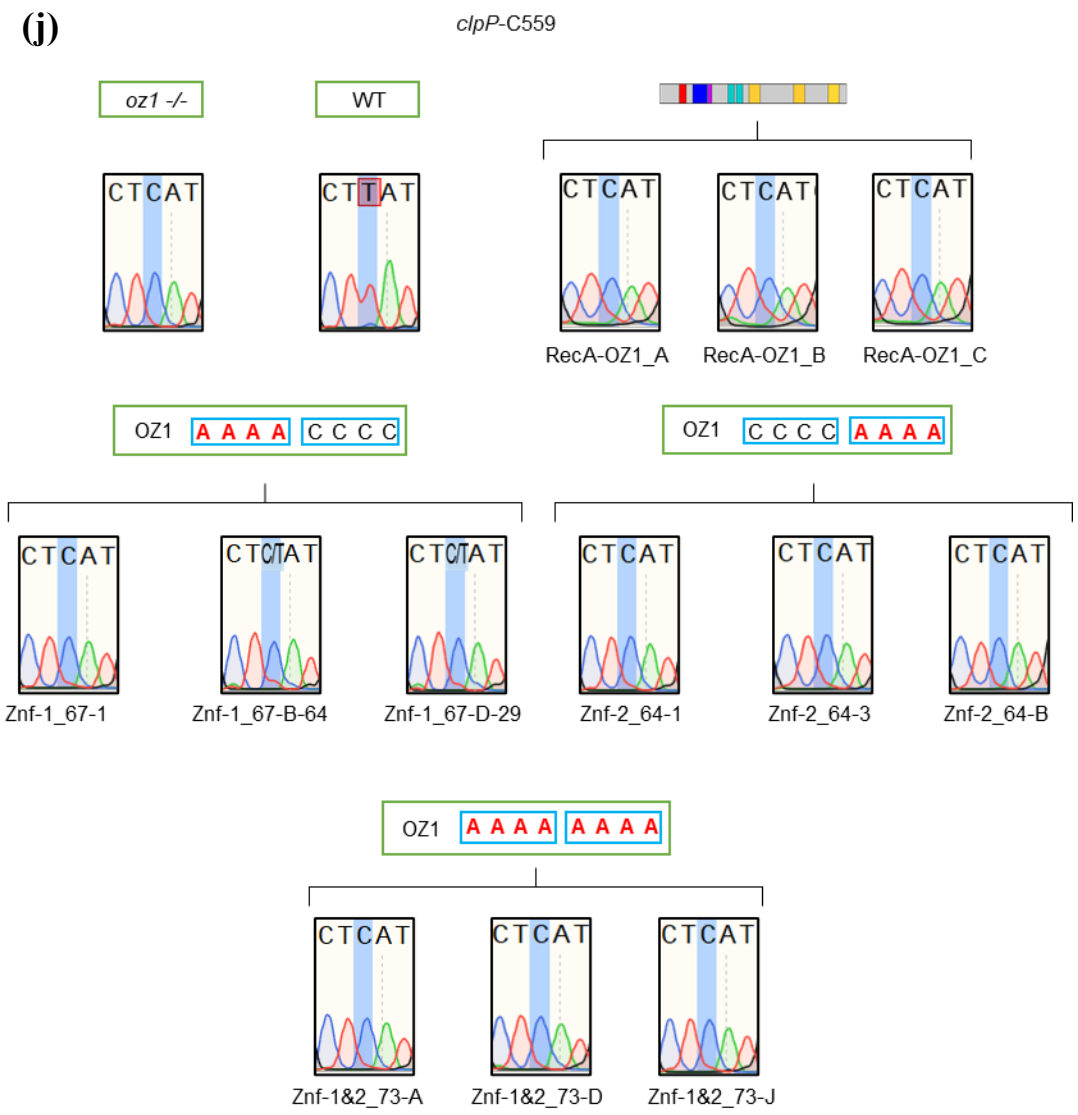


Figure Apx3.2 cont.



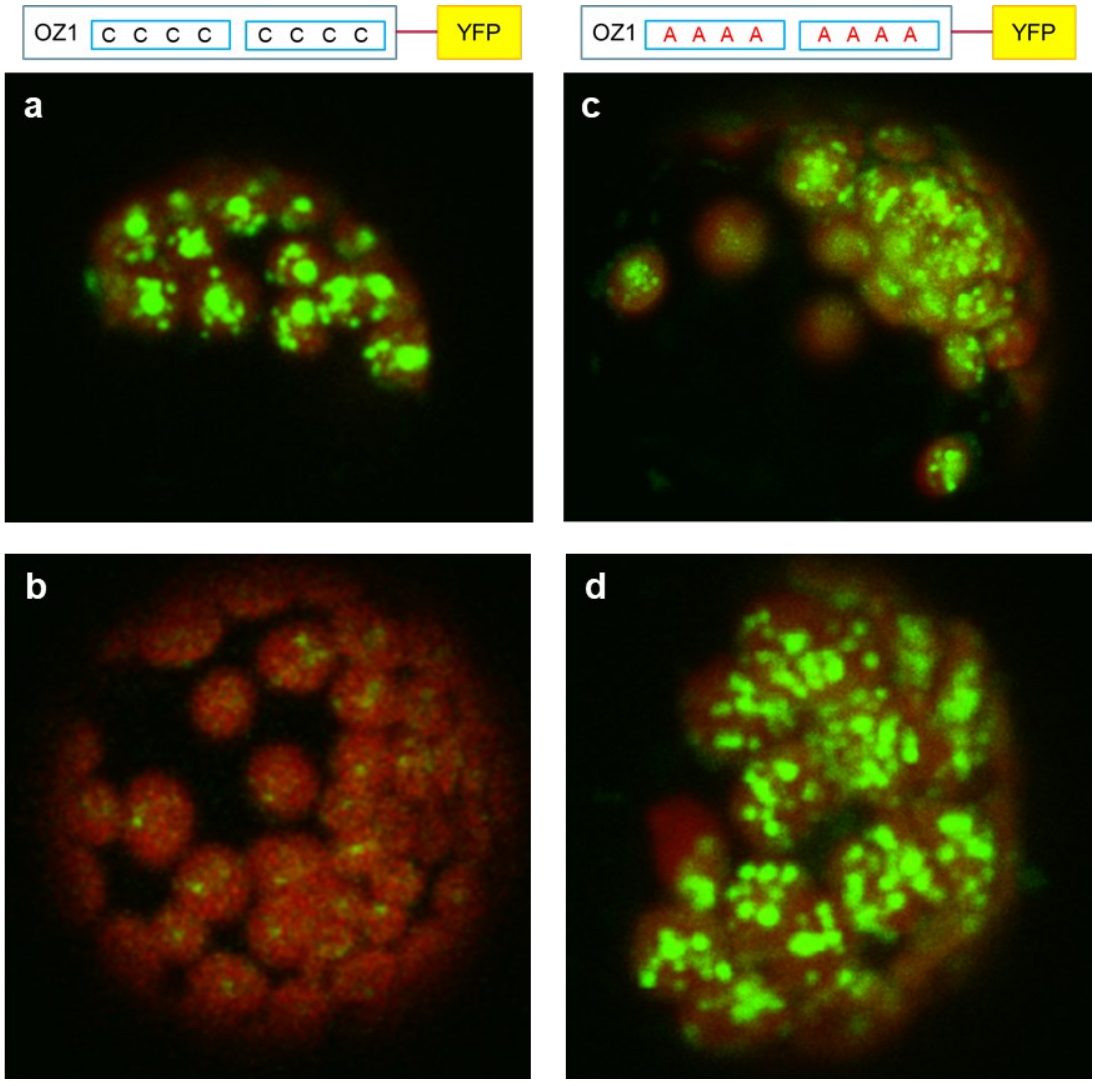


Figure Apx3.3. A-to-C mutation of the putative zinc-binding residues of the OZ1 zinc fingers does not prevent accumulation of OZ1 in the chloroplasts. Fluorescent micrographs of Arabidopsis protoplasts transformed with pEXSG plasmids expressing OZ1 WT sequence (**a**, **b**) or OZ1_Znf-1&2 (**c**, **d**) with a C-terminal fusion of full-length EYFP. Red = chlorophyll autofluorescence, green = YFP. pEXSG-OZ1-YFP and pEXSG-OZ1_Znf-1&2-YFP were transfected into Arabidopsis protoplasts using the method outlined in [2], using 3.0×10^5 cells per transformation.

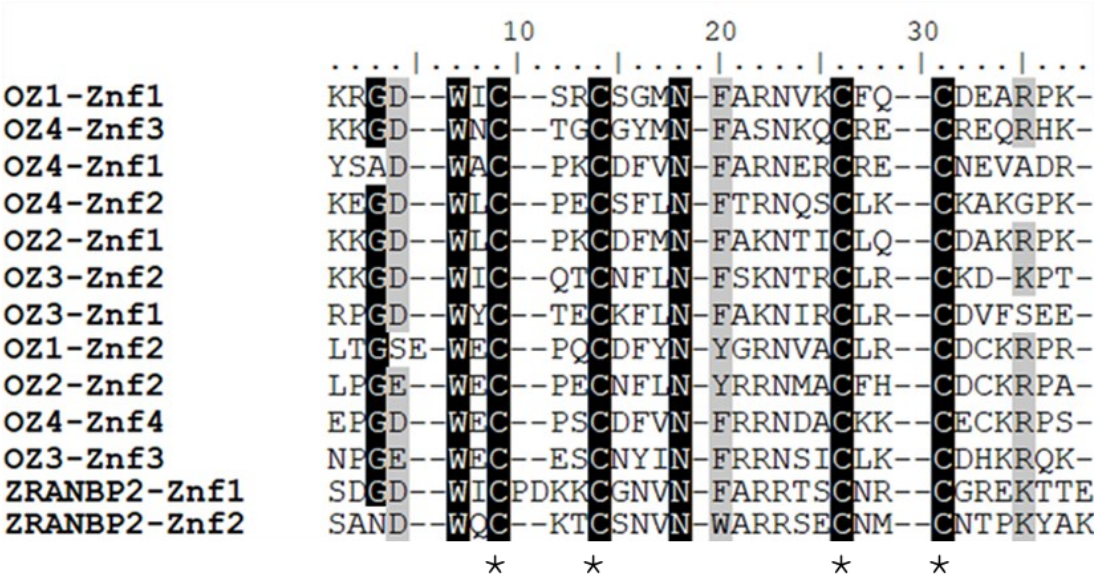


Figure Apx3.4. MUSCLE alignment of RanBP2 Znf domains from Arabidopsis OZ proteins and the human splicing factor ZRANB2. Asterisk = conserved putative zinc-binding cysteines; black shading = 85% identical amino acids; gray = similar amino acids.

Table Apx3.1. Summary of results from yeast two-hybrid assays between OZ1 and truncations of PPR proteins. Numbers in parentheses indicate the nucleotide sequence range of the truncation. PPR proteins were divided into the PPR tract, the E domains, and the DYW domain, and truncations were prepared as either individual domains or combinations (e.g., "EDYW" = E domains plus DYW domain).

PPR Protein Construct	OZ1 Interaction?
RARE1 (100-stop)	Yes
RARE1_PPR (100-1674)	No
RARE1_PPRE (100-1869)	Yes
RARE1_EDYW (1651-stop)	Yes
RARE1_E (1651-1869)	No
RARE1_DYW (1870-stop)	No
OTP82 (292-stop)	Yes
OTP82_PPR (292-1614)	Yes
OTP82_PPRE (292-1818)	Yes
OTP82_EDYW (1609-stop)	Yes
OTP82_E (1609-1818)	No
OTP82_DYW (1813-stop)	Yes
CRR28 (121-stop)	Yes
CRR28_PPR (124-1275)	Yes
CRR28_PPRE (124-1500)	Yes (weak)
CRR28_EDYW (1270-stop)	Yes (weak)
CRR28_E (1270-1500)	No
CRR28_DYW (1501-stop)	No
CRR4 (253-stop)	No
CLB19 (162-stop)	No
DYW1 (151-stop)	Yes (weak)
DYW2 (97-stop)	Yes

REFERENCES

1. 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.* **11**, 1–23 (2015).

2. Yoo, S.-D., Cho, Y.-H. & Sheen, J. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* **2**, 1565–1572 (2007).