ulm university universität ulm

Several proteins form variable RNA editing complexes in plant organelles

Dissertation

zur Erlangung des Doktorgrades Dr. rer. nat.

der Fakultät für Naturwissenschaften

der Universität Ulm

vorgelegt von

Eszter Bayer-Császár

aus Szeged

Ulm 2016

Amtierender Dekan der Fakultät für Naturwissenschaften:
Prof. Dr. Peter Dürre

Erstgutachter:
PD. Dr. habil. Mizuki Takenaka

Zweitgutachter:
Prof. Dr. Anita Marchfelder

Datum der Promotion: 01.02.2017

Index

1 Summary

In RNA editing in plant organelles, PPR (Pentatricopeptide Repeat) proteins with an E or E and DYW domains at the C-terminus are important participants for recognizing the specific *cis*-elements. These proteins are able to connect to their target RNA sequence via the PPR domains. Recognition of the specific target RNA sequence by PPR proteins is based on the specific binding of one PPR motif to one nucleotide.

There are several possibilities for the function of the E-domains in the PPR RNA editing factors: they may bind to the RNA as well as the PPR motifs, or provide the interacting surface for other proteins involved also in the RNA editing process. It is also not excluded that E domains perform these two possible tasks simultaneously. The DYW domain contains a conserved motif ($HxEx_nCxxC$) of cytidine deaminases and shows zinc-binding capability, suggesting the domain works as the so far missing deaminase enzyme. But its cytidine deaminase activity has not yet been proofed.

The MORFs (Multiple Organellar RNA Editing Factors) also termed as RIPs (RNA editing factor Interacting Proteins) are also involved in several C-to-U substitution events in flowering plants mitochondria and/or in chloroplasts. Nine members of this family contain the conserved 100 amino acids long MORF box and one protein contains only the C-terminally half of that. The exact function of these proteins is unknown yet, but it is suggested that they work together with PPR proteins to perform certain RNA editing events maybe through their direct connections to PPR proteins at each MORF box.

To understand further detailed functions of the three domains PPR, E and DYW in RNA editing factors, I employed three different approaches. The first one is functional complementation of the *mef28-1* RNA editing mutants with chimeric PPR type RNA editing factors. The MEF28 is involved in the editing of two neighboring cytidines at the *nad2*-89 and 90 sites. We are interested in how this rarely event is performed by a single PPR protein. The DYW domain of MEF28 is not able to substitute to other DYW domains to edit the downstream cytidine in the two *nad2* sites, suggesting its requirement for the flexible targeting function. In the second approach I used Y2H analyses for mapping the binding sites between MEFs (Mitochondrial RNA Editing Factors) and MORFs. In the most cases the bait wild type and partial MEF proteins could connect to the MORF1, MORF2, MORF8 and MORF9, rarely with MORF3 and very rarely with MORF4, MORF5, MORF6 and MORF7. In contrast, the wild type and the partial MEF8 bait constructs (except the MEF8_ECE+) could bind to almost all MORF preys, indicates that MEF8 is a very unique PPR protein. After Y2H analyses of several different partial E domain constructs, I found that the MEF21 is able to bind with the MORF1 through the N-terminal part of the E domain. Finally, I established the *Pichia pastoris* expression system to express recombinant PPR type RNA editing proteins. The system is now available for further molecular functional analysis of PPR proteins that have been difficult to be expressed in *E.coli* system.

1 Zusammenfassung

Beim RNA-Editing in Pflanzenorganellen sind PPR-Proteine, die eine E- oder E- und DYW-Domäne am C-Terminus besitzen, wichtig für die Erkennung spezifischer *cis*-Elemente. Diese Proteine können mit Hilfe ihrer PPR-Domänen an ihre Ziel-RNA-Sequenz binden. Die Erkennung der spezifischen RNA-Sequenz mittels PPR-Proteinen basiert auf der spezifischen Bindung von einem PPR-Motiv an jeweils ein Nukleotid.

Es gibt mehrere mögliche Funktionen von E-Domänen in PPR-RNA-Editing-Faktoren: Zum einen können sie an RNA binden als auch die PPR-Motive. Zum anderen stellen sie die interagierende Oberfläche für andere Proteine zur Verfügung, die im RNA-Editing-Prozess involviert sind. Es ist auch nicht auszuschließen, dass E-Domänen alle diese Funktionen gleichzeitig erfüllen können. Die DYW-Domäne besitzt ein konserviertes Motiv ($HxEx_nCxxC$) aus Cytidin-Deaminasen und ist fähig Zink zu binden, was vermuten lässt, dass diese Domäne als Deaminase-Enzym fungiert. Diese Cytidindeaminase-Aktivität konnte bisher jedoch noch nicht bewiesen werden.

Die MORFs (Multiple Organellar RNA Editing Factors), die auch als RIPs (RNA editing factor Interacting Proteins) bezeichnet werden, sind auch an mehreren C-zu U- Substitutions-Prozessen beteiligt, die in Mitochondrien und/oder in Chloroplasten von Blütenpflanzen stattfinden. Neun Mitglieder dieser Familie besitzen die konservierte 100 Aminosäuren-lange MORF-Box und ein Protein besitzt nur die C-terminale Hälfte der MORF-Box. Die genaue Funktion dieser Proteine ist bis jetzt noch unklar. Es wird vermutet, dass sie durch eine direkte Bindung der PPR-Proteine an die MORF-Box am RNA-Editingprozess beteiligt sind. Um die genauen Funktionen der PPR-Motive sowie der E- und DYW-Domänen von RNA-Editingfaktoren zu untersuchen, wurden verschiedene Analysen durchgeführt. Zum einen wurden funktionale Komplementation von *mef28-1* RNA-Editing-Mutanten mit rekombinanten PPR RNA-Editing-Faktoren durchgeführt. MEF28 ist am Editing von zwei benachbarten Cytidinen an der *nad2*-89 und -90 Stelle involviert. Wir sind daran interessiert, wie dieser seltene Prozess durch ein einzelnes PPR-Protein katalysiert werden kann. Das Editing an der Position 90 im *nad2*-Transkript kann nicht durch die DYW-Domänen anderer PPR-Proteine katalysiert werden, was auf die Notwendigkeit der nativen MEF28-DYW-Domäne für die flexible Editierung der zwei benachbarten Cytidine hindeutet.

Im zweiten Projekt konnte ich anhand von Y2H Untersuchungen gezeigt werden, wo Bindungsstellen zwischen MEFs (Mitochondrial RNA Editing Factors) und MORFs sind. In den meisten Fällen konnte das Wildtype Köderprotein und partial MEF-Proteine mit MORF1, MORF2, MORF8 und MORF9 binden, aber selten mit MORF3 und sehr selten mit MORF4, MORF5, MORF6 und MORF7. Im Gegensatz dazu, konnte das Wildtype und partial MEF8 Zielkonstrukt (außer das MEF8_ECE+) an fast alle MORF-Proteine binden. Das zeigt, dass MEF8 ein sehr einzigartiges PPR-Protein ist. Nach Y2H Analysen von einigen verschiedenen partialen E-Domän-Konstrukten habe ich herausgefunden, dass MEF21 fähig ist mit dem MORF1 durch die N-terminale E-Domäne zu binden. Schließlich wurde im Zuge dieser Arbeit ein *Pichia pastoris*-basiertes Expressionssystem für rekombinante PPR-Proteine etabliert. Dieses System ermöglicht die biochemische Analyse von PPR-Proteinen, welche bislang nicht mit bakteriellen Expressionssystemen hergestellt werden konnten.

2 Introduction

2.1 RNA editing

2.1.1 Discovery of RNA editing

The central dogma declares that the genetic information is coded into the DNA and this genetic code is then transcribed into mRNA, which is carrying the information for the translational product, the protein. RNA editing is a post-transcriptionally process that can modify this information by altering the primary transcript with two different mechanisms: insertion/deletion or substitution, affecting the mRNA, tRNA or rRNA and by the type of A to I conversion in miRNA and besides that some viral RNA can become also as editing targets (Gray Michael 2012; Brennicke et al. 1999). RNA editing occurs not just in the nucleus, it is also present occurrence in the organelles (Brennicke et al. 1999).

RNA editing was firstly described in 1986 in the mitochondria of kinetoplastids (e.g. in *Trypanosoma brucei* a parasitic protozoan) recognizing extra inserted nucleotides in the transcript of the *coxII* (cytochrome oxidase subunit II) gene (Benne et al. 1986). This type of editing is an insertion/deletion event of the uridines with cleavage-ligation steps. Today also known, that these cells are having not just insertion/deletion editing, but also C-to-U modification in the mitochondria (Simpson et al. 2000; Gray Michael 2009).

In 1984, two types of apolipoprotein B proteins in human were detected using monoclonal antibodies, the APOB-48 (48KDa) and the APOB-100 (100KDa) (Tercé et al. 1985). In 1987, the next discovery was the C-to-U RNA editing in mammals by a catalytic deaminase (APOBEC-1 enzyme) resulting a stop codon by CAA→UAA codon change. This amino acid change is important for producing the shorter version of the apolipoprotein B, the APOB-48 protein, which is synthesized in the small intestine. The apoB gene encodes another protein, the APOB-100 produced in the liver in human and secreted into the blood plasma. Both of these proteins are responsible for lipid binding and transport (Gray Michael 2009; Brennicke et al. 1999). The 5-HT_{2C} pre-mRNA, a serotonin receptor transcript is target for the ADAR2, but also non coding RNAs as the small interfering RNAs (siRNAs) or micro RNAs (miRNAs) could be also substrates for RNA editing in mammal (Brennicke et al. 1999; Gray Michael 2012).

In 1989, three groups namely Mike Gray, Jean-Michel Grienenberger and Axel Brennicke groups reported the plant mitochondrial RNA editing (Brennicke 2009, Covello et al. 1989; Gualberto et al. 1989; Hiesel et al. 1989). Same C-to-U RNA editing was also reported in chloroplast (Hoch et al. 1991). But until today no cytoplasmic RNA editing has been reported in plants (Takenaka et al. 2013b). RNA editing in plants will be further mentioned in details in the next sections.

RNA editing is also appeared in prokaryotes: A to I conversion with a specific deamination enzyme (tadA) occurs in *Escherichia coli* $tRNA^{Arg2}$ (Wolf et al. 2002) and C-to-U editing in Archaea transfer RNA was also reported (Gray Michael 2012; Randau et al. 2009).

Some of the *trans*-acting factors have been already identified. The guide RNAs (gRNAs) are small RNA molecules (55-70 nucleotides) playing role in deletion/insertion process of uridylate residues by their base-pairing antisense sequence (10-15 nucleotides match) in the mitochondrial pre-mRNAs of kinetoplastid protists (e.g. *Trypanosoma brucei*) (Brennicke et al. 1999; Leung and Koslowsky 2001). The APOBEC-1 cytidine deaminase enzyme catalyzes the C-to-U editing in the apoB

mRNA in mammals (Brennicke et al. 1999). The mammalian ADARs are adenosine deaminases processing A-to-I editing on the double-stranded RNAs (dsRNAs) and do not need any cofactors since they have N-terminal target RNA binding domain (dsRBD) and a C-terminal catalytic domain (Blanc and Davidson 2002; Shikanai 2006). On the other hand, the APOBEC1 enzyme requires RNA-binding motifs containing ACFs (apobec-1 complementation factor) as co-factors to bind downstream sequence of the target C.

2.1.2 RNA editing in plants

RNA editing frequency and sites in the chloroplast and mitochondria differ between the plant species. C-to-U editing occurs from the mosses to the seed plants and the reverse U-to-C occurs from the hornworts to the ferns. In the lycopod *Isoetes engelmannii* mitochondria there are approximately 1200 C-to-U substitution events, however in the same organelle of the moss *Physcomitrella patens* 11 C-to-U changes has been reported (Takenaka et al. 2013b). In Angiosperms this number increases to more than 400, for example *Arabidopsis thaliana* has more than 450 mitochondrially and 30 chloroplast C-to-U modification, respectively (Zehrmann A. et al. 2009). In land plants mitochondria and/or chloroplast C-to-U and U-to-C substitution affect mainly the coding region of mRNAs, or modify transfer RNAs (tRNAs), but also introns (in certain cases important for the splicing occurrence) and other non-translated regions (5` and 3`UTRs) are undergoing this modification (Takenaka et al. 2008; Takenaka et al. 2013b). In ribosomal RNAs (rRNAs) it seems to be very rarely event (*Oenothera* 26S rRNA shows C-to-U and U-to-C alterations) (Takenaka et al. 2013b; Schuster et al. 1991). Most of the RNA editing events affect the first or the second nucleotide in the codon, which can result a different amino acid than was encoded by the genome. They are not just altering the amino acid sequence, but also in the case of C-to-U editing creating initiation codon (from ACG→AUG change) or even termination codon (from CAA→ UAA or CGA→ UGA) and U-to-C modification can remove a stop codon (from UAA→ CAA) (Takenaka et al. 2013b). How this process exactly operates and what is the editing enzyme is still unclear. It was suggested already in 2001 that the site-specific *trans*-factors of plant RNA editing are proteins and not RNA (the latter idea came from the Trypanosoma guide RNA editing system) (Hirose and Sugiura 2001).

The C-to-U editing process could be a deamination or transamination of the C-4 cytosine amid, transglycolysation in the ribosyl residue or can occur via deletion/insertion per the deletion of cytidine 5`monophosphate (CMP) and the insertion of uridine 5`monophasphate (UMP). To decide what kind of biochemical reaction is this editing process there are several possibilities including the radiolabeling methods too. The deletion/insertion reaction could be excluded after the detection of radiolabeled UMP when the alpha phosphate group was labeled with phosphorus-32 (^{32}P) radioactive isotope in CMP, because the persistence of the radioactivity through the alpha phosphate group means that the phosphodiester group was not cleaved (Yu and Schuster 1995; Rajasekhar and Mulligan 1993).

After the uses of tritium, - the radioactive hydrogen isotope (^{3}H), to labeling the cytosine base, allowed the detection of labeled UMP after *in vitro* reaction, further excluded the possibility of the reaction as a transglycolysation event, because in a transglycolysation event the glycosil bound would be cleaved and formed again and the cytosine base would be replaced by uracil (Yu and Schuster 1995; Rajasekhar and Mulligan 1993).

Transamination means that an amino group is transferred to a ketoacid to create a new amino acid. α-ketoglutarate and oxaloacetate are important acceptor molecules in this reaction. Adding them in an *in vitro* RNA editing system did not enhance the reaction which strongly suggests that the RNA editing is a deamination process (where the amino group is removed from the molecule

releasing ammonia during this reaction). However, the U-to-C reverse reaction is not easy to be explained with the deaminase reaction since never has been found a classic cytidine deaminase to be able to catalyze it and those acceptors show no influence to the system (Takenaka et al. 2007; Takenaka and Brennicke 2003).

2.1.3 The PPR family proteins

The first discovered protein participating in RNA editing in plant plastids is the CRR4 (CHLORORESPIRATORY REDUCTION 4) in *Arabidopsis thaliana* (Kotera et al. 2005). CRR4 protein is involved in RNA editing of *ndhD* transcripts (*ndhD*-1) which encodes NAD(P)H dehydrogenase (NDH) subunit. These proteins are PPR (pentatricopeptide repeat) proteins with a C-terminal so-called E domain (Okuda et al. 2009; Salone et al. 2007). The first mitochondrially RNA editing protein in plants, MITOCHONDRIAL RNA EDITING FACTOR1 (MEF1) is identified by mapping of the nuclear locus, which is responsible for the different RNA editing rate in reciprocal crosses between ecotypes C24 and Columbia (Col). MEF1 is also a PPR protein with E and DYW domain and involved in the RNA editing in 3 mitochondrial mRNAs (*rps4*-956, *nad7*-963 and *nad2*-1160) (Zehrmann et al. 2009).

PPR family members are containing degenerate 35-amino-acid repeats, the PPR motifs. The sequence similarities between these motifs and the TPR (tetratrico peptide repeat: degenerate 34 amino acid) motifs and structure predictions suggested that the PPRs are forming helix-turn-helix structures and one PPR motif forms two antiparallel α-helices (Small and Peeters 2000; Delannoy et al. 2007). Contrast to the TPR proteins which are solenoid proteins and to be involved in protein-protein interactions (O`Toole et al. 2008; D`Andrea and Regan 2003), PPRs are RNA binding proteins (confirmed by *in vitro* experiments (Ruwe and Schmitz-Linneweber 2011).

The PPR genes are highly expanded in land plants with more than 450 members in *Arabidopsis thaliana* representing approx. the 2% of the protein coding genes in this species (Nakamura et al. 2012) and with 650 PPR proteins in *Oryza sativa* PPR family, against other eukaryotes with <30 members (e.g. humans have only 6 PPR proteins and *Saccharomyces cerevisiae* only 5 PPRs) and just some of these genes are also reported in prokaryotes e.g. in the plant pathogen, *Ralstonia solanacearum* (Manna 2015; Yagi et al. 2013*;* Shikanai 2006; Lurin et al. 2004). The majority of the PPR proteins in the PPR family are predicted to be targeted to either mitochondria or chloroplast or both organelles (Lurin et al. 2004). PPR proteins are involved in many processes in plant organelles also except for RNA editing: e.g. HCF152 from the P subfamily is involved splicing of *petB* RNA and also for the stability of the spliced product after this process, RNA cleavage (like the CRR2 protein with an additional DYW domain is important for *ndhB* transcript cleavage), and the PGR3 protein plays an important role in the translation of *petL* gene (Shikanai 2006; Lurin et al. 2004).

PPR family proteins are divided into two subfamilies, P and PLS subfamilies (Fig. 2.1). The P subfamily proteins have the classical 35 amino acid motifs, the P motifs. The PLS domain are named after their unique PPR structures. They contains in addition to the 35 amino acids long P motifs, PPR-like longer (L) motifs, the L1 (35 amino acids) and the L2 (36 amino acids), and the PPR-like short (S) motifs (31-32 amino acids). The PLS subfamily divided into further 4 subgroups (Fig. 2.1) after following the classifications of Lurin et al. From these subgroups the PLS subgroup is the only one without any C-terminal extension. The other 3 subgroups are named after their extensions, namely: the members of the E subgroup have an E-domain (extension domain), the E+ subgroup proteins have an E+ extension after the E-domain, and the DYW subgroup proteins have the DYW domain (named after the canonical C-terminal Asp(D)-Tyr(Y)-Trp(W) conserved amino acid triplet). A new redefined classification for PLS class PPR proteins was proposed in 2016 (Cheng et al. 2016) after structure

predictions and computational analysis (Fig. 2.1). The E-domain has been divided to two PPR motifs E1 and E2 since their structure shows similarity with the PPR motif and classified into the E1 and E2 subgroups, respectively (Fig. 2.1). The E+ motif containing proteins stayed in the E+ subgroup but the E+ motif is defined as a part of the DYW domain. In this thesis, I use conventional (following Claire et al) classification for the PLS class PPR proteins.

The exact role of E domain is still unknown. The E and the E+ domains are degenerated (Lurin et al. 2004), therefore it is unlikely to be the missing catalytic domain. The similar structure of E domain to PPR motifs implies that this extension motif can be also involved in RNA binding. Alternatively, this domain has binding surfaces for other proteins (Chateigner-Boutin et al. 2013). It has been observed that the E domain is shorter (without E+ motif) by the chloroplast targeted PPR proteins and the mitochondrial targeted PPRs can have shorter E domain, or EE+ domains. It can suggest that they have different functions (Chateigner-Boutin et al. 2013). Therefore the EE+ domains have been swapped between those PPRs that are targeted in different organelles. After this domain swapping the recombinated proteins were not able to carry out their function in RNA editing anymore (Chateigner-Boutin et al. 2013).

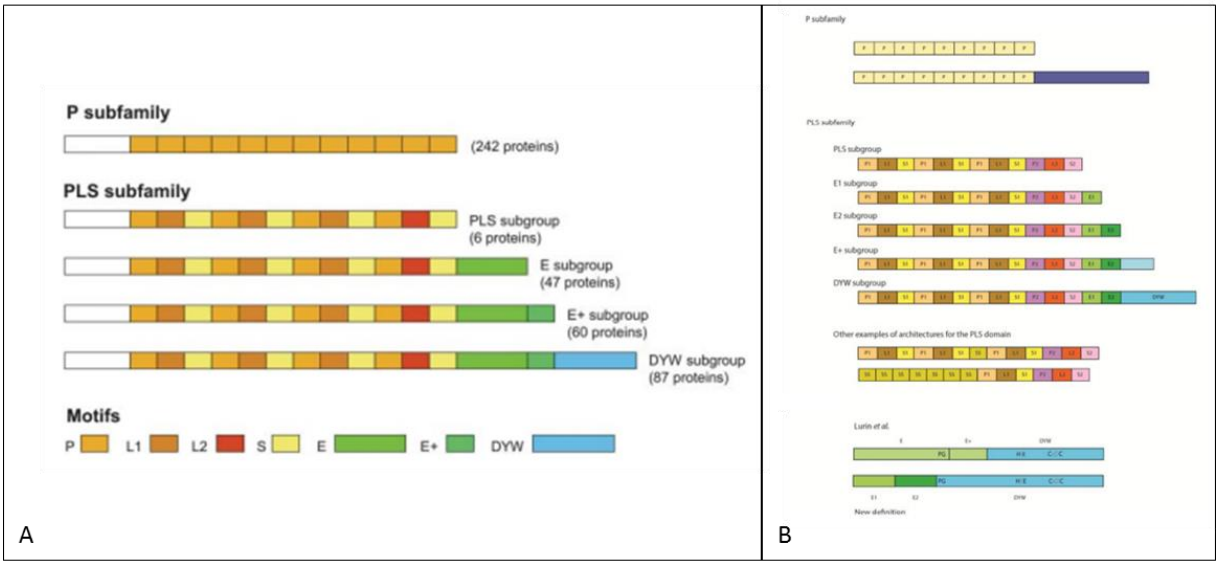


Figure 2.1: Two classifications of PPR proteins. (A) Classification by Lurin et al. (Lurin et al. 2004) and (B) is the new definition by Cheng et al. (Cheng et al. 2016). The classification by Lurin is followed in this thesis.

2.1.4 PPR proteins as site specific trans-factors for RNA editing

The *cis*-element for each RNA editing site in plant organelles has been analyzed by *in vivo* and *in vitro* methods in tobacco and cauliflower. It has been shown, that the substitution of 16 nucleotides upstream but also the 9 nucleotides downstream sequence around the editing site abolished the editing in the chloroplast *psbL* mRNAs. In the *ndhB* chloroplast mRNA just upstream region of editing sites seemed to be important for the editing reaction. If the upstream sequence of the editing site of *psbL* (pL5) or *ndhB* (nB5) has been used as competitors for the *psbL* and *ndhB* mRNAs, respectively, the editing was arrested of the latter two mRNAs, but it was not, if the downstream competitors of the *psbL* (pL3) or *ndhB* (nB3) was used. If the nB5 was used as a competitor for the *psbL* or the pL5 for the *ndhB* mRNAs, it did not have influence on the editing of the *psbL* or on the *ndhB* mRNAs. These

results are suggesting the importance of the upstream sequences of the investigated mRNAs (Hirose and Sugiura 2001).

Many PPR proteins have been isolated as specific factors for defined RNA editing site(s) implies that PPR proteins bind and recognize cis-elements of each editing site (Okuda et al. 2006; Kobayashi et al. 2011; Takenaka et al. 2013a). The molecular mechanism of RNA recognition by PPR domain has been unclear for long time. The similar structures between the TPR proteins and the PPRs, both of which contain two antiparallel α helices pro motif, suggests a similar motif character. It was proposed that these hairpins are able to create a superhelical structure using the PPR5 protein (consists of 10 PPR motifs) about which has been also reported that can bind to ssRNA with higher affinity, like the CRP1 protein which is also a chloroplast targeted protein with 14 PPR motifs (Williams-Carrier et al. 2008).

By the biochemical analysis and structural modeling of HCF152 PPR motifs surface, the RNA binding of the HCF152 was suggested as a nucleobase-motif specificity manner. It was also suggested that 5 amino acids can be possible binding surface to the target RNA sequence: the 1st, 4th, 8th 12th and ii where the latest one is two amino acids before the next helix of the next motif and the others are located next to each other on the surface of each PPR motif (Kobayashi et al. 2011).

Furthermore, *in silico* comparison of the amino acid sequence in each PPR motif in RNA editing factors and several P-class PPR proteins and their putative targeting sequence revealed defined the two amino acids corresponding to the nucleotide specificity in P and S motifs. These two amino acids are located at the position 6 on helix A (the first helix in a PPR motif) and at the position 1` of the next motif in helix A` (Barkan et al. 2012). The PPR code with amino acids at position 6 and 1' suggests a PPR motif in each PPR type RNA editing factor binds to a nucleotide. The motif-nucleotide correspondences begin with the S2 motif assigned to the nucleotide at position -4 of the editing site, can properly position the PPR proteins to the target RNA sequence (Fig. 2.2).

L2 and S2 motifs, which locate at the N terminal of E domain of the editing factor, have similar sequence to the L and S motifs, respectively, but they differ slightly in their sequence. The L, L2 and S2 motifs for RNA editing, of which for example the L motif has been suggested as a spacer for relaxing the structure (Barkan et al. 2012). The correlation between amino acid and nucleotide preferences between the L, L2, S2 motifs and the target sequence suggests that these motifs also contribute to the specific binding to the RNA, although their connections seem to be weaker than P and S motifs (Takenaka et al. 2013a).

Almost all PPR proteins required for RNA editing belong to the PLS subfamily in the C-terminally extended subgroups. In longer PPR proteins which have more than 10 repeats e.g. the MEF12 has 25 PPR elements (Härtel et al. 2013c), probably not all of the motifs have to bind to the RNA sequence for RNA editing reaction, but rather serve as a spacer and the position of the spacer may be changeable in the case of PPR with more target sites (Takenaka et al. 2013a). How the short PPR domains in some RNA editing factors (e.g. MEF8, MEF8s which has only 5 PPR motifs) are involved in the editing at specific target site remained unanswered, since the number of PPR motifs seem not to be sufficient for sequence specific RNA binding.

Structural analysis of PLS class RNA editing factors with target RNA is indispensable to understand precise molecular mechanism of their specific RNA binding. However, so far there has been no report on it due to the difficulty of protein expression in *E.coli*. In the chapter III of this thesis I described a novel approach with *Pichia pastoris* to express PPR type RNA editing factors.

2.1.5 The enzyme catalyzing the RNA editing events in plants

So far the deaminase catalyzing enzyme has not yet identified. A most frequently proposed hypothesis is that the editing enzyme is the C-terminal DYW domain of the PPR type editing factors, which contains the active site of cytidine deaminases (except the proline before the first cysteine) namely the residues $HxEx_nCxxC$ (Salone et al. 2007).

The crystal structure of Escherichia coli cytidine deaminase (CDA) revealed that this enzyme is a zinc binding enzyme coordinated by one histidine and two cysteine residues and the glutamate is involved in protonation of the nitrogen group and removing amino group. APOBEC-1 enzyme is an important cytidine deaminase for producing APOB48 protein in mammals and it is homologous to other deaminases at the zinc binding domain (MacGinnitie et al. 1995). The zinc dependency of the plant RNA editing reaction was tested in pea mitochondrial extract using the 1.10-*O*-phenanthroline, a Zn^{2+} ion chelator. The editing reaction was insensitive to the chelator suggesting that in this process other type of enzymes is involved rather than the classical cytidine deaminases (Takenaka and Brennicke 2003).

The zinc ion affinity of DYW1 at the residues $HxEx_nCxxC$ was indicated by several point mutants and deletion mutants (Boussardon et al. 2014) with the Inductively Coupled Plasma Mass Spectrometry (ICP-MS). Addition to the zinc binding domain, the DYW domains have a highly conserved short sequence at the C-termini, namely the canonical DYW triplet, which has been found by some proteins as EYW or DFW triplet. This amino acid triplet has been suggested playing a role as a docking signal as similarly observed characteristics in some other proteins (Boussardon et al. 2014).

The importance of the DYW domain differs in the PPR proteins. The DYW domain of MEF11, CRR22 and CRR28 does not essential for the function, against to MEF1 and CRR2 proteins, where the MEF11, CRR22 and CRR28 are involved in RNA editing and the CRR2 has endonucleolytic activity (Okuda et al. 2009; Zehrmann et al. 2011). This domain is not interchangeable between certain PPRs, it has been proven e.g. between CRR22-CRR2 and CRR28-CRR2 proteins, while it is in others e.g. between CRR22-CRR28 (Okuda et al. 2009).

2.1.6 The MORF proteins

Several proteins that belong to the same gene family were isolated as necessary for efficient RNA editing in both plant organelles. The multiple organellar RNA editing factor 1 (MORF1) (Takenaka et al. 2012) and RNA editing factor interacting protein 1 (RIP1) (Bentolila et al. 2012) were identified by a forward genetic screen of an ethyl methanesulfonate (EMS) mutated *Arabidopsis thaliana* line (Takenaka et al. 2012) or by tandem mass spectrometry (MS/MS) analysis (Bentolila et al. 2012) of CoIP purified PPR type RNA editing factor binding proteins, respectively. Discovery of the MORF/RIP proteins suggests that the C-to-U deamination process in plant organelles require more complex protein network than previously thought.

The MORF family consists of 9 genes encoding the MORF1-MORF9 with a 100 amino acids conserved region (MORF box) in the central part and of one potentially pseudogene (At1g53260) which encodes just the C-terminal half of the MORF box.

MORF proteins are specific to flowering plants (Takenaka et al. 2012). The MORF1 and the MORF4 proteins show similarities (out of the MORF box) to the Pat protein domain (Takenaka 2014a) what are members from the PAT (perlipin, adipophilin, and TIP47) family and playing a role in lipid metabolisms. In vivo fluorescence tagged protein analyses identified the MORF2 and MORF9 (the one pseudogene encoded protein just predicted by Predotar) are located in chloroplast, MORF5 and MORF8 are targeted to both organelles and all of the rest are located in the mitochondria (Zehrmann et al. 2015; Takenaka et al. 2012).

MORFs/RIPs are necessary for effective RNA editing in mitochondria at many sites: e.g. *morf3* mutant plants show effect for more than 40 mitochondrial RNA editing sites, and the RIP1/MORF8 controls 75 % of the mitochondrial sites, and in chloroplast at almost all sites e.g. mutations in MORF2 and MORF9 proteins cause loss of editing or lower editing in the investigated mutants (Bentolila et al. 2013; Takenaka et al. 2012).

The MORF proteins are required for some editing sites, which are known as targets for specific PPR type RNA editing factors, like MORF1 and MEF19 required at site *ccmB*-566, MORF1 and MEF21 at site *cox3*-257, MORF3 and MEF13 at identical 7 sites. Furthermore some sites require more than one MORF for effective RNA editing, e.g. MORF2 and MORF9 at sites *ndhB*-1255, *petL*-5, *psbZ*-50 and *ndhD*-2 the last one was also affected in the mutant of a PPR type editing factor *crr4* (Shi et al. 2016 a; Bentolila et al. 2013; Takenaka et al. 2012). These common shared sites between the specific MORF-PPR proteins and the MORF-MORF proteins suggest a direct connection between these editing factors (Fig. 2.2). In Y2H assays the chloroplast located MORF2 and MORF9 form heterodimers and almost all of the MORFs are able to form homodimer and several MORF and PPR type RNA editing factors can also interact with each other. Pull-down analysis also showed that MORF1 form MORF1-MORF1 homodimer, the dual targeted MORF8 form heterodimer with the chloroplast targeted MORF2 or MORF9 proteins, and MORF1-MEF19 and MORF1-MEF21, which share the target sites are also able to bind to each other (Zehrmann et al. 2015; Takenaka et al. 2012). These results suggest that more than two MORF proteins are responsible for effective RNA editing in some RNA editing sites (Zehrmann et al. 2015; Bentolila et al. 2013).

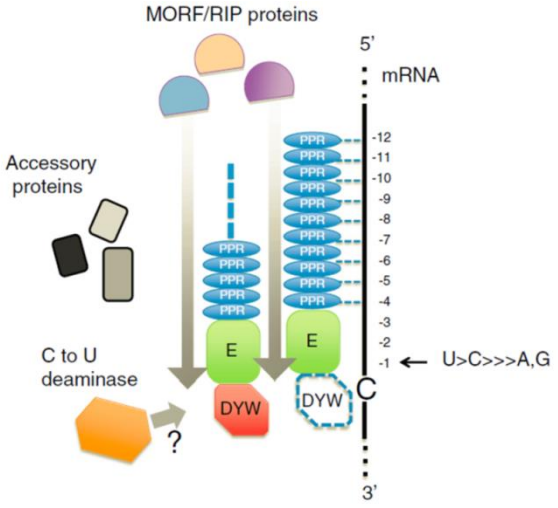


Figure 2.2: Model for the flowering plants editosome complex (Takenaka et. al. 2014b).

Although there are several hypothesis for the roles of MORFs in RNA editing complexes, (e.g. assisting by the PPR-RNA connection, or bringing the zinc ions to the active site of the enzyme), how they work in the complexes is still unclear (Takenaka 2014a). In this thesis, I analyzed the possible binding sites between MEF and MORF RNA editing factors using the Y2H method.

2.1.7 Further possible members of the editosome complex

Chloroplast ribonucleoprotein family (cpRNPs) members are stromal proteins (Hirose and Sugiura 2001) and involved in the stabilization of plastid transcripts and RNA editing (Tillich et al. 2009). The cpRNPs bind to chloroplast mRNAs *in vitro*. One of the members of the CpRNPs, CP31 protein has two

RNA Recognition Motifs (RRM) and an N-terminal acidic domain (Tillich et al. 2009; Hirose and Sugiura 2001) and is involved not only in RNA editing but also other post-transcriptional processes. The two paralogous of CP31 are encoded in *Arabidopsis thaliana,* the CP31A and the CP31B, which are involved in RNA editing in plastids (Tillich et al. 2009). However, the Cp31 proteins are not a part of editosome. They bind to mRNA and facilitate the accessibility of PPR editing factors to the editing site (Hirose and Sugiura 2001; Tillich et al. 2009).

Other RRM domain containing proteins, Organelle RNA Recognition Motif-containing (ORRM) proteins are also listed as members for the editosome. The ORRM1 is plastid targeted protein containing a pair of truncated RIP domains at the N-terminal and one RRM domain at the C-terminal, affecting 12 sites in *Arabidopsis thaliana* (Sun et al. 2013). The ORRM1 has been identified as a RIP1 homologue with blastp search and the duplicated N-terminal part of the MORF box has been identified with the MEME software (Sun et al. 2013). The ORRM2 and ORRM3 (Shi et al. 2015) are involved in mitochondrial editing, like the ORRM4 protein (Shi et al. 2016 b). Mutants of ORRM2 and ORRM3 affect ca. 6% of the mitochondrial editing sites in *Arabidopsis thaliana,* respectively. In Y2H experiments the RIP1 (MORF8) can interact with ORRM3 (Shi et al. 2015) and ORRM1. Complementation analysis suggested that the N-terminal part of the MORF box which is duplicated in the ORRM1 protein is not required for the editing (Sun et al. 2013).

3 Results

3.1 Editing factor specific function by targeting adjacent cytidines

MEF28 is a mitochondrial RNA editing factor involved in the editing events of two adjacent cytidines in the *nad2* transcript at positions 89 and 90. This type of editing with the substitutions of two neighboring Cs are very rarely events. In addition to the MEF28 target sites the SLO2 is another mitochondrial located PPR protein which is described to edit also two adjacent cytidines, namely at the nucleotides 144 and 145 in the *mttB* transcript in *Arabidopsis thaliana* (Zhu et al. 2012). Another example is when from the two target cytidines one is going to be edited by a PPR protein and the other one by another factor (Arenas-M et al. 2013). How this process occurs in MEF28 target sites is unknown. There are three possible explanations. The first is that the PPR protein is able to shift along the sequence to edit both adjacent cytidines, but the computational prediction of the PPR binding site shows high scores only with the cis-element of the first cytidine but lower scores with downstream site, therefore it would be unlikely (Takenaka M.: 2016 personal notification). The second explanation is that the E and DYW domains can directly access the two cytidines, without shifting. **The third possibility is that the MEF28 interact with other cofactor(s) and the editing is done in a complex**. To investigate how the MEF28 is involved in the editing at two adjacent sites, I generated several domain swapped constructs with other RNA editing factors.

3.1.1 The MEF28 protein

The editing efficiency in Columbia (Col) WT plants reaches to~94% at the *nad2*-89 site and to ~88% at the *nad2*-90 site. In the *mef28-1* T-DNA mutant line, a T-DNA is inserted in the DYW domain and this insertion disrupts the function in the At5g06540 gene but still ~7% of the editing effect remained in these mutants at *nad2*-89 site. Overexpression of the At5g06540 gene in the *mef28-1* restores the editing level to ~100% confirmed the function of the MEF28 protein at these two sites of the *nad2* transcript (Fig. 3.1) (Härtel 2013b).

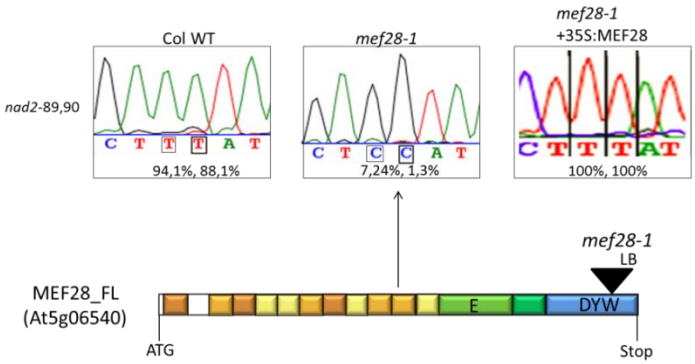


Figure 3.1: The At5g06540 gene encoded MEF28_FL protein respond for the *nad2*-89 and *nad2*-90 sites in the Columbia (wt) plants. The complementation analysis for the *mef28-1* T-DNA mutants has been done under the control of the CaMV-35S promoter via *Agrobacterium* transformation by Härtel Barbara.

The DYW domain contains the active site ($HxEx_nCxxC$) of cytidine deaminases and it supports the idea that this domain is responsible for the catalytic activity (Salone et al. 2007). Deletion of DYW domains gives different effect on each PPR type RNA editing factors. Some proteins lose their editing ability (Okuda et al. 2009; Zehrmann et al. 2011), but the others are still functional after eliminating the domain (Okuda et al. 2009; Zehrmann et al. 2011). The MEF28 protein without this domain (Härtel 2013b), the MEF28ΔDYW was not able to restore the editing in the *mef28-1* mutants after protoplast complementation analysis (Fig. 3.2) (Härtel 2013b).

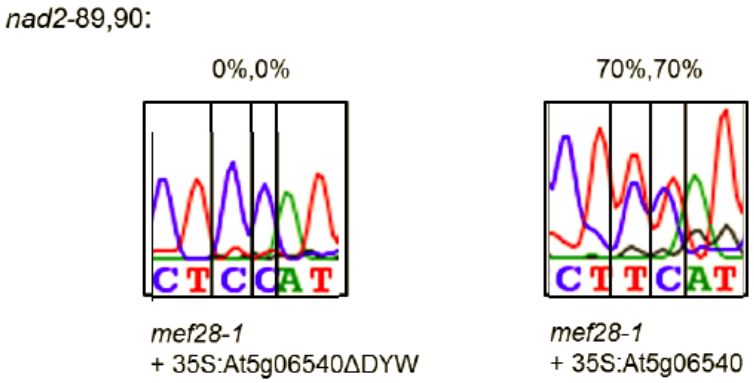


Figure 3.2: The *mef28-1* T-DNA could not be complemented in protoplasts using the At5g06540ΔDYW sequence. As a control the full length of the At5g06540 gene has been used which was able to sets back the editing ratio to 70% (figures from: Härtel 2013b).

This result suggested that the C-terminal domain of MEF28 is specifically important for the ability to edit two adjacent cytidines. Therefore, I analyzed the function of MEF28 C-terminal domains by domain substitutions to other C-terminal domains in various editing factors. MEF11, CRR22, MEF25, MEF13 and MEF19, E and/or DYW domain were chosen as the swapping partner for the MEF28, since these five proteins have at least one cytidine (two at the *cox3*-422 and *ndhD*-887 sites, and three at the *nad2*-59 site) around the target editing site and just one of these Cs is edited (Fig. 3.3). The MEF19 is a member of the E subgroup, the MEF25 and MEF13 are mitochondrially located factors from the E+ subgroup, the MEF11 is also a mitochondrial RNA editing factor but belongs to

the DYW subgroup as well as the MEF28, and the CRR22 is a chloroplast RNA editing factor also with a DYW domain at the C-termini.

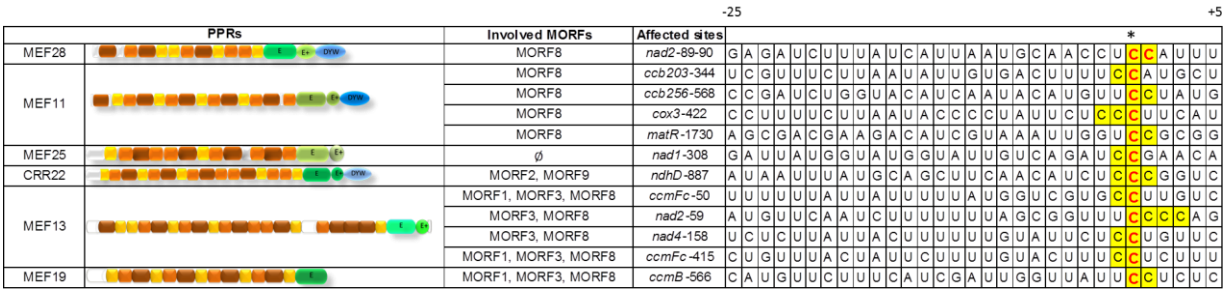


Figure 3.3: PPR proteins used for designing the MEF28 chimera constructs with their domain structures (classificated after Lurin et al. 2004). The Cs colored with red represent the targets of each PPRs. The Cs in yellow boxes are non-edited cytidines adjacent to the target editing sites. The table includes the site-sharing MORF proteins (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). There is no common nucleotides between the aligned mRNA sequences, except for the target editing site (*).

3.1.2 Complementation of *mef28-1* with MEF28ΔDYW

To confirm the previous protoplast complement assays for MEF28ΔDYW protein, we analyzed stable transformants with MEF28ΔDYW (Fig. 3.4).

The results show lower editing efficiency at both the *nad2*-89 and *nad2*-90 sites, confirming that the DYW domain is indispensable for the editing function. The second site is almost 0% edited, while at the *nad2*-89 site the editing is a bit higher than in the *mef28-1* mutant (which shows about 7% editing) in all 6 complemented lines, suggesting that the DYW domain of the MEF28 protein is needed for editing at both sites and specially indispensable for the *nad2*-90 site.

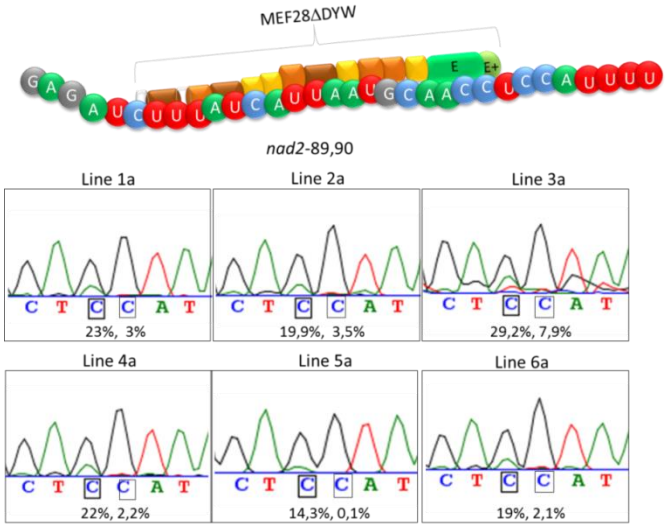


Figure 3.4: MEF28 sequence until the E+ domain MEF28ΔDYW is overexpressed under the control of the CaMV-35S promoter. The results show the importance of the MEF28_DYW domain for the successfully RNA editing at both investigated sites.

3.1.3 Exchanging the DYW domain of MEF28

The DYW domain of MEF11 is not necessary for restoration of the investigated editing sites, which are affected in the *mef11-1* mutant plants (Zehrmann et al. 2011). On the other hand, the DYW domain of MEF1 is required for the editing at sites *nad2*-1160, *nad7*-963 *and rps4*-956 (Zehrmann et al. 2011). The result of DYW deleted MEF28 suggest that the unique ability to edit two adjacent sites derives from the DYW domain. To analyze the possibility, we tried to complement the *mef28-1* mutants with the MEF28 chimera clones with swapping the DYW domains from the MEF11 (Fig. 3.5.A) and from the CHLORORESPIRATORY REDUCTION22 (CRR22) (Fig. 3.5.B), which do not edit adjacent cytidines of their target editing site (Fig. 3.3).

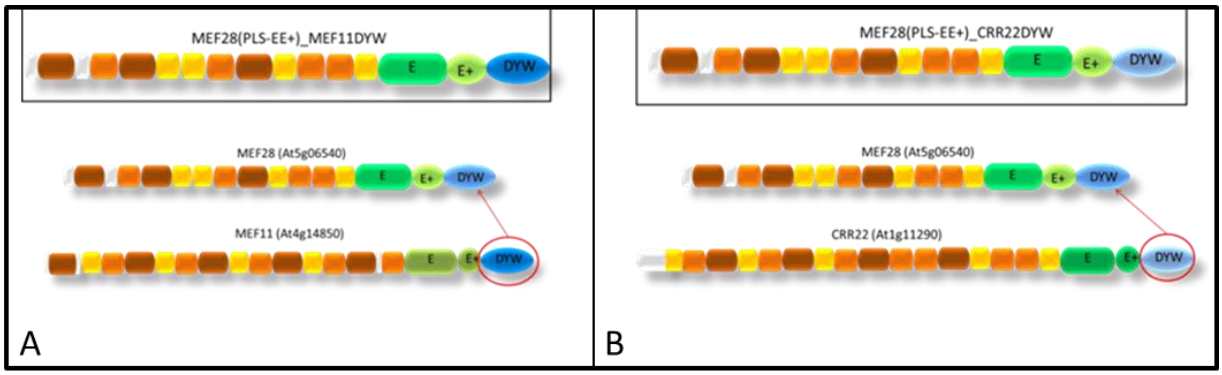


Figure 3.5: Investigation of the role of DYW motifs, using MEF28 PPRs together with the DYW domains of MEF11 (A) and CRR22 (B) editing factors. These chimera genes under the control of the CaMV-35S promoter are transformed into the *mef28-1 Arabidopsis thaliana* T-DNA mutant line.

The chimera construct MEF28(PLS-EE+)_MEF11DYW restore the RNA editing at the *nad2*-89 site while did not at the *nad2*-90 site (Fig. 3.6). The results with the MEF28(PLS-EE+)_CRR22DYW are very similar to the MEF28(PLS-EE+)_MEF11DYW construct (Fig. 3.6). The first site is complemented until maximally to ~65% and the downstream one is not.

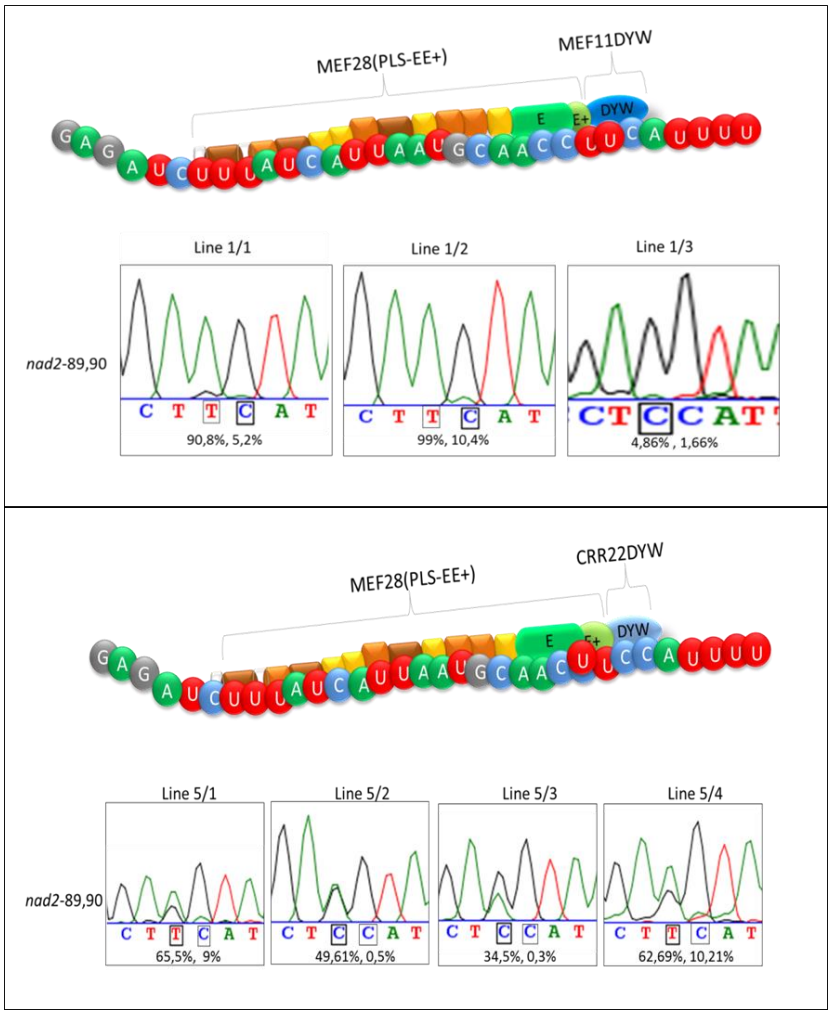


Figure 3.6: Transformants with the MEF28(PLS-EE+)_MEF11DYW and MEF28(PLS-EE+)_CRR22DYW constructs. Top panel: Transformants with the MEF28(PLS-EE+)_MEF11DYW construct under the control of the CaMV-35S promoter. In the lines 1 and 2 the C at the *nad2*-89 is edited against the site *nad2*-90. Bottom panel: Four different lines transformed with MEF28(PLS-EE+)_CRR22DYW construct. The first site is edited in the line 1 and 4 up to ~60%, in other two lines ~49% and ~34%, respectively. The second site is not edited in the 2 and 3 lines.

3.1.4 Complementation assays with the MEF28PLS motifs fused to different EE+DYW domains

The loss of editing ability at the downstream sites in the DYW substituted lines can be due to the incompatibility with the MEF28 EE+ domain. We used the DYW domain from the MEF11 and from the CRR22 proteins with the bona fide EE+ domains of them, MEF28PLS-MEF11EE+DYW (Fig. 3.7.A) and the MEF28PLS-CRR22EE+DYW chimeric constructs (Fig. 3.7.B).

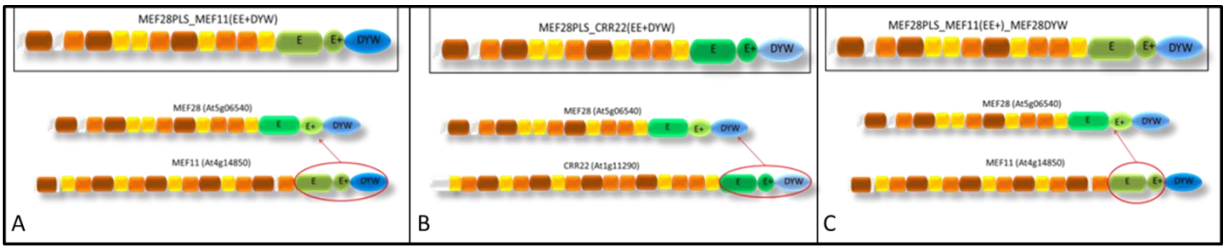


Figure 3.7: Chimera genes by the fusion of MEF28PPRs and EE+DYW sequence of MEF11 (A) or CRR22 proteins (B) and a chimera construct (C) using the MEF11_EE+ domain and the MEF28 PLS and DYW domains.

Substitution of MEF28EE+DYW domain to MEF11EE+DYW domain resulted in loss of editing at both sites (Fig. 3.8), suggesting the EE+ domain in the MEF11 is not compatible to edit both MEF28 target sites.

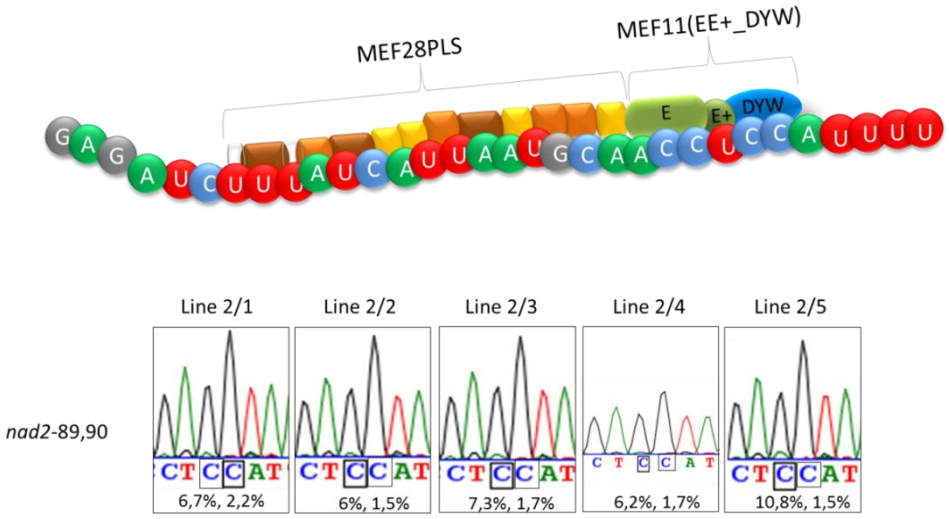


Figure 3.8: Transformants with the MEF28PLS_MEF11_EE+DYW construct. The MEF11_EE+DYW domains cannot restore the editing at the *nad2*-89 and the *nad2*-90 sites in the respective *mef28-1* mutants.

The transformation of MEF28PLS_CRR22EE+DYW into *mef28-1* restored editing at both *nad2* sites in 3 lines (Fig. 3.9).

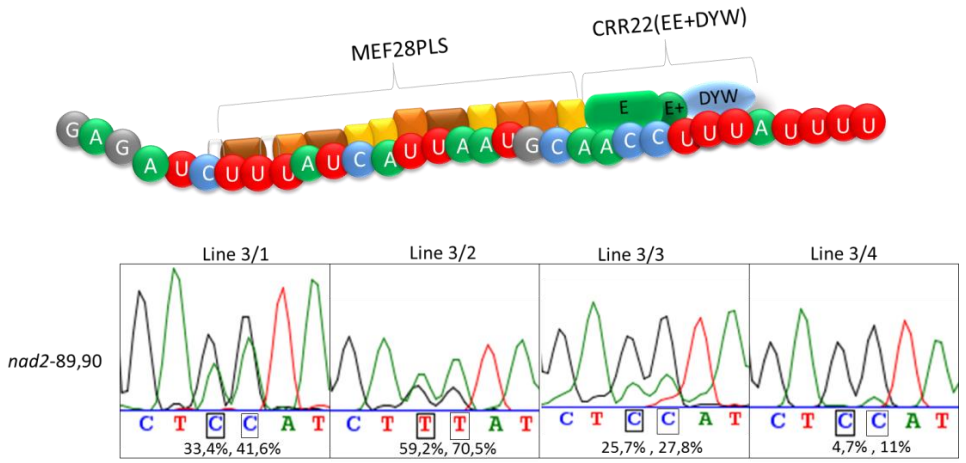


Figure 3.9: Complementation analysis using the CRR22_EE+DYW domain added to the MEF28_PPR repeats. Among the investigated 4 lines, three of them complementRNA editing at both sites.

This result shows that the EE+DYW domain swapping is possible between the two DYW subclass protein, the mitochondrially located MEF28 protein and the CRR22 chloroplast targeted PPR protein.

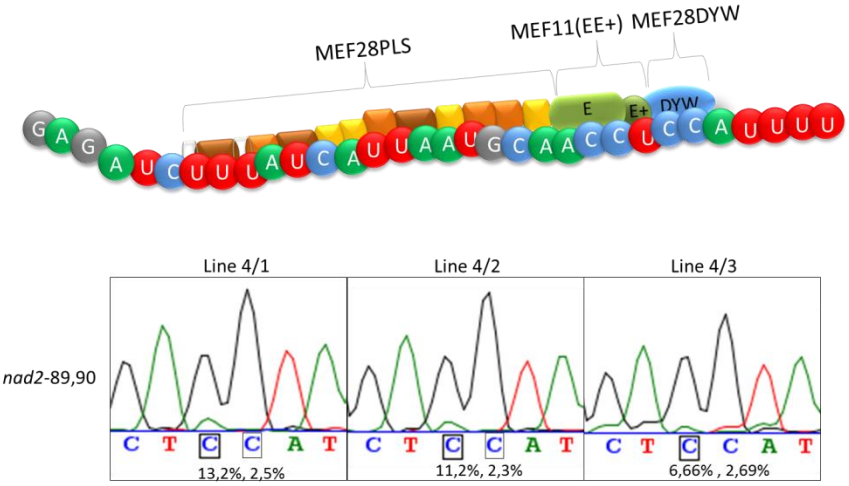


Figure 3.10: Transformants with the MEF28PLS_MEF11(EE+)_MEF28DYW construct. The fusion protein could not restore the editing at the two *nad2* sites in the *mef28-1* T-DNA mutants.

On the other hand, transformed *mef28-1* mutant plants with the MEF28PLS_MEF11(EE+)_MEF28DYW chimera, in which only the EE+ domain is substituted by the MEF11_EE+ motifs (Fig. 3.7.C), cannot complement the function of MEF28 (Fig. 3.10), suggesting **the relevance of the adequate combination of the EE+ domain and PPR domain, since swapping of DYW domains between MEF28 and MEF11 still can edit at least one site.**

3.1.5 Complementation analyses with MEF28 substituted the E or EE+ domains to other E+ and E subclass PPR proteins

The MEF28, MEF11 and the CRR22 are members of the DYW subclass PPR proteins. As same as these DYW subclass PPR proteins, some E subclass PPR proteins can also edit a specific cytidine in adjacent cytidines. I analyzed the effect of substitions of MEF28 EE+DYW domain to the EE+ motifs of the MEF25, MEF13, and the E motif of MEF19 (Fig. 3.11.A, B, C), all of which are mitochondrial editing factor and have unedited cytidines adjacent to the C to be edited.

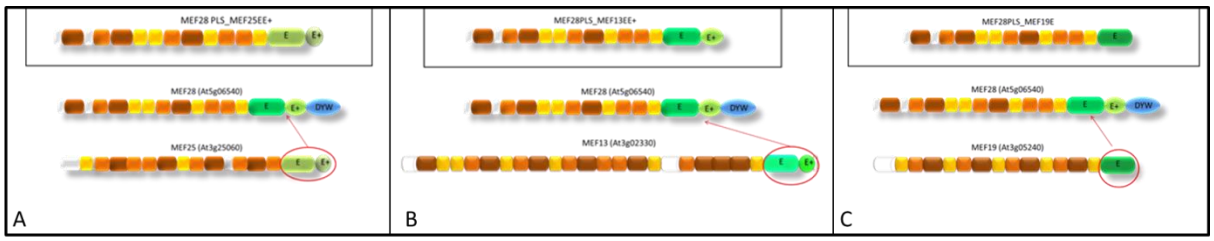


Figure 3.11: Fused proteins keeping the MEF28_PLS domain but the EE+ domains of the MEF25 (A) ,MEF13 (B) proteins and the E domain of the MEF19 factor (C) are added, respectively.

The *mef25* mutant show no editing at one of the two adjacent cytidines, *nad1*-308 but the neighboring *nad1*-307 site is still edited in the mutant, suggesting that another factor is involved in the *nad1*-307 (Arenas-M. et al. 2013) editing event.

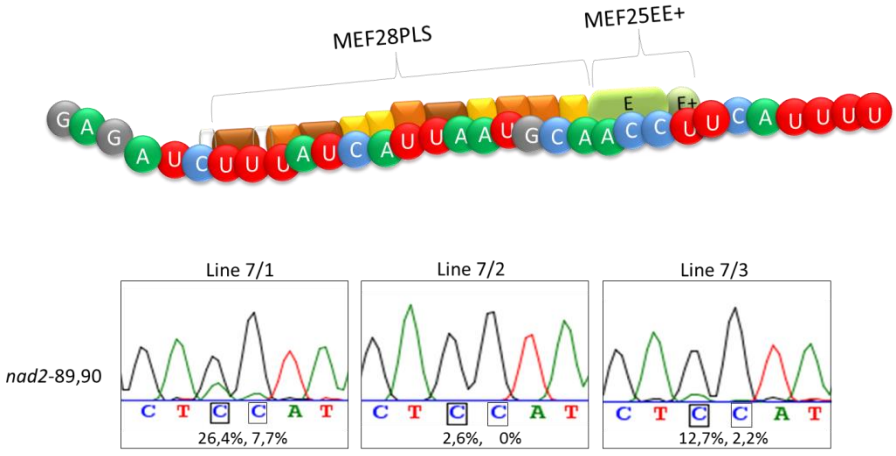


Figure 3.12: Transformation with the MEF28PLS_MEF25EE+ construct. The fused protein for the complementation assay is illustrated above. The three selected lines (1, 2, 3) are shown below. One of the three lines show slightly increased RNA editing at the *nad2*-89 site.

One of the transfomed lines with the MEF28PLS_MEF25EE+ chimera (Fig. 3.12), line 1 shows the editing efficiency ~26% at the *nad2*-89 site and the *nad2*-90 site was not affected in the line 2 and 3 and it was ~7% in the first line. The results are very similar to the instance of the MEF28ΔDYW (Fig. 3.4).

The MEF13 protein is a member of the E+ subgroup and has eight target sites. Four of the eight target sites has unedited cytidine in the neighbour (Fig. 3.3).

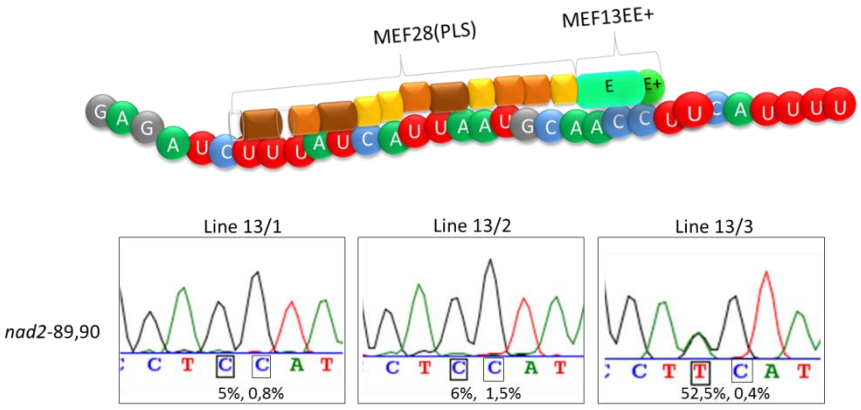


Figure 3.13: Transformants with the MEF28PLS_MEF13EE+ construct. Above is shown the used recombinated protein. Only the first site is edited in one of three transformed lines.

The MEF28(PLS)_MEF13EE+ chimera edits up to 52 % at the *nad2*-89 site but not the *nad2*-90 site in the line 3 (Fig. 3.13). None of the two sites are edited in the other two lines.
The MEF19 is also a member of the E subclass PPR proteins but contains only an E domain without E+ domain. The *mef19* mutant lines with a T-DNA insertion in the E domain have loss of editing at the *ccmB*-566 site (Takenaka et al. 2010) which has an unedited cytidine at one nucleotide downstream. The MEF19_E domain fused to the MEF28_PPR motifs was transformed to the *mef28-1* mutants. One of the transformed lines shows 98% and 99% editing at position 89 and 90 and none of them was affected in the second selected line (Fig. 3.14).

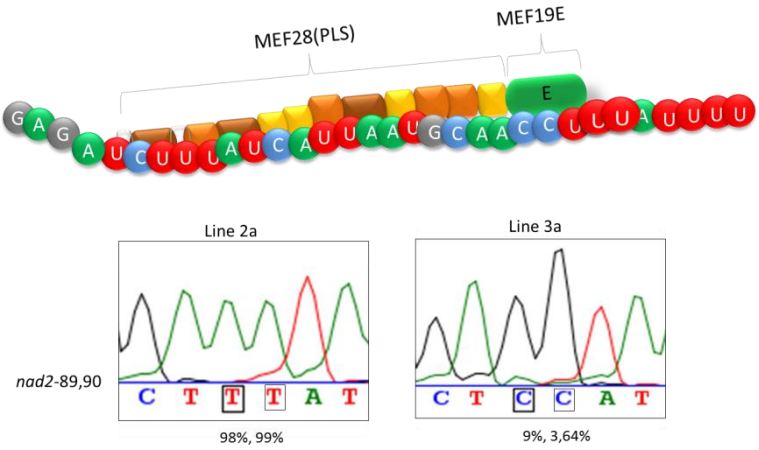


Figure 3.14: Transformants with the MEF28PLS_MEF19E construct. Above is the chimera clone. This fused protein complements editing of *mef28-1* mutant at the *nad2*-89, 90 sites in one transformed line (the chimera has been made by Romer D. BSc).

3.1.6 RNA editing at positions nad2-89 and 90 in different plant species

Comparison of the *nad2* transcript sequences in different species from the angiosperms indicates there is no RNA editing at the positions 89 and 90 in *Beta vulgaris*, *Oryza sativa* and *Triticum aestivum* (Fig. 3.15.A). The position 89 is T in these mitochondrial genomes suggesting that RNA editing at this position is not necessary in these species. Indeed there are no MEF28 homologues in these three plant species. There is one editing target cytidine at position corresponding to *nad2*-89 in *Oenothera berteriana* (Fig. 3.15.A). Two adjacent targeted Cs are present in *Brassica napus* similarly to *Arabidopsis thaliana* (Fig. 3.15.A). I searched for MEF28 homologues in *Oenothera berteriana* and in *Brassica napus*. In the previous I found no homologues due to the lack of available genome data and in the latter I found a homologue protein (Fig. 3.15.A, B).

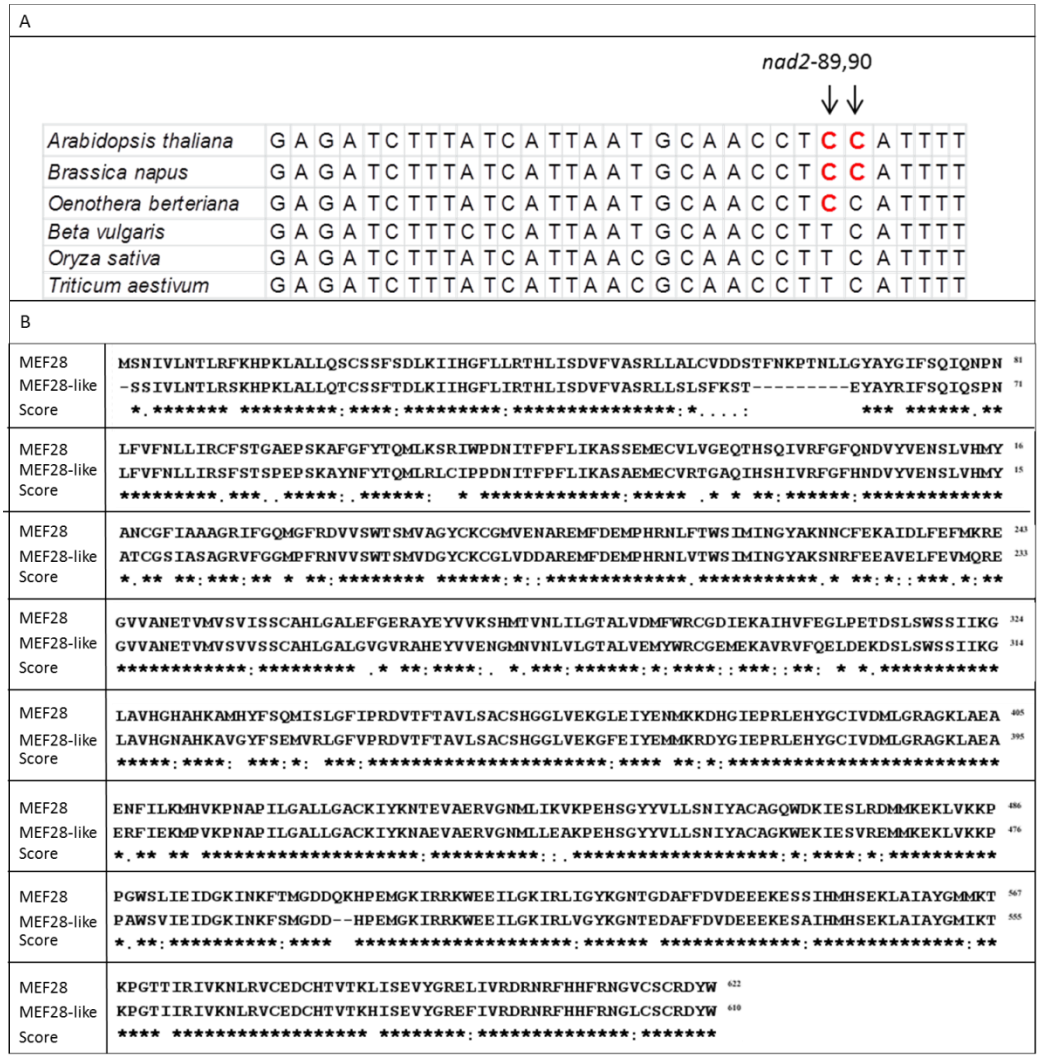


Figure 3.15: Comparison of mitochondrial nad2 genes and MEF28 homologues in several angiosperms. (A) An alignment of *nad2* gene sequences in different plant species. Red colored C shows the targets for RNA editing. (B) Comparison of MEF28 in Arabidopsis and its homologue in Brassica. The matches (*) and similarities are indicated under the alignment (made with DNADynamo program).

3.2 MORF protein interacting surfaces in the PPR type RNA editing factors

3.2.1 Background for the Y2H experiments

Shared target site(s) between PPR type RNA editing factors and MORF RNA editing factors (Tab. 3.1) suggests that the proteins in these two gene families are specifically connected to each other for their tasks in the RNA editing process. Previous pull-down assays and Y2H results with full-length PPR and MORF proteins support this model (Takenaka et al. 2012). My second project was mapping the possible binding site(s) between PPR type RNA editing factors and MORF proteins to reveal more detailed structure of the protein complexes. I investigated these interaction surfaces using the Y2H system with different deletion constructs.

PPR type RNA editing factors consist of three domains. The PPR domain of RNA editing factors binds to the 5' cis-element sequence of the target C(s) (Takenaka et al. 2013a) and the DYW domain is suggested to be involved in the enzymatic reaction since it contains the similar sequence of the active site in the cytidine deaminases (Salone et al. 2007) although its enzymatic function is not documented yet (Takenaka et al. 2014b). The E domain consists of helical structures which is similar to the PPR motifs (Okuda et al. 2006) implies the possiblity of the binding also to the RNA (Takenaka et al. 2014b). However, there is no evidence of the RNA binding. The E-domain can be only a spacer between the target RNA sequence and the target C or a protein binding site (Takenaka et al. 2014b).

To investigate the interaction between parts of PPR proteins and MORFs I cloned short fragments of PPR domain and C-terminal domains (E, E+ and DYW) of PPR editing factors into Y2H vectors. The main focus of the assay is to examine the interaction between those proteins which share the target site(s), but I also included other MORF proteins in my assays to evaluate the correlation between the *in vivo* functional connection and the binding specificity. The Y2H assay results can give us more insigth into the interaction since this *in vivo* method allows the proteins more likely to be in their native forms (Matchmaker™ GAL4 Two-Hybrid System 3 & Libraries User Manual). It is one of the very valid methods to reveal the interaction between two components, although the results have to be confirmed with other methods biochemically e.g. immunoprecipitation, pulldown, photoreactive crosslinking, BiFC etc. or/and even biophysically e.g. fluorescence resonance energy transfer (FRET).

Table 3.1: The tables (A,B) show in the Y2H assay investigated PPR proteins and with which proteins they significantly share targets. Figure A shows the significant and also the non significant effects for editing extents in *morf8* and *morf3-1* mutants (Shi et al. 2016 a; Bentolila et al. 2013). Figure B is summarizing all data for the proteins that share targets involving also MORF1 (Shi et al. 2016 a, 2015; Bentolila et al. 2013; Takenaka et al. 2012; Bentolila et al. 2012). Markered with orange color those MEFs which share their target with at least one of the MORFs (from MORF1-to MORF9).

PPR proteins	Gene	Affected sites	Editing extent			
			rip1(morf8)	**rip1 (MORF8) wt**	***rip3-1 (morf3-1)***	**rip3-1 (morf3-1)wt**
MEF1	*nad7*	963	0,19 (S)	0,45 (S)	1,00 (NS)	1,00 (NS)
	nad2	1160	0,15 (S)	0,91 (S)	1,00 (NS)	1,00 (NS)
	rps4	956	0,01 (S)	0,31 (S)	0,85 (NS)	0,83 (NS)
MEF8	*nad6*	95	0,78 (S)	0,96 (S)	0,99 (NS)	0,99 (NS)
	nad5	676	0,88 (S)	0,99 (S)	1,00 (NS)	1,00 (NS)
MEF10	*nad2*	842	0,18 (S)	0,95 (S)	1,00 (NS)	1,00 (NS)
MEF11	*ccmFc*	378	0,11 (S)	0,15 (S)	0,05 (S)	0,12 (S)
	cox3	422	0,73 (S)	0,99 (S)	0,98 (NS)	0,97 (NS)
	matR	1730	0,14 (S)	0,71 (S)	0,61 (NS)	0,63 (NS)
	ccmFn-2	344	0,50 (S)	0,95 (S)	0,92 (NS)	0,93 (NS)
	ccmC	568	0,09 (S)	0,88 (S)	0,95 (NS)	0,95 (NS)
MEF13	*cox3*	314	0,40 (S)	0,98 (S)	0,71 (S)	0,98 (S)
	nad7	213	0,40 (S)	0,91 (S)	0,18 (S)	0,95 (S)
	ccmFc	50	0,50 (S)	0,94 (S)	0,40 (S)	0,99 (S)
	ccmFc	415	0,45 (S)	0,93 (S)	0,60 (S)	0,95 (S)
	nad2	59	0,60 (S)	0,99 (S)	0,38 (S)	1,00 (S)
	nad4	158	0,85 (S)	0,99 (S)	0,92 (NS)	1,00 (NS)
	nad5	1916	0,94(NS)	1,00 (NS)	0,74 (S)	1,00 (S)
MEF19	*ccmB*	566	0,08 (S)	0,75 (S)	0,76 (S)	0,87 (S)
MEF21	*cox3*	257	0,58 (S)	0,96 (S)	0,97 (NS)	0,94 (NS)
MEF26	*cox3*	311	0,66 (S)	0,87 (S)	0,97 (NS)	0,97 (NS)

B

PPR proteins	C terminal domain(s)	Affected sites	MORF1	MORF3	MORF8
MEF1	**EE+DYW**	*rps4*-956	-	-	+
		nad7-963	-	-	+
		nad2-1160	-	-	+
MEF10	**EE+DYW**	*nad2*-842	-	-	+
MEF26	**EE+DYW**	*cox3*-311	-	-	+
		nad4-166	-	-	+
MEF8	**EE+DYW**	*nad5*-676	-	-	+
		nad6-95	-	-	+
MEF11	**EE+DYW**	*cox3*-422	-	-	+
		ccmC-568/*ccb256*-568	-	-	+
		ccb203-344/*ccmFN2*-344	-	-	+
		matR-1730	-	-	+
		ccmFc-378/*ccb452*-378	-	+	-
		nad4-124	-	-	-
MEF13	**EE+**	*ccmFc*-50/*ccb452*-50	+	+	+
		ccmFc-415/ccb452-415	+	+	+
		nad7-213	-	+	+
		cox3-314	-	+	+
		nad2-59	-	+	+
		nad4-158	-	+	+
		nad5-1916	-	+	-
		nad5-1665	-	-	-
MEF21	**EE+**	*cox3*-257	+	-	+
MEF18	**E**	*nad4*-1355	-	-	-
MEF20	**E**	*rps4*-226	-	-	+
MEF19	**E**	*ccb206*-566/*ccmB*-566	+	+	+
MEF9	**E**	*nad7*-200	-	-	-

3.2.2 Autoactivation

In the Y2H assay, some polypeptides can directly interact to target DNA and activate the transcription from the reporter genes without binding of activation domain (autoactivation). To eliminate such polypeptides clones from the partial MEF proteins, I tested the Y2H with MEF-pGBKT7 constructs with the pGADT7-empty vector (Fig. 3.16, 3.17, 3.18). If the partial MEF proteins directly interact with promoter region,, they give us growing colonies on the SD-TLHA (SD-tryptophan-leucine-histidine-adenine) plates.

Among the inspected 42 deletion-constructs, 6 constructs showed autoactivation (Fig. 3.16, 3.17, 3.18). Among these 6 the MEF13_EE+, MEF13_ECE+ and the MEF13_E+ proteins containing cells were able to grow also on the SD-TLHA + 3AT plates (Fig. 3.17), suggesting very strong autoactivation with these constructs. Therefore I removed them from further experiments. The other 3 constructs, the MEF9_E, the MEF19_EN and the MEF21_EN could not grow on the 3AT containing plates (Fig. 3.16, 3.18), implying that these autoactivation is not so strong. I included these 3 constructs in further analyses but evaluated with careful consideration of the autoactivation effect.

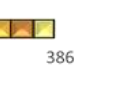
		SD-TLHA	SD-TLHA + 3AT
	pGBKT7	pGADT7-empty	
1 386	MEF1_PLS		
E DYW 387 588	MEF1_EE+DYW		
1 424	MEF8_PLS		
E 425 520	MEF8_EE+		
425 455	MEF8_EN		
456 520	MEF8_ECE+		
1 579	MEF9_PLS		
E 580 656	MEF9_E		
1 255	MEF10_PLS-N		
256 427	MEF10_PLS-C		
E 428 503	MEF10_E		
E DYW 428 623	MEF10_EE+DYW		
1 482	MEF11_PLS		
E 29 561	MEF11_PLS-EE+		

Figure 3.16: Results for autoactivation test with pGBKT7-MEF and pGADT7-empty vectors. In all charts on the left the shorter protein sequences are included (classification of subgroups after Lurin et al. 2004). The squares framed in red indicate the positives colonies growing on SD-TLHA or SD-TLHA and 3AT plates.

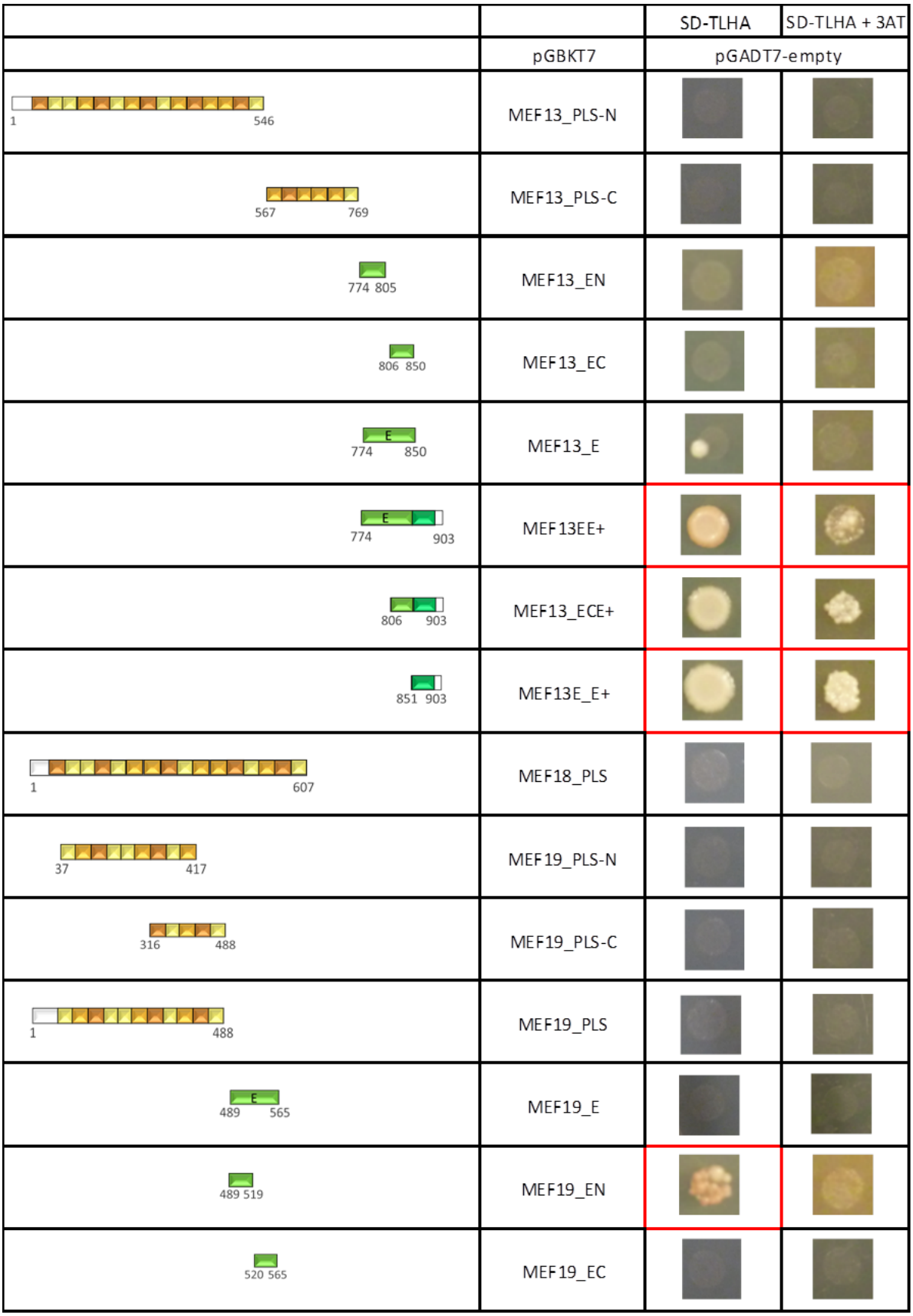


Figure 3.17: Results for autoactivation test with pGBKT7-MEF and pGADT7-empty vectors. In all charts on the left the shorter protein sequences are included (classification of subgroups after Lurin et al. 2004). The squares framed in red indicate the positives colonies growing on SD-TLHA or SD-TLHA and 3AT plates.

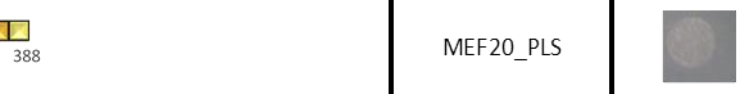
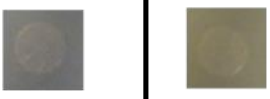
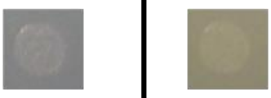
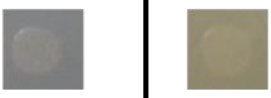
		SD-TLHA	SD-TLHA + 3AT
	pGBKT7	pGADT7-empty	
1 388	MEF20_PLS		
41 340	MEF21_PLS-N		
240 411	MEF21_PLS-C		
E 412 534	MEF21_EE+		
412 442	MEF21_EN		
412 425	MEF21_EN-N		
443 488	MEF21_EC		
E 412 488	MEF21_E		
443 534	MEF21_ECE+		
1 472	MEF26_PLS-N		
473 675	MEF26_PLS-C		
E 680 756	MEF26_E		
E DYW 680 882	MEF26_EE+DYW		

Figure 3.18: Results for autoactivation test with pGBKT7-MEF and pGADT7-empty vectors. In all charts on the left the shorter protein sequences are included (classification of subgroups after Lurin et al. 2004). The squares framed in red indicate the positives colonies growing on SD-TLHA or SD-TLHA and 3AT plates.

3.3 Y2H assay between PPRs from the DYW subgroup and MORF proteins

3.3.1 MEF1

MEF1 is necessary for editing at three positions in three different transcripts: *rps4*-956, *nad7*-963 and the *nad2*-1160 (Zehrmann et al. 2011), which share target sites with the MORF8 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). For the Y2H assays, I prepared two MEF1 deletion clones, MEF1_PLS having all of the PLS motifs and MEF1_EE+DYW with just the E, E+ and DYW domains. The full length MEF1 protein (MEF1_FL) shows weak interaction with MORF8, stronger interaction with MORF1 and no with any of other MORFs (Fig. 3.19). In contrast, the MEF1_PPR domain shows interaction with the MORF8 and also with the MORF1, MORF2, MORF3, MORF9 and weakly with the MORF6 (Fig. 3.19). The EE+DYW domain shows interaction with the MORF1 and very weak with MORF2, but not with MORF8 and other MORFs (Fig. 3.19). None of them grew on the 3AT plates indicating the interactions are not strong enough to survive with the additive 3-aminotriazole (Fig. 3.19).

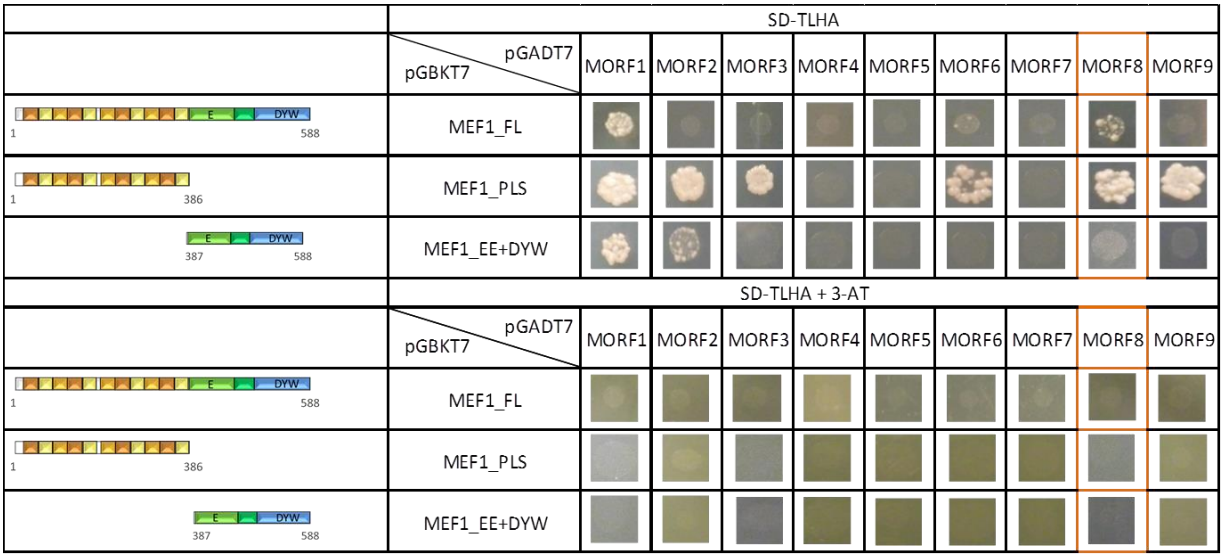


Figure 3.19: Y2H results for the parts of MEF1 in pGBKT7 and MORFs in pGADT7. The orange colored rectangles indicate the colonies with MORF8 protein that shares the target sites with the MEF1 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF1_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequences are indicated with the initial and the last amino acid numbers.

3.3.2 MEF10

The MEF10 protein is necessary for RNA editing at *nad2*-842 site (Härtel et al. 2013a) which is also affected in the *morf8* mutant (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data; Härtel et al. 2013a). MEF10 and MORF8 interact in Y2H analysis (Härtel et al. 2013a). The MEF10 establishes binding also with the MORF1 and MORF2 (Fig. 3.20). The N terminal part of the whole PPR motifs and the E domain shows similar pattern but slightly weaker binding with MORF8 (Fig. 3.20). The C-terminal motifs of the PPR repeats interestingly show a stronger interaction with MORF6 and weakly with MORF3 addition to MORF1, MORF2 and MORF8 (Fig. 3.20). None of them grew on the 2,5mM 3AT containing media (Fig. 3.20).

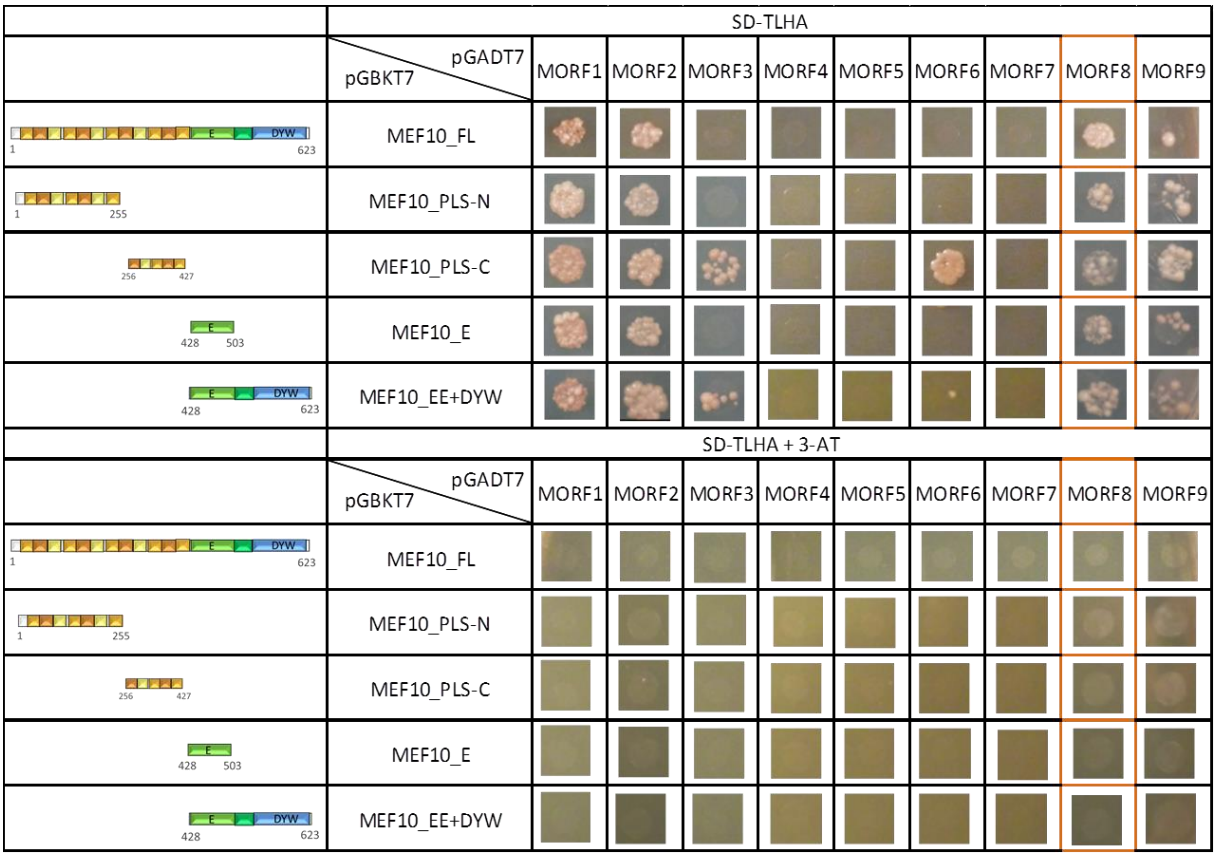


Figure 3.20: Y2H results between the MEF10 constructs and the MORF proteins. The orange framed rectangles indicate the colonies with MORF8 protein which share target sites with the MEF10 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF10_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequences are indicated with the initial and the last amino acid numbers.

3.3.3 MEF26

In the *mef26-2* T-DNA mutant, the editing is absent at the *cox3*-311 site and decreased until 60% at the *nad4*-166 site (Arenas-M. et al. 2014).. Both MEF26 target sites are also targets of the MORF8 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). In the Y2H assay, the MEF26_FL, MEF26_PLS-N, MEF26_PLS-C and the MEF26 EE+DYW create connection with the MORF8, but the only E domain cannot (Fig. 3.21). The entire investigated constructs bind to MORF1 and to the plastid targeted MORF2 (Fig. 3.21). The MEF26_FL, MEF26_PLS-N and MEF26_EE+DYW interact also with MORF9, one of the chloroplast targeted RNA editing factors (Fig. 3.21). With the MORF3 the N-

terminal deletion of the PLS motifs and the EE+DYW domains together establish contact (Fig. 3.21). No colonies are visible with the MORF4, MORF5 (Fig. 3.21). The MORF6 shows a very weak interaction with the MEF26_PLS-N and the MORF7 interact stronger with the MEF26_EE+DYW (Fig. 3.21). The connection between the MEF26_PLS-N and MORF1 is so strong that the yeast cells were able to grow on the 3AT plate (Fig. 3.21). Also the yeast cells with MORF7 and the MEF26_EE+DYW show weak growing on the 3AT plate (Fig. 3.21).

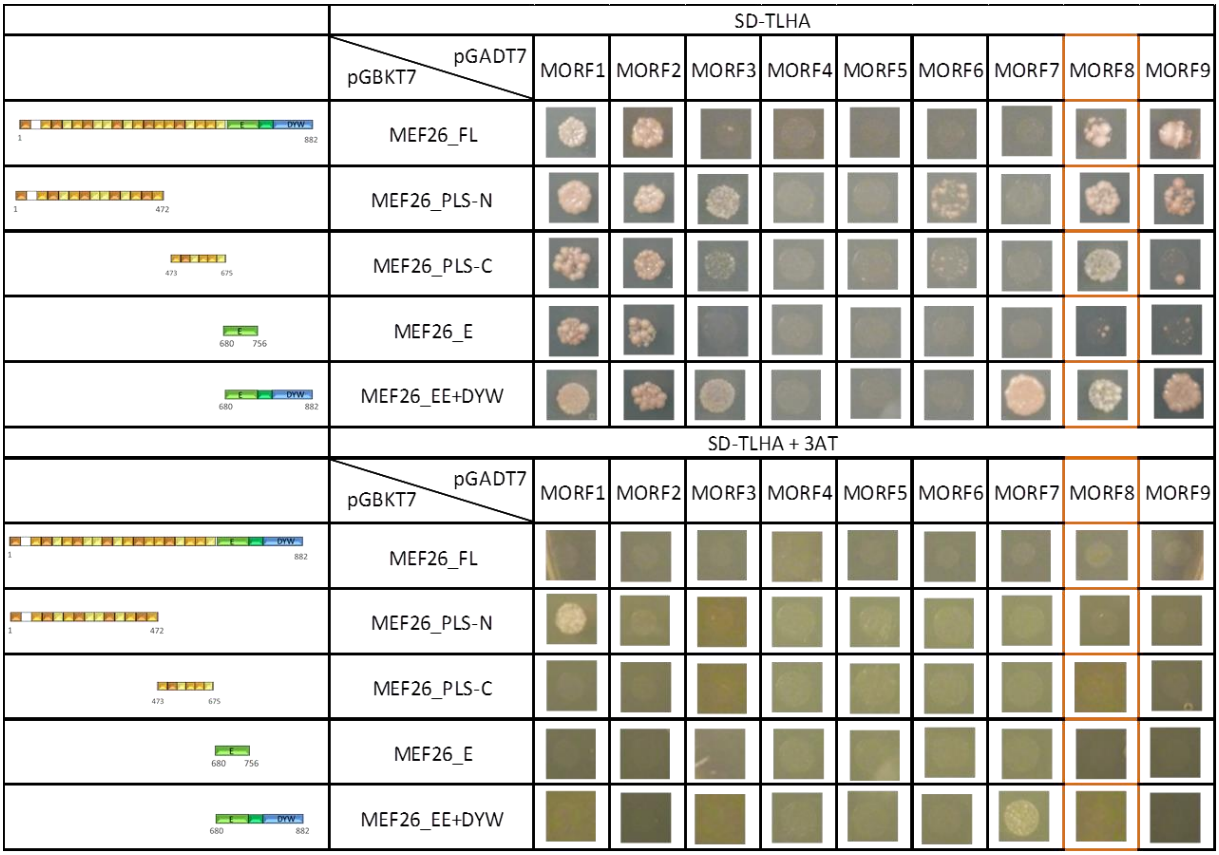


Figure 3.21: Y2H results between the pGBKT7-MEF26 and the pGADT7-MORF constructs. The deletions are indicated on the left side. In the lowermost table show the affected site of the MEF26_FL. The orange framed rectangles indicate the colonies with MORF8 protein which share target sites with the MEF26 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF26_FL pictures: Jörg A. unpublished data).

3.3.4 MEF8

The MEF8 has just 5 PPR repeats, what are statistically not sufficient for specific recognition of the target sites (Verbitskiy et al. 2012). The E domain of this protein is different from the other PPR type RNA editing factors, because it does not have conserved 11 amino acids at the C-terminal part of the E domain, implying that the complete third helix deduced from the predicted secondary structure is missing (http://bioinf.cs.ucl.ac.uk/psipred/). These features suggest that MEF8 does not recognize specific RNA sequence but supply DYW domain to other E or E+ subclass PPR proteins as same as DYW1 which has only a part of E domain and DYW domain and no PPR motifs (Boussardon et al. 2012). MEF8 has two target sites, the *nad5*-676 and the *nad6*-95, which are also affected in the *morf8* mutant (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). The MEF8 is unique not just considering its structure, but also in the results from the Y2H investigation. Almost all of the MEF8 deletion constructs show interaction not just with the MORF8 but also with other 8 MORFs (Fig. 3.22). The only exception is the ECE+, which contains the C-terminal part of the E domain and the E+

domain (Fig. 3.22). Maybe it is because the MEF8 protein has an 11 amino acids gap in this region compared to other PRR type RNA editing factors. The strongest connection was observed between the N-terminal part of the E-domain and the MORFs, suggesting that there is a suitable MORF binding surface in the N-terminus (Fig. 3.22). Strong binding between the MEF8_EN and MORFs are also confirmed by the examination on the SD-TLHA + 3AT medium (Fig. 3.22). The C-terminal part of the E and the E+ domain show positive interaction only with the MORF8 factor (Fig. 3.22).

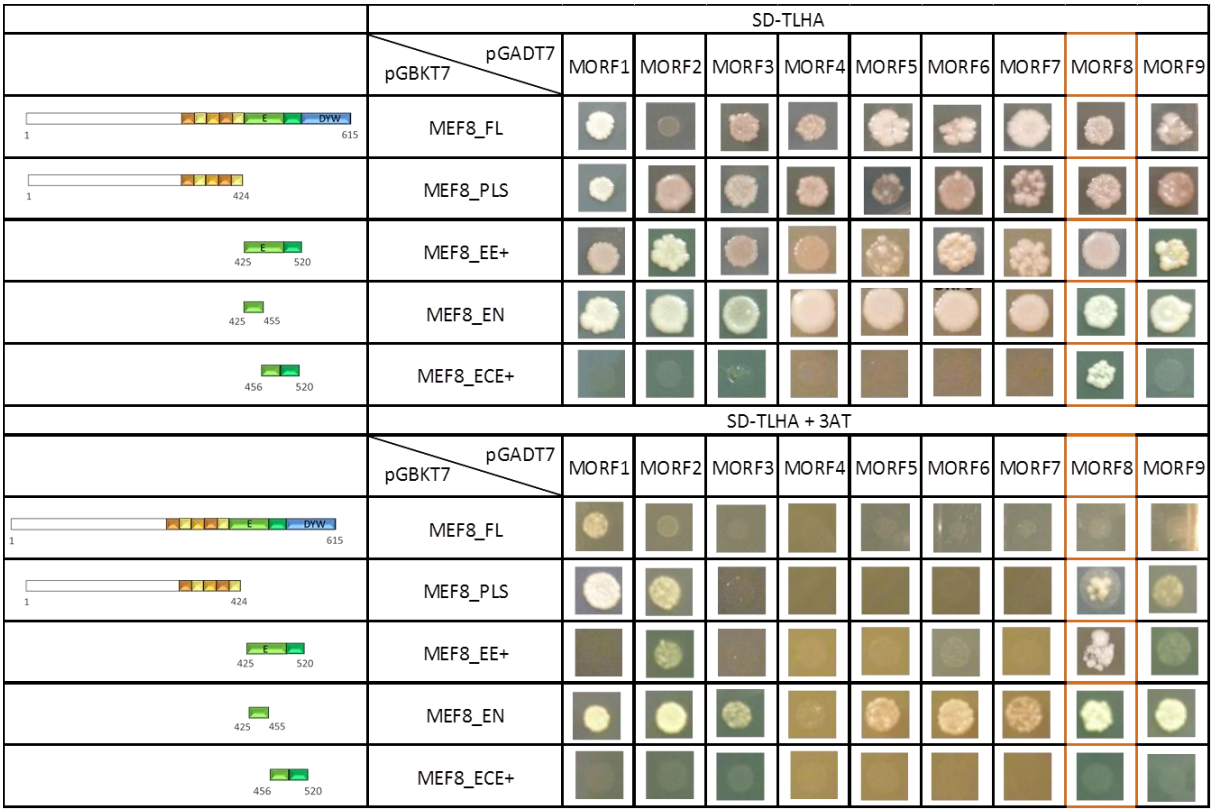


Figure 3.22: Y2H results between the MEF8 constructs and the MORF proteins. The orange framed rectangles indicate the colonies with MORF8 protein which share target sites with the MEF8 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF8_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequences are indicated with the initial and the last amino acid numbers.

3.3.5 MEF8 related proteins

The *MEF8S,* At2g34370 and the At1g29710 genes encoded proteins are similar to the MEF8 protein and at the At2g15690 loci encoded protein is also related but it is farther represented in the family tree (Verbitskiy et al. 2012: Fig. 3.23.A). All of these four PPR proteins have only 5 or 6 PPR motifs as similar to MEF8. The MEF8S, the most related protein of MEF8, which affects the same two target editing sites of MEF8 in pollen, shows very similar interaction results to MEF8 (Jörg A. unpublished data: Fig. 3.23.C). The other three investigated short PPR proteins gave similar results. All of them interact with MORF1, MORF2, MORF3, MORF8 and MORF9 (Fig. 3.23.C). At the locus At2g34370 and At2g15690 encoded proteins have a connection also with the MORF6 (Fig. 3.23.C). The At2g15690 gene encoded protein moreover gives positive interaction also with the MORF4 (Fig. 3.23.C). Bindings between the MORF1 and three short PPRs are so strong that the yeast cells appear also in the presence of the 3-AT (Fig. 3.23.C).

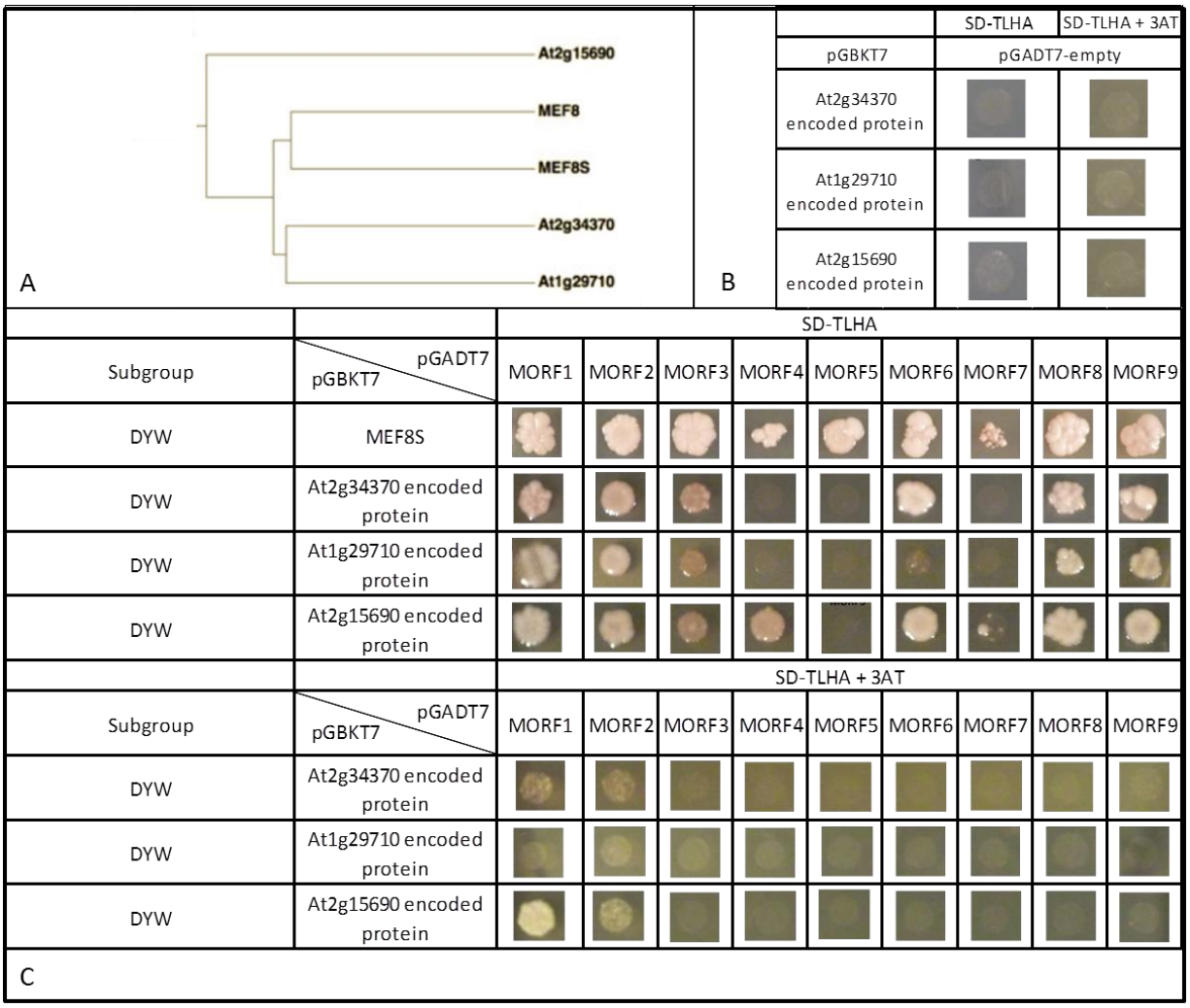


Figure 3.23: Y2H assay between the MEF8 related proteins and the MORF proteins. (A) Similarities between the investigated proteins are represented in family tree (Verbitskiy et al. 2012). (B) Autoactivation test results. (C) Y2H colonies are indicated between the MEF8 related proteins and the MORFs. The orange framed rectangle indicates the colonies with MORF8 protein which share target sites with the MEF8S (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (MEF8S_FL pictures: Jörg A. unpublished data). On the left side the subgroup of the investigated PPR proteins are indicated.

3.3.6 MEF11

The MEF11 has six target sites (Takenaka et al. 2013a; Verbitskiy et al. 2010) and four of them, the *cox3*-422, *ccb256*-568, *ccb203*-344 and the *matR*-1730 are also targeted by the MORF8 and the *ccb452*-378 is targeted by MORF3 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). The *nad4*-124 does not show any editing reduction in any of the MORF mutants. With the MORF3 the MEF11 protein shows a very weak interaction and binds to the MORF8, MORF1 and MORF2 (Fig. 3.24). If the DYW domain or the DYW domain together with the EE+ domains were deleted, the connection to the MORF8 factor is weaker and to the MORF3 is disappared (Fig. 3.24), suggesting that the DYW domain is needed for the connections. All of the MEF11 constructs could interact with MORF1, MORF2 and weakly with MORF9 (Fig. 3.24). None of them was strong enough to make the cells able to survive with addition of the 3AT into the selective media (Fig. 3.24).

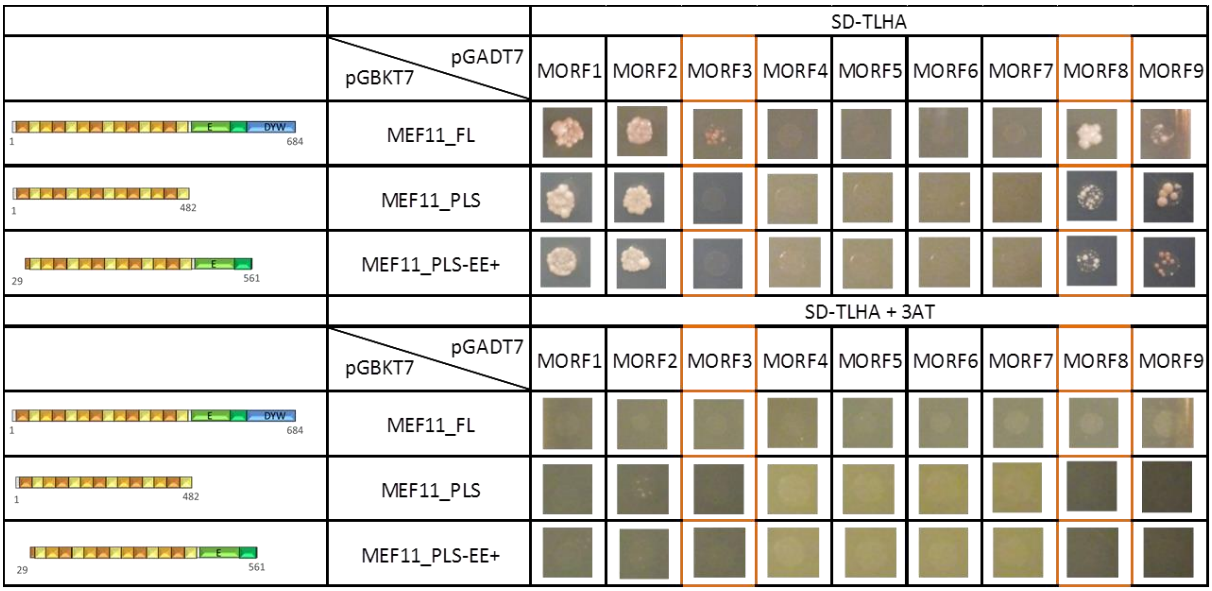


Figure 3.24: Y2H results between the MEF11 constructs and the MORF proteins. The orange framed rectangles indicate the colonies with MORF3 which shares one target site and MORF8 proteins which shares four target sites with the MEF11 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF11_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequences are indicated with the initial and the last amino acid numbers.

3.4 Y2H assay between the E or E+ subgroup PPR proteins and MORFs

3.4.1 MEF13

The MEF13 has 8 target sites (Glass et al. 2015) and from them significantly one transcript at the same two positions (*ccb452*-50, -415) are affected by three MORFs, the MORF1, MORF3 and the MORF8, five sites (*cox3*-314, *nad2*-59, *nad4*-158, *nad5*-1916, *nad7*-213) are also significantly affected by MORF3 and MORF8, and in the *nad5* transcript at position 1665 seems to be not significantly affected by other MORFs (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). The MEF13 binds to the MORF1 and to MORF8, but not to the MORF3 (Fig. 3.25) in spite that many MEF13 targets editing sites are affected in *morf3* mutant lines. The MEF13 also connects to the MORF2 and MORF9 chloroplast targeted MORFs (Fig. 3.25). It is not excluded that the MEF13 is functional also in the chloroplast since this protein is possibly dual targeted (Glass et al. 2015). The MEF13_PLS-N and MEF13 _PLS-C interact with MORF1, MORF2 and MORF9 (Fig. 3.25). The MEF13_PLS-C, the MEF13_E domain and the N or C parts of the E domain are able to connect with MORF3 and MORF8 (Fig. 3.25). None of the partial MEF13 constructs interact with MORF4, MORF5, MORF6 and MORF7 except for a combination with MEF13 C-terminal PPR repeats and MORF4 (Fig. 3.25). By the MEF13_PLS-C + MORF6 and the MEF13_E + MORF9 constructs, weak interactions are observed (Fig. 3.25). The MEF13_PLS-C + MORF4 and the MEF13 full-length + MORF2 containing cells showed some colonies on the SD-TLHA + 3AT plates (Fig. 3.25).

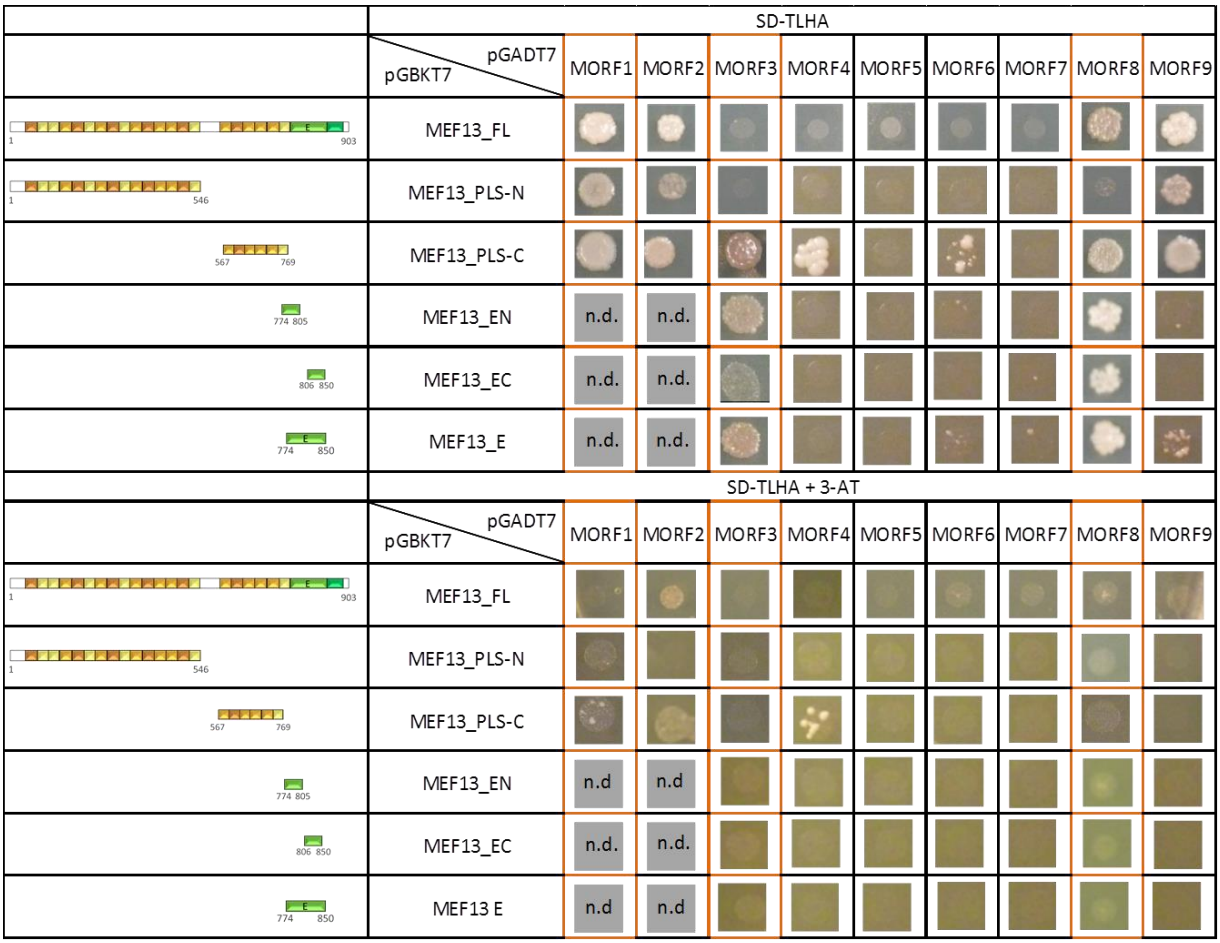


Figure 3.25: Y2H results between the MEF13 constructs and the MORF proteins. The n.d. indicates the not analyzed constructs. The orange framed rectangles indicate the colonies with MORF1, MORF3 and MORF8 proteins which share target sites with the MEF13 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF13_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequences are indicated with the initial and the last amino acid numbers.

3.4.2 MEF21

A cytidine in *cox3* transcript at position 257 needs the MEF21 protein to be edited (Takenaka et al. 2010). MORF1 and the MORF8 are also involved in the RNA editing at the site (Shi et al. 2016 a supplemental data; Glass et al. 2015; Bentolila et al. 2013 supplemental data). All of the MEF21 partial constructs bind to the MORF1 and MORF2 (Fig. 3.26). All of the partial E domain clones interact not just with the MORF1, but also with the MORF8 (Fig. 3.26), suggesting that the C-terminal part of the E domain also can be a binding surface for the MORF1 and MORF8 proteins. In counter to this, the MEF21 establish a connection with the MORF3 just if the E domain is present (we have to be careful with the EN partial protein data because it shows autoactivity on the SD-TLHA plates (Fig. 3.18)). On the other hand MEF21_EN-N is enough to establish connection with MORF1, MORF2 and MORF8 (Fig. 3.26). On the 3AT containing plates only the MEF21_EN/MEF21_EE+/MEF21_ECE+ with MORF2 cotransformed cells grow (Fig. 3.26).

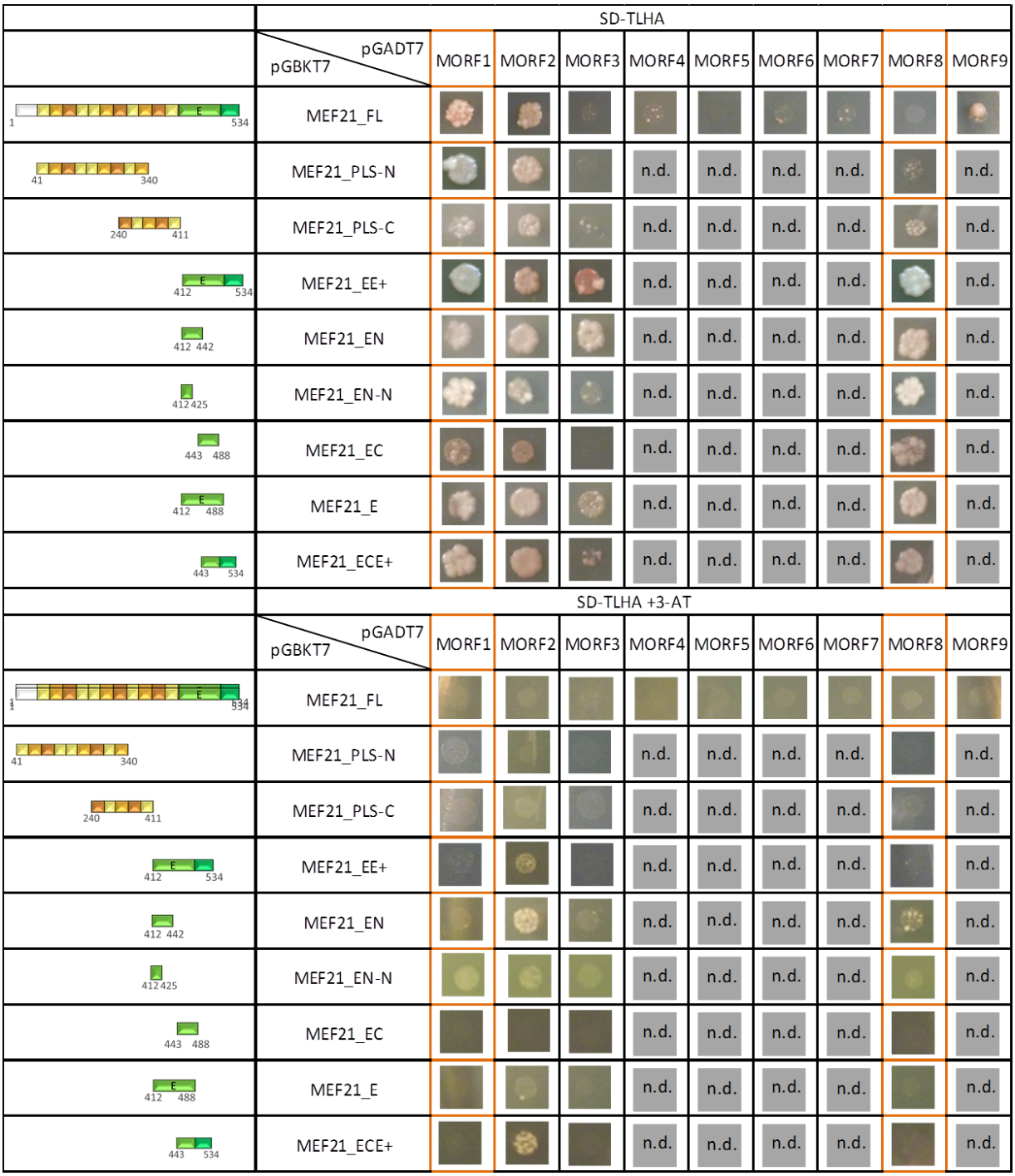


Figure 3.26: Y2H results between the MEF21 constructs and the MORF proteins. The n.d. indicates not analyzed constructs. The orange framed rectangles indicate the colonies with MORF1 and MORF8 proteins, which share target site with the MEF21 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF21_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequences are indicated with the initial and the last amino acid numbers.

3.5 Y2H assay between PPRs from the E subgroup and MORFs

3.5.1 MEF18

The MEF18 has a rather degenerated E-domain (Takenaka et al. 2010). The target RNA editing site of the MEF18, *nad4*-1355 is not affected in any MORF mutants (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). Despite this fact I examined the interactions between these MEF18 constructs and the MORFs. The MEF18 (MEF18_FL) and the MEF18_PLS could establish binding with MORF1, MORF2 and MORF8 (Fig. 3.27). For binding to the MORF3 the E domain is necessary (Fig. 3.27). None of them could grow on the SD-TLHA + 3AT selective media (Fig. 3.27).

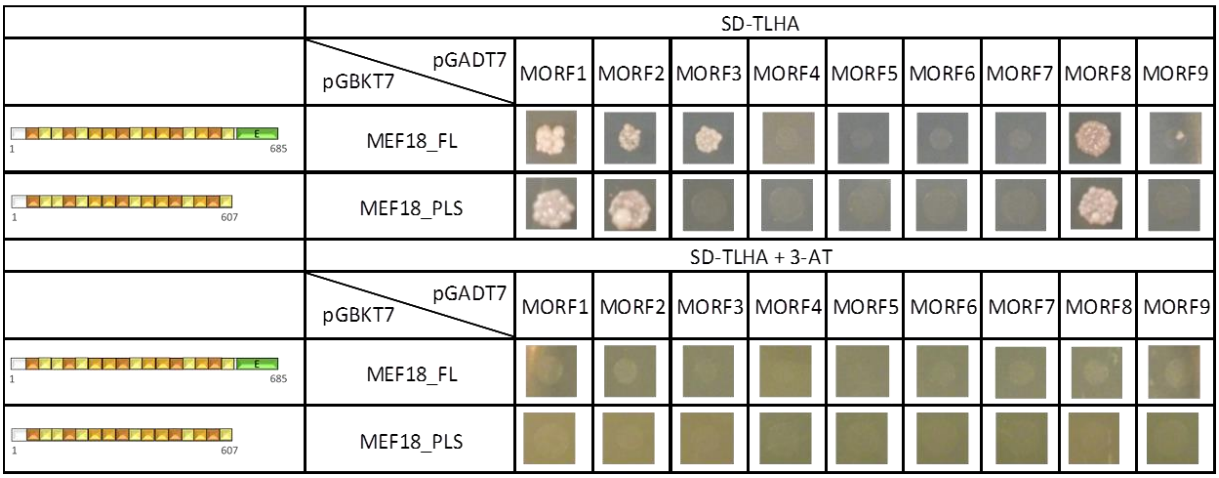


Figure 3.27: Y2H results between the MEF18 constructs and the MORF proteins. The MEF18 affected site is not shared between the MORF proteins (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF18_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequence is indicated with the initial and the last amino acid numbers.

3.5.2 MEF20

The MEF20 has also rather a degenerated E-domain (similarly to the MEF18) which is a few amino acids shorter than the E domain of MEF19 (Takenaka et al. 2010). The MEF20 shares *rps4*-226 RNA editing site with the MORF8 protein as target (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data), but in the Y2H assay they do not connect to each other (Fig. 3.28). MORF8 does not interact also with the MEF20_PLS construct (Fig. 3.28). The MEF20_FL and its PPR domain interact just with MORF1 in this assay (Fig. 3.28). The only one positive result on the SD-TLHA 3-AT plates is weak connection between the MEF20_FL and the MORF1 (Fig. 3.28).

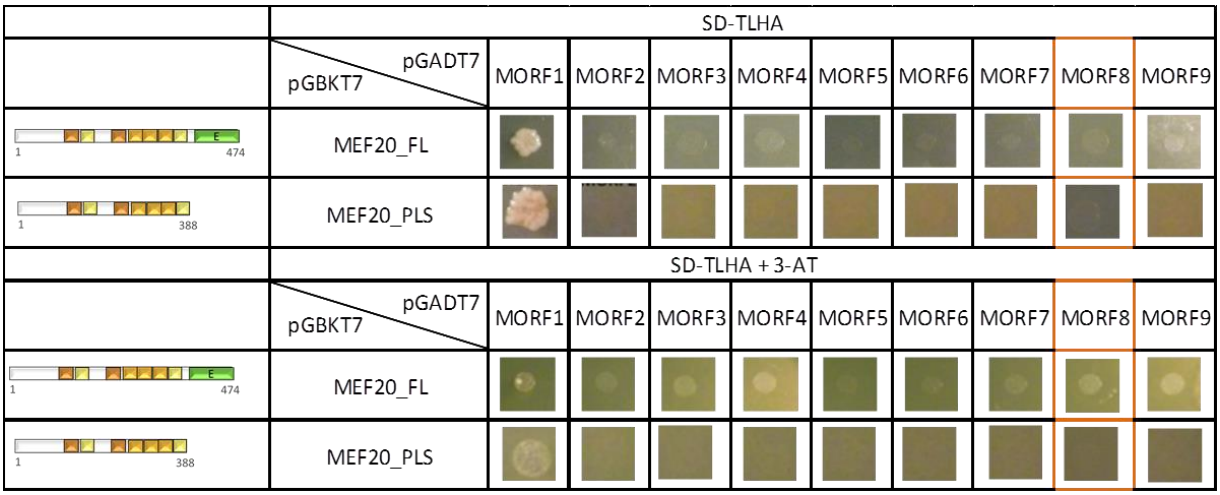


Figure 3.28: Y2H results between the MEF20 constructs and the MORF proteins. The orange framed rectangles indicate the colonies with MORF8 protein which share target site with the MEF20 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF20_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequence is indicated with the initial and the last amino acid numbers.

3.5.3 MEF19

E domain of the MEF19 is a few amino acids longer than those of the MEF18 and MEF20 but same size as the E-domain of the MEF9 protein (Takenaka et al. 2010). The MEF19 is necessary for the editing at site *ccb206*-566 (Takenaka et al. 2010), which is affected also in the mutant of MORF1, MORF3 and MORF8 genes (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). The MEF19 (MEF19_FL), MEF19_PLS, MEF19 just with its E-domain (MEF19_E), the N-terminally E-domain (MEF19_EN) and the C-region of that (MEF19_EC) interact with MORF1 and MORF8 and also with MORF2 (Fig. 3.29). The same partial constructs (except the MEF19_PLS, which has not yet been investigated with MORF9) interact also with MORF9, although the MEF19_EC shows only very weak interaction like with MORF8 (Fig. 3.29). MEF19_E/MEF19_EN also connect with MORF3 (Fig. 3.29). The PLS-N partial MEF19 connects to MORF1 and MORF2 and the MEF19_PLS-C interacts with the MORF2 and MORF8 proteins (Fig. 3.29). We have to be careful with the MEF19_EN partial protein data because it shows autoactivity on the SD-TLHA plates (Fig. 3.17). Binding with the MORF1, MORF2, MORF8 and MORF9 seem to be true since the colonies grow on the 3AT containing plates (Fig. 3.29) while autoactivation is not observed on the medium. (Fig. 3.17). The MEF19_PLS-C and MORF2 RNA editing factors containing cells also grow on the 3AT containing media (Fig. 3.29).

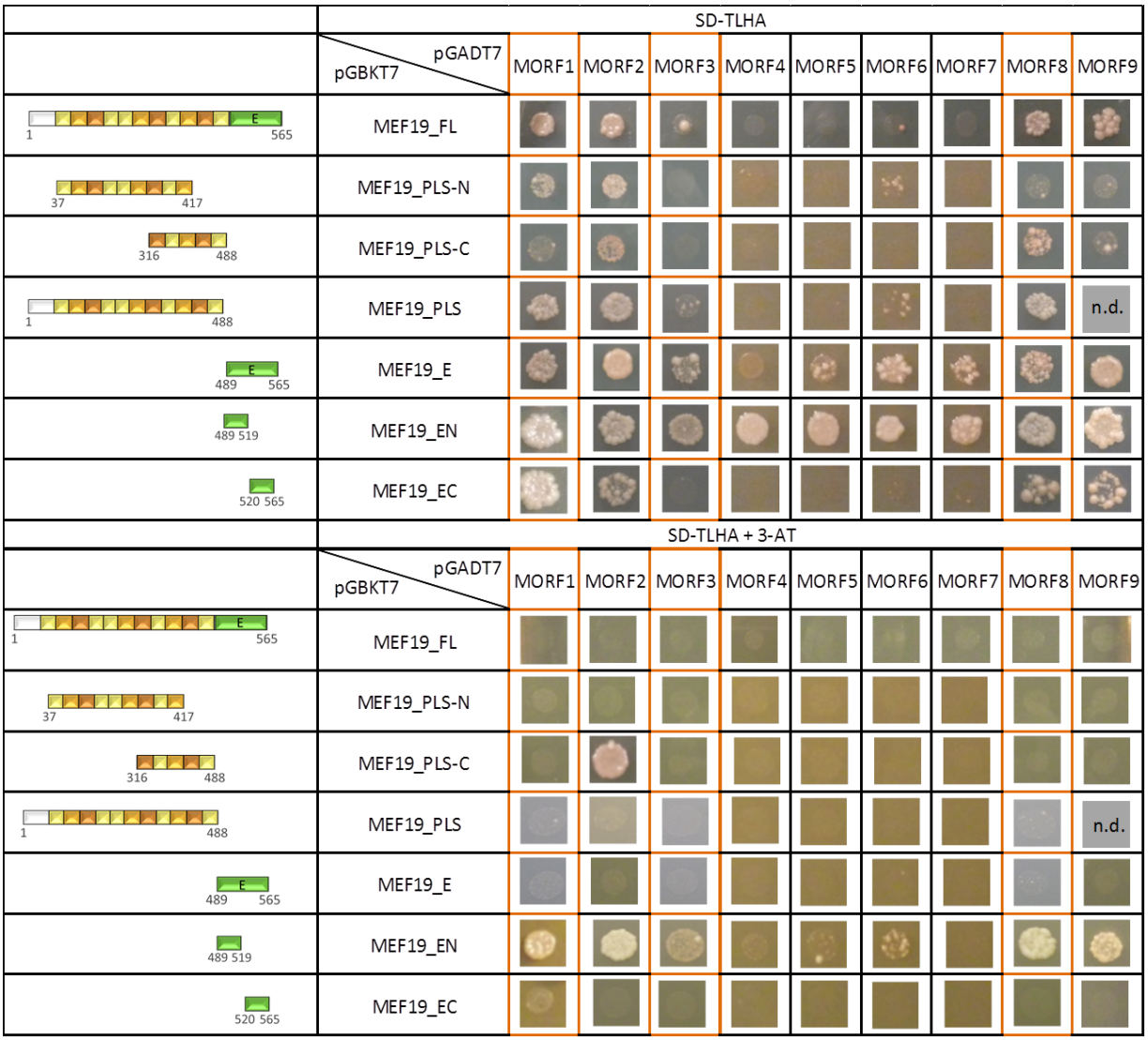


Figure 3.29: Y2H results between the MEF19 constructs and the MORF proteins. The n.d. indicates the not analyzed constructs. The orange framed rectangles indicate the colonies with MORF1, MORF3 and MORF8 proteins which share target site with the MEF19 (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF19_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequences are indicated with the initial and the last amino acid numbers.

3.5.4 MEF9

The MEF9 has an E domain, which is in the same size as the MEF19 E domain, and is necessary for the editing at position 200 of the *nad7* transcript (Takenaka et al. 2010). The MEF9 does not share its target site with any other MORFs (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). In the Y2H assay the MEF9 (MEF9_FL), _PLS and _E establish binding with MORF1, MORF2, MORF8, and MORF9 (Fig. 3.30). The MEF9_FL and MEF9_PLS show interaction very weakly with MORF3 (Fig. 3.30). The E domain of MEF9 could interact with all MORFs (Fig. 3.30), but we have to be careful with the data because it shows autoactivity on the SD-TLHA plates (Fig. 3.16). On the SD-TLHA + 3AT selective medium the MEF9 E domain shows very strong interaction with MORF2 and weaker with MORF1, MORF3, MORF6 and MORF9 (Fig. 3.30) while autoactivation is not occurred on it (Fig. 3.16).

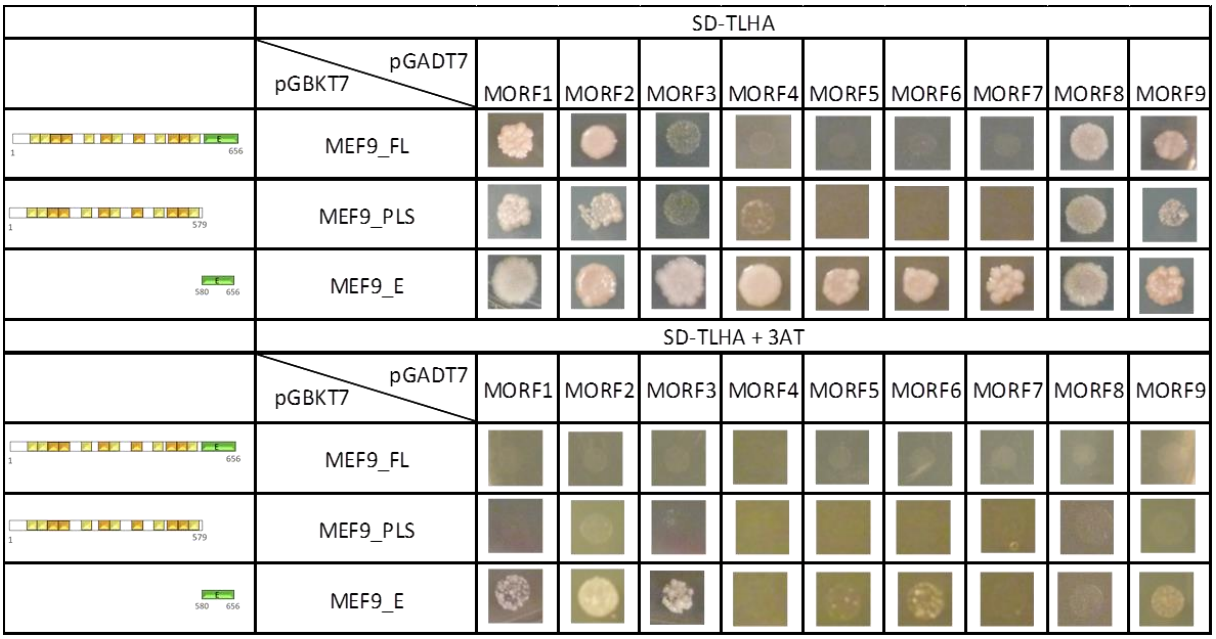


Figure 3.30: Y2H results between the MEF9 constructs and the MORF proteins. The MEF9 affected site is not affected in any MORF mutant plants (Shi et al. 2016 a and Bentolila et al. 2013 supplemental data). (The MEF9_FL pictures: Jörg A. unpublished data). On the left the MEF deletion sequences are indicated with the initial and the last amino acid numbers.

3.6 Investigation of interaction between different deletion constructs

3.6.1 Connections between MEF19/MEF13 + MORF3 deletion constructs

Y2H assays with partial fragments of PPR type RNA editing factors suggested that E domains especially the N-terminal of them have higher affinity to MORF proteins than the other parts. MORF proteins have homo- and hetero- interaction at the C-terminus at least in the case of MORF1 and MORF3 (Zehrmann et al. 2015). Therefore, in which parts MORF proteins interact to PPR proteins are important to figure out in the structure of RNA editing complexes. Interaction between partial E domain fragments of several PPR editing factors and various MORF fragments are analyzed with Y2H assay.

The MORF3_N2 and N3 that contains the N-terminal MORF3 protein interact with both the MEF13_EN and the MEF19_EN (Fig. 3.31). The MEF13_EN does not interact with any other MORF3 fragments (Fig. 3.31) suggesting that the N-terminal MORF3 are important for the binding to the N-terminal part of the E-domain. The MEF13_EC can also bind to the N-terminal MORF3 protein (MORF3_N2 and _N3) like the MEF13 E-domain without the E+ motif (Fig. 3.31), which suggests that not just the N-terminal part but also the C-terminus of the E-domain can bind to the N-terminal MORF3 constructs. Furthermore, the entire MEF13_E domain binds also to MORF3_C1 and very weakly to MORF3_C2 (Fig. 3.31). The interaction between MEF13_E and MORF3 is stronger at amino acid positions 1-104 of MORF3, weaker between amino acids 105-159 and not observed after position 160 (Fig. 3.31). For MEF19_EN, we can take only MORF3_N2 and _N3 as significant positive interaction, since it is autoactivated with empty vector on the SD-TLHA medium (Fig. 3.17). The MEF19_EC binds to MORF3_C1 but not to other parts (Fig. 3.31).

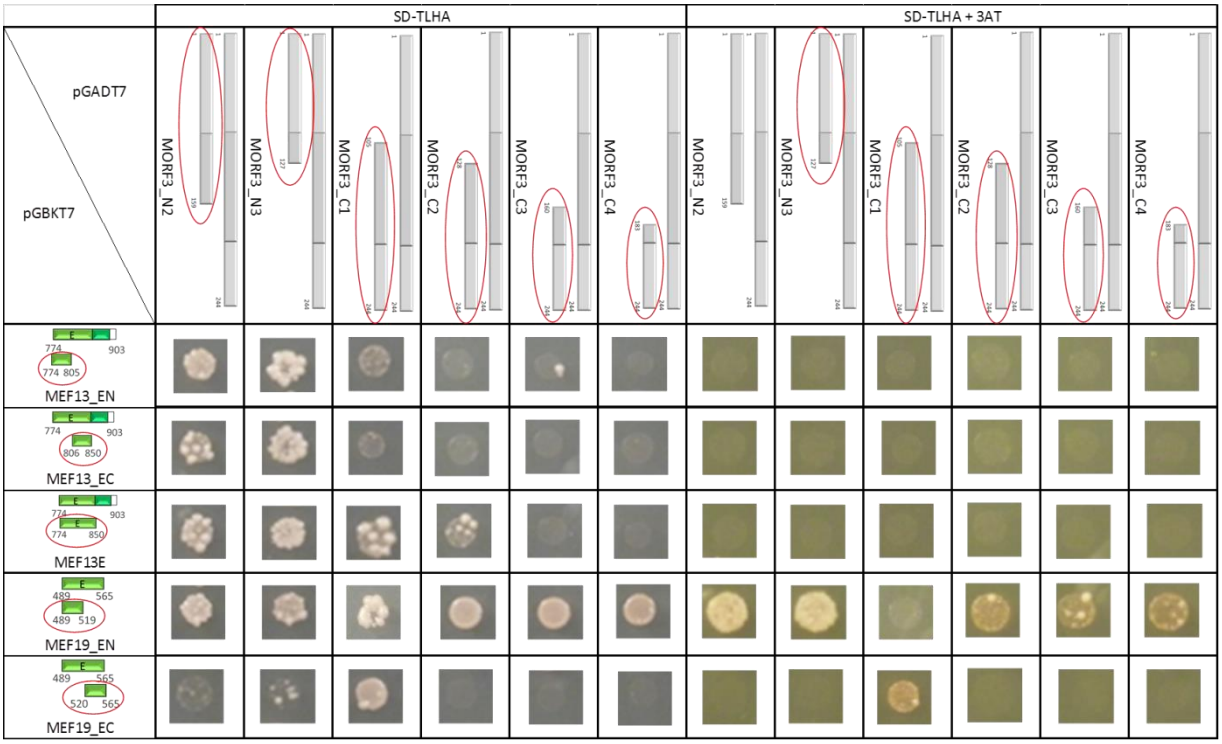


Figure 3.31: The Y2H assay with the bait, E domain fragments of MEF13 and MEF19 (left) and the prey, MORF3 protein fragments (top).

3.6.2 Mapping the binding site(s) between MEF19/MEF21 + MORF1 deletion constructs

Interaction between MORF1 fragments and E domain fragments of MEF19 or MEF21 are analyzed because MORF1 share sites with these two MEFs (Takenaka et al. 2012) and the E domain is a possible the interacting surface (Chateigner-Boutin et al. 2013) . It is important not to forget, that both MEF19_EN and MEF21_EN interact with an empty construct (Fig. 3.17, 3.18). However, autoactivated yeast cells containing them cannot grow on 3AT media (Fig. 3.17, 20), suggesting that their interactions are not so strong. Therefore, we counted only the results of the MEF19_EN and MEF21_EN on a medium with 3AT as positive interactions. The MEF21 E domain and the N-termini of the MEF21_EN (MEF21_EN-N) could interact just with the MORF1_N2 (Fig. 3.32), which suggests that this shorter N terminal part of the E domain is sufficient as a binding surface of MORF1_N2. The MEF21_EC does not bind any MORF1 partial proteins and the MEF21_ECE+ interacts with the MORF1_C2 and _C4 (Fig. 3.32). These results suggest that the 13 amino acid long sequence at the N-terminal of MEF21 E domain can bind to the N-terminal part of the MORF box in the MORF1. The MEF21_EC, ECE+ results suggest that the E+ motif is also able to provide a binding surface for MORF1 but for the C-terminus of the MORF box. It is contrast to the MEF21-EN. None of the constructs show any positive results on the 3AT media, just the autoactive MEF19_EN and MEF21_EN proteins containing cells could slowly grow with MORF1_N2 (Fig. 3.32).

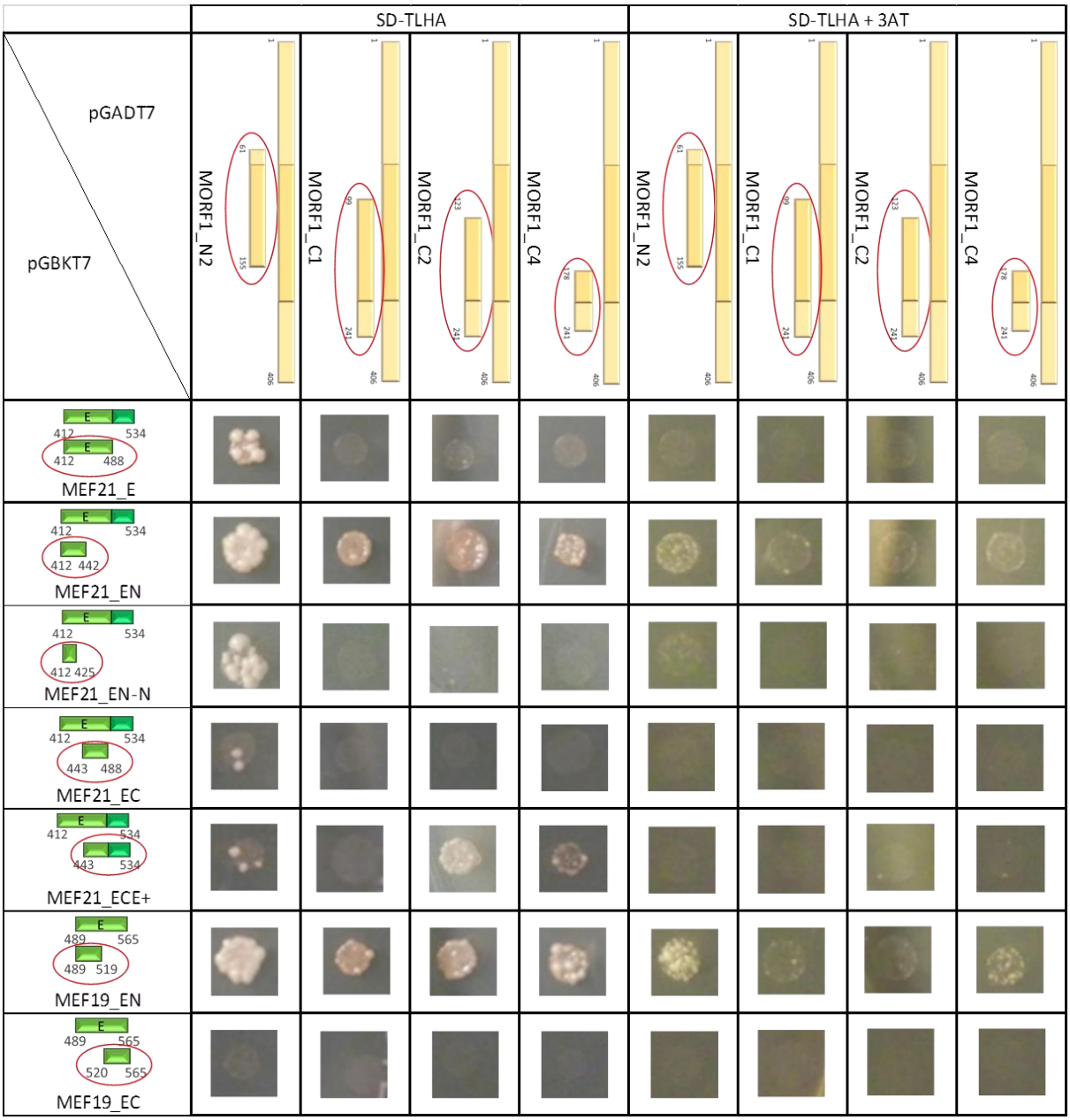


Figure 3.32: Results of the Y2H analysis after using several different E-domain and MORF1 deletions. The sequences of the proteins are signed on the left and above.

3.7 Expression of MEF21 using the *Pichia pastoris* expression system

3.7.1 General overview

To obtain a recombinant protein is important for its molecular functional analysis. Without any heterologous overexpressing systems it is hard to have sufficient amount of selected endogenous protein(s) in some organisms like Arabidopsis. To overexpress *Arabidopsis thaliana* proteins using the *E.coli* system is the most popular method but often accompanied with difficulties. The expressed protein may be toxic to the cell before or after the induction and can cause to form inclusion bodies (Rosano and Ceccarelli 2014). The reason for the inclusion bodies can be inappropriate pH, the protein folding strategies, or missing of required cofactors or disulfide bonds (Rosano and Ceccarelli 2014). It is not possible to predict whether the protein will be expressed or not, and soluble or insoluble. There are several strategies to solve these problems but it often requires complex optimization process. One of the alternative strategies is to use other hosts for the further investigations, like a eukaryotic expression system, like the *Pichia pastoris*. If the choice of protein to be expressed is a eukaryotic protein and requires certain posttranslational modifications for correct folding and structure, it makes sense to try to express it with a eukaryotic system.

3.7.2 The *Pichia pastoris* expression system

The *Pichia* expression system was introduced more than 40 years ago by Phillips Petroleum (Ahmad et al. 2014). Difficulties of using this system are increased in the 1970s after the oil crisis, because the expression system used by him required methanol of which price was increased at the time (Ahmad et al. 2014). Later however the use of this system was again more extended and several different inducing systems became available. Since 1990s, a successful overexpression of a plant enzyme in a 20 gram per liter amount pro culture has been reported (Ahmad et al. 2014). This is showing us that it is possible to produce a large amount of proteins with this organism. There are many other advantages for this system, the availability of several different inducible promoter systems, the possibility to express the protein intracellularly or even extracellularly, the eukaryotic posttranslational modifications like glycosylation, producing disulphide bonds, proline isomerization etc. (Daly and Hearn 2005). The *Pichia* system provides the opportunity to express those proteins, which seem to be toxic for the prokaryotic system, because in the *Pichia* system cells are growing in a proper medium at first and transferred to the other medium for inducing the protein expression.
I employed this system to express the MEF21 plant RNA editing factor since it is difficult to be expressed in the *E.coli* expression system and the *Pichia* system provides the eukaryotic modifications which are not obtained in *E.coli* system.

3.7.3 Prediction of possible glycosylation site

To use the *Pichia* system, it is important to predict the possible N- and O-glycosylation site(s) (Fig. 3.33.C,D) on the expressed proteins by an online tool in advance, because the hyperglycosylation which affect the protein folding can happen, despite it is not as prominent as in *S. cerevisiae* (Ahmad et al. 2014). Very little but some O-glycosylation has been detected in *Pichia*. Glycosylation is that glycan is added to the protein. The N-glycosylation is addition of the oligosaccharides to the nitrogen atom of asparagine in a specific pattern (Asn-X-S/T) (Fig. 3.33.C), and the O-glycosylation (Fig. 3.33.D) is that the sugar molecule is added to the oxygen atom of the serine or threonine residue (https://www.thermofisher.com). The O-glycosylation sites do not have specific amino acid pattern,

but are predicted through the structural motifs. Both types of glycosylation are predicted for some candidate proteins for *Pichia* expression system. In MEF21, I found only one possible N-glycosylation site (Fig. 3.33.A) and no O-glycosylation site (Fig. 3.33.B). MEF1, MEF13 and MEF19 are also neither hyperglycosylated nor O-glycosylated. It is also same in the case of some MORF proteins, MORF1, MORF2, MORF3, MORF8 and MORF9. But by the use of the NetOGlyc server, several possible O-glycosylation sites are found in these MORFs, therefore this system is not appropriate to overexpress them (MORF1, 2, 3, 8, 9).

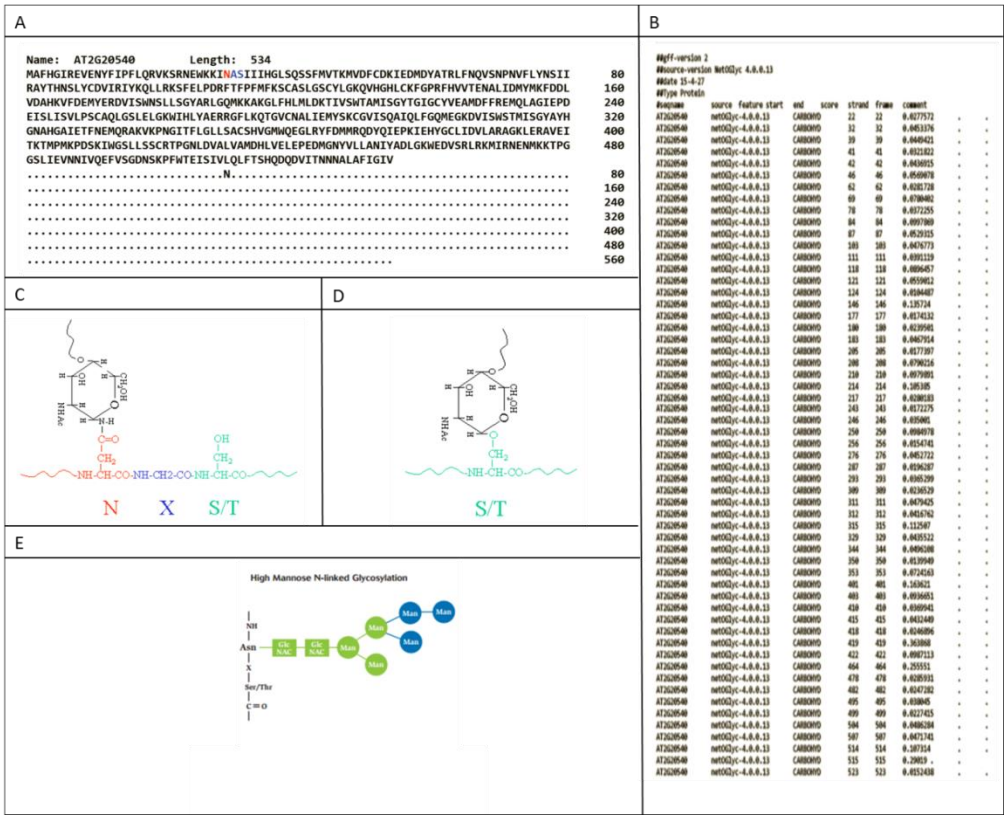


Figure 3.33: Predictions and illustrations of glycosylations for MEF21. (A) N-glycosylation site prediction for MEF21 (http://www.cbs.dtu.dk/services/NetNGlyc/). (B) Predicted O-glycosylation for MEF21 (http://www.cbs.dtu.dk/services/NetOGlyc/). (C) Illustration for N-glycosylation in a specific pattern (http://www.ionsource.com/Card/carbo/nolink.htm). (D) Illustration for O-glycosylation (http://www.ionsource.com/Card/carbo/nolink.htm). (E) Illustration for hyperglycosylation (https://tools.thermofisher.com/content/sfs/appendix/Expression/Post-Translational%20Modifications.pdf).

3.7.4 Selection of vector systems

At first, I selected extracellular vector system for expressing the PPR proteins. For an extracellular expression, cell lyses is not required, therefore the cellular contents which may destroy expressed proteins (e.g. proteases) will not be released. Furthermore, since *Pichia pastoris* has low amount of secreted native proteins, the expressed proteins by extracellular system can be easily recovered. We chose the pPICZα (Fig. 3.34) vector for the extracellular expression. In this vector there is an alcohol oxidase (AOX1) high inducible promoter (P_{AOX1}), which is inducible with methanol and repressed by glucose or glycerol and the most commonly used promoter for the *Pichia* expression system (Daly and Hearn 2005). This vector allows for selection both in bacterial and *Pichia* cells via by Zeocin™ (Fig. 3.34). The vector expresses the α-factor signal sequence (from *S. cer.*) and tags for the recombinant

protein detection and purification (c-myc and His-tag) (Fig. 3.34). The alfa-mating factor (α-MF) is a pre-pro peptide and it is important for the extracellular secretion. The pre-sequence is the signal sequence (19 amino acids) required for the translocation into the ER and the following pro-region (further 60 amino acids) will be cleaved in the Golgi (Daly and Hearn 2005).

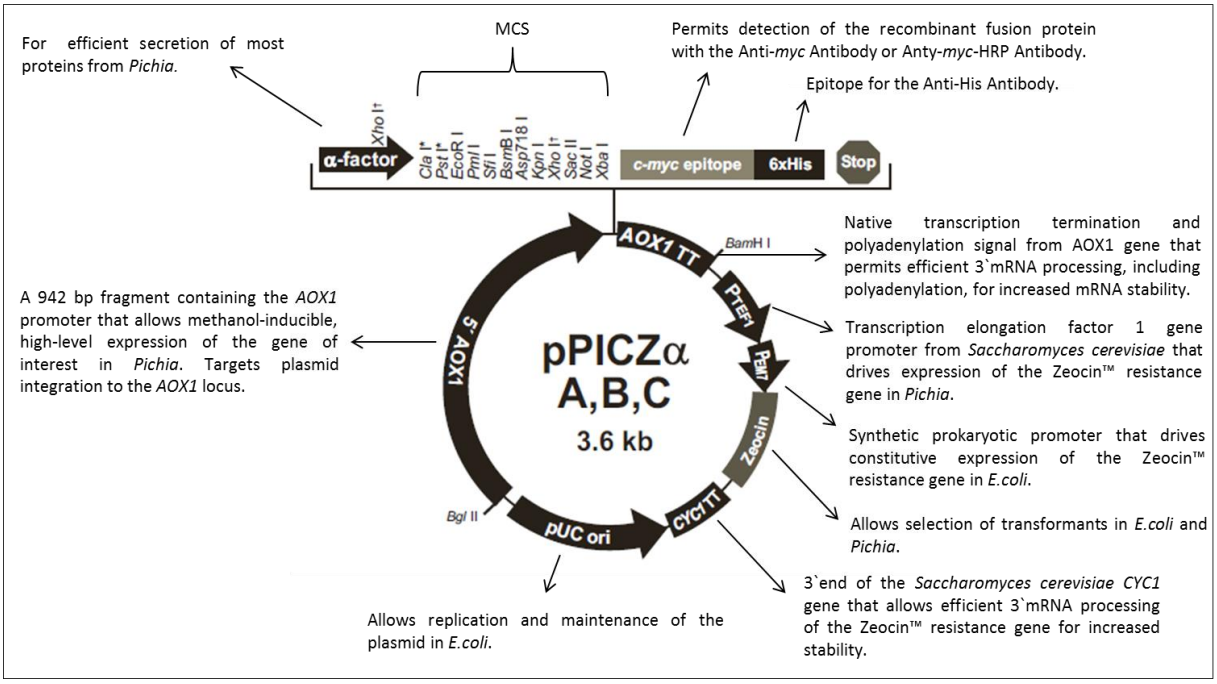


Figure 3.34: Properties of the pPICZα vector (Invitrogen™ V195-20: 21,22).

3.7.5 Introduction into the experimental process

First of all I cloned the PPR type RNA editing factor, MEF21 encoded sequence into the pPICZα (Fig. 3.35). The plasmid DNAs were linearized and transformed by electroporation (Fig. 3.35) after the suggestion of Invitrogen™ to lead multicopy integration into *Pichia* genome (Invitrogen™: K1740-01). The expression cassette is integrated in the *AOX1* locus of the *Pichia* genome through a single crossover event (Fig. 3.35) in the host strain X33. This crossover can happen between the AOX1 locus and the AOX1 promoter region of the vector or through the termination region (TT) (Invitrogen™: K1740-01). After the transformation, the genomic DNA of the *Pichia* cells has to be isolated and sequenced (Fig. 3.35). If everything is correct then the Mut phenotype should be analyzed to check whether the *AOX1* locus remains in intact during a replacement event (Fig. 3.35). If the *AOX1* locus remains in intact, the phenotype will be Mut^+ (Methanol utilization plus) and these strains can metabolize methanol in the wild type form. The Mut^+ strains can grow normally on the MD (Minimal Dextrose) and also on the MM (Minimal Methanol) plates. In contrast, the Mut^S (Methanol utilization slow) strains utilize methanol inefficiently, therefore they cannot express the protein of interest in a large amount. These strains do not grow or just very slowly on the MM plates and show normally growth rates on the MD plates (Invitrogen™: K1740-01).

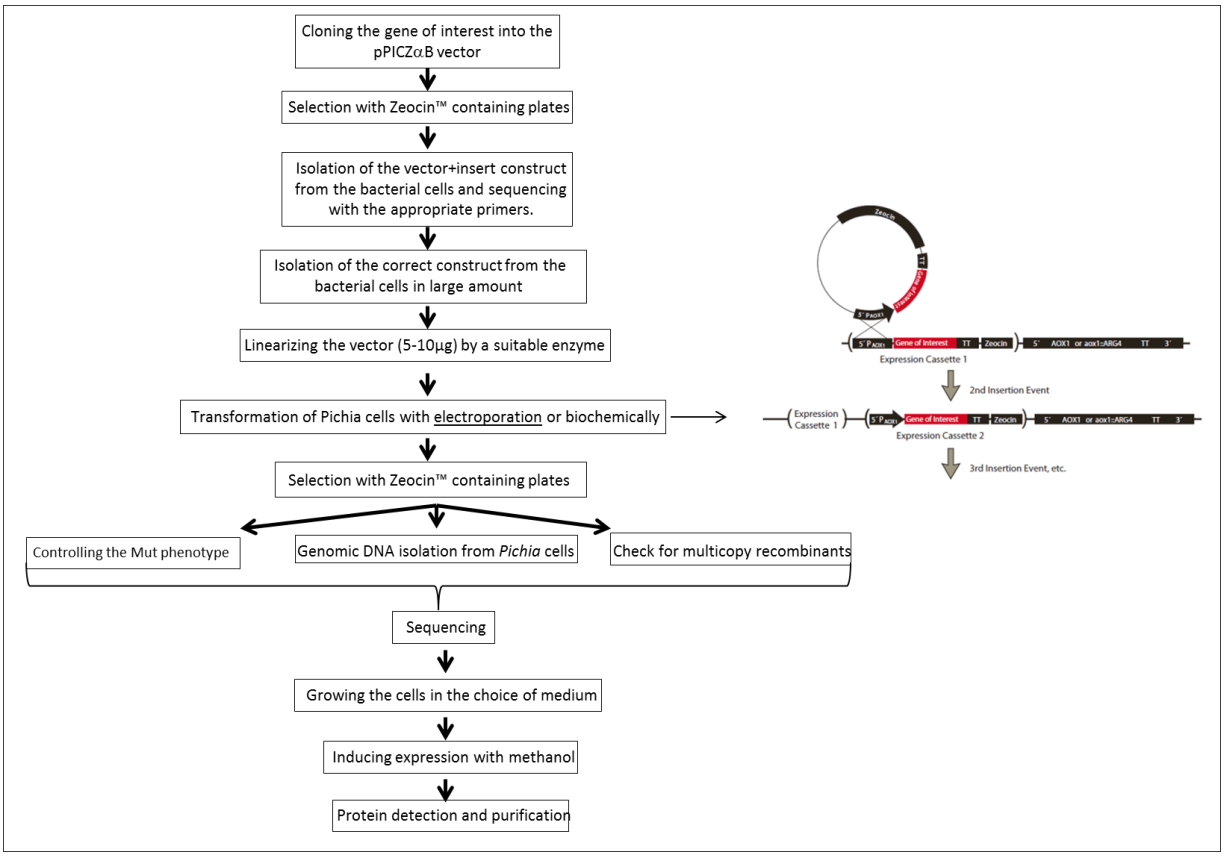


Figure 3.35: The experimental process step by step, using the *Pichia pastoris* expression system.

3.7.6 Estimation of integrated cassette copy number

It is expected that multiple genome integration of the expression cassettes occurs especially in the case of electroporation. The yeast colonies which have more than one expression cassettes are screened with increased amount of Zeocin™ in the plates: 100µg/mL, 500µg/mL, 1000µg/mL and 2000µg/mL (Daly and Hearn 2005), since the higher expression level is expected from the higher copy numbered cells. Colonies on the 100µg/mL Zeocin™ containing plates have one copy of the expression cassette, on the 500µg/mL Zeocin™ two copies, on the 1000µg/mL Zeocin™ three copies and on the 2000 µg/mL Zeocin™ four copies (Daly and Hearn 2005) of integration are expected, respectively. The copy number can increase, since 15-25 copies of integration were also detected in the colonies grown on 1000µg/mL Zeocin™ containing media (Daly and Hearn 2005).

3.7.7 Selection for single and multicopy recombinants

Transformed cells were plated on 100µg/mL Zeocin™ containing YPDS (Yeast Extract Peptone Dextrose) medium (Fig. 3.36.B). Resistance colonies were further streaked and patched on 100µg/mL, 500µg/mL, 1000µg/mL and 2000µg/mL Zeocin™ containing mediums to select multiple integrated expression cassettes containing cells (Fig. 3.36.D). The transformed cells grew on all of these plates but showed the lowest growth rates on the 2000µg/mL Zeocin™ plate (Fig. 3.36.D) suggesting that all of the colonies contain several expression cassettes in the genome and possibly express high level of MEF21 protein. It was not possible to decide which colonies would give more protein amounts from those which grew on the 2000µg/mL Zeocin™ plate, because all of them seem to be similar(Fig. 3.36.D). Therefore I randomly chose two colonies for further investigations.

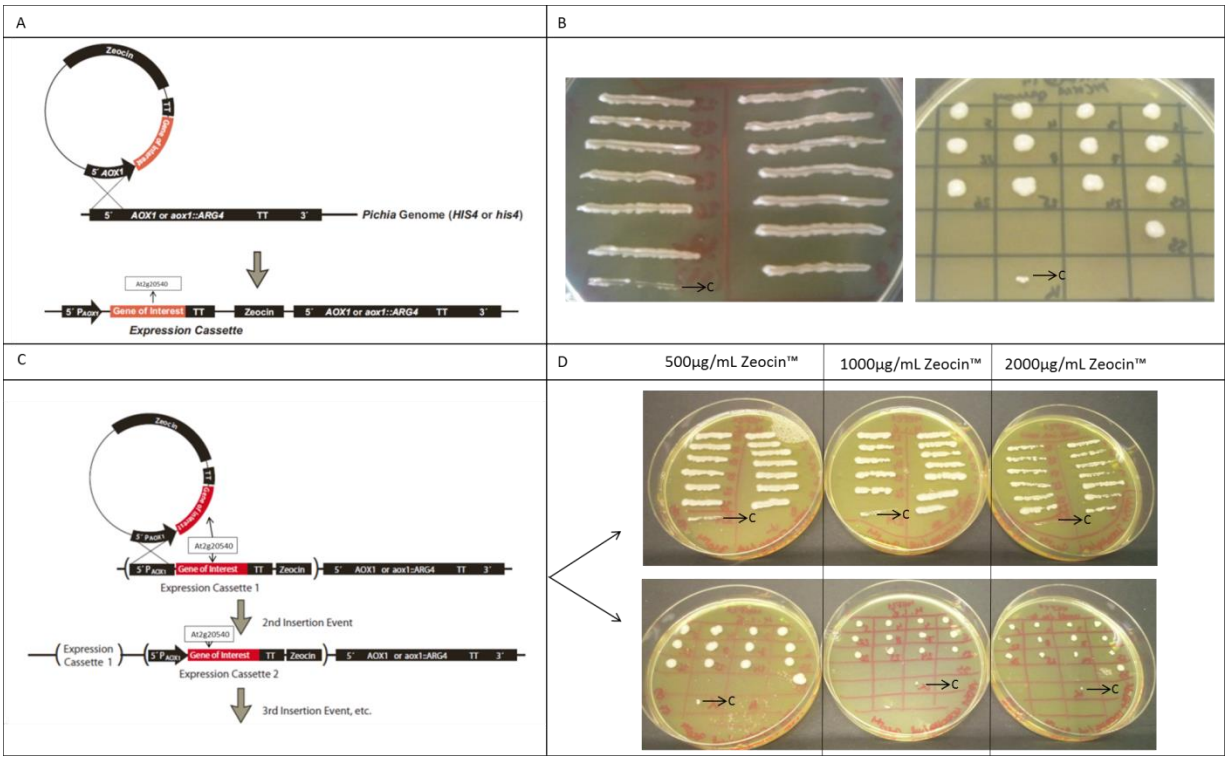


Figure 3.36: Integration event(s) of the expression cassette. (A) Illustration of the integration of the expression cassette (Invitrogen™ K1740-01: 62). (B) Colonies growing on 100µg/mL Zeocin™ containing plates. (C) Illustration of the multiple integration events (Invitrogen™ K1740-01: 63). (D) Streaked and patched colonies on three different plates, containing three different amounts of Zeocin™. Control=c (untransformed X33 cells).

3.7.8 Analysis of the isolated genomic DNA from Pichia and investigation of the Mut phenotype

Mut Phenotypes of these colonies were analyzed on the MD and MM plates. All of the colonies grew normally on the MD (Fig. 3.37.A) plate but slower on the MM plate (Fig. 3.37.B) as same as control 33 strain (Mut^+) shows both on the MD and MM plates similar growing to the other colonies (Fig. 3.37.A, B), suggesting that all of the MEF21 containing cells are Mut+ phenotype. The AOX1 locus and the present of the MEF21 gene in the expression cassette in the cells were analyzed by PCR (Fig. 3.37.C, D, E).

I used MEF21 gene specific primers (MEF21F and MEF21R) and AOX1 specific primers (5`AOX1 and 3`AOX1) to decide whether the MEF21 gene is in the expression cassette or not (Fig. 3.37.C, D, E). PCR with several colonies showed MEF21 locus specific amplification with both 5`AOX1 and MEF21R and MEF21F 3`AOX1 primer sets (Fig. 3.37.D, E) suggesting that the cells are containing at least one copy of MEF21 in the AOX1 locus. PCR with the 5`AOX1 and 3`AOX1 amplified 2.2 kb and about 600bp amplifications (Fig. 3.37.C). The 2.2 kb DNA can derive from AOX1 or MEF21 and the smaller bands correspond to the amplified products from endogenous AOX locus, ~592bp, respectively. These results suggest most of these cell lines contain at least a copy of the MEF21 gene expression cassette(s) and also cassette(s) which do not contain this gene. I continued further processes since the presence of MEF21 expression cassette in the *Pichia* cell lines was confirmed.

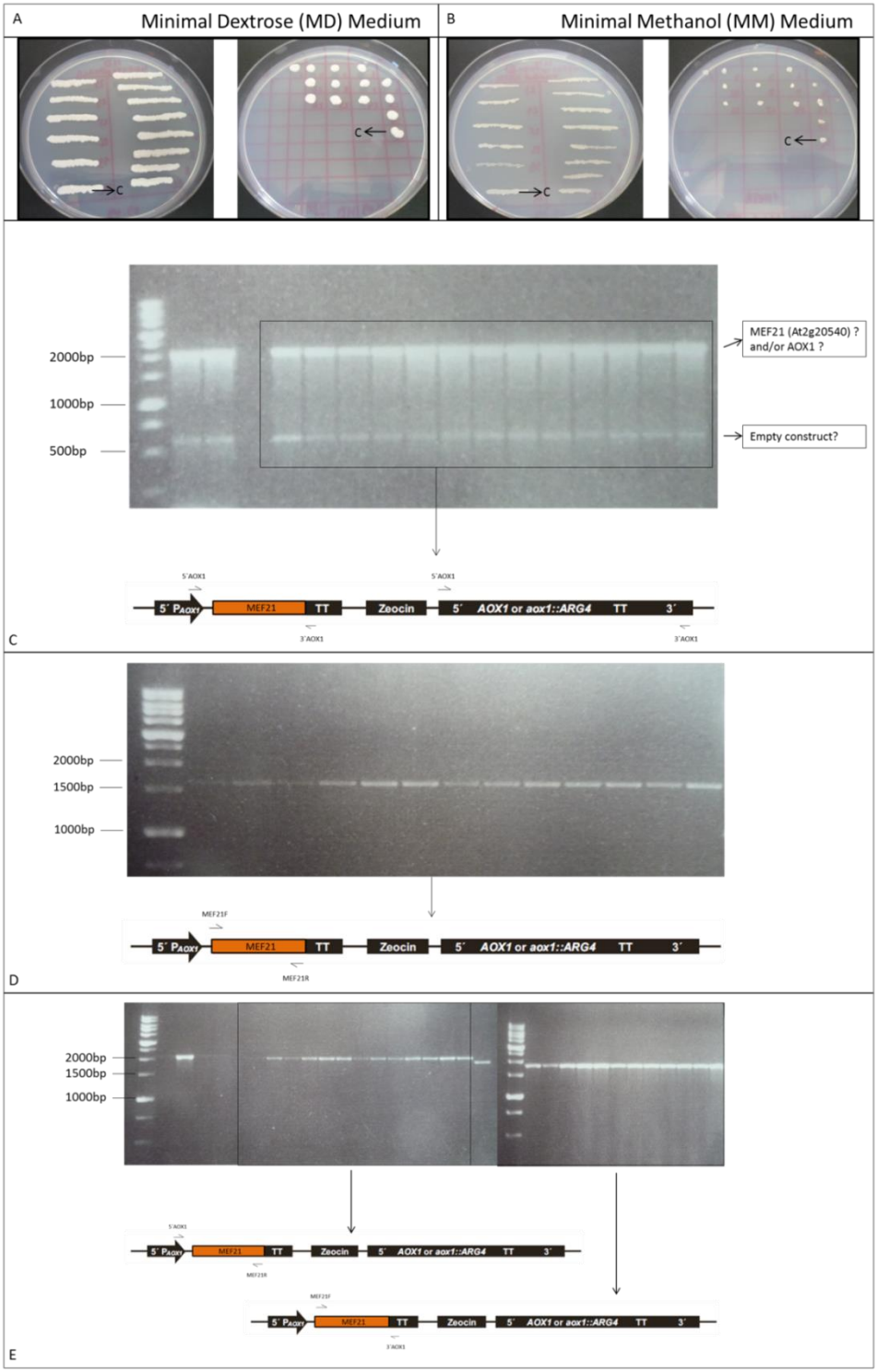


Figure 3.37: Analyzing the Mut phenotype and the expression cassette. The uppermost figures represent the streaked (A) and patched (B) colonies on MD and MM mediums. (C,D,E) Gel pictures using different AOX1 and/or gene specific primers.

3.7.9 Detection of protein expression after methanol induction

The MEF21 gene cassettes which were integrated *Pichia* cells were cultured in 50ml and 5ml liquid cultures, respectively. To grow and induce the MEF21 sequence containing *Pichia* cells I used two protein induction medium systems. At first cells were grown in the BMGY medium, where the glycerol inhibit the expression of the gene and then transferred to the BMMY medium, which

contains methanol instead of glycerol for inducing the protein expression. For the second induction system, I used the MGY (Minimal Glycerol) and MM (Minimal Methanol) mediums to escape from the proteases that are active at a neutral pH. Since *Pichia pastoris* is resistant to the lower pH, the cell growth rate is not affected (Invitrogen™: K1740-01). The MGY medium contains glycerol to inhibit the induction and the MM contains methanol for protein induction.

I chose the colony 33 for the larger-scale, and colony 33 and 4 for the small-scale (Chart 10.3). To optimize incubation time, I collected 1ml of the cultures at different time points (Chart 10.3) and stored in -80^{o}C. The *Pichia* culture samples were centrifuged to separate the pellet and the supernatant to separate the intra-, and extracellularly expressed proteins.

Each fraction was analyzed by SDS ponceau stain and western blot techniques to detect expressed MEF21 proteins (Fig. 3.38). In all cases the expressed proteins were detected in the cell pellets at around 70 kDa or between 25-40 kDa with the anti-myc-HRP antibody. However these proteins are not detected with the anti-His antibody (Fig. 3.38). Predicted size for the MEF21 protein is ~60,89 kDa (http://www.sciencegateway.org/tools/proteinmw.htm), for the tags ~2,5 kDa (Invitrogen™: V195-20) and for the α-factor signal sequence ~9,3 kDa (Invitrogen™: V195-20), respectively. Since processed MEF21 protein should be ~60,89kDa (~63,4 with C-terminal *myc* epitope tag), detected around 70 kDa proteins are likely to contain the α-factor signal sequence but no C-terminal His tag.

To analyze whether expressed proteins are soluble or not, I separated the soluble and insoluble fractions isolated from cultured MEF21-MM colony 33 (larger-scale) and analyzed with SDS ponceau stain and Western blot. We detected the protein of interest just from the insoluble fraction.

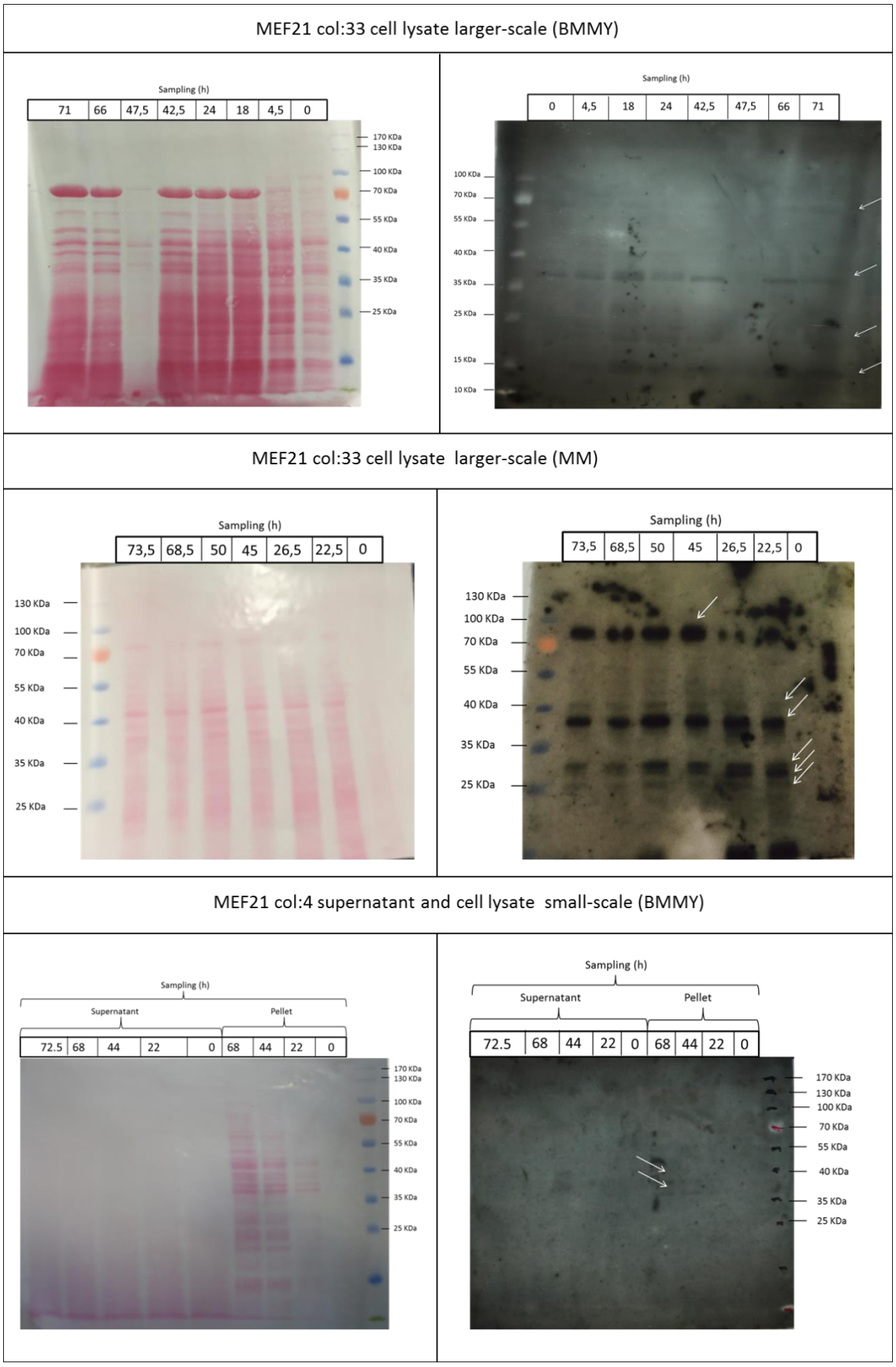


Figure 3.38: SDS ponceau stain (on the left) and Western blot analysis (on the right). (The SDS and the Western blot have been made by Daniil Verbitskiy and Sascha Haag.)

4 Discussion

4.1 Distinct functions of C-terminal domains in RNA editing factors

4.1.1 The DYW domain of MEF28 is indispensable for RNA editing

The DYW domain of MEF28 is indispensable for the editing reaction for both, the *nad2*-89 and *nad2*-90 sites, turned out from complementation results using the MEF28ΔDYW sequence (Chart 10.1). Losing RNA editing function in the DYW deleted proteins are observed in QED1 and lower recovery of editing function in the MEF1 protein (Wagoner et al. 2014; Zehrmann et al. 2010). On the other hand, deletion of DYW domain in MEF11, ELI1, OTP82, CRR28 and CRR22 did not show significant effect on the target editing sites in mitochondria and in chloroplasts, respectively (Hayes et al. 2013; Zehrmann et al. 2011; Okuda et al. 2010, 2009). The DYW domains in MEF11 and CRR22 give editing function to at least at one site when fused with MEF28(PLS-EE+) suggests that these DYW domain are also functional for RNA editing (Chart 10.1).

4.1.2 Relevance of suitable DYW domain for the dual targeting function in MEF28

How do editosome complexes distinguish target RNA editing sites between two or more Cs next to each other? In the case of the *nad2* transcript, cytidines at 89 and 90 are edited with MEF28, counter to the MEF11, CRR22, MEF25, MEF13 and MEF19, which edit just one cytidine of the two or more adjacent cytidines.

The compatibility of DYW domains of the MEF11 and CRR22 with the MEF28_PLS-EE+ sequence for editing at *nad2*-89 but not for the *nad2*-90 site (Zehrmann et al. 2011; Okuda et al. 2009) suggests the differece derives from the distinct accessibility of the DYW domains to the downstream site. The DYW domain has been proposed to be the editing enzyme because of its characteristic amino acid sequence which matches the active site of cytidine deaminases (Salone et al. 2007), though this deaminase activity has not been documented yet (Takenaka 2014a). Possibly DYW domain containing proteins need help of other partner(s) to be functional. The unique feature of MEF28 DYW domain maybe through its possible co-factor(s) come true.

In the possible interacting partners of the MEF28 DYW domain, the MORF8 protein can be included. In the *RIP1* (*MORF8*) mutants the *nad2*-89 site is edited (~50%) and the *nad2*-90 site is not (~0%) (Bentolila et al. 2013 supplemental datas) suggesting that the MEF28 protein needs a protein complex including MORF8, or the RNA binding is stabilized by MORF8 to provide sufficient time for editing downstream site.

4.1.3 Supplied DYW domains can be different between each E containing PPR

E, EE+ or EE+DYW domains from different PPR factors fused to the PLS repeats of MEF28 show distinct effects at the two respective sites (Fig. 3.9, 3.10 and Chart 10.1).

The PLS motifs of MEF28 fused with the EE+ domains of MEF28, MEF25 and MEF13, where the characteristic $HxEx_nCxxC$ deaminase motif is not there, still show a lower effect at the *nad2*-89 site. Up to 29% using the MEF28_EE+, up to 26.4% with MEF25EE+ and up to 52.5% with MEF13EE+ construct (Chart 10.1). Furthermore, the MEF28_PLS fused with the MEF19_E domain could completely edit both sites in the *nad2* transcript (Chart 10.1).

These RNA editing activities in these chimera clones can be explained by the trans supplied DYW domains to distinct E domains. The E subclass PPR protein, CRR4 and the DYW1 which has a C-terminal part of E+ and an entire DYW domain are connecting and acting together for RNA editing at *ndhD*-1 site (Boussardon et al. 2012) implies that missing DYW domains in the E or EE+ subclass PPRs might be complemented by other DYW subclass PPRs. Strong candidates of the DYW domain supplier for the E subclass PPR proteins are the MEF8, MEF8s and other shorter PPR proteins as the At1g29710, At2g34370 or the At2g15690 (Verbitskiy et al. 2012; Takenaka 2014a) which show affinities to many MORF proteins in Y2H (Fig. 3.26). However, other DYW subclass PPRs with longer PPR domains, are also possible to be the partner of the E subclass PPR proteins. The E domain of each E subclass PPR may determine further members of each editosome. In the case of the MEF28ΔDYW protein, a DYW protein may connect with the EE+ domain only very unstably therefore cannot properly function as a full length MEF28. Maybe the situation is similar by MEF28PLS_MEF25EE+ and MEF28PLS_MEF13EE+ proteins (Fig. 10.1). In contrast, the specific complex with MEF28PLS_MEF19E and MEF28_CRR22 maybe a DYW containing protein interacted through the E domain can edit two continuous cytidines as well as MEF28 probably with a support of other proteins (Fig. 10.1).

MEF28PLS_MEF11EE+DYW and the MEF28PLS_MEF11EE+_MEF28DYW lose editing ability for the two *nad2* sites. It is possible that the EE+ domain of MEF11 cannot associate with *nad2* sequnece around the MEF28 target sites or the complex needed for the editing at the *nad2*-89, 90 sites cannot be formed through the MEF11EE+ domain (Fig. 10.1).

4.1.4 Involvement of other factors in the selection of target cytidines

It is surprising that C domains of CRR22 and MEF19 confer editing ability for two adjacent sites with a MEF28 PPR domain despite that neither of them edit one nucleotide downstream cytidine of the bona fide target sites. These results suggest the two continuous RNA editing sites cannot be explaind by only the difference of C-terminal domains. One of other explanations can be RNA sequence around the editing sites. Sequence alignment between the MEF28, MEF11, MEF25, CRR22, MEF13 and MEF19 target mRNA sequences revealed no shared nucleotide sequence around the MEF11, MEF25, CRR22, MEF13, MEF19 target sites (Fig. 3.3). However, it is not excluded that RNA secondary structures including the target editing sites influence the selective editing in respective adjacent cytidines. Alternatively, difference of co-factors between MEF28 and other PPR proteins may affect the target ability to the neighbouring cytidines. The other explanation is disctinct co-factors. I searched for the distinct involvements of MORF proteins which share the affected site(s) with the respective PPR proteins (Fig. 3.3). Almost all of them share site(s) with MORF8, except for CRR22 and MEF25 (Shi et al. 2016 and Bentolila et al. 2013 supplemental data). The CRR22 chloroplast targeted protein shares site only with MORF2 and MORF9 (Härtel 2013b) and the MEF25 does not share its target with any MORFs. Interaction with specific MORF protein may alter the editing complex formation and affect the accessibility to neighbouring cytidines.

In Plants there are also RRM containing proteins which are involved in RNA editing. The CP31A and CP31B, members of the cpRNP family are nonspecific stabilizers for the transcripts in plastids (Tillich et al. 2009). The ORRM (Organelle RRM) protein family (with one RRM domain) includes the ORRM1 which is plastid-targeted and the ORRM2, ORRM3 and ORRM4 are involved at mitochondrial editing sites (Shi et al. 2016). Due to the verstaility of this domain, it is able to provide a variable surface whereby it would be possible to distinguish between RNA ligands and different interacting partners (Shi et al. 2015). Y2H analysis showed that the ORRM3 and ORRM4 can interact with each other, with the MORF8 (RIP1) protein and with themselves (Shi et al. 2016, 2015; Zehrmann et al. 2015). The adjacent target sites in the *nad*2 transcript of MEF28 are shared by

MORF8 and so far there are no ORRM containing proteins whose mutants show editing defect at these sites (Shi et al. 2016 and Bentolila et al. 2013 supplemental datas). **It will be interesting to analyze whether there are more not yet documented co-factors of the MEF28.**

4.1.5 RNA editing in different angiosperms at the *nad2*-89 and 90 sites

There is no RNA editing in *Beta vulgaris*, *Oryza sativa* and *Triticum aestivum* at the *nad2*-89, 90 sites. There is one editing event at position 89 in the *nad2* transcript in *Oenothera berteriana* and interestingly both neighboring sites are editing targets in *Brassica napus.* In *Beta vulgaris*, *Oryza sativa*, *Triticum aestivum and Oenothera berteriana* no MEF28 homologues I found just in *Brassica napus.* Highly conserved DYW domains in *Arabidopsis* and *Brassica* **suggest functional conservation** of the two MEF28 homologues **to edit adjacent cytidines in the *nad2* transcript**.

4.2 Analysing the results by the use of the Y2H system

4.2.1 Y2H results investigating all MEF-MORF_FL constructs

In the most cases in the Y2H assay, the investigated MEFs interact with the MORF1, MORF2 and with the MORF8. The MEF constructs often established binding with MORF9 and more rarely with MORF3. Also in the most cases the MEF baits and the MORF4,5,6,7 preys could not make connection with each other. These general observations are surprising because as bait I used several different deletion constructs (Chart 10.2).

4.2.2 Y2H results investigating MEF + MORF1, MORF3 and MORF8 connections

The MORF1 shares target site(s) with MEF13, MEF21 and MEF19 (Tab. 3.1). These proteins bind in all deleted versions (and also the full-length protein) to the MORF1 in the Y2H assay (Chart 10.2). The MEF13 and the MEF19 sites are affected by MORF1 and are also affected by MORF3 and MORF8 (Tab. 3.1). The MEF13_FL and MEF19_FL could not bind to the MORF3 but additionally to their binding with MORF1, they could also bind to MORF8 (Fig. 3.25, 31). It has been suggested that the MORF8 possibly can enhance the binding between the MEF13 and MORF3 (Glass et al. 2015). It is suggested that the MEFs and MORFs interact *in vivo* (Takenaka et al. 2012). It is also a question, that if they bind together, than the MEFs can bind just one MORF or they are able to bind more MORFs in a specific complex? After we do not have any crystallized proteins, therefore it will stay as a question.
From the shorter proteins, the MEF19 shorter proteins do not interact with the MORF3 (Fig. 3.29). In contrast the MEF13_PLS-C, E, EN and EC could connect to the MORF3 except the MEF13_PLS-N (Fig. 3.25). The most of the investigated proteins have at least one site what they share with MORF8 (Tab. 3.1). The MORF8 could bind to almost all Y2H constructs (Chart 10.2): in the most cases it could bind to the full length proteins (except the MEF20_FL and the MEF21_FL), to the PLS motifs (except the MEF20_PLS), to the C-terminal part of the PLS motifs, to the E-domain, to the EE+DYW motifs, to the C-terminal part of the E-domain (the MEF19_EN, MEF21_EN, MEF9_E, MEF13_EE+, MEF13_E+, MEF13_ECE+ results are not included in this summarizing after they are autoactive); could bind to the MEF21_EN-N (the shortest one E-domain construct) and also to the MEF21_ECE+ and MEF8_ECE+. It could not bind to the MEF11_PLS-EE+ and to the most PLS-N constructs. From the MEF baits + MORF preys just the MEF8_EE+ and the MEF8_EN with the MORF8 seemed to be strong enough to survive the 3AT in the medium (Chart 10.2).

The MEF8 is special not just in terms of its structure with its short PPR motifs and degenerated E domain (Verbitskiy et al. 2012), but also in the Y2H assay. Surprisingly it was able to interact with fast every deletion constructs (except the ECE+) with almost all MORF proteins, which makes the MEF8 very unique after investigating with this method too (Fig. 3.22, Chart 10.2). The positive results between the mitochondrially located MEFs (the only one predicted exception is the MEF13 which can be dual targeted: Glass et al. 2015) and MORF2 or MORF9 are also surprising because these proteins are locating in the chloroplast.

It is important to confirm the protein interactions with other methods too. Also if a result is negative, we can be never sure that was a true negative or it was a false negative. Therefore it is also important to repeat the Y2H assays more times as it is true to the other methods too.

4.2.3 MORF1 binds to the N-terminal part of the E domain

I investigated several shortened proteins from MEF21 and MEF13 E-domains with MORF partial proteins, to have more insight for the possible binding structure(s). Here I summarize the results got from MEF21 + MORF1 and MEF13 + MORF3 constructs. The MEF19 results I did not include because the MEF19_EN was autoactive (Fig. 3.17) and therefore the MEF19_EC gives not enough data.
The MEF21_EE+, MEF21_E, MEF21_EN-N, MEF21_EC and also the MEF21_ECE+ could bind to the MORF1 (Fig. 3.26), suggesting that the MORF1_FL possibly can bind to the whole EE+ domain of MEF21. From the MORF1 deletions constructs, the MORF1_N2 could bind to the MEF21_E,_EN and _EN-N but not the others (Fig. 3.32). **These results suggest that the MORF1 can bind through its N-terminal part between the 61. and 99. amino acids to also the N-terminal of E domain (as a minimal binding surface to the shortest one E-domain protein is MEF21_EN-N).** The MEF13_E, MEF13_EN and the MEF13_EC could interact with the MORF3 (Fig. 3.25). The MEF13_E could also interact with two more MORF3 constructs: with the MORF3_C1 and MORF3_C2, but none of the shortened MEF13_E-domains could establish any connection with the MORF3_C3 and MORF3_C4 (Fig. 3.31). **These results suggest that the possible binding site of MEF13_E-domain in the MORF3 protein can be from the first amino acid until the 127. amino acid.**

4.3 *Pichia pastoris* expression system

Pichia protein expression system requires much more time for optimization than the e.g. *E.coli* prokaryotic expression system, however this eukaryotic system gives many advantages especially for expressing eukaryotic originate proteins with similar posttranslational modifications (e.g. disulphide bonds, glycosylation etc.) and protein folding.

In our experiments, MEF21 protein was successfully expressed (Fig. 3.38) but stocked in the cells. Western blot analysis (Fig. 3.38) of the expressed proteins in *Pichia* suggested the MEF21 still has the pre-pro peptide or just the pro peptide. In *Saccharomyces cerevisiae,* the pre-pro peptide is required for the translocation into the ER and will be cleaved in the ER. The pro-sequence, which is 60 amino acids long, is important for the translocation into the Golgi, where this part of the signal sequence will be also cleaved. The expressed MEF21 proteins seem to be stocked in the ER or in the Golgi apparatus without cleavage of pre-pro or pro peptides.

The MEF21 is a mitochondrially targeted protein necessary for the editing at the *cox3*-257 site (Takenaka et al. 2010). Since we expressed full-length MEF21, the remained mitochondrial targeting

sequence may affect the folding, subcellular localization of the protein and it can inhibit cleavage of the α-factor signal sequence.

The expressed MEF21 in Pichia is insoluble. It may be possible to optimize the conditions to get them in soluble form. If we could get soluble proteins, we can purify the proteins using the Myc tag antibody and use for further analyses.

Highly expressed and purified proteins with this approach can be applied for further analysis including crystallization or NMR (Nuclear Magnetic Resonance) spectroscopy to get more information about the protein structure and dynamics. High-level expression of proteins is possible using the *Pichia* expression system (Ahmad et al. 2014) therefore it would be worth to continue the investigations using this system with the MEF21 gene without mitochondrial target signal, or with other RNA editing factors.

5 Materials and methods

5.1 Materials

5.1.1 Machines

PCR thermal cyclers: Flex Cycler^{2} (Analytik Jena),VWR. Balances: SI-2002A (Denver Instrument), R200D Electronic Semi-Microbalance (Sartorius research), MP-3000 (Chyo Electronic Balance). PH electrode: BlueLine 14 pH (SI analytics). Benchtop centrifuges: Mikro 185 (Hettich), Mikro 200R (Hettich), Biofuge primo R (Thermo Scientific). High speed centrifuges: Avanti J-E (Beckman Coulter), J2-MC (Beckman). Vortex: Genie 2^{TM} (Zurich, Switzerland). Shakers: IKA® Rocher 2D digital rocking shaker, GFL 3020 orbital shaker, innova®43 incubator shaker (New Brunswick™). Incubators, tube thermostat, thermocycler: TCR 100 tube thermostat (ROTH), Microbiological incubators (Heraeus Instruments), Biomed 60 thermocycler. Magnetic stirrer: IKAMAG® RCT. Eppendorf mixer: 5432 Eppendorf mixer (Berlin GmbH), 5436 Eppendorf thermomixer (Germany). Electrophoresis Power Supplies: EPS 301 (Amersham pharmacia biotech), EPS 600 (Amersham pharmacia biotech). Heaters: Dri-Block® BD•2D (Techne), Dri-Block® BD•3 (Techne). Waving platform shaker: Polymax 1040 (Heidolph). Electroporator micropulser: BIO-RAD MicroPulser™. Clean Benches: Lamin Air HB 2448 (Heraeus Instruments, Unity Lab Services), Gelaire Flow Laboratories.

5.1.2 Kits

NucleoSpin Gel and PCR clean-up (Macherey-Nagel), NucleoSpin Plasmid (Macherey-Nagel), RNA Isolation (GE Healthcare), Invitrogen™™:Easyselect *Pichia* Expression Kit: For Expression of Recombinant Proteins Using pPICZ and pPICZα in *Pichia pastoris* , Invitrogen™™: pPICZα A,B and C *Pichia* expression vectors for selection on Zeocin™ and purification of secreted, recombinant proteins, In-Fusion® HD cloning system (Clontech Laboratories, Mountain View, USA).

5.1.3 Sequencing

Macrogen Europe (Amsterdam, Netherlands) and LGC Genomics GmbH (Berlin, Germany).

5.1.4 Chemicals

Duchefa Biochemie, ROTH, BD, Applichem GmbH, Sartorius, Promega, Fluka, VWR, Sigma.

5.1.5 Enzymes, dNTPs

Enzymes: Fermentas, Thermo Scientific, Promega, Clontech. Deoxynucelotide (dNTP) solutions: Genaxxon.

5.1.6 Oligonucleotides

Constructs for complementation assay				
Product		**Forward primer (5`->3`)**	**Reverse primer (3`->5`)**	
MEF28PLS-MEF25EE+	MEF28S2-MEF25E_F	GAAGCCAAACGCTCCTATATGGGTTGCTCTTCTTTCAG	At3g25060iFRstop	GTGCGGCCGCAAGCTTCTATTCGACTCCGGAACATAC
MEF28PLS-EE+_MEF11DYW	MEF28E-MEF11DYW_F	GAAGCCAAACGCTCCTGCTGCAGGGTACAAACCC	At4g14850iFRstop	GTGCGGCCGCAAGCTTGTTACCAATAATCCTTACAAGAACATATCC
MEF28PLS-EE+_CRR22DYW	MEF28E-CRR22DYW	TTGGGGAAGATAAGAGAAGCTGGTTATGTTCCTGAC	At1g11290iFRstop	GTGCGGCCGCAAGCTTGTCACCAGTAATCTCCGC
MEF28PLS_MEF11EE+DYW	MEF28S2_MEF11EDYW_F	AAGCCAAACGCTCCTGTTTGGGGTGCTCTTCAG	At4g14850iFRstop	GTGCGGCCGCAAGCTTGTTACCAATAATCCTTACAAGAACATATCC
MEF28PLS_CRR22EE+DYW	MEF28S2-CRR22EDYW	AAGCCAAACGCTCCTGTATATGGTGCCATGTTAGG	At1g11290iFRstop	GTGCGGCCGCAAGCTTGTCACCAGTAATCTCCGC
MEF28PLS_MEF11EE+_MEF28DYW	MEF28S2_MEF11EDYW_F	AAGCCAAACGCTCCTGTTTGGGGTGCTCTTCAG	MEF11ER	TTCCATTTCGTTCCTTAGCTT
	MEF11E_MEF28DYW_F	AGGAACGAAATGGAATTGATAGGGTATAAAGGGAAC	At5g06540iFRstop	GTGCGGCCGCAAGCTTTCACCAATAATCTCTGCAGGAA
Constructs for Y2H assay				
Product		**Forward primer (5`->3`)**	**Reverse primer (3`->5`)**	
At2g34370	At2g34370ATGiFF	CGAATTCTGTACAGGCATGGTGAGACTAGTTTGCAGC	At2g34370stopiFR	GTGCGGCCGCAAGCTTACCAGTAATCTTTGCAGGAGC
At1g29710	At1g29710ATGiFF	CGAATTCTGTACAGGCATGGTGCGATTGTGGTGTG	At1g29710stopiFR	GTGCGGCCGCAAGCTTACCAAAGATTGTTGCACCGG
At2g15690	At2g15690ATGiFF	CGAATTCTGTACAGGCATGTCTTCTCTAATGGCCATTC	At2g15690stopiFR	GTGCGGCCGCAAGCTTACCAGTAATCCCCGCAAGAAC
MEF1_PLS	MEF1iFFATG	CGAATTCTGTACAGGCATGGCTCTGAATTCTTCAGCT	At5g52630iFRS2stop	GTGCGGCCGCAAGCTTAAGATTCCGTGGGATCAATCG
MEF8_PLS	At2g25580ATGiFF (2)	CGAATTCTGTACAGGCATGTATACGAAGTTATCATTGTTCCC	At2g25580S2iFR	GTGCGGCCGCAAGCTTATCAACATTTGGCTCCATCG
MEF9_PLS	At1g62260ATGiFF	CGAATTCTGTACAGGCATGATCCGATCACGAAGTGTC	At1g62260S2iFR	GTGCGGCCGCAAGCTTCGTTTTGTCTGGCTCAAAAG
MEF11_PLS	At4g14850iFFATG	CGAATTCTGTACAGGCATGAGCCTCCTCTCCGCGGAC	At4g14850S2iFR	GTGCGGCCGCAAGCTTCGAAATTGTCGGTTGAATTGG
MEF18_PLS	At5g19020ATGinFusionF	CGAATTCTGTACAGGCATGATCAAATTGATTCGTTTTTTCCG	At5g19020S2iFR	GTGCGGCCGCAAGCTTCTTTACAGGCATCTTCTTGATCA
MEF19_PLS	At3g05240ATGiFF	CGAATTCTGTACAGGCATGATGAAGAAGCACTACAAACC	At3g05240S2iFR	GTGCGGCCGCAAGCTTTTAATTAACATTGGGTTTTACCGGCATAG
MEF20_PLS	At3g18970iFFATG2	CGAATTCTGTACAGGCATGAGCTCTGTTTTTCCAGGA	At3g18970S2iFR	GTGCGGCCGCAAGCTTCTTTATAGGCATAGCTAATATAAACTGATAA
MEF10_PLS-N	At3g11460ATGiFF	CGAATTCTGTACAGGCATGATCGTCGTTACAAGCTTTG	At3g11460PPR6iFR	GTGCGGCCGCAAGCTTAGGACAAACACCCGATGA
MEF13_PLS-N	At3g02330ATGiFF	CGAATTCTGTACAGGCATGGCGGAGAGTCTCAGATTATTG	At3g02330PPR7iFR	GTGCGGCCGCAAGCTTTCGTTGAAAGAATCGACTATG
MEF19_PLS-N	At3g05240iFFnoTS	CGAATTCTGTACAGGCAACGTCATTCCTCTAAGCCG	At3g05240P2iFR	GTGCGGCCGCAAGCTTAGGTGTAGCATTACCTTTCTCT
MEF21-PLS-N	At2g20540iFFnoTS	CGAATTCTGTACAGGCAGCAGCTTCATGGTCACTAA	At2g20540P2iFR	GTGCGGCCGCAAGCTTTTAAGGCTTCACCTTCGCTC
MEF26_PLS-N	InFuAt3g03580 up (MEF28ATGiFF)	CGAATTCTGTACAGGCATGCAGACTAGAGTGTCATCACC	At3g03580PPR7iFR	GTGCGGCCGCAAGCTTCCCAGTTCCCATACTACTG
MEF10_PLS-C	At3g11460PPR5iFF	CGAATTCTGTACAGGCGATCCTTTTACATTGGTTAGTG	At3g11460S2iFR	GTGCGGCCGCAAGCTTAGCACCGTCCGGTTCAACA
MEF13_PLS-C	At3g02330PPR6iFF	CGAATTCTGTACAGGCATGTGTGTTTCATGGAATTC	At3g02330S2iFR	GTGCGGCCGCAAGCTTTTCAAATGGCATCTCTCG
MEF19_PLS-C	At3g05240P1098iFF	CGAATTCTGTACAGGCGATAAAGTAACCTTTTTGAGTGTTATCAG	At3g05240S2iFR	GTGCGGCCGCAAGCTTTTAATTAACATTGGGTTTTACCGGCATAG
MEF21_PLS-C	At2g20540L718iFF	CGAATTCTGTACAGGCGATGAGATCAGTCTCATTTCTGTT	At2g20540S2RiFR	GTGCGGCCGCAAGCTTTTACTTCGAGTCGGGTTTCATCG
MEF26_PLS-C	At3g03580PPR6iFF	CGAATTCTGTACAGGCGATACGGTAACATGGAACACG	At3g03580S2iFR	GTGCGGCCGCAAGCTTTTTTATCGGCATTGCTTGG
MEF11_PLS-EE+	At4g14850iFFnoTS	CGAATTCTGTACAGGCGCAAGAATCGTCAAAACCCTC	At4g14850iFRstopE+	GTGCGGCCGCAAGCTTATTCCATTTCGTTCCTTAGCTTG
MEF1_EE+DYW	At5g52630iFFE	CGAATTCTGTACAGGCGTATGGGGAGCTTTATTAACGAG	At5g52630iFRstop	GTGCGGCCGCAAGCTTACCAATAGTCATTACAAGAACACTTG
MEF10_EE+DYW	At3g11460EiFF	CGAATTCTGTACAGGCGTTTGGGGTGCTCTATTGG	At3g11460stopiFR	GTGCGGCCGCAAGCTTGTCACCAGTAATCTTTGCACGA
MEF26_EE+DYW	At3g03580EiFF	CGAATTCTGTACAGGCATATGGGCATCTGTGCTAAG	InFuAt3g03580 rev (MEF26 stopiFR)	GTGCGGCCGCAAGCTTGTCACCACCGATCTTTGCAG
MEF8_EE+	At2g25580EiFF	CGAATTCTGTACAGGCGTTTGGGAAACATTGATGAATCTTTCTC	At2g25580 EnostopiFR	GTGCGGCCGCAAGCTTCACCATATGCATCTTTAAATTCCTC
MEF21_EE+	At2g20540EiFF	CGAATTCTGTACAGGCATATGGGGTTCATTGTTGAGC	At2g20540stopiFR2	GTGCGGCCGCAAGCTTGCTAAACTATGCCTATAAATGCAAG
MEF13_EE+	At3g02330EiFF	CGAATTCTGTACAGGCATATGGAGAACCCTGTTAGGAG	At3g02330iFR	GTGCGGCCGCAAGCTTTTAGCAGTAACACCATTGGTCC
MEF10_E	At3g11460EiFF	CGAATTCTGTACAGGCGTTTGGGGTGCTCTATTGG	At3g11460EiFR	GTGCGGCCGCAAGCTTCTTGTGTTCCACATAACTATACCC
MEF13_E	At3g02330EiFF	CGAATTCTGTACAGGCATATGGAGAACCCTGTTAGGAG	At3g02330EiFR	GTGCGGCCGCAAGCTTTTTTAGCTCAACCCAACTGC
MEF19_E	At3g05240EiFF	CGAATTCTGTACAGGCATATGGGGTGCATTGTTAAATGG	At3g05240stopiFR	GTGCGGCCGCAAGCTTGAAGTAGGACCCTGAATTAGAAC
MEF21_E	At2g20540EiFF	CGAATTCTGTACAGGCATATGGGGTTCATTGTTGAGC	At2g20540EiFR	GTGCGGCCGCAAGCTTGTTCACTTCTATCAAACTACCACC
MEF26_E	At3g03580EiFF	CGAATTCTGTACAGGCATATGGGCATCTGTGCTAAG	At3g03580EiFR	GTGCGGCCGCAAGCTTACCGACTTCGATCCAGCT
MEF9_E	At1g62260EiFF	CGAATTCTGTACAGGCGTATGGGGTGCATTATTGGATG	At1g62260stopiFR	GTGCGGCCGCAAGCTTGTTACGTTGAGGAGTCAACCC
MEF8_EN	MEF8_EiFF	CGAATTCTGTACAGGCGTTTGGGAAACATTGATGAATCTTTCTC	At2g25580iFR455T	GTGCGGCCGCAAGCTTTGTAGGATCTAAAAATTCAACAACCTCA
MEF13_EN	At3g02330EiFF	CGAATTCTGTACAGGCATATGGAGAACCCTGTTAGGAG	At3g02330iFR805Q	GTGCGGCCGCAAGCTTTTGTGGATCCAATCTCAACAG
MEF19_EN	At3g05240EiFF	CGAATTCTGTACAGGCATATGGGGTGCATTGTTAAATGG	At3g05240iFR519L	GTGCGGCCGCAAGCTTAAGTTCTTCTGGCTCAGCAAC
MEF21_EN-N	At2g20540EiFF	CGAATTCTGTACAGGCATATGGGGTTCATTGTTGAGC	MEF21delE-1-1iFR	GTGCGGCCGCAAGCTTATTGCCTGGTGTTCTGCA
MEF13_EC	At3g02330iFF806D	CGAATTCTGTACAGGCGACTCTTCAGCTTATACACTTTTATC	At3g02330EiFR	GTGCGGCCGCAAGCTTTTTTAGCTCAACCCAACTGC
MEF19_EC	At3g05240iFF520G	CGAATTCTGTACAGGCGGTAGCGGAATCTATGTTCTTC	At3g05240stopiFR	GTGCGGCCGCAAGCTTGAAGTAGGACCCTGAATTAGAAC
MEF21_EC	MEF21delE-1-2fw	CGAATTCTGTACAGGCCTTGATGTTGCTCTGGTTGC	At2g20540iFR422E	GTGCGGCCGCAAGCTTCTCCGGCTCTAGCTCCAC
MEF8_ECE+	At2g25580iFF456R	CGAATTCTGTACAGGCCGGTTGAACAAACAATCTCG	At2g25580EiFRnostop	GTGCGGCCGCAAGCTTCACCATATGCATCTTTAAATTCCTC
MEF21_ECE+	At2g20540iFF443D	CGAATTCTGTACAGGCGATATGGGAAATTATGTTTTGCTTGC	At2g20540stopiFR2	GTGCGGCCGCAAGCTTGCTAAACTATGCCTATAAATGCAAG
MEF13_ECE+	At3g02330iFF806D	CGAATTCTGTACAGGCGACTCTTCAGCTTATACACTTTTATC	At3g02330nostopiFR	GTGCGGCCGCAACTTGTTAGCAGTAACACCATTGGTCC
MEF13_E+	At3g02330E+iFF	CGAATTCTGTACAGGCGATGAGCTTCATGTGTTTCTCG	At3g02330nostopiFR	GTGCGGCCGCAACTTGTTAGCAGTAACACCATTGGTCC
Construct for *Pichia pastoris* expression system				
Product		**Forward primer (5`->3`)**	**Reverse primer (3`->5`)**	
MEF21	At2g20540_pPICZαB-fw (MEF21F)	GAAGCTGCAGGAATTCACATGGCGTTTCATGGAATCAGAG	At2g20540_pPICZαB-nostop 2 (MEF21R)	CCGCCGCGGCTCGAGGAACTATGCCTATAAATGCAAGAGC

5.1.7 Vectors

For complementations: Gateway vectors: pENTRA41b, pMpGWB102. For transformation of *Pichia pastoris*: pPICZαB. For Y2H system: pGADT7, pGBKT7, pGBKT7-p53, pGBKT7-Lam, pGADT7T.

Vectors	System	Vector type	Vector size (kb)	Reference
pENTR41b : pENTR™1A vector (Invitrogen) with the MCS of pET41 (Merck Millipore Novagen)	Gateway recombination system	Entry vector	2.7	Gateway® pENTR™ Dual Selection Vectors: Invitrogen™
pMpGWB102	Gateway recombination system	Destination vector	12	Ishizaki K. *et al.*, 2015
pPICZαB	Pichia pastoris expression system	Expression vector for secretion	3.6	pPICZα A,B and C, Invitrogen™ User Manual
pGBKT7	Y2H assay	Bait vector	7.3	Clontech Laboratories, Inc., Protocol No. PT3248-5
pGADT7 AD	Y2H assay	Prey vector	7.9	Clontech Laboratories, Inc., Protocol No. PT3249-5
pGBKT7-53	Y2H assay	Positive control (bait)	8.3	Clontech Laboratories, Inc., Protocol No. PT4084-1
pGADT7-T	Y2H assay	control prey vector	10	Clontech Laboratories, Inc., Protocol No. PT4084-1
pGBKT7-Lam	Y2H assay	Negative control (bait)	7.9	Clontech Laboratories, Inc., Protocol No. PT3955-1

Vectors	Properties	Reference
pENTR41b : pENTR™1A vector (Invitrogen) with the MCS of pET41 (Merck Millipore Novagen)	Kan^r, *ccd*B, *att*L1, *att*L2,	Gateway® pENTR™ Dual Selection Vectors: Invitrogen™
pMpGWB102	Spc^r, HygB^r, *ccd*B, *att*R1, *att*R2, CaMV 35S promoter, LB, RB,	Ishizaki K. *et al.*, 2015
pPICZαB	α-factor signal sequence, C-terminal c-Myc epitope and 6xHis tag, Zeo^r	pPICZα A,B and C, Invitrogen™ User Manual
pGBKT7	*TRP1*, Kan^r, GAL4 DNA-BD, N-terminal c-Myc epitope tag	Clontech Laboratories, Inc., Protocol No. PT3248-5
pGADT7 AD	*LEU2*, Amp^r, SV40 NLS, GAL4 AD, N-terminal HA tag	Clontech Laboratories, Inc., Protocol No. PT3249-5
pGBKT7-53	*TRP1*, Kan^r, GAL4 DNA-BD, N-terminal c-Myc epitope tag, Murine p53 insert	Clontech Laboratories, Inc., Protocol No. PT4084-1
pGADT7-T	*LEU2*, Amp^r, SV40 NLS, GAL4 AD, N-terminal HA tag, SV40 large T-antigen	Clontech Laboratories, Inc., Protocol No. PT4084-1
pGBKT7-Lam	*TRP1*, Kan^r, GAL4 DNA-BD, N-terminal c-Myc epitope tag, Human Lamin C	Clontech Laboratories, Inc., Protocol No. PT3955-1

Abbreviations	
Km^r	Kanamycin resistance gene
Spc^r	Spectinomycin resistance gene
HygB^r	Hygromycin B resistance gene
Zeo^r	Zeocin™ resistance gene
Amp^r	Ampicillin resistance gene
TRP1	Tryptophan nutritional marker
LEU2	Leucine nutritional marker
GAL4 DNA-BD	GAL4 DNA binding domain (amino acids 1-147)
GAL4 AD	GAL4 activation domain (amino acids 768-881)
SV40 NLS	Simian virus 40 nuclear localization signal
*att*L1, L2	Recombination sites for the LR reaction with *att*R1, R2
*att*R1, R2	Recombination sites for the LR reaction with *att*L1, L2
LB, RB	Left border and right border sequences for *Agrobacterium*-mediated transformation

5.1.8 Organisms

5.1.8.1 Plant lines

We used the *Arabidopsis thaliana* Ecotype Columbia (Col-0) to analyze the RNA editing at the two *nad2* sites in wild type plants and the *mef28-1* (SAIL_77E03) T-DNA mutant line for complementation**.** The SAIL line seeds have been ordered from NASC (European Arabidopsis Stock Center). The conditions in the growth chambers were 16h light and 8h dark cycles, 21°C and 65% humidity.

5.1.8.2 Bacteria

We used *Escherichia coli* K12 strain (*rec*A, *end*A) and *Agrobacterium tumefaciens s*train GV2260 (Deblaere et al. 1985) containing the pGV2260 non-oncogenic Ti Plasmid with rifamipicin and carbenicillin resistance sequences.

5.1.8.3 Yeasts

For the Y2H assay we used *Saccharomyces cerevisiae* strain PJ69-4A (James et al. 1996) and for the expression analysis *Pichia pastoris* X33 strain with Mut^+ phenotype (Invitrogen™ cat.no. V195-20).

5.1.9 Used websites

- PSIPRED Protein Sequence Analysis Workbench is a tool to predict the secondary structure of proteins from the amino acid sequence:
  - http://bioinf.cs.ucl.ac.uk/psipred/
    - http://bioinf.cs.ucl.ac.uk/psipred/result/a6869cce-03cf-11e3-8bb1-00163e110593 (13.08.2013)
- The Arabidopsis Information Resource:
  - http://www.arabidopsis.org/ (Swarbreck et al. 2007)
- Uniprot Knowledgebase I used for analyzing the motifs of PPR proteins:
  - http://www.uniprot.org/ (Apweiler et al. 2004)
- Protein size calculator:
  - http://www.sciencegateway.org/tools/proteinmw.htm (12.10.2015)
- NetNGlyc and NetOGlyc servers to predict the N- and O-Glycosylation sites in proteins:
  - http://www.cbs.dtu.dk/services/NetNGlyc/ (26.04.2015)
  - http://www.cbs.dtu.dk/services/NetOGlyc/ (27.04.2015)
- Illustration for N-glycosylation and O-glycosylation in a specific pattern:
  - http://www.ionsource.com/Card/carbo/nolink.htm (22.06.2016)
- Illustration for hyperglycosylation:
  - https://tools.thermofisher.com/content/sfs/appendix/Expression/Post-Translational%20Modifications.pdf (22.06.2016)
- Predotar: to predict N-terminal target sequence:
  - https://urgi.versailles.inra.fr/predotar/predotar.html (08.05.2015)
- TargetP 1.1: to predict protein subcellular localization:
  - http://www.cbs.dtu.dk/services/TargetP/ (08.05.2015)
- $FLAGdb^{++}$: integrative plant genome database: (Samson et al. 2004)
  - http://urgv.evry.inra.fr/projects/FLAGdb++/HTML/index.shtml
- For redifining the PPR domains:
  - http://www.plantppr.com/ (24.06.2016)

5.1.10 Software's for DNA, RNA and protein sequence analysis

DNA Dynamo and MEGA 5.2.

5.2 Methods

5.2.1 Standard molecular biology techniques

For Polymerase Chain Reaction (PCR), Gel electrophoresis, cloning, bacterial transformation, etc. for everyone available protocols were used.

5.2.2 Y2H assay

Deletion constructs of Mitochondrial RNA editing factors (MEFs) were cloned into the bait (pGBKT7) vector using the In-Fusion® HD cloning system (Clontech Laboratories, Mountain View, USA). Full length of MEF proteins and Multiple Organellar RNA editing Factors (MORFs) were already cloned into the prey (pGADT7) vector (Takenaka et al. 2012; Zehrmann et al. 2015).

For cotransformation of PJ69-4A yeast cells, the LiAc/single-stranded carrier DNA/PEG method was used (Shigeyuki et al. 2010; Matchmaker™ Gold Yeast-Two-Hybrid System 3 & Libraries User Manual). By the pGBKT7-53 plasmid encoded murine p53 protein and by the pGADT7-T plasmid encoded SV40 large T antigen served as positive control and pGBKT7-Lam plasmid encoded human Lamin C and the pGADT7-T plasmid encoded SV40 large T antigen served as negative control. The bait protein was tested for possible autoactivation using the empty pGADT7 prey vector for the cotransformation (Matchmaker™ Gold Yeast-Two-Hybrid System 3 & Libraries User Manual).

For selection of the successfully contransformed vectors the SD-Leu-Trp double dropout medium, for selection for possible protein interactions the SD-Leu-Trp-Ade-His quadruple dropout medium was used and to suppress background of leaky *HIS* reporter genes 2,5mM 3-AT (3-amino-1,2,4-triazole) was added to the SD-Leu-Trp-Ade-His medium (Saghbini et al. 2001) Matchmaker™ Gold Yeast-Two-Hybrid System 3 & Libraries User Manual). For the drop assay 5 µl was dropped onto the mediums of the overnight liquid culture with a diluted cell density for OD_{600} 0,1 and 0,3 (Zehrmann et al. 2015).

5.2.3 *Pichia pastoris* expression system

After linearizing the pPICZαB vector by EcoRI and KpnI restriction enzymes for cloning by the In-Fusion cloning system the full length of the MEF21 (Mitochondrial RNA Editing Factor 21) sequence was used. For putative secreted expression the MEF21 protein was fused to the α-factor secretion signal sequence. The native termination codon of MEF21 sequence was changed to allow expressing the gene with C-terminal *myc* epitope and polyhistidine tags for potential detection of the fusion protein. To the LB (pH7.5) agar plates for the selection of successfully transformed bacterial cells 25µg/ml Zeocin™ was used. (Invitrogen™™: V195-20).

For the transformation of *Pichia pastoris* X33 (Mut^+) strain the electroporation method I choose. The electroporation was done by 8,55µg linearized pPICZαB vector + MEF21_FL insert construct using the MssI enzyme. For selection of transformed yeast cells 100µg/ml Zeocin™ was added into the YPDS plates (Invitrogen™™: V195-20). Putative multi-copy recombinants I selected using 500µg/ml, 1000 µg/ml and 2000 µg/ml Zeocin™. The methanol utilization phenotype with Minimal Methanol medium (MM) and Minimal Dextrose medium (MD) was confirmed. (Invitrogen™™: K1740-01). As control the X33 (Mut^+) host train was used.

To confirm the integration of the expression cassette into the *Pichia* genome the genomic DNA was isolated (Harju et al. 2004) and verified by PCR using vector specific and MEF21 gene specific primers. For small scale and large scale expression putative higher expressers and Mut^+ phenotype having recombinant strains has been selected. *Pichia pastoris* recombinant strains were growing in MGY (Minimal Glycerol) medium before induce expression in MM (Minimal Methanol) medium and were growing in BMGY (Buffered Glycerol-Complex) medium before methanol induction in BMMY (Buffered Methanol-Complex) medium (Invitrogen™™: K1740-01). For Western blot analysis Anti-*myc*-HRP and/or Anti-His-HRP Antibody was used.

5.2.4 Creating chimeric libraries

The inserts were cloned by the In-Fusion® HD cloning system (Clontech Laboratories, Mountain View, USA) into the vector-insert construct which has been linearized by Inverse PCR. MEF28ΔDYW constructs has been created by PCR and construct of MEF28ATG-S2-MEF19E by overlap extension PCR (Romer 2013). Shuttling of constructs from the pENTRA41b entry vector to the pMpGWB102 destination vector to creating expression clones was made with LR Clonase II enzyme mix in the LR reaction using the instructions of the Gateway™ Cloning system (Invitrogen™™).

5.2.5 Isolation of genomic DNA from Arabidopsis thaliana leaves

After crushing the leaves 100µl Edward solution (200 mM Tris-HCl (pH 7.5), 250 mM NaCl, 25 mM EDTA, 0.5% SDS) than gDNAs were precipitated using 200µl absolute ethanol. It was followed by a centrifugation step (20 min. 13.000 rpm) than samples had to let dry at room temperature for 30 minutes. The gDNA was dissolved in 100µl 1xTE buffer (pH 7.5). For the polymerase chain reaction 1µl gDNA was added to a final volume of 10µl. (Edwards et al. 1991).

5.2.6 Complementation assay of Arabidopsis thaliana T-DNA mutant lines

Agrobacterium cells were transformed with 1µg of vector-insert constructs by heat shock (Höfgen and Willmitzer 1988). Successfully transformed cells were able to grow on selection media using the correct antibiotics corresponding to the vector and the helper plasmid. The pMpGWB102 vector has spectinomycin and the helper plasmid has carbenicillin resistance. Therefore the pMpGWB102-chimeric clones containing *Agrobacterium* cells could grow on spectinomycine and carbenicillin selection media. The infiltration medium for *Agrobacterium*-mediated transformation consists of 5% Sucrose (wt/vol) and 0,05% Silwet L-77 (vol/vol) (Zhang et al. 2006). The flower buds of the appropriate T-DNA mutant lines were transformed using the floral dip method (Clough and Bent 1998).

Constructs for complementation analysis:

- The *mef28-1* (SAIL_77_E03) T-DNA insertion mutants has been complemented with:
  - pAD41-MEF28_FL (Härtel 2013b)
  - pAD41-MEF28ΔDYW (Härtel 2013b)
  - pMpGWB102-MEF28ATG-S2-MEF13E (Romer 2013)
  - pMpGWB102-MEF28ATG-S2-MEF19E (Romer 2013)
  - pMpGWB102-MEF28S2_MEF25EE+
  - pMpGWB102-MEF28EE+_CRR22DYW
  - pMpGWB102-MEF28S2_CRR22EDYW
  - pMpGWB102-MEF28EE+_MEF11DYW
  - pMpGWB102-MEF28S2_MEF11EDYW
  - pMpGWB102-MEF28S2_MEF11EE+_MEF28DYW

RNA editing analysis:

Before RNA isolation, seeds were stored for two weeks at 4C^{o}, than selected on spectinomycin containing Murashige and Skoog (MS) medium for recombinant constructs regarding to the pMpGWB102 vector. RNA was isolated using the RNA isolation Kit from GE Healthcare. Synthesis of cDNA was made using M-MLV Reverse Transcriptase (Promega). Determination of RNA editing levels was established by Zehrmann et al. (2009). For the sequencing primers has been used in the same orientation for the cDNA products derived from the complemented lines.

6 Literature

Ahmad M., Hirz M., Pichler H. and Schwab H. Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production. Appl. Microbiol. Biotechnol. 98: 5301–5317 (2014).

Apweiler R., Bairoch A., Wu C., Barker W., Boeckmann B., Ferro S., Gasteiger E., Huang H., Lopez R., Magrane M., Martin M.J., Natale D., O`Donovan C., Redaschi N. and Yeh L. UniProt: the Universal Protein knowledgebase. Nucleic Acids Research 32: 115-119 (2004).

Arenas-M. A., Takenaka M., Moreno S., Gómez I. and Jordana X. Contiguous RNA editing sites in the mitochondrial nad1 transcript of Arabidopsis thaliana are recognized by different proteins. FEBS Letters 587: 887-891 (2013).

Arenas-M. A., Zehrmann A., Moreno S., Takenaka M. and Jordana X. The pentatricopeptide repeat protein MEF26 participates in RNA editing in mitochondrial *cox3* and *nad4* transcripts. Mitochondrion 19: 126-134 (2014).

Barkan A., Rojas M., Fujii S., Yap A., Seng Chong Y., Bond Ch. and Small I. A combinatorial Amino Acid Code for RNA Recognition by Pentatricopeptide Repeat Proteins. PLOS Genetics 8: 1-8 (2012).

Benne R., Van Den Burg J., Brakenhoff J., Sloof P., Van Boom J. and Tromp M. Major transcript of the Frameshifted *coxII* Gene From Trypanosome Mitochondria Contains Four Nucleotides That Are Not Encoded in the DNA. Cell Press 46: 819-826 (1986).

Bentolila S., Heller W., Sun T., Babina A., Friso G., van Wijk K. and Hanson M. RIP1, a member of an Arabidopsis protein family, interacts with the protein RARE1 and broadly affects RNA editing. PNAS Early Edition: 1-9 (2012).

Bentolila S., Oh J., Hanson M. and Bukowski R. Comprehensive High-Resolution Analysis of the Role of an Arabidopsis Gene Family in RNA Editing. PLOS Genetics 9: 1-13 (2013).

Blanc V. and Davidson N. C-to-U RNA Editing: Mechanisms Leading to Genetic Diversity. JBC 278: 1395-1398 (2002).

Boussardon C., Salone V., Avon A., Berthomé R., Hammani K., Okuda K., Shikanai T., Small I. and Lurina C. Two Interacting Proteins Are Necessary for the Editing of the NdhD-1 Site in Arabidopsis Plastids. The Plant Cell 24: 3684-3694 (2012).

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

Brennicke A., Marchfelder A. and Binder S. RNA editing. FEMS Microbiology Reviews 23: 297-316 (1999).

Brennicke A. The Long and Winding Road to RNA Editing in Plant Mitochondria. IUBMB Life 61(12): 1105–1109 (2009).

Chateigner-Boutin A., Colas des Francs-Small C., Fujii S., Okuda K., Tanz S. and Small I. The E domains of pentatricopeptide repeat proteins from different organelles are not functionally equivalent for RNA editing. The Plant Journal 74: 935-945 (2013).

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

Clough S. and Bent A. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant Journal 16: 735-743 (1998).

Covello P. and Gray M. RNA editing in plant mitochondria. Nature 341: 662-666 (1989).

Daly R. and Hearn M. Expression of heterologous proteins in Pichia pastoris: a useful experimental tool in protein engineering and production. J. Mol. Recognit. 18: 119-138 (2005).

Deblaere R., Bytebier B., De Greve H., Deboeck F., Schell J., Van Montagu M. and Leemans J. Efficient octopine Ti plasmid-derived vectors for *Agrobacterium-mediated* gene transfer to plants. Nucleic Acids Research 13: 4777-4788 (1985).

Delannoy E., Stanley W., Bond C. and Small I. Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. Biochemical Society Transactions 35: 1643-1647 (2007).

Edwards K., Johnstone C. and Thompson C. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Research 19: 1349 (1991).

Glass F., Härtel B., Zehrmann A., Verbitskiy D. and Takenaka M. MEF13 Requires MORF3 and MORF8 for RNA Editing at Eight Targets in Mitochondrial mRNAs in *Arabidopsis thaliana*. Molecular Plant 8: 1466-1477 (2015).

Gray M. W. RNA Editing in Plant Mitochondria: 20 Years Later. IUBMB Life 61(12): 1101-1104 (2009).

Gray M. W. Evolutionary Origin of RNA Editing. Biochemistry 51: 5235-5242 (2012).

Gualberto J., Lamattina L., Bonnard G., Weil J. and Grienenberger J. RNA editing in wheat mitochondria results in the conservation of protein sequences. Nature 341: 660-662 (1989).

Harju S., Fedosyuk H. and Peterson K. Rapid isolation of yeast genomic DNA: Bust n' Grab. BMC Biotechnology 4**:** 8 (2004).

Härtel B., Zehrmann A., Verbitskiy D., van der Merwe J., Brennicke A. and Takenaka M. MEF10 is required for RNA editing at nad2-842 in mitochondria of *Arabidopsis thaliana* and interacts with MORF8. Plant Mol. Biol. 81: 337-346 (2013a).

Härtel B. MEFs und MORFs: Charakterisierung von Vertretern zweier Arten von RNA-Editing-*trans*-Faktoren. Dissertation, Universität Ulm (2013b).

Härtel B., Zehrmann A., Verbitskiy D. and Takenaka M. The longest mitochondrial RNA editing PPR protein MEF12 in Arabidopsis thaliana requires the full-length E-domain. RNA Biology 10: 1-6 (2013c).

Hayes M., Giang K., Berhane B. and Mulligan M. Identification of Two Pentatricopeptide Repeat Genes Required for RNA Editing and Zinc Binding by C-terminal Cytidine Deaminase-like Domains. JBC 288:36519-36529 (2013).

Hiesel R., Wissinger B., Schuster W. and Brennicke A. RNA Editing in Plant Mitochondria. Science 246: 1632 (1989).

Hirose T. and Sugiura M. Involvement of a site-specific *trans*-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast *in vitro* RNA editing system. The EMBO Journal 20: 1144-1152 (2001).

Hoch B., Maier R., Appel K., Igloi G. and Kössel H. Editing of a chloroplast mRNA by creation of an initiation codon. Nature 353: 178-180 (1991).

Höfgen R. and Willmitzer L. Storage of competent cells for Agrobacterium transformation. Nucleic Acids Research 16: 9877 (1988).

Ishizaki K., Nishihama R., Ueda M., Inoue K., Ishida S., Nishimura Y., Shikanai T. and Kohchi T. Development of Gateway Binary Vector Series with Four Different Selection Markers for the Liverwort *Marchantia polymorpha*. PLOS ONE DOI:10.1371/journal.pone.0138876 (2015).

James P., Halladay J. and Craig E. Genomic Libraries and a Host Strain Designed for Highly Efficient Two-Hybrid Selection in Yeast. Genetics 144: 1425-1436 (1996).

Kobayashi K., Kawabata M., Hisano K., Kazama T., Matsuoka K., Sugita M. and Nakamura T. Identification and characterization of the RNA binding surface of the pentatricopeptide repeat protein. Nucleic Acids Research doi: 10.1093/nar/gkr1084: 1-12 (2011).

Kotera E., Tasaka M. and Shikanai T. A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature 433: 326-30 (2005).

Leung S. and Koslowsky D. RNA editing in Trypanosoma brucei: characterization of gRNA U-tail interactions with partially edited mRNA substrates. Nucleic Acids Research 29: 703-709 (2001).

Lurin C., Andrés Ch., Aubourg S., Bellaoui M., Britton F., Bruyére C., Caboche M., Debast C., Gualberto J., Hoffmann B., Lecharny A., Le Ret M., Martin-Magniette M., Mireau H., Peeters N., Renou J., Szurek B., Taconnat L. and Small I. Genome-Wide Analysis of Arabidopsis Pentatricopeptide Repeat Proteins Reveals Their Essential Role in Organelle Biogenesis. The Plant Cell 16: 2089-2103 (2004).

MacGinnitie A., Anant S. and Davidson N. Mutagenesis of apobec-1, the Catalytic Subunit of the Mammalian Apolipoprotein B mRNA Editing Enzyme, Reveals Distinct Domains That Mediate Cytosine Nucleoside Deaminase, RNA Binding, and RNA Editing Activity. JBC 270: 14768-14775 (1995).

Manna S. An overview of pentatricopeptide repeat proteins and their applications. Biochimie 113: 93-99 (2015).

Nakamura T., Yagi Y. and Kobayashi K. Mechanistic Insight into Pentatricopeptide Repeat Proteins as Sequence-Specific RNA-Binding Proteins for Organellar RNAs in Plants. Plant Cell Physiol. 53(7): 1171-1179 (2012).

Okuda K., Nakamura T., Sugita M., Shimizu T. and Shikanai T. A Pentatricopeptide Repeat Protein Is a Site Recognition Factor in Chloroplast RNA Editing. JBC 281: 37661-37667 (2006).

Okuda K., Chateigner-Boutin A., Nakamura T., Delannoy E., Sugita M., Myouga F., Motohashi R., Shinozaki K., Small I. and Shikanai T. Pentatricopeptide Repeat Proteins with the DYW Motif Have Distinct Molecular Functions in RNA Editing and RNA Cleavage in Arabidopsis Chloroplasts. The Plant Cell 21: 146-156 (2009).

Okuda K., Hammani K., Tanz S., Peng L., Fukao Y., Myouga F., Motohashi R., Shinozaki K., Small I. and Shikanai T. The pentatricopeptide repeat protein OTP82 is required for RNA editing of plastid ndhB and ndhG transcripts. The Plant Journal 61: 339–349 (2010).

O`Toole N., Hattori M., Andres Ch., Iida K., Lurin C., Schmitz-Linneweber Ch., Sugita M. and Small I. On the expansion of the Pentatricopeptide Repeat Gene Family in Plants. Mol. Biol. Evol. 25: 1120-1128 (2008).

Rajasekhar V. and Mulligan M. RNA Editing in Plant Mitochondria: α-Phosphate Is Retained during C-to-U Conversion in mRNAs. The Plant Cell 5: 1843-1852 (1993).

Randau L., Stanley B., Kohlway A., Mechta S., Xiong Y. and Söll D. A Cytidine Deaminase Edits C to U in Transfer RNAs in Archaea. Science 324(5927): 657-659 (2009).

Romer D. Synthese von PPR-Protein-Konstrukten mit substituierten C-terminalen Domänen für Funktionsanalysen. Bachelorarbeit, Universität Ulm (2013).

Rosano G. and Ceccarelli E. Recombinant protein expression in Escherichia coli: advances and challenges. Frontiers in microbiology doi:10.3389/fmicb.2014.00172 (2014).

Ruwe H. and Schmitz-Linneweber C. Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? Nucleic Acids Research doi:10.1093/nar/gkr1138 (2011).

Saghbini M., Hoekstra D. and Gautsch J. Media Formulations for Various Two-Hybrid Systems. Methods in Molecular Biology 177: 15-39 (2001).

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

Samson F., Brunaud V., Duchene S., De Oliveira Y., Caboche M., Lecharny A and Auburg S. FLAGdb++: a database for the functional analysis of the Arabidopsis genome. Nucleic Acids Research 32: 347-350 (2004).

Schuster W., Ternes R., Knoop V., Hiesel R., Wissinger B. and Brennicke A. Distribution of RNA editing sites in *Oenothera* mitochondrial mRNAs and rRNAs. Current Genetics 20: 397-404 (1991).

Shi X., Hanson M. and Bentolila S. Two RNA recognition motif-containing proteins are plant mitochondrial editing factors. Nucleic Acids Research doi: 10.1093/nar/gkv245: 1-12 (2015).

Shi X., Germain A., Hanson MR. and 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 a).

Shi X., Bentolila S. and Hanson M. Organelle RNA recognition motif-containing (ORRM) proteins are plastid and mitochondrial editing factors in Arabidopsis. Plant Signaling & Behavior doi: 10.1080/15592324.2016.1167299 (2016 b).

Shikanai T. RNA editing in plant organelles: machinery, physiological function and evolution. Cell. Mol. Life Sci. 63: 698-708 (2006).

Simpson L., Thiemann O., Savill N. and Alfonzo J. Evolution of RNA editing in trypanosome mitochondria. PNAS 97: 6986–6993 (2000).

Small I. and Peeters N. The PPR motif – a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci. 25: 46-47 (2000).

Sun T., Germain A., Giloteaux L., Hammani K., Barkan A., Hanson M. and Bentolila S. An RNA recognition motif-containing protein is required for plastid RNA editing in Arabidopsis and maize. PNAS 110: 1169-1178 (2013).

Swarbreck D., Wilks C., Lamesch P., Berardini T., Garcia-Hernandez M., Foerster H., Li D., Meyer T., Muller T., Ploetz L., Radenbaugh A., Singh S., Swing V., Tissier C., Zhang P. and Huala E. The Arabidopsis Information Resource (TAIR): gene structure and function annotation. Nucleic Acids Research 36: 1009-1014 (2007).

Takenaka M., Verbitskiy D., van der Merwe J., Zehrmann A. and Brenncike A. The process of RNA editing in plant mitochondria. Mitochondrion 8: 35-46 (2008).

Takenaka M., Verbitskiy D., Zehrmann A. and Brennicke A. Reverse Genetic Screening Identifies Five E-class PPR Proteins Involved in RNA Editing in Mitochondria of *Arabidopsis thaliana*. JBC 285: 27122-27129 (2010).

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

Takenaka M., Zehrmann A., Brennicke A. and Graichen K. Improved Computational Target Site Prediction for Pentatricopeptide Repeat RNA Editing Factors. PLOS ONE 8: 1-12 (2013a).

Takenaka M., Zehrmann A., Verbitskiy D., Härtel B. and Brennicke A. RNA Editing in Plants and Its Evolution. Annu. Rev. Genet. 47: 335-352 (2013b).

Takenaka M. How Complex Are the Editosomes in Plant Organelles? Molecular Plant 7: 582-585 (2014a).

Takenaka M., Verbitskiy D., Zehrmann A., Härtel B., Bayer-Császár E., Glass F. and Brennicke A. RNA editing in plant mitochondria —Connecting RNA target sequences and acting proteins. Mitochondrion 19: 191-197 (2014b).

Tercé F., Milne R., Weech P., Davignon J. and Marcel Y. Apolipoprotein B-48 and B-100 Very Low Density Lipoproteins Comparison in Dysbetalipoproteinemia (Type III) and Familial Hypertriglyceridemia (Type IV). ATVB 5: 201-211 (1985).

Tillich M., Hardel S., Kupsch Ch., Armbruster U., Delannoy E., Gualberto J., Lehwark P., Leister D., Small I. and Schmitz-Linneweber C. Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. PNAS 106: 6002-6007 (2009).

Verbitskiy D., Zehrmann A., van der Merwe J., Brennicke A. and Takenaka M. The PPR protein encoded by the LOVASTATIN INSENSITIVE 1 gene is involved in RNA editing at three sites in mitochondria of *Arabidopsis thaliana*. The Plant Journal 61: 446-455 (2010).

Verbitskiy D., Zehrmann A., Härtel B., Brennicke A. and Takenaka M. Two related RNA editing proteins target the same sites in mitochondria of *Arabidopsis thaliana*. JBC 287: 38064-72 (2012).

Wagoner J., Sun T., Lin L. and Hanson M. Cytidine Deaminase Motifs within the DYW Domain of Two Pentatricopeptide Repeat-containing Proteins Are Required for Site-specific Chloroplast RNA Editing. JBC 290: 2957-2968 (2014).

Williams-Carrier R., Kroeger T. and Barkan A. Sequence-specific binding of a chloroplast pentatricopeptide repeat protein to its native group II intron ligand. RNA 14: 1930-1941 (2008).

Wolf J., Gerber A. and Keller W. tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. The EMBO Journal 21: 3841-3851 (2002).

Yagi Y., Tachikawa M., Noguchi H., Satoh S., Obokata J. and Nakamura T. Pentatricopeptide repeat proteins involved in plant organellar RNA editing. RNA Biology 10: 1-7 (2013).

Yu W. and Schuster W. Evidence for a Site-specific Cytidine Deamination Reaction Involved in C to U RNA Editing of Plant Mitochondria. JBC 270: 18227-18233 (1995).

Zehrmann A., Verbitsky D., van der Merwe J., Brennicke A. and Takenaka M. A DYW Domain–Containing Pentatricopeptide Repeat Protein Is Required for RNA Editing at Multiple Sites in Mitochondria of Arabidopsis thaliana. The Plant Cell 21: 558-567 (2009).

Zehrmann A., Verbitskiy D., Härtel B., Brennicke A. and Takenaka M. RNA editing competence of trans-factor MEF1 is modulated by ecotype-specific differences but requires the DYW domain. FEBS Letters 584: 4181-4186 (2010).

Zehrmann A., Verbitskiy D., Härtel B., Brennicke A. and Takenaka M. PPR proteins network as site-specific RNA editing factors in plant organelles. RNA Biology 8: 67-70 (2011).

Zehrmann A., Härtel B., Glass F., Bayer-Császár E., Obata T., Meyer E., Brenncike A. and Takenaka M. Selective Homo- and Heteromer Interactions between the Multiple Organellar RNA Editing Factor (MORF) Proteins in *Arabidopsis thaliana*. JBC 290: 6445-6456 (2015).

Zhang X., Henriques R., Lin S., Niu Q. and Chua N. Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nature Protocols 1: doi:10.1038/nprot.2006.97 (2006).

Zhu Q., Dugardeyn J., Zhang C., Takenaka M., Kühn K., Craddock C., Smalle J., Karampelias M., Denecke J., Peters J., Gerats T., Brennicke A., Eastmond P., Meyer E. and Van Der Straeten D. SLO2, a mitochondrial pentatricopeptide repeat protein affecting several RNA editing sites, is required for energy metabolism. The Plant Journal 71: 836-849 (2012).

7 List of figures

8 List of charts

9 List of abbreviations

A	adenine
ACF	apobec-1 complementation factor
AOX	alcohol oxidase
ADAR	adenosine deaminase acting on RNA
APOB-48	apolipoprotein B (48kDa)
APOB-100	apolipoprotein B (100KDa)
APOBEC-1	apolipoprotein B mRNA editing enzyme
A-to-I	adenosine-to-inosine
3-AT	3-amino-1,2,4-triazole
bp	base pair
BiFC	Bimolecular fluorescence complementation
C	cytosine
CaMV-35S	cauliflower mosaic virus promoter
ccb203/ccmFN2	cytochrome c biogenesis orf203 (encodes a protein of the mitochondrial membrane)
ccb256/ccmC	cytochrome c biogenesis orf256 (encodes a protein of the mitochondrial membrane)
ccb452/ccmFc	cytochrome c biogenesis orf452 (mitochondrial encoded)
ccb206/ccmB	cytochrome c biogenesis orf206 (mitochondrial encoded)
CDA	cytidine deaminase
cDNA	complementary DNA

CMP	cytidine 5´-monophosphate
CoIP	Co-Immunoprecipitation
Col	Columbia ecotype of *Arabidopsis thaliana*
coxII	encodes cytochrome c oxidase subunit 2 in mitochondria (in *Trypanosoma brucei)*
cox3	encodes cytochrome c oxidase subunit 3 (in the mitochondrial respiratory chain complex IV)
CP31	CHLOROPLAST RNA-BINDING PROTEIN 31 (31kDa)
CP31A	CHLOROPLAST RNA-BINDING PROTEIN 31A (31kDa)
CP31B	CHLOROPLAST RNA-BINDING PROTEIN 31B (31kDa)
cpRNP	Chloroplast ribonucleoprotein family
CRP1	chloroplast-targeted RNA-binding protein 1
CRR	CHLORORESPIRATORY REDUCTION
C24	C24 ecotype of *Arabidopsis thaliana*
C-to-U	cytidine-to-uridine
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide
dsRBD	double-stranded RNA-binding domain
dsRNA	double-stranded RNA
ELI1	EDITING LACKING INSERTIONAL MUTANT 1 (chloroplast located pentatricopeptide repeat-containing protein)
EMS	ethyl methanesulfonate
FL	full-length
FRET	fluorescence resonance energy transfer
G	guanine

GR domain	glycine rich domain
gRNA	guide RNA
^{3}H	tritium
HCF152	HIGH CHLOROPHYLL FLUORESCENCE 152 (pentatricopeptide repeat-containing protein, located in the chloroplast)
5-HT_{2C}	serotonin (5-hydroxytryptamine) receptor
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
kb	kilobase
kDa	kilodalton
MEF	Mitochondrial RNA editing factor
mRNA	messenger ribonucleic acid
nad1	encodes subunit 1 of mitochondrial NAD(P)H dehydrogenase (in the mitochondrial respiratory chain complex I)
nad2	encodes subunit 2 of mitochondrial NAD(P)H dehydrogenase (in the mitochondrial respiratory chain complex I)
nad4	encodes subunit 4 of mitochondrial NAD(P)H dehydrogenase (in the mitochondrial respiratory chain complex I)
nad5	encodes subunit 5 of mitochondrial NAD(P)H dehydrogenase (in the mitochondrial respiratory chain complex I)
nad6	encodes subunit 6 of mitochondrial NAD(P)H dehydrogenase (in the mitochondrial respiratory chain complex I)
nad7	encodes subunit 7 of mitochondrial NAD(P)H dehydrogenase (in the mitochondrial respiratory chain complex I)
ndhD	encodes subunit of the chloroplast NAD(P)H dehydrogenase complex (photosystem I)
matR	encodes intron maturase (located in mitochondrion)
MD	minimal dextrose (medium)

miRNA	microRNA
MM	minimal methanol (medium)
MORF	multiple sites organellar RNA-editing factors
MS/MS	tandem mass spectrometry
mttB	membrane targeting and translocation (plastid encoded)
Mut^+	methanol utilization plus
Mut^S	methanol utilization slow
ORRM	Organelle RNA Recognition Motif-containing protein
OTP82	ORGANELLE TRANSCRIPT PROCESSING 82 (chloroplast located pentatricopeptide repeat-containing protein)
^{32}P	phosphorus-32 radioactive isotope
PAT	perlipin, adipophilin, and TIP47
PCR	polymerase chain reaction
petB	encodes subunit of the cytochrome b6f complex in chloroplast
petL	encodes subunit of the cytochrome b6f complex in chloroplast
PGR3	PROTON GRADIENT REGULATION 3 (chloroplast located pentatricopeptide repeat-containing protein)
PPR	pentatricopeptide repeat
psbL	encodes a subunit of the PSII in chloroplast
psbZ	encodes a subunit of the PSII in chloroplast
RIP	RNA editing factor interacting protein
RNA	ribonucleic acid
rps4	encodes 30S ribosomal protein S4 (part of the 30S ribosomal subunit in the chloroplast)

RRM	RNA Recognition Motif
rRNA	ribosomal ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
SD-TLHA	synthetically defined medium lacking tryptophan, leucine, adenine and histidine
siRNA	small interfering RNA
SLO2	SLOW GROWTH 2 (pentatricopeptide repeat-containing protein, located in the mitochondria)
ssRNA	single-stranded RNA
T	thymine
tadA	tRNA arginine adenosine deaminase (A-to-I editing enzyme)
TPR	tetratricopeptide repeat
tRNA	transfer ribonucleic acid
T-DNA	transfer DNA
U	uracil
UMP	uridine 5´-monophosphate
UTR	untranslated region
U-to-C	uridine-to-cytidine
wt	wild type
YPDS	yeast extract peptone dextrose (media)
Y2H	Yeast two-hybrid

10 Appendix

Chart 10.1: The *mef28-1* mutant lines stably transformed with several different chimera genes. The effectiveness of RNA editing in the transformed lines are shown by color-intensity.

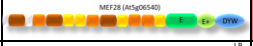
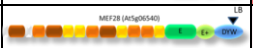
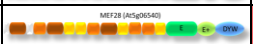
		***nad2*-89 (%)**			***nad2*-90 (%)**		
Col WT		94,1			88,1		
mef28-1		7,24			1,3		
mef28-1+MEF28		100			100		
mef28-1+MEF28ΔDYW (=MEF28EE+)		23	19,9	29,2	3	3,5	7,9
		22	14,3	19	2,2	0,1	2,1
mef28-1+MEF28PLS_MEF25EE+		26,4	2,6	12,7	7,7	0	2,2
mef28-1+MEF28PLS_MEF13EE+		5	6	52,5	0,8	1,5	0,4
mef28-1+MEF28(PLS-EE+)_MEF11DYW		90,8	99	4,86	5,2	10,4	1,66
mef28-1+MEF28(PLS-EE+)_CRR22DYW		65,5	49,61	34,5	9	0,5	0,3
		62,69			10,21		
mef28-1+MEF28PLS_MEF11(EE+DYW)		6,7	6	7,3	2,2	1,5	1,7
		6,2	10,8		1,7	1,5	
mef28-1+MEF28PLS_CRR22(EE+DYW)		33,4	25,7	4,7	41,6	27,8	11
		59,2			70,5		
mef28-1+MEF28PLS_MEF11EE+_MEF28DYW		13,2	11,2	6,66	2,5	2,3	2,69
mef28-1+MEF28PLS_MEF19E		98			99		
		9			3,64		

Chart 10.2: Summarizing all results in the Y2H assay between the MEF baits and MORF_FL proteins, except the autoactive constructs which are not included. The MORFs also involved in the same editing event affected by a MEF protein are indicated in orange squares. The n.d. indicates the not analyzed constructs.

	SD-TLHA									SD-TLHA + 3AT								
pGBKT7 \ pGADT7	MORF1	MORF2	MORF3	MORF4	MORF5	MORF6	MORF7	MORF8	MORF9	MORF1	MORF2	MORF3	MORF4	MORF5	MORF6	MORF7	MORF8	MORF9
MEF1_FL	+	-	-	-	-	-	-	(+)/-	-	-	-	-	-	-	-	-	-	-
MEF8_FL	+	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
MEF9_FL	+	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
MEF10_FL	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
MEF11_FL	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
MEF13_FL	+	+	-	-	-	-	-	+	+	-	(+)	-	-	-	-	-	-	-
MEF18_FL	+	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
MEF19_FL	+	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
MEF20_FL	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MEF21_FL	+	+	-	-	-	-	-	-	(+)/-	-	-	-	-	-	-	-	-	-
MEF26_FL	+	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
At2g34370 encoded protein	+	+	+	-	-	+	-	+	+	(+)/-	(+)/-	-	-	-	-	-	-	-
At1g29710 encoded protein	+	+	+	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
At2g15690 encoded protein	+	+	+	+	-	+	-	+	+	+	(+)/-	-	-	-	-	-	-	-
MEF1_PLS	+	+	+	-	-	(+)	-	+	+	-	-	-	-	-	-	-	-	-
MEF8_PLS	+	+	+	+	(+)	+	(+)	+	+	-	-	-	-	-	-	-	-	-
MEF9_PLS	+	+	-	-	-	-	-	+	(+)	-	-	-	-	-	-	-	-	-
MEF11_PLS	+	+	-	-	-	-	-	(+)/-	(+)/-	-	-	-	-	-	-	-	-	-
MEF18_PLS	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
MEF19_PLS	+	+	-	-	-	-	-	+	n.d.	-	-	-	-	-	-	-	-	n.d.
MEF20_PLS	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MEF10_PLS-N	+	+	-	-	-	-	-	(+)	(+)/-	-	-	-	-	-	-	-	-	-
MEF13_PLS-N	+	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
MEF19_PLS-N	(+)	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MEF21_PLS-N	+	+	-	n.d.	n.d.	n.d.	n.d.	-	n.d.	-	-	-	n.d.	n.d.	n.d.	n.d.	-	n.d.
MEF26_PLS-N	+	+	(+)	-	-	(+)/-	-	+	+	+	-	-	-	-	-	-	-	-
MEF10_PLS-C	+	+	(+)/-	-	-	+	-	(+)/-	+	-	-	-	-	-	-	-	-	-
MEF13_PLS-C	+	+	+	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-
MEF19_PLS-C	-	(+)	-	-	-	-	-	(+)	-	-	+	-	-	-	-	-	-	-
MEF21_PLS-C	(+)	+	-	n.d.	n.d.	n.d.	n.d.	(+)	n.d.	-	-	-	n.d.	n.d.	n.d.	n.d.	-	n.d.
MEF26_PLS-C	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
MEF11_PLS-EE+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MEF1_EE+DYW	+	(+)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MEF10_EE+DYW	+	+	-	-	-	-	-	(+)	-	-	-	-	-	-	-	-	-	-
MEF26_EE+DYW	+	+	+	-	-	-	+	+	+	-	-	-	-	-	-	(+)/-	-	-
MEF8EE+	+	+	+	+	(+)	+	+	+	+	-	(+)	-	-	-	-	-	(+)	-
MEF21EE+	+	+	+	n.d.	n.d.	n.d.	n.d.	+	n.d.	-	-	-	n.d.	n.d.	n.d.	n.d.	-	n.d.
MEF10_E	+	+	-	-	-	-	-	(+)	-	-	-	-	-	-	-	-	-	-
MEF13_E	n.d.	n.d.	+	-	-	-	-	+	-	n.d.	n.d.	-	-	-	-	-	-	-
MEF19_E	+	+	(+)	-	(+)/-	+	(+)	(+)	+	-	-	-	-	-	-	-	-	-
MEF21_E	+	+	(+)	n.d.	n.d.	n.d.	n.d.	+	n.d.	-	-	-	n.d.	n.d.	n.d.	n.d.	-	n.d.
MEF26_E	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MEF8_EN	+	+	+	+	+	+	+	+	+	+	+	(+)	-	+	+	(+)	+	+
MEF13_EN	n.d.	n.d.	+	-	-	-	-	+	-	n.d.	n.d.	-	-	-	-	-	-	-
MEF21_EN-N	+	+	-	n.d.	n.d.	n.d.	n.d.	+	n.d.	-	-	-	n.d.	n.d.	n.d.	n.d.	-	n.d.
MEF13_EC	n.d.	n.d.	(+)	-	-	-	-	+	-	n.d.	n.d.	-	-	-	-	-	-	-
MEF19_EC	+	(+)	-	-	-	-	-	(+)/-	(+)/-	-	-	-	-	-	-	-	-	-
MEF21_EC	+	+	-	n.d.	n.d.	n.d.	n.d.	+	n.d.	-	-	-	n.d.	n.d.	n.d.	n.d.	-	n.d.
MEF8_ECE+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
MEF21_ECE+	+	+	-	n.d.	n.d.	n.d.	n.d.	+	n.d.	-	-	-	n.d.	n.d.	n.d.	n.d.	-	n.d.

Chart 10.3: Sampling and methanol induction time points by large- and small-scale expressions. It is also shown how much methanol I added to the mediums. The MEF21 colony 4 inoculated in MM medium has not been tested.

	Culture	Sampling time points (h)	Induction: added methanol amounts	Added meth time points (h)
Larger-scale expression	MEF21-BMMY col 33	0; 4,5; 18; 24; 42,5; 47,5; 66;71	2ml (to 400ml)	25
				24
				24
	MEF21-MM col 33	0; 22,5; 26,5; 45; 50; 68,5;73,5	500 µl (to 100ml)	25,5
				24
				22
Small-scale expression	MEF21-BMMY col 33	0; 22; 44; 68; 72,5	220,7µl to 44,1ml	22
	MEF21-BMMY col 4	0; 22; 44; 68; 72,5	166,2µl to 33,2ml	22
	MEF21-MM col 33	0; 22; 44; 68; 72,5	125µl to 25ml	22
	MEF21-MM col 4	0; 22; 44; 68; 72,5	90µl to 18ml	22
	MEF21-BMMY col 33	0; 22; 44; 68; 72,5	185µl to 37ml	23,5
	MEF21-BMMY col 4	0; 22; 44; 68; 72,5	135µl to 27ml	23,5
	MEF21-MM col 33	0; 22; 44; 68; 72,5	105µl to 21ml	23,5
	MEF21-MM col 4	0; 22; 44; 68; 72,5	80µl to 16ml	23,5
	MEF21-BMMY col 33	0; 22; 44; 68; 72,5	170µl to34ml	24
	MEF21-BMMY col 4	0; 22; 44; 68; 72,5	120µl to 24ml	24
	MEF21-MM col 33	0; 22; 44; 68; 72,5	90µl to 18ml	24
	MEF21-MM col 4	0; 22; 44; 68; 72,5	65µl to 13ml	24

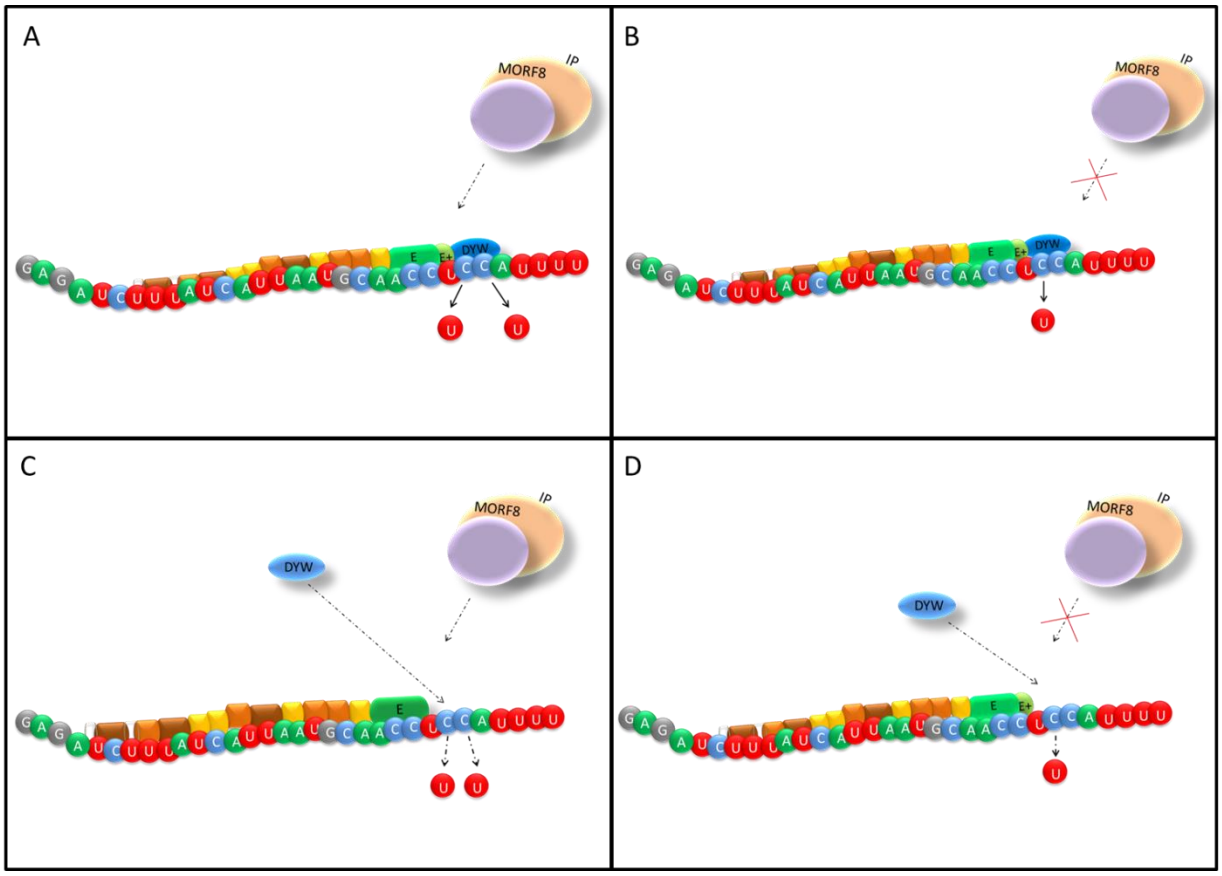


Figure 10.1: Models to explain the results got from the complementation assay for *nad2*-89 and *nad2*-90 RNA editing sites. (A) If the MORF8 and IP can bind, both sites are edited. (B) If the MORF8 and IP cannot bind, the first site is edited. (C) Both sites are edited because the MORF8 and IP can bind and the E domain is supported by a DYW domain of an other PPR protein. (D) The E domain is supported by a DYW domain of an other PPR protein but the MORF8 and IP cannot bind, therefore the first site is edited. IP = Interacting Partner.

11 Curriculum Vitae

Personal Details:

Name:	**Eszter Bayer-Császár**
Place of birth:	Szeged/Hungary
Date of birth	30.04.1985
Nationality:	Hungarian

Education:

01.09.2004 – 15.06.2009	Diploma with master degree University of Szeged (Hungary) Faculty of Science and Informatics *Biology major* **Specialization**: molecular biology **Principal subjects**: molecular biology, genetics, bioinformatics

Research experience:

16.05.2012 – **today**	**Ph.D. student** University of Ulm (Germany) Institute of Molecular Botany **Thesis topic**: Investigation of the potential function of the E and DYW domains of PPR proteins in *Arabidopsis thaliana*
26.10.2015 – 25.11.2015	**Research worker** University of Mar del Plata (Argentina) Institute for Biological Research **Thesis topic**: DIC microscopy
01.09.2009 – 28.02.2010	**Ph.D. student** University of Szeged (Hungary) Department of Genetics **Thesis topic**: Investigation of the membrane transport process in *Drosophila melanogaster*
01.09.2007 – 01.06.2009	**Diploma work** University of Szeged (Hungary) Department of Medical Biology, Faculty of Medicine **Thesis topic**: Analysis of the dominant negative female sterile *Himca* mutation in *Drosophila Melanogaster*
Languages:	Hungarian, English, German

12 List of Publications

12.1 Papers

Brehme. N., Bayer-Császár E., Glass F. and Takenaka M. The DYW Subgroup PPR Protein MEF35 Targets RNA Editing Sites in the Mitochondrial rpl16, nad4 and cob mRNAs in Arabidopsis thaliana. PLOS ONE DOI:10.1371/journal.pone.0140680 (2015).

Zehrmann A., Härtel B., Glass F., Bayer-Császár E., Obata T., Meyer E., Brennicke A. and Takenaka M. Selective Homo- and Heteromer Interactions between the Multiple Organellar RNA Editing Factor (MORF) Proteins in Arabidopsis thaliana. JBC 290:6445-6456 (2015).

Takenaka M., Verbitskiy D., Zehrmann A., Härtel B., Bayer-Császár E., Glass F. and Brennicke A. RNA editing in plant mitochondria - Connecting RNA target sequences and acting proteins. Mitochondrion 19:191-197 (2014).

12.2 Posters

„Approaches to identify novel proteins involved in RNA editing in plant mitochondria"
Daniil Verbitskiy, Anja Zehrmann, Härtel, Eszter Császár and Mizuki Takenaka
17th Annual Meeting of the International Society of Endocytobiology - German Section (ISE-G)
Germany, Berlin, 21.09.2015 – 24.09.2015

„Deciphering the interaction interface of MORF proteins"
Franziska Glass, Eszter Bayer-Császár, Anja Zehrmann, Mizuki Takenaka
XVI. Annual Meeting of the International Society of Endocytobiology - German Section (ISE-G)
Herzogenhorn, Germany, 21.07.2014 – 24.07.2014

„Refinement and Verification of the RNA Recognition Code for PPR RNA editing factors"
Mizuki Takenaka, Anja Zehrmann, Franziska Glass, Nadja Brehme, Eszter Bayer-Császár, Helena Stegherr, Daniil Verbitskiy, Knut Graichen and Axel Brennicke
XVI. Annual Meeting of the International Society of Endocytobiology - German Section (ISE-G)
8th Regio Plant Science Meeting Stuttgart-Tübingen-Ulm
Ulm, Germany, 20.02.2015
Herzogenhorn, Germany, 21.07.2014 – 24.07.2014

„Interactions between MEF and MORF RNA editing factors"
Anja Zehrmann, Härtel, Franziska Glass, Eszter Bayer-Császár, Daniil Verbitskiy, Axel Brennicke and Mizuki Takenaka
7th Regio Plant Science Meeting Stuttgart-Tübingen-Ulm
Stuttgart, Germany, 13.02.2014

„Identification of the binding sites between MEF and MORF proteins"
Eszter Császár, Daniil Verbitskiy, Härtel, Mizuki Takenaka and Anja Zehrmann

8th International Conference for Plant Mitochondrial Biology (ICPMB), Rosario, Argentina, 12.05.2013-16.05.2013

„Mapping the binding sites between MEF and MORF RNA editing factors"
Eszter Császár, Anja Zehrmann, Daniil Verbitskiy, Härtel Barbara and Mizuki Takenaka
Endosymbiosis from Prokaryotes to Eukaryotic Organelles - International Meeting, München, Germany, 10.10.2012 – 13.10.2012
XV. Annual Meeting of the German Section of the International Society for Endocytobiology, Martinsried, 08.10.2012 – 10.10.2012

„Interactions between various editing factors"
Anja Zehrmann, Daniil Verbitskiy, Härtel, Eszter Császár, Axel Brennicke, Mizuki Takenaka
Endosymbiosis from Prokaryotes to Eukaryotic Organelles - International Meeting München, Germany, 10.10.2012 – 13.10.2012
XV. Annual Meeting of the German Section of the International Society for Endocytobiology
Martinsried, 08.10.2012 – 10.10.2012

13 Acknowledgements

I want to thank to Prof. Dr. Axel Brennicke and to Dr. Mizuki Takenaka to have given me an opportunity to work in this group and for their supports.

Furthermore I want to thank to Dr. Mizuki Takenaka for his help in the solution of scientific questions and open-minded discussion of my ideas that kept my enthusiasm and creativity.

I want to thank to Prof. Dr. Anita Marchfelder to be the second consultant.

I would like to thank to Bianca Wolf and Dr. Daniil Verbitskiy for their friendship and kindness. Thanks to Bianca Wolf also for the positive atmosphere what she created around her and for her amazing job in the Y2H system.

Thanks to Dr. Anja Jörg for introducing me in the Y2H assay and to Dr. Daniil Verbitskiy and Sascha Haag for their kind help in protein detection.

I also want to thank to Angelika Müller and Dagmar Pruchner for their great work and to Evelyn Laible-Schmid for the nice conversations what we had.

I am grateful for the kind help and valuable discussion to all members in the group and also to everybody in the institute.

Finally I would like to thank to my family for their support and love.

14 Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig angefertigt habe und keine anderen Hilfsmittel und Quellen, als die angegebenen benutzt, sowie wörtlich und inhaltlich übernommene Stellen als solche kenntlich gemacht habe.

Ulm den_____ _____

Eszter Bayer-Császár