**Title-**

# Determining targeting specificity of nuclear-encoded organelle proteins with the self-assembling split-fluorescent protein toolkit

Mayank Sharma[1]*, Carola Kretschmer[2], Christina Lampe[1], Johannes Stuttmann[2] and Ralf Bernd Klösgen[1]

[1]Institute of Biology-Plant Physiology, Martin Luther University Halle-Wittenberg, Weinbergweg 10, 06120 Halle (Saale), Germany
[2]Institute of Biology-Genetics, Martin Luther University Halle-Wittenberg, Weinbergweg 10, 06120 Halle (Saale), Germany

1. Mayank Sharma (* Corresponding author) (mayank796@gmail.com)
   Phone: +49 - 345 - 55 26 204
2. Carola Kretschmer (kretschmer@genetik.uni-halle.de),
3. Christina Lampe (chrstnlampe@gmail.com),
4. Johannes Stuttmann (johannes.stuttmann@genetik.uni-halle.de)
5. Ralf Bernd Klosgen (klosgen@pflanzenphys.uni-halle.de)

Number of Tables- 1; Number of Figures- 9 (all color); Word Count- (5280)
Supplementary File- 1 (Supplementary Figures- 4, Supplementary Tables- 3)

***Running title:***
Determining targeting specifictity with *sa*split-GFP

***Keywords***
Self-assembling split-GFP, protein transport, mitochondria, chloroplast, dual targeting, *in vivo* imaging, Golden Gate cloning

**Highlight-**

Several mono-specific proteins showed dual targeting to plastids and mitochondria with the self-assembling split-GFP system. A *Golden Gate*-based vector toolkit was constructed to facilitate easy cloning and subsequent determination of protein targeting specificity.

**Abstract-**

A large number of nuclear-encoded proteins are targeted to the organelles of endosymbiotic origin, namely mitochondria and plastids. To determine the targeting specificity of these proteins, fluorescent protein tagging is a popular approach. However, ectopic expression of fluorescent protein fusions commonly results in considerable background signals and often suffers from the large size and robust folding of the reporter protein, which may perturb membrane transport. Among the alternative approaches that have been developed in recent years, the self-assembling split-fluorescent protein (*sa*split-FP) technology appears particularly promising to analyze protein targeting specificity *in vivo*. Here, we have improved this technology with respect to sensitivity and systematically evaluated its utilization to determine protein targeting to plastids and mitochondria. Furthermore, to facilitate high throughput screening of candidate proteins we have developed a *Golden Gate*-based vector toolkit, named PlaMiNGo (Plastid and/or Mitochondria targeted proteins N-terminally fused to GFP11 tags via *Golden Gate* cloning). As a result of these improvements, dual targeting could be detected for a number of proteins, which had earlier been characterized as being targeted to a single organelle only. These results were independently confirmed with a plant phenotype complementation approach thus demonstrating the sensitivity and robustness of the *sa*split-FP-based method to analyze the targeting specificity of nuclear-encoded proteins.

1. Introduction-

Many nuclear-encoded proteins are targeted across membranes to reach their final destination in the cell, i.e. a sub-cellular compartment or cell organelle. To determine the specificity of such targeting, fluorescent protein tagging (FP tagging), that is, the fusion of the candidate protein with a fluorescent reporter (e.g., Green Fluorescent Protein, GFP) and subsequent *in vivo* imaging by fluorescence microscopy, is a popular approach. However, this widely utilized method carries some inherent pitfalls (reviewed by Moore and Murphy 2009). First, the large size and intrinsic folding properties of the reporter protein might perturb membrane transport of the candidate protein (Marques et al., 2004). Second, considerable background fluorescence signals due to protein overexpression and hence saturation of the organelle transport machinery is sometimes observed (Sharma et al., 2018a). And finally, proteins targeted to an organelle at low amounts are difficult to visualize using this approach due to weak fluorescence signals originating from a limited number of protein molecules.

These constraints become particularly prominent when it comes to the analysis of proteins targeted to mitochondria and/or plastids (Tanz et al., 2013). These organelle-targeted proteins usually carry an N-terminal transport signal called 'transit peptide' to facilitate transport of the passenger protein across the membranes via organelle specific translocation machineries. Amongst a repertoire of proteins encoded in the plant cell nucleus, more than 3000 are transported into mitochondria and/or plastids (Van Wijk and Baginsky 2011; Rao et al., 2017). While most of these proteins are targeted to only one type of organelle, a group of proteins exists with dual targeting specificity, i.e., they carry an 'ambiguous' transit peptide capable of translocating the passenger protein into both mitochondria and plastids (reviewed in Sharma et al., 2018b). Despite having dual targeting properties some of these proteins are still preferentially targeted to one of the two organelles. In this case, determining targeting specificity using a conventional FP-tagging approach is particularly tedious and error-prone, as highly intensive fluorescence signals coming from one organelle can mask signals of low-intensity coming from the other one (Duchêne et al., 2005, Sharma et al., 2018b).

One possibility to circumvent these problems is to detect the fluorescence signals, coming from each organelle, in separate cells. This has become technically feasible by the invention of the self-assembling green fluorescent protein (*sa*split-GFP) technology, a popular approach developed recently to determine targeting specificity of proteins *in vivo* (Cabantous et al., 2005). Spontaneous self-assembly of split-GFP relies on two highly engineered fragments derived from the 'superfolder' GFP variant, sfGFP. The large fragment, GFP1-10 OPT (further referred to as GFP1-10), comprises ten N-terminal antiparallel β-sheets of GFP. The smaller fragment, GFP11 M3 (further referred to as GFP11), comprises only 16 amino acids and represents the C-terminal $11^{th}$ β-sheet of GFP. When brought into close proximity, the two fragments can assemble spontaneously to reconstitute the functional fluorophore without the need of an additional interacting partner (Figure 1a) (Cabantous et al., 2005). For subcellular protein localization studies, GFP1-10 is fused to a transport signal of known organelle specificity and analyzed together with a chimeric protein comprising the candidate protein and GFP11. Fluorescence complementation is achieved specifically and exclusively if the transport signal of the candidate mediates transport of GFP11 into the organelle housing GFP1-10. The fluorescence signal remains limited to the compartment containing GFP1-10, irrespective of whether the GFP11-fused candidate protein is targeted to further subcellular compartments or not (Figure 1b). Thus, the *sa*split-GFP technology allows selective *in vivo* imaging of a protein of interest in a respective compartment with enhanced signal-to-noise

ratio. Furthermore, the small GFP11 tag appears to have a lower propensity for interfering with membrane transport in comparison to the full length fluorescent proteins used for direct FP-tagging. Smaller size tags were previously shown to be advantageous over large FP-tags for live cell imaging (Andresen et al., 2004; Giepmans et al., 2006).

The *sa*split-GFP system has already been applied to a wide range of organisms and adapted to elucidate a variety of cellular functions or processes including, for example, *in vivo* protein solubility, sub-cellular localization of pathogen effectors, endogenous protein labeling for *in vivo* imaging, protein-protein interaction studies, and membrane protein topology determination (Cabantous and Waldo 2006; Van Engelberg and Palmer, 2010; Machettira et al., 2011; Cabantous et al., 2013; Kamiyama et al., 2016; Henry et al., 2017). The principal suitability of the *sa*split-GFP system for *in vivo* imaging of protein targeting in plant cells has also been demonstrated recently (Park et al., 2017). In this case, transgenic *Arabidopsis thaliana* lines expressing the organelle-targeted GFP1-10 receptor were transiently transformed with constructs expressing a candidate protein fused to GFP11 tag. However, the requirement of transgenic 'receptor' plant lines and low transient transformation efficiency hampers the utilization of this approach for high throughput protein targeting studies. Along with a limited yield of transformed cells, the relatively low brightness of the *sa*split-GFP prevented the visualization of proteins targeted in low amounts into the organelles.

Therefore, we have here systematically evaluated and optimized the *sa*split-GFP system for analysis of protein targeting specificity in plant cells and assessed the effect of multimerization of the GFP11 tag on the intensity of the fluorescence signals inside mitochondria and plastids. A *Golden Gate*-based vector toolkit named PlaMiNGo (Analysis of Plastid and/or Mitochondrial targeted proteins N- terminally fused to GFP11 tags via *Golden Gate* cloning) was developed to facilitate high-throughput analyses of candidate proteins with high transformation efficiency and enhanced signal-to-noise ratio. With this approach, dual targeting to mitochondria and plastids could be detected for several proteins that were previously characterized as being targeted to a single organelle only, which demonstrates its improved sensitivity. Importantly, plastid targeting of two of these proteins was independently confirmed using a phenotype complementation-based approach in stable transgenic *Arabidopsis* plants, which proves the *in vivo* relevance of results obtained with the PlaMiNGo system.

## 2. Results-

### *2.1 The* sa*split-GFP system to determine protein targeting specificity*

In order to study protein targeting into plastids, the transit peptide of a chloroplast protein, ferredoxin-NADP$^+$-oxidoreductase of spinach ($FNR_{1-55}$; Zhang et al., 2001), was fused to the GFP1-10 receptor (the large fragment of the *sa*split-GFP system) to facilitate its localization into plastids. For mitochondrial localization of GFP1-10, the N-terminal 100 amino acid residues comprising the presequence of the mitochondrial Rieske Fe/S protein of potato ($mtRi_{1-100}$; Emmermann et al., 1994) were used. Both transport signals had previously been characterized for their targeting specificity to a single organelle with *in vivo* and *in vitro* approaches (Rödiger et al., 2011). For the initial experiments evaluating the suitability of the *sa*split-GFP system for our purposes, each transport signal was likewise combined with the $GFP11_{x7}$ tag (small fragment of *sa*split-GFP system). A seven-fold repeat of this GFP11 tag

($GFP11_{x7}$; separated via a 5 amino acid linker) was used, as such multiple GFP11 tags had been reported to intensify the fluorescence signals in mammalian cells (Kamiyama et al., 2016). When these constructs were co-expressed with the $FNR_{1-55}$/GFP1-10 and $mtRi_{1-100}$/GFP1-10 gene constructs, fluorescence signals were exclusively obtained in those instances in which the same transport signal was present in both chimeras, i.e. in plastids after co-expression of $FNR_{1-55}$/GFP1-10 and $FNR_{1-55}$/$GFP11_{x7}$ and in mitochondria after co-expression of $mtRi_{1-100}$/GFP1-10 and $mtRi_{1-100}$/$GFP11_{x7}$ (Figure 2). When infiltrated alone, neither of theses constructs generated any detectable fluorescence signal (Suppl. Figure 1). This demonstrates the suitability and specificity of the *sa*split-GFP system for the analysis of protein targeting.

Next, we wanted to evaluate the performance of this system by determining the protein targeting behavior of several dually targeted proteins. For this purpose, we have selected three previously characterized nuclear-encoded organelle proteins from *Arabidopsis thaliana* with proven dual targeting characteristics, namely TyrRS (Tyrosine-tRNA synthetase; At3g02660), GrpE (co-chaperone GrpE1; At5g55200), and PDF (peptide deformylase 1B; At5g14660) (Berglund et al., 2009; Baudisch et al., 2014). Earlier experiments employing eYFP (enhanced Yellow Fluorescent Protein) fusions had shown that all three proteins are targeted to both endosymbiotic organelles, either in comparable amounts ($TyrRS_{1-91}$/eYFP) or preferentially to either mitochondria ($GrpE_{1-100}$/eYFP) or chloroplasts ($PDF_{1-100}$/eYFP) (Sharma et al., 2018a and Figure 3). The respective N-terminal amino acid sequences carrying the organelle transport signals of the candidate proteins were fused to $GFP11_{x7}$ tags and analyzed in our system. All three candidates ($TyrRS_{1-91}$/$GFP11_{x7}$, $GrpE_{1-100}$/$GFP11_{x7}$ and $PDF_{1-100}$/$GFP11_{x7}$) showed targeting to both plastids and mitochondria when co-transformed with the respective organelle-targeted receptors (Figure 3). Even plastid targeting of $GrpE_{1-100}$/$GFP11_{x7}$ was clearly visible, which is remarkable considering the vague fluorescence signals obtained in this organelle with the 'standard' fluorescent reporter fusion (Figure 3b). Thus, separation of the fluorescence signals for the two organelles into different cells proved to be advantageous to determine the low plastid targeting properties of GrpE.

### 2.2 *Multimerization of the GFP11 tag leads to fluorescence signal enhancement in plastids but not in mitochondria*

In the initial experiments described above, we have used $GFP11_{x7}$, the seven-fold repeat of the GFP11 tag. However, the requirement or benefits of such multiple GFP11 tags to enhance fluorescence signals in mitochondria and plastids of plant cells were not systematically assessed. Hence, we have compared the fluorescence signal intensity obtained in the two organelles with either seven ($GFP11_{x7}$), three ($GFP11_{x3}$) or a single repeat ($GFP11_{x1}$) of the GFP11 tag when fused to the dually targeted $TyrRS_{1-91}$ peptide as protein transport signal. It turned out that the use of a single GFP11 tag yields only very faint signals in plastids while the $GFP11_{x3}$ and $GFP11_{x7}$ tags significantly enhance the fluorescence signals (Figure 4a) supporting the assumption that multiple GFP11 repeats can boost fluorescence signal intensity in this organelle. In contrast, for mitochondria no such correlation of number of GFP11 tag repeats and fluorescence signal intensity was found. Instead, the signal intensities were largely similar for all three constructs (Figure 4b). However, since the fluorescence signals obtained with the single GFP11 tag in mitochondria were brighter than in plastids, they are usually sufficient for proper visualization of organelle.

### 2.3 *Construction of the PlaMiNGo toolkit*

To facilitate easy combination of the *sa*split-GFP technology and fluorescence signal enhancement for high-throughput screening of protein targeting specificity, we have developed a set of *Golden Gate*-based vectors. Destined to analyze the targeting specificity of candidate proteins to plastids and mitochondria, these vectors facilitate easy cloning of the candidate proteins upstream of single or multiple GFP11 tags. In this PlaMiNGo toolkit, we have utilized highly efficient gene regulatory elements, namely a 'long' 35S promoter (Engler et al., 2014) and *ocs* or *rbcs E9* transcriptional terminators (De Greve et al., 1982; Coruzzi et al., 1984), to control the expression of the chimeric genes, (Figure 5a and Suppl. Figure 3). Moreover, to avoid the requirement for performing co-transformation using two plasmids, we have utilized a single T-DNA expression system (Grefen and Blatt, 2012; Hecker et al., 2015) comprising two gene expression cassettes. This has the advantage that each transformed cell expresses both chimeras simultaneously. Consequently, the final vectors contain expression cassettes, for GFP1-10 gene chimeras to be targeted to either plastid or mitochondria by fusion with $mtRi_{1-100}$ or $FNR_{1-55}$ and another for one, three or seven times repeat of GFP11 tags. The latter expression cassette furthermore carries a *ccd*B cassette upstream of the GFP11 tag, which can be replaced by the candidate gene in a single *Golden Gate* reaction step allowing for background-free selection of positive clones. As a result, six vectors, three destined to analyze potential plastid targeting and three for mitochondria targeting analysis, were generated (Figure 5a and Suppl. Figure 3).
To evaluate the functionality of these vectors, the dual targeting transport signal of TyrRS (N terminal 1-91 amino acids) was cloned upstream of the GFP11 tags of all six vectors and analyzed via *Agrobacterium*-mediated transient transformation of *Nicotiana benthamiana* leaves. In all instances, tremendously improved signal intensities were observed for both organelles (Figure 5b and Suppl. Figure 2). Now, even in plastids a single copy of the GFP11 tag was sufficient for reliable detection of the reconstituted GFP fluorescence, although considerable signal enhancement could still be observed with $GFP11_{x3}$ and $GFP11_{x7}$ tags. This suggests that the $GFP11_{x7}$ tag, when co-expressed with the plastid targeted GFP1-10 receptor, should allow for detection even of minute amounts of plastid-localized proteins. Likewise, upon imaging of mitochondria the fluorescence signals obtained with the single GFP11 tag were more than 3-fold stronger than the signals obtained in the previous experiments using separate vectors. Still, as observed earlier, no further improvement of signal intensities by multimerization of the GFP11 tag could be obtained in mitochondria (Suppl. Figure 2b). Instead, artificial protein aggregates were observed in some cells expressing constructs with multiple GFP11 tags ($GFP_{11x3}$ and $GFP_{11x7}$) (Suppl. Video 1).

### *2.4 Multicolor imaging of dual protein targeting to two organelles*

A further objective of the study was to establish simultaneous multicolor imaging with the *sa*split-FP system. For this purpose, we have modified the GFP1-10 receptor to generate a yellow shifted variant (YFP1-10) using a single amino acid substitution (T203Y) as reported earlier (Kamiyama et al., 2016). To test if this variant can assemble with GFP11 to generate a functional fluorophore inside the organelles, the plastid ($FNR_{1-55}$) or mitochondria ($mtRi_{1-100}$) transport signals were separately fused to the N-terminus of YFP1-10. These fusions were co-infiltrated with a gene construct encoding the dual targeting transport signal $TyrRS_{1-91}$ fused to either $GFP11_{x1}$ or $GFP11_{x7}$. When imaged with an YFP-specific filter-set, fluorescence signals were solely obtained in mitochondria (with $mtRi_{1-100}$/YFP1-10 and $TyrRS_{1-}$

$_{91}$/GFP11$_{x1}$) or in plastids (with FNR$_{1-55}$/YFP1-10 and TyrRS$_{1-91}$/GFP11$_{x7}$), demonstrating that the YFP1-10 fragment can indeed assemble with GFP11 in both organelles (Figure 6a, b).
Next, we have tested if multicolor imaging, i.e. the simultaneous labeling of plastids and mitochondria with YFP1-10 and GFP1-10, respectively, within the same cell is possible. For this purpose, the FNR$_{1-55}$/YFP1-10 fusion was co-infiltrated with a vector comprising mtRi$_{1-100}$/GFP1-10 and TyrRS$_{1-91}$/GFP11$_{x1}$. This should result in transformed cells co-expressing the dually targeted GFP11 (via fusion with TyrRS$_{1-91}$) and two different receptors targeted to two different organelles, namely YFP1-10 to plastids (FNR$_{1-55}$/YFP1-10) and GFP1-10 to mitochondria (mtRi$_{1-100}$/GFP1-10). Indeed, the resulting transformed cells emitted fluorescence signals of different spectra in the two organelles due to reassembly of the GFP11 tag with both, the GFP1-10 and YFP1-10 receptors within mitochondria and plastids, respectively (Figure 6c). However, in all transformed cells a certain degree of "bleed through" of signals, i.e., the appearance of YFP fluorescence signals in GFP channel and *vice versa,* could be detected. The adjustment of the filter-sets to avoid such “bleed through” inevitably led to significant reduction of the fluorescence signal intensity. In summary, the YFP1-10 derivative of *sa*split-GFP system is not yet perfect for multicolor imaging but represents a promising basis for development of such tools.

### 2.5 Analysis *of protein targeting specificity with PlaMiNGo*

Finally, we have examined the suitability of the PlaMiNGo toolkit using eight candidates with presumed targeting specificity either to plastids, mitochondria, or to both organelles (Table 1). Three of these proteins, namely Gtred (Monothiol glutaredoxin-S15), GCS (Glycine cleavage system H protein 1) and GAPDH (Glyceraldehyde-3-phosphate dehydrogenase B) had earlier been reported to be dually targeted (Baudisch et al., 2014). Two other candidates, namely FNR and RbcS (small subunit of Rubisco from pea), are well-characterized plastid proteins (Highfield and Ellis 1978; Zhang et al., 2001), while the residual three candidates, namely mtRi, ATPS (ATP synthase subunit beta-3) and CoxIV (Cytochrome *c* oxidase subunit IV) of yeast, are known for their mitochondrial targeting specificity (Maarse et al., 1984; Emmermann et al., 1994; Baudisch et al., 2014). The protein fragments comprising the transport signals of these proteins were cloned as fusions with GFP11 tags into the PlaMiNGo vectors, additionally comprising the organelle targeted receptor fusions FNR$_{1-55}$/YFP1-10 or mtRi$_{1-100}$/YFP1-10.
Five of the eight candidates showed in our assay system the same targeting behavior as reported in the literature (Table 1): mtRi and FNR showed exclusive transport into either mitochondria or plastids, respectively (Figure 7a-b) (Rodiger et al., 2011), and the dual targeting candidates GCS, Gtret and GAPDH showed transport into both organelles (Baudisch et al., 2014) (Figure 7c-e). It should be noted though that in the case of GAPDH mitochondrial targeting was rather weak in our assays and could be observed only in few transformed cells. However, it was rather unexpected that the remaining three monospecific candidates, namely ATPS$_{1-100}$, RbcS$_{1-79}$ and CoxIV$_{1-29}$, showed dual targeting in our experiments (Figure 8). In the literature, mitochondrial targeting of RbcS has already been described once (Rudhe et al., 2002) but these results were solely based on *in vitro* assays. In addition, dual targeting of the yeast mitochondria presequence, CoxIV, could be assumed considering high degree of freedom of non-plant mitochondria transport signals (Staiger et al., 2009). However, this dual targeting was entirely unexpected for the plant mitochondrial protein, ATPS$_{1-100}$, which never showed any targeting to plastids when analyzed with *in vivo*

fluorescent protein tagging and *in vitro* protein transport experiments (Baudisch et al., 2014 and Suppl. Figure 4).

### 2.6 *Phenotype complementation confirms the plastid targeting properties of ATPS*

These unexpected results demanded for independent confirmation. Thus, to re-evaluate the plastid targeting properties of the transport signal of mitochondrial ATPS we have applied a phenotype complementation approach. In this approach, we made use of the *immutans* mutant of *Arabidopsis thaliana* that shows a white-green sectored (variegated) leaf phenotype and stunted growth when grown under daylight conditions. The mutant phenotype is due to the absence of a functional nuclear-encoded plastid protein, namely plastid terminal oxidase (PTOX; Carol et al., 1999; Aluru et al., 2001). Complementation of the variegated phenotype of *immutans* requires a functional PTOX protein in plastids (Fu et al., 2005 and Figure 9c). To adapt PTOX for our analysis, we have replaced the authentic transit peptide of PTOX with the N-terminal 100 AA residues of two candidate proteins, namely ATPS and GCS (a validated dually targeted protein) generating $ATPS_{1-100}$/mPTOX and $GCS_{1-100}$/mPTOX, respectively. Two further constructions encoding either the authentic PTOX precursor or the transit peptide-free mature PTOX protein (mPTOX) were generated for comparison. These gene chimeras were expressed under the control of the CaMV35S promoter in the *immutans* mutant plants. As expected, both the authentic PTOX precursor and $GCS_{1-100}$/mPTOX were able to complement the variegated phenotype, as expected. In contrast, the expression of mPTOX cannot complement the mutant phenotype (Figure 9). These results confirm that targeting of PTOX into plastids is essential for complementation of the *immutans* variegated phenotype. Remarkably, also the expression of $ATPS_{1-100}$/mPTOX resulted in mutant phenotype complementation (Figure 9f). However, such complementation was observed in only four from six transgenic lines analyzed suggesting that plastid targeting of $ATPS_{1-100}$/mPTOX is, in principle, possible but apparently less efficient than with typical plastid targeted transit peptides. These results clearly underline the basic plastid targeting properties of mitochondrial protein ATPS and thus re-confirm the results obtained with the *sa*split-GFP technology established here.

## 3. Discussion-

The goal of this study was to assess the suitability of the signal enhanced *sa*split-GFP system to determine the targeting specificity of nuclear-encoded organelle proteins and to develop tools for rapid cloning and subsequent analysis of their targeting behavior. The characterization of the targeting specificity of organelle proteins is crucial (i) to elucidate the functional properties of the organelle transport machineries, (ii) to study evolutionary aspects of protein targeting, and (iii) for biotechnological applications employing organelles. The application of the *sa*split-GFP system, as demonstrated in this study, provides a novel toolbox to quickly determine the targeting properties of candidate proteins with high sensitivity.

### 3.1 *Selective imaging and fluorescence signal enhancement with* sa*split-GFP technology*

#### 3.1.1 Selective Imaging

The selective imaging of organelles is one of the major advantages of the *sa*split-GFP system in comparison to ‘standard’ fluorescent protein tagging approaches. The requirement for the presence of the non-fluorescing GFP1-10 receptor in a specific subcellular location is the key

for selective imaging (Kaddoum et al., 2010). In this study, two transport signals, namely $FNR_{1-55}$ and $mtRi_{1-100}$, were selected for localization of the receptor specifically within two sub-cellular locations, the plastid stroma and the mitochondrial matrix, respectively. As a result, fluorescence signals will appear only if the GFP11 tagged protein is completely imported into the same sub-cellular location and not if a protein is merely binding to the organelle surface. This was otherwise difficult to distinguish with FP-based approaches, specifically for mitochondria due to their small size.

### 3.1.2 Fluorescence signal enhancement

The self-assembling split-GFP molecules have been reported to produce fluorescent signals of lower intensity than 'standard' fluorescent proteins (Kökker et al., 2018). This problem can be circumvented with the use of multiple GFP11 tags. However, the two organelles respond differently to this modification. While signal enhancement with multiple GFP11 tags works well in plastids, fluorescence signal enhancement could not be observed in mitochondria (Figure 4). One possible reason might be the size difference between these organelles. Plastids are much larger in size and thus the proteins in the plastid stroma are present in lower concentration. Consequently, the chances for self-assembly of *sa*split-GFP fragments in this organelle are lower than in mitochondria and thus fluorescence signal enhancement could be observed by increasing number of GFP11 tags in plastids. Furthermore, differences in the physicochemical properties of the two organelles, e.g. pH, might likewise contribute to this phenomenon.

The use of more efficient gene regulatory elements, i.e. promoter and terminator, in the PlaMiNGo toolkit also led to significant enhancement in fluorescence signal intensity. However, in combination with multiple GFP11 tags such increased gene expression can lead to the formation of aggregations in transformed cells, particularly if these tags are combined with mitochondria targeting transport signals. The intrinsic property of the *sa*split-GFP fragments to form dimers and aggregates (Cabantous et al., 2005) and comparatively less efficient unfoldase activity of the mitochondrial protein translocation machinery (Agarraberes and Dice 2001) could be one of the possible reasons for this phenomenon. Since, the protein unfolding prior to translocation is apparently more efficient in plastids, the multiple GFP11 tags can efficiently be imported into this organelle. On the other hand, high expression of plastid targeted GFP1-10 alone could result in appearance of faint fluorescence signals, probably due to formation of dimers. In most instances, these faint signals are clearly distinguishable from 'actual' fluorescence signals obtained via self-assembly of split-GFP but still it should considered as an experimental control to avoid the misinterpretation of results

## *3.2 High sensitivity of* sa*split-fluorescence protein system*

Fluorescent signal enhancement in combination with selective imaging makes the *sa*split-GFP system highly sensitive with respect to targeting specificity determination. In consequence, dual targeting of several proteins was newly detected with the PlaMiNGo toolkit developed here, which had previously been missed due to inherent limitations of 'standard' fluorescent protein tagging approaches. For example, GAPDH shows dual targeting with the *sa*split-GFP system, in line with the results of *in vitro* import experiments (Baudisch et al., 2014). In contrast, with 'classical' *in vivo* approaches using FP-tagging, GAPDH appeared to be solely transported into plastids (Baudisch et al., 2014). Similarly, as shown here, the transit peptide of RbcS is able to translocate the GFP11 tag into mitochondria, but this property remained undetected with the FP-tagging approach. Remarkably, such mitochondria targeting properties of the RbcS transit peptide were also found in a recent study employing

sulfadiazine-resistant plants (Tabatabaei et al., 2018). In contrast, dual targeting of the transport signal of yeast CoxIV had never been reported earlier. Dual targeting of yeast mitochondrial transport signal might be a consequence of the fact that yeast does not contain plastids and thus the transport signal of yeast mitochondria had not 'learned' to distinguish between the two endosymbiotic organelles (Staiger et al., 2009). The plastid targeting of a yeast mitochondria transport signal has also been reported earlier (Huang et al., 1990). However, such dual targeting was most unexpected for the transport signal of ATPS because neither *in vitro* nor *in vivo* approaches gave any hint of the plastid targeting properties of this transport signal (Baudisch et al., 2014). Even transgenic plants expressing $ATPS_{1-100}$/eYFP did not show any plastid localization (Suppl. Fig. 4). The fact, that it was clearly detectable here and that this result could be independently confirmed by a phenotype complementation approach using *immutans* mutants (Figure 8a, 9f), underlines the sensitivity of the *sa*split-GFP technology.

### *3.3 Modularity of PlaMiNGO toolkit*

The vector toolkit is based on the principle of modular cloning (Weber et al., 2011). Hence, the components of the PlaMiNGo toolkit can easily be rearranged for *in vivo* imaging of proteins targeted to various other sub-cellular compartments. The vectors constructed in module 1 (Suppl. Figure 3) are binary vectors and can be utilized for plant cell transformation via *Agrobacterium* or via several other methods, e.g. protoplast transformation or particle bombardment. The gene of interest can be cloned upstream to one of the GFP11 tags with a single *Golden Gate* cloning reaction. Similarly, GFP1-10 can be targeted to the different sub-cellular or even sub-organellar compartments via cloning of the specific transport signal N-terminally in a *Golden Gate* 'ready' vector (pTEI176 or pTEI177). These vectors carry a *ccd*B negative selection cassette upstream of GFP1-10 with *Bsa*I restriction sites A|ATG at the 5´ end and T|TCG at the 3´ end. Consequently, the two vectors carrying GFP11 and GFP1-10 gene chimeras should be co-expressed in a single cell in order to determine protein targeting specificity to the organelle of interest.

### *3.4 Significance of high dual targeting frequency*

The results obtained in this study strongly suggest that the number of dually targeted proteins is much higher than assumed so far. More than just increasing the list of such proteins, these findings also highlight the evolutionary conserved nature of organelle translocation machineries in plant cells. Even if some of these proteins are 'mistargeted' to the 'wrong' organelle, for example, due to protein overexpression, it shows the fundamental targeting properties of the respective transport signal. The reason of such widespread dual targeting is still open. On the one hand, it supports the hypothesis that dual targeting might be an evolutionary remnant (Staiger et al., 2009). On the other hand, the acquisition of dual targeting properties might still be ongoing process, which allows for the development of completely new biochemical pathways in an organelle (Martin 2010, Xu et al., 2013). It is, therefore, important to determine the targeting properties of protein transport signals in order to understand its physiological role in the plant cell. The *sa*split-GFP based approach and the PlaMiNGo toolkit developed here provide important tools to determine the targeting properties of a protein.

## 4. Materials and method

## *4.1 Molecular Cloning-*

### 4.1.1 Generation of vectors for co-infiltration

The GFP1-10 and GFP11$_{x7}$ fragments were amplified by PCR from plasmids pcDNA3.1-GFP (1-10) and pACUH-GFP11$_{x7}$-mCherry-ß-tubulin (a gift from Bo Huang lab; Addgene#70218 and #70219) and cloned into pRT100mod-based vectors (Baudisch et al. 2014) either with digestion/ligation or with Restriction Free cloning (Bond and Naus, 2012). Primers for gene amplification are summarized in Suppl. Table 1. The gene sequence coding for the N-terminal 91 amino acids of TyrRS was amplified from a vector provided by E. Glaser and cloned accordingly (Baudisch et al. 2014). The above constructions, comprising promoter and terminator regions (*CaMV35S::Gene of interest:GFP1-10/GFP11$_{x7}$::t35S*), were later sub-cloned into a *Golden Gate* compatible pLSU4GG binary vector with *Bsa*I restriction/ligation reaction (Erickson et al., 2017). *Golden Gate* cloning was performed in 15 µl reaction with following conditions: 2.5 units of T4 DNA ligase (Thermo Scientific), 5 units of *Bsa*I (Thermo Scientific), 1X BSA, 20 cycles of incubation at 37°C for 2 min and 16°C for 5 min, final deactivation and denaturation at 50°C for 10 min and 80°C for 10 min respectively.

### 4.1.2 Construction of the PlaMiNGo toolkit

The modular cloning principle and DNA fragments of the Plant Parts I and II toolkits were used for vector construction (Weber et al., 2011; Engler et al., 2014; Gantner et al., 2018). The modules utilized for cloning of *Golden Gate*-based vectors are illustrated in Supplementary Figure 3. *Golden Gate* reactions were performed with 20 fmol of each DNA module with the following conditions: 2.5 units of T4 DNA ligase (Thermo Scientific), 5 units of *Bsa*I or *Bpi*I (New England Biolabs), 30 cycles of incubation at 37°C for 2 min and 16°C for 5 min, final denaturation at 80°C for 10 min. When required, the restriction/ligation reactions were subsequently supplemented with fresh ligation buffer and ligase for terminal ligation, and incubated for ≥ 3h at 16°C. Ligation mixtures were transformed into Dh10b or *ccd*B survival II cells (Thermo Scientific) and grown on plates with appropriate selective medium. PCR primers for Level 0 modules and gene sequences are summarized in Suppl. Table 2 and 3.

### 4.1.3 Cloning of candidate proteins into PlaMiNGo vectors

The candidate targeting signals (TyrRS$_{1-91}$, FNR$_{1-55}$, mtRi$_{1-100}$, ATPS$_{1-100}$, GCS$_{1-100}$, GAPDH$_{1-100}$, Gtred$_{1-100}$, RbcS$_{1-79}$ and CoxIV$_{1-29}$) were amplified from the corresponding cDNA templates (Nelson et al., 2007; Berglund et al., 2009; Baudisch et al., 2014) and cloned via standard *Golden Gate* reaction (see above) into PLaMiNGo vectors in exchange for a ccdB negative selection cassette. Overhangs of the fragment to be cloned were A|ATG at the 5´ end and T|TCG at the 3´ end. Two additional nucleotides were inserted in some of the fusions to maintain the reading frame resulting in an additional codon for an alanine residue.

### 4.1.4 Cloning for complementation of the *immutans* phenotype

The gene sequence carrying mature PTOX$_{57-295}$ was amplified from a cDNA clone provided by Steven Rodermel (Iowa State University, USA) and sub-cloned via restriction free cloning into pRT100mod-based vectors downstream of the gene sequences coding for the transport signals of either GCS$_{1-100}$ or ATPS$_{1-100}$. Additionally, the PTOX full-length gene and mature PTOX (lacking the transit peptide) were cloned into empty pRT100mod vectors under control of CaMV35 promoter and terminator. The candidate gene constructions with promoter and terminator (*CaMV35S::Candidate:PTOXmat::t35S*) were then cloned into a *Golden Gate* compatible pLSU4GG binary vector with *Bsa*I cut/ligation reaction.

### *4.2 Agrobacterium infiltration*

The constructs were transformed into electro-competent cells of *Agrobacterium tumefaciens* GV3101 (pMP90) (Koncz and Schell, 1986). *Agrobacterium* infiltrations of 6 - 8 weeks old fully expanded *Nicotiana benthamiana* leaves were performed as described by Sharma et al. (2018a). For co-infiltration, each bacterial strain was adjusted to $OD_{600}$ = 0.8 and mixed in a 1:1 ratio prior to infiltration.

### *4.3 Microscopy and imaging -*

Confocal laser scanning microscopy was carried out as described by Sharma et al. (2018a). For GFP/YFP dual channel imaging, the 493-518 (GFP) and 519-620 (YFP) filter ranges were used. When required, brightness and contrast of the images were equally adjusted for each image to avoid any discrepency in visualization of signal intensities.

### *4.4 Signal Quantification-*

For quantification of signal intensities, infiltration of all relevant constructs was carried out on different spots of the same leaf. At least three images from each infiltration spot were used for quantification. Quantification of the signals was performed with raw images using the Fiji program (Schindelin et al., 2012). For the purpose of quantification, image acquisition was done with the 20x objective in 7 to 8 Z-stacks covering the epidermal cell layer and later stacked to project the maximum intensities. The mean gray values of stacked images were calculated using the 'Measure' option of Fiji and further utilized for comparision of the fluorescence signal strenght in arbitrary units (A.U.).

### *4.5 Growth and transformation of immutans plants-*

The *Arabidopsis thaliana immutans* seeds (provided by Steven Rodermel, ISU) were germinated at 5 µmol $m^{-2}$ $s^{-1}$ lights and 8/16 h light-dark cycles. After three weeks of germination, the seedlings were transferred to 50 µmol $m^{-2}$ $s^{-1}$ lights. For induction of flowering, the plantlets were transferred to >150 µmol $m^{-2}$ $s^{-1}$ light and incubated at a 16/8 h light/dark cycle. Floral-dip transformation was performed as described (Davis et al. 2009). At least three independent T2 transgenic lines were selected in each case for the analysis of phenotype. After germination at >150 µmol $m^{-2}$ $s^{-1}$ light in a 8/16 h light/dark cycle, these plants were transferred to 16/8 h light/dark cycle. Pictures presented here are from 5 week old plants grown in >150-µmol $m^{-2}$ $s^{-1}$ light.

Competing Interests-

The authors declare no competing interests.

Acknowledgement-

We would like to thank Theresa Ilse for assistance in cloning, Elzbieta Glaser (Stockholm University, Sweden), Martin Schattat (Martin Luther University Halle-Wittenberg, Germany) and Bo Huang (University of California, San Francisco, USA; via Addgene) for kindly providing constructs and Steven Rodermel (Iowa State University, USA) for providing *immutans* mutants and the PTOX clone. MS was supported by a fellowship from the BRAVE project funded by the ERASMUS MUNDUS Action 2 program of the European Union.

References-


1. Agarraberes, F.A. and Dice, J.F. (2001) Protein translocation across membranes. Biochimica et Biophysica Acta (BBA)-Biomembranes. 1513:1-24.

2. Aluru, M.R., Bae, H., Wu, D. and Rodermel, S.R. (2001) The Arabidopsis immutans mutation affects plastid differentiation and the morphogenesis of white and green sectors in variegated plants. Plant Physiology. 127:67-77.

3. Andresen, M., Schmitz-Salue, R. and Jakobs, S. (2004) Short tetracysteine tags to β-tubulin demonstrate the significance of small labels for live cell imaging. Molecular Biology of The Cell. 15:5616-5622.

4. Baudisch, B., Langner, U., Garz, I. and Klösgen, R.B. (2014) The exception proves the rule? Dual targeting of nuclear- encoded proteins into endosymbiotic organelles. New Phytologist. 201:80-90.

5. Berglund, A.K., Pujol, C., Duchene, A.M. and Glaser, E. (2009) Defining the determinants for dual targeting of amino acyl-tRNA synthetases to mitochondria and chloroplasts. Journal of Molecular Biology. 393:803-814.

6. Bond, S.R. and Naus, C.C. (2012) RF-Cloning.org: an online tool for the design of restriction-free cloning projects. Nucleic Acids Research. 40:W209-W213.

7. Cabantous, S., Terwilliger, T.C. and Waldo, G.S. (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nature Biotechnology. 23:102.

8. Cabantous, S. and Waldo, G.S. (2006) In vivo and in vitro protein solubility assays using split GFP. Nature Methods. 3:845.

9. Cabantous, S., Nguyen, H.B., Pedelacq, J.D., Koraïchi, F., Chaudhary, A., Ganguly, K., Lockard, M.A., Favre, G., Terwilliger, T.C. and Waldo, G.S. (2013) A new protein-protein interaction sensor based on tripartite split-GFP association. Scientific Reports, 3:2854.

10. Carol, P., Stevenson, D., Bisanz, C., Breitenbach, J., Sandmann, G., Mache, R., Coupland, G. and Kuntz, M. (1999) Mutations in the Arabidopsis gene IMMUTANS cause a variegated phenotype by inactivating a chloroplast terminal oxidase associated with phytoene desaturation. The Plant Cell. 11:57-68.

11. Cheng, N.H. (2008) AtGRX4, an Arabidopsis chloroplastic monothiol glutaredoxin, is able to suppress yeast grx5 mutant phenotypes and respond to oxidative stress. FEBS Letters. 582:848-854.

12. Coruzzi, G., Broglie, R., Edwards, C. and Chua, N.H. (1984) Tissue- specific and light- regulated expression of a pea nuclear gene encoding the small subunit of ribulose- 1, 5- bisphosphate carboxylase. The EMBO Journal, 3:1671-1679.

13. Davis, A.M., Hall, A., Millar, A.J., Darrah, C. and Davis, S.J. (2009) Protocol: Streamlined sub-protocols for floral-dip transformation and selection of transformants in Arabidopsis thaliana. Plant Methods. 5:3.

14. De Greve H., Dhaese P., Seurinck J., Lemmers M, Van Montagu M, Schell J. (1982) Nucleotide sequence and transcript map of the Agrobacterium tumefaciens Ti plasmid-encoded octopine synthase gene. Journal of Molecular and Applied Genetics. 1:499–511

15. Duchêne, A.M., Giritch, A., Hoffmann, B., Cognat, V., Lancelin, D., Peeters, N.M., Zaepfel, M., Maréchal-Drouard, L. and Small, I.D. (2005) Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. Proceedings of the National Academy of Sciences. 102:16484-16489.

16. Emmermann, M., Clericus, M., Braun, H.P., Mozo, T., Heins, L., Kruft, V. and Schmitz, U.K. (1994) Molecular features, processing and import of the Rieske iron sulfur protein from potato mitochondria. Plant molecular biology. 25:271-281.

17. Engler, C., Youles, M., Gruetzner, R., Ehnert, T.M., Werner, S., Jones, J.D., Patron, N.J. and Marillonnet, S. (2014) A golden gate modular cloning toolbox for plants. ACS synthetic biology. 3:839-843.

18. Erickson, J. L., Kantek, M., and Schattat, M. H. (2017) Plastid-nucleus distance alters the behavior of stromules. Frontiers in Plant Science. 8:1135.

19. Fu, A., Park, S. and Rodermel, S. (2005) Sequences required for the activity of PTOX (IMMUTANS), a plastid terminal oxidase in vitro and in planta mutagenesis of iron-binding sites and a conserved sequence that corresponds to exon 8.Journal of Biological Chemistry. 280:42489-42496.

20. Van Engelenburg, S.B. and Palmer, A.E. (2010) Imaging type-III secretion reveals dynamics and spatial segregation of Salmonella effectors. Nature Methods. 7:325.

21. Gantner, J., Ordon, J., Ilse, T., Kretschmer, C., Gruetzner, R., Löfke, C., Dagdas, Y., Bürstenbinder, K., Marillonnet, S. and Stuttmann, J. (2018) Peripheral infrastructure vectors and an extended set of plant parts for the Modular Cloning system. PloS One. 13:0197185.

22. Giepmans, B.N., Adams, S.R., Ellisman, M.H. and Tsien, R.Y. (2006) The fluorescent toolbox for assessing protein location and function. Science. 312:217-224.

23. Grefen, C. and Blatt, M.R. (2012) A 2in1 cloning system enables ratiometric bimolecular fluorescence complementation (rBiFC). Biotechniques. 53:311-314.

24. Hecker, A., Wallmeroth, N., Peter, S., Michael, R.B., Harter, K. and Grefen, C. (2015) Binary 2in1 vectors improve in planta (co-) localisation and dynamic protein interaction studies. Plant Physiology 168:776-87.

25. Henry, E., Toruño, T.Y., Jauneau, A., Deslandes, L., Coaker, G.L. (2017) Direct and indirect visualization of bacterial effector delivery into diverse plant cell types during infection. Plant Cell. 29: 1555–1570

26. Highfield, P.E. and Ellis, R.J. (1978) Synthesis and transport of the small subunit of chloroplast ribulose bisphosphate carboxylase. Nature. 271:420.

27. Huang, J., Hack, E., Thornburg, R.W. and Myers, A.M. (1990) A yeast mitochondrial leader peptide functions in vivo as a dual targeting signal for both chloroplasts and mitochondria. The Plant Cell. 2:1249-1260.

28. Kaddoum, L., Magdeleine, E., Waldo, G.S., Joly, E. and Cabantous, S. (2010) One-step split GFP staining for sensitive protein detection and localization in mammalian cells. Biotechniques, 49:727-8.

29. Kamiyama, D., Sekine, S., Barsi-Rhyne, B., Hu, J., Chen, B., Gilbert, L.A., Ishikawa, H., Leonetti, M.D., Marshall, W.F., Weissman, J.S. and Huang, B. (2016) Versatile protein tagging in cells with split fluorescent protein. Nature Communications. 7:11046.

30. Koncz, C. and Schell, J. (1986) The promoter of T L-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. Molecular and General Genetics. 204:383-396.

31. Köhler, R.H., Zipfel, W.R., Webb, W.W. and Hanson, M.R. (1997) The green fluorescent protein as a marker to visualize plant mitochondria in vivo. The Plant Journal. 11:613-621.

32. Köker, T., Fernandez, A. and Pinaud, F. (2018) Characterization of Split Fluorescent Protein Variants and Quantitative Analyses of Their Self-Assembly Process. Scientific reports. 8:5344.

33. Maarse, A.C., Van Loon, A.P., Riezman, H., Gregor, I., Schatz, G. and Grivell, L.A. (1984) Subunit IV of yeast cytochrome c oxidase: cloning and nucleotide sequencing of the gene and partial amino acid sequencing of the mature protein. The EMBO Journal. 3:2831-2837.

34. Machettira, A.B., Gross, L.E., Sommer, M.S., Weis, B.L., Englich, G., Tripp, J. and Schleiff, E. (2011) The localization of Tic20 proteins in Arabidopsis thaliana is not restricted to the inner envelope membrane of chloroplasts. Plant Molecular Biology. 77:381.

35. Marques, J.P., Schattat, M.H., Hause, G., Dudeck, I. and Klösgen, R.B. (2004) In vivo transport of folded EGFP by the ΔpH/TAT-dependent pathway in chloroplasts of Arabidopsis thaliana. Journal of Experimental Botany. 55:1697-1706.

36. Martin W (2010) Evolutionary origins of metabolic compartmentalization in eukaryotes. Philos Trans R Soc B. 365:847–855

37. Moseler, A., Aller, I., Wagner, S., Nietzel, T., Przybyla-Toscano, J., Mühlenhoff, U., Lill, R., Berndt, C., Rouhier, N., Schwarzländer, M. and Meyer, A.J. (2015) The mitochondrial monothiol glutaredoxin S15 is essential for iron-sulfur protein maturation in Arabidopsis thaliana. Proceedings of the National Academy of Sciences. 112:13735-13740.

38. Moore, I. and Murphy, A., (2009) Validating the location of fluorescent protein fusions in the endomembrane system. The Plant Cell. 21:1632-1636.

39. Nelson, B.K., Cai, X. and Nebenführ, A. (2007) A multicolored set of in vivo organelle markers for co - localization studies in Arabidopsis and other plants. The Plant Journal, 51:1126-1136.

40. Park, E., Lee, H.Y., Woo, J., Choi, D. and Dinesh-Kumar, S.P. (2017) Spatiotemporal monitoring of Pseudomonas effectors via type III secretion using split fluorescent protein fragments. The Plant Cell. 29:1571-1584.

41. Rao, R.S.P., Salvato, F., Thal, B., Eubel, H., Thelen, J.J. and Møller, I.M. (2017) The proteome of higher plant mitochondria. Mitochondrion. 33:22-37.

42. Rödiger, A., Baudisch, B., Langner, U. and Klösgen, R.B. (2011) Dual targeting of a mitochondrial protein: the case study of cytochrome c1. Molecular Plant. 4:679-687.

43. Rudhe, C., Chew, O., Whelan, J. and Glaser, E., 2002. A novel in vitro system for simultaneous import of precursor proteins into mitochondria and chloroplasts. The Plant Journal. 30:213-220.

44. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. and Tinevez, J.Y. (2012) Fiji: an open-source platform for biological-image analysis. Nature Methods. 9:676.

45. Sharma, M., Bennewitz, B. and Klösgen, R.B. (2018) Dual or Not Dual?—Comparative Analysis of Fluorescence Microscopy-Based Approaches to Study Organelle Targeting Specificity of Nuclear-Encoded Plant Proteins. Frontiers in Plant Science. 9:1350

46. Sharma, M., Bennewitz, B. and Klösgen, R.B. (2018) Rather rule than exception? How to evaluate the relevance of dual protein targeting to mitochondria and chloroplasts. Photosynthesis research, 138(3), 335-343.

47. Staiger, C., Hinneburg, A. and Klösgen, R.B. (2009) Diversity in degrees of freedom of mitochondrial transit peptides. Molecular Biology and Evolution. 26:1773-1780.

48. Tabatabaei, I., Dal Bosco, C., Bednarska, M., Ruf, S., Meurer, J. and Bock, R. (2018) A highly efficient sulfadiazine selection system for the generation of transgenic plants and algae. Plant Biotechnology Journal. doi: 10.1111/pbi.13004. [Epub ahead of print]

49. Tanz, S.K., Castleden, I., Small, I.D. and Millar, A.H. (2013) Fluorescent protein tagging as a tool to define the subcellular distribution of proteins in plants. Frontiers in Plant Science. 4:214.

50. Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. PLoS ONE 6:e16765.

51. Wetzel, C.M., Jiang, C.Z., Meehan, L.J., Voytas, D.F. and Rodermel, S.R. (1994) Nuclear—organelle interactions: the immutans variegation mutant of Arabidopsis is plastid autonomous and impaired in carotenoid biosynthesis. The Plant Journal. 6:161-175.

52. van Wijk, K.J. and Baginsky, S. (2011) Plastid proteomics in higher plants: current state and future goals. Plant Physiology. 155:1578-1588.

53. Xu, L., Carrie, C., Law, S.R., Murcha, M.W. and Whelan, J. (2013) Acquisition, conservation, and loss of dual-targeted proteins in land plants. Plant Physiology. 161:644-662.

54. Zhang, H., Whitelegge, J.P. and Cramer, W.A., (2001) Ferredoxin: NADP+ oxidoreductase is a subunit of the chloroplast cytochrome b6f Complex. Journal of Biological Chemistry. 276:38159-38165.

**Table 1- Sub-cellular localization of candidate proteins as determined with different approaches.**

| **Candidate** | ***sa*split-GFP** | ***in vitro* import¶** | **FP Localization** | **Gene Accession** |
|---|---|---|---|---|
| FNR | Chloro | Chloro[2] | Chloro[2] | M86349.1 |
| mtRi | Mito | Mito[2] | Mito[2] | X79332.1 |
| GCS | Mito and Chloro | Mito and Chloro[2] | Mito and Chloro[2] | At2g35370 |
| GAPDH | Mito and Chloro | Mito and Chloro[2] | Chloro[2] | At1g42970 |
| Gtred | Mito and Chloro | Mito and Chloro[2] | Mito and Chloro[2], Chloro[4], Mito[5] | At3g15660 |
| CoxIV | Mito and Chloro | NA | Mito[6] | SGD:S000003155 |
| ATPS | Mito and Chloro | Mito[2] | Mito[2] | At5g08680 |
| SSU | Mito and Chloro | Mito and Chloro[1] | Chloro[3] | XM_016585367 |

¶Import studies were performed with the authentic precursor proteins; *NA*- not avialable ; [1]Rudhe et al., 2002; [2] Baudisch et al., 2014; [3]Nelson et al., 2007; [4]Cheng, 2008; [5]Moseler et al., 2015; [6]Köhler et al., 2003; *Mito*- mitochondria; *Chloro*- Chloroplasts or plastids.

**Figures-**

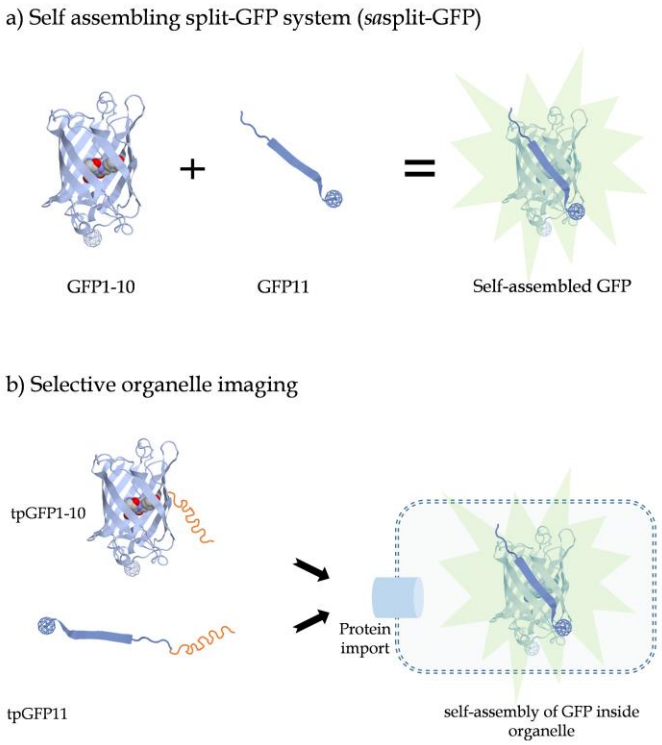


**Figure 1-** Schematic drawing illustrating the principle of the *sa*split-GFP system. **(a)** Two non-fluorescing fragments, GFP1-10 and GFP11, can self-assemble to generate a fluorescing GFP molecule. **(b)** Transport of both GFP chimeras into the same organelle is essential to achieve fluorescence signals.

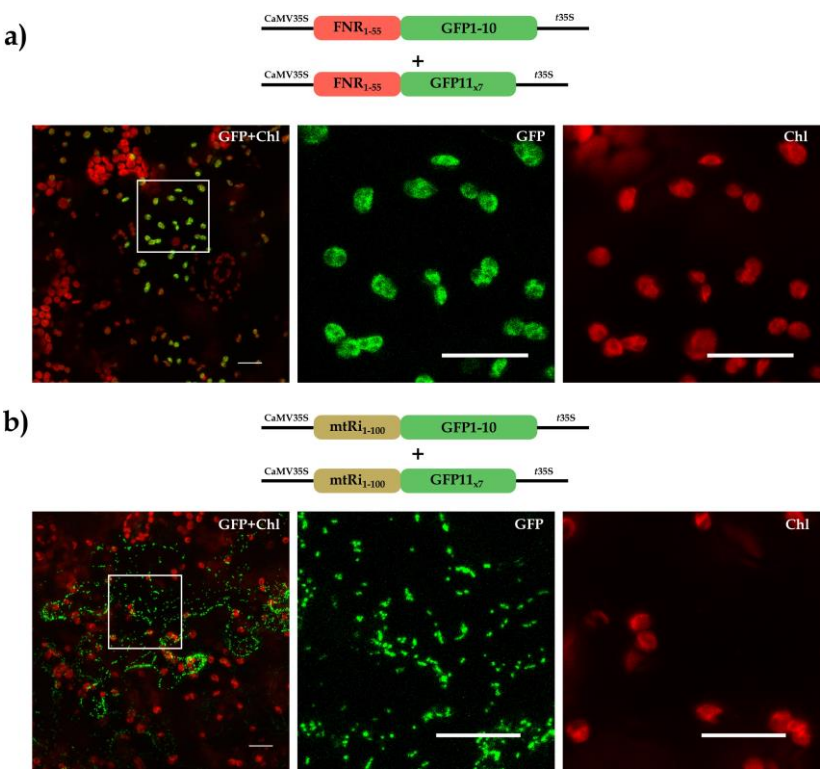


**Figure 2-** Establishment of the *sa*split-GFP system for *in vivo* organelle imaging. The coding sequences of **(a)** $FNR_{1-55}$/GFP1-10 and $FNR_{1-55}$/$GFP11_{x7}$ **(b)** $mtRi_{1-100}$/GFP1-10 and $mtRi_{1-100}$/$GFP11_{x7}$ were transiently co-expressed after *Agrobacterium* co-infiltration into the lower epidermis of *Nicotiana benthamiana* leaves and analyzed by confocal laser scanning microscopy (CLSM). Image acquisition of transformed cells was done with 20-x objective in several Z-stacks, which were subsequently stacked for maximum intensity projection. Representative cells (*left panels*) are

presented as overlay images of the chlorophyll channel (displayed in red) and the GFP channel (displayed in green). The strong chlorophyll signals in the background are derived from the larger chloroplasts of untransformed mesophyll cells underneath the epidermal cell layers. The *squares* highlight areas of the transformed cells that are shown in higher magnification separately for the chlorophyll channel (*middle panels*) and the GFP channel (*right panel*) as indicated. Scale bars correspond to 20 µm.

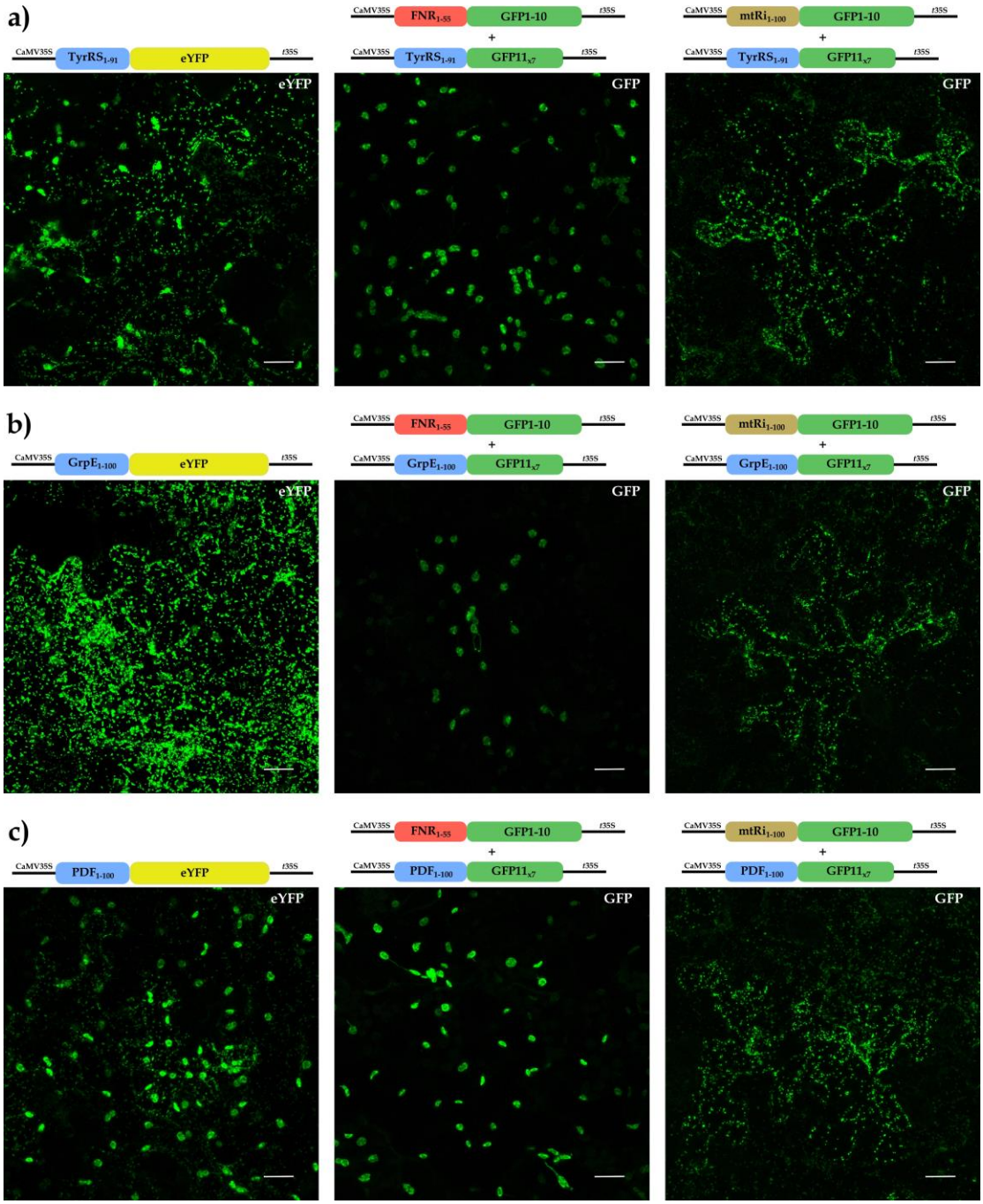


**Figure 3-** Comparison of FP-tagging and sasplit-GFP approaches to analyze the targeting specificity candidate proteins with proven dual-targeting properties. The coding sequences of

**(a)** $TyrRS_{1-91}/GFP11_{x7}$, **(b)** $GrpE_{1-100}/GFP11_{x7}$ and, **(c)** $PDF_{1-100}/GFP11_{x7}$ were transiently expressed with either $FNR_{1-55}$/GFP1-10 (*middle panels*) or $mtRi_{1-100}$/GFP1-10 (*right panels*) via *Agrobacterium* co-infiltration into the lower epidermis of *Nicotiana benthamiana* leaves and analyzed by CLSM. For comparison the respective eYFP fusions namely **(a)** $TyrRS_{1-91}$/eYFP, **(b)** $GrpE_{1-100}$/eYFP, and **(c)** $PDF_{1-100}$/eYFP were transiently expressed as well and analyzed by CLSM (*left panels*). For further details see the legend of Figure 2. Scale bars correspond to 20 µm.

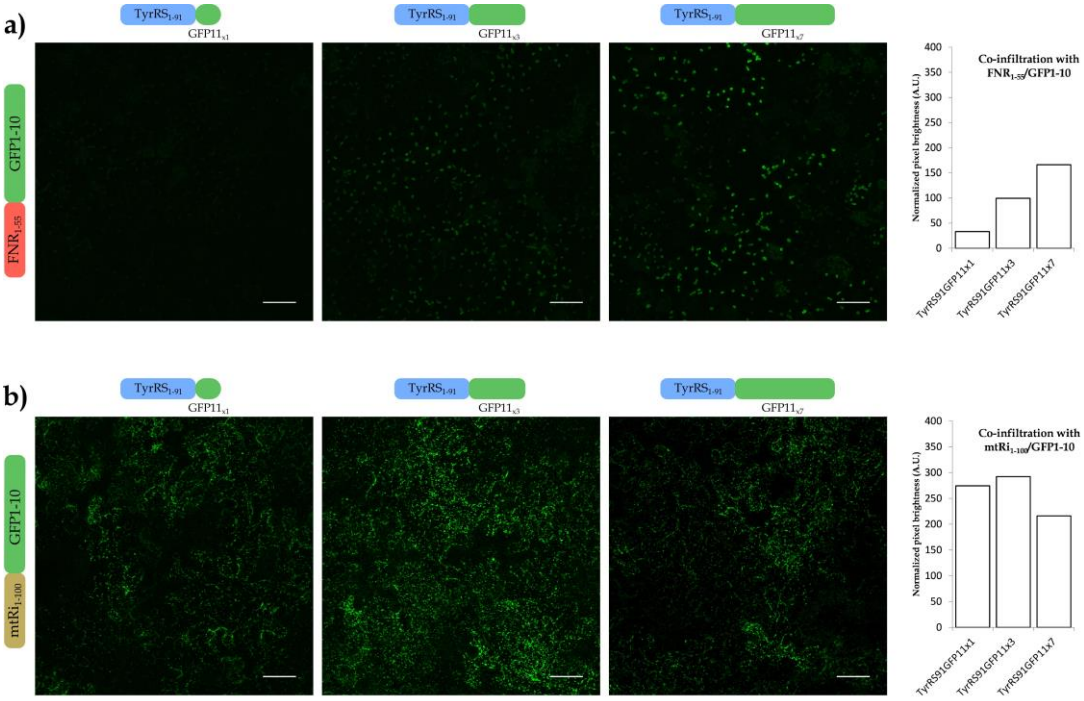


**Figure 4-** Effect of multiple GFP11 tags on fluorescence signal intensity. The gene coding sequence of the dual targeting transit peptide $TyrRS_{1-91}$ was fused with different GFP11 variants and co-expressed with either **(a)** $FNR_{1-55}$/GFP1-10 or **(b)** $mtRi_{1-100}$/GFP1-10 in *Nicotiana benthamiana* lower epidermal cells. The graph represents normalized average pixel brightness obtained after subtraction of background fluorescence signals in arbitrary units (A.U.) from 9 maximum intensity projected images for each experiment, obtained from two plant replicates (see Suppl. Figure 2 for absolute quantification). Scale bars correspond to 50 µm.

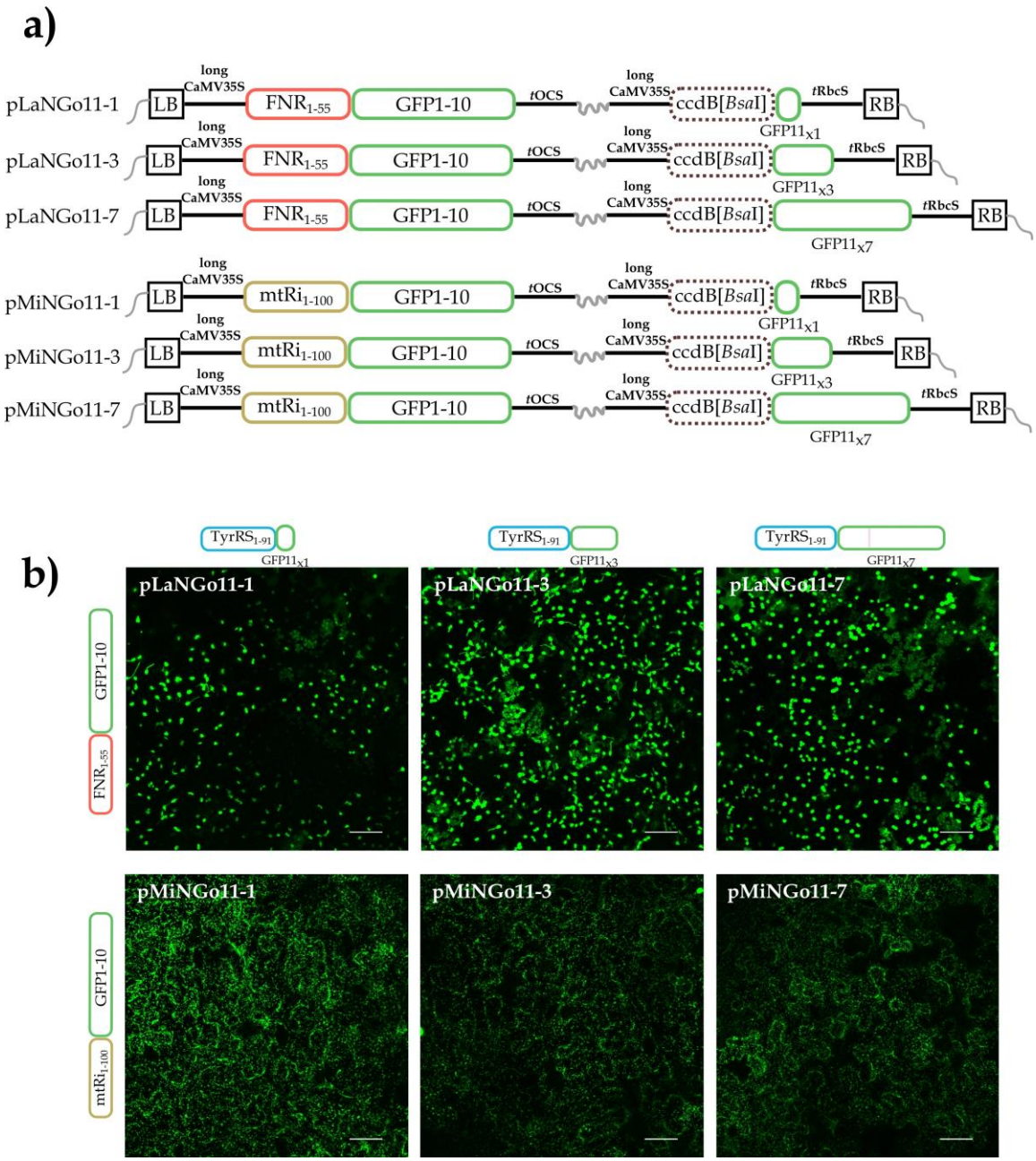


**Figure 5**- PlaMiNGo toolkit **(a)** Schematic representation of the vectors generated (see Supplementary Figure 3 for details) **(b)** PlaMiNGo vectors carrying the dually targeted $TyrRS_{1\text{-}19}$ transit peptide at the N-terminus of GFP11 tags were used to transform *Nicotiana benthamiana* lower epidermis cells via *Agrobacterium* infiltration. The quantification of signals is presented in supplementary Figure 2. Scale bars correspond to 50 µm.

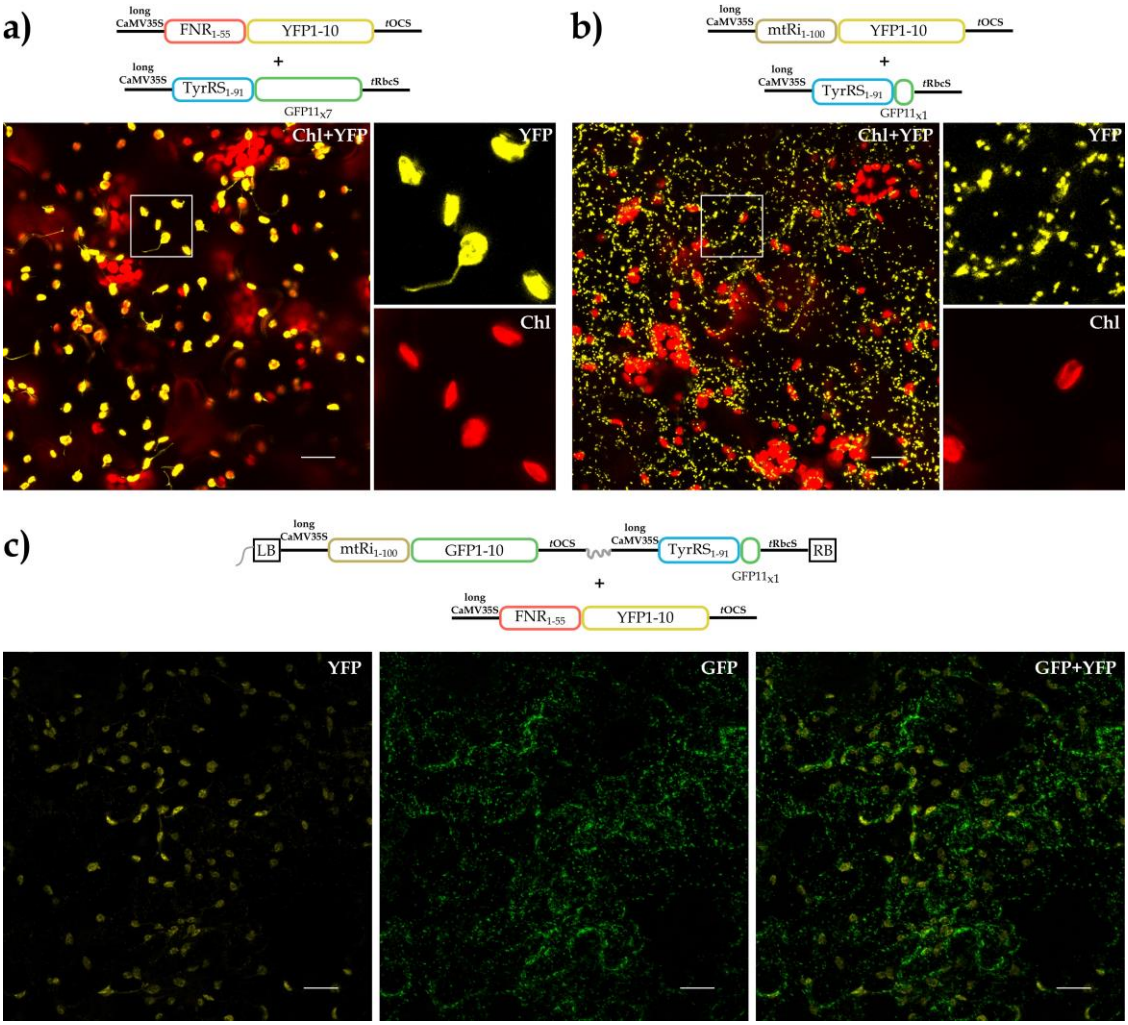


**Figure 6-** Multicolor imaging with *sa*split-YFP. The gene coding sequences of **(a)** $FNR_{1-55}$/YFP1-10 and $TyrRS_{1-91}$/$GFP11_{x7}$, or **(b)** $mtRi_{1-100}$/YFP1-10 and $TyrRS_{1-91}$/$GFP11_{x1}$ were transiently co-expressed in *Nicotiana benthamiana* leaf epidermis and analyzed by CLSM. **(c)** For the purpose of simultaneous multicolor imaging within the same cell, the PlaMiNGO vector comprising $mtRi_{1-100}$/GFP1-10 and $TyrRS_{1-91}$/$GFP11_{x1}$ was co-infiltrated with $FNR_{1-55}$/YFP1-10 resulting in self-assembly of GFP in mitochondria and YFP in plastids. Scale bars correspond to 20 µm.

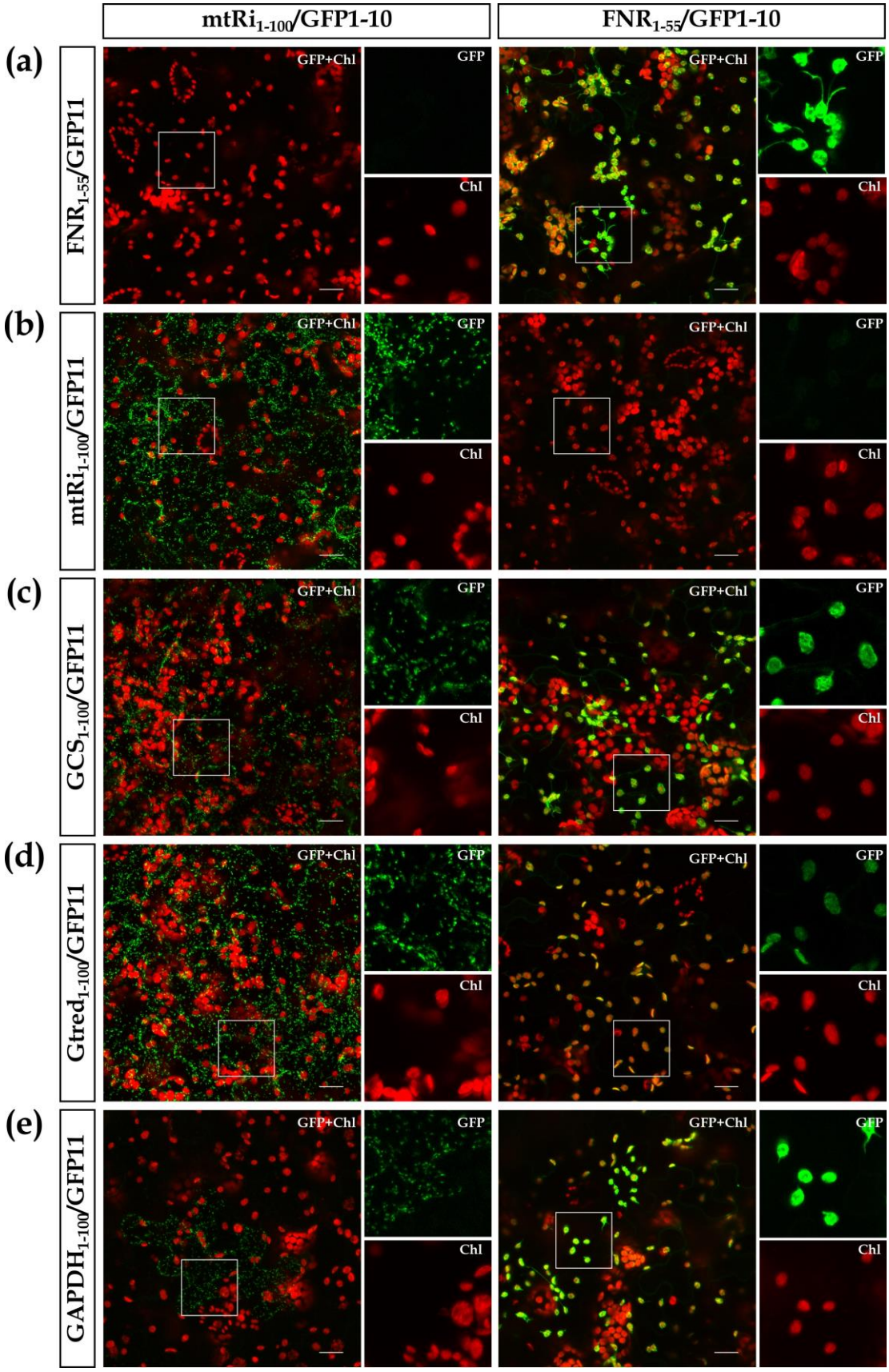
mtRi1-100/GFP1-10
FNR1-55/GFP1-10
(a)
FNR1-55/GFP11
(b)
mtRi1-100/GFP11
(c)
GCS1-100/GFP11
(d)
Gtred1-100/GFP11
(e)
GAPDH1-100/GFP11
GFP+Chl
GFP
Chl

**Figure 7-** Analysis of candidate proteins with the PLaMiNGo toolkit. Gene fragments encoding the transport signals of either **(a)** FNR, **(b)** mtRi, **(c)** GCS, **(d)** Gtred or **(e)** GAPDH, were cloned upstream of $GFP11_{x1}$ and $GFP11_{x7}$ tags in the two PlaMiNGo vectors comprising either $mtRi_{1-100}$/GFP1-10 (*left panels*) or $FNR_{1-55}$/GFP1-10 (*right panels*) respectively. The resulting constructs were used to transform the lower epidermis of *Nicotiana benthamiana* leaves via *Agrobacterium* infiltration. For further details see the legend of Figure 2. Scale bars correspond to 20 µm.

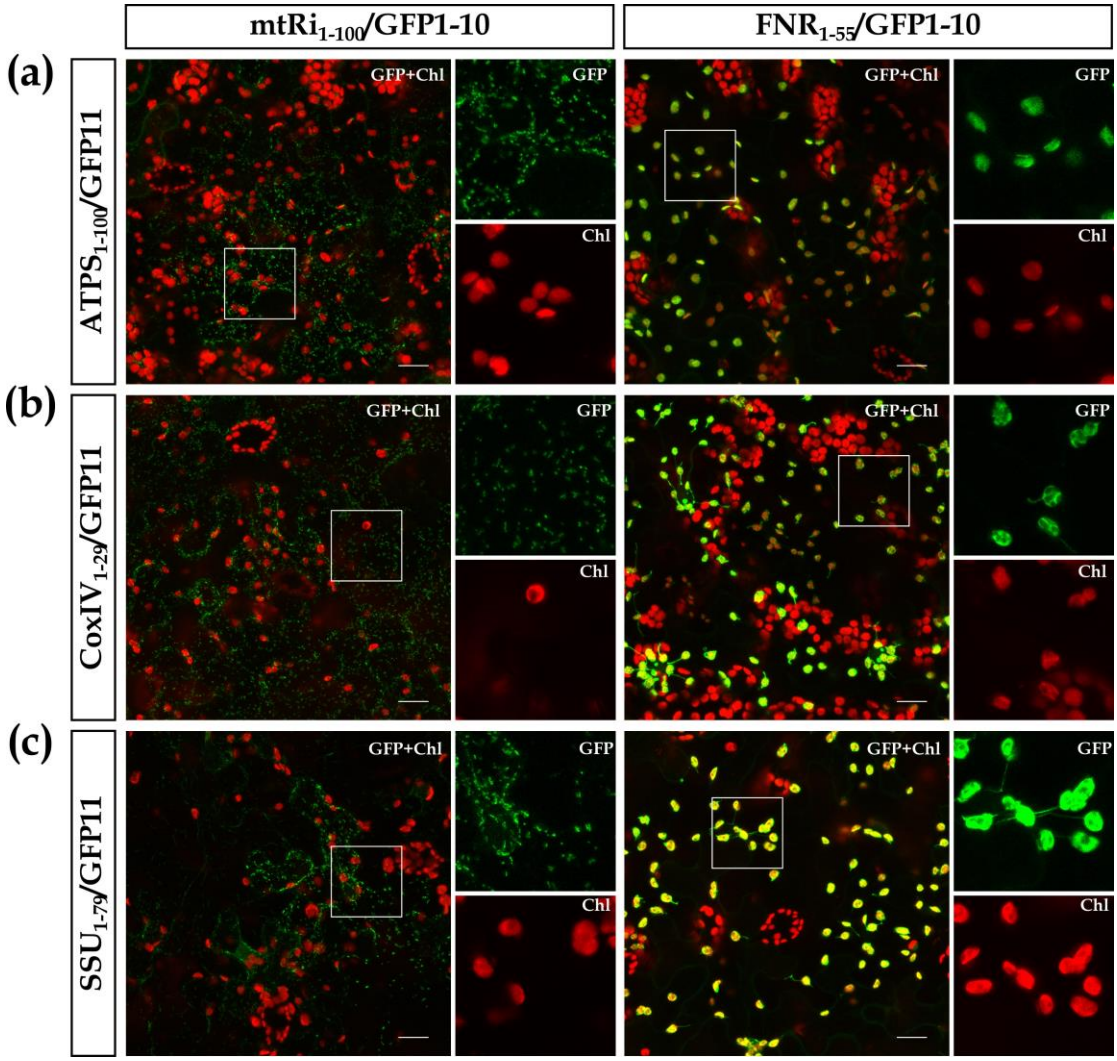


**Figure 8-** Dual localization of presumed monospecific candidate proteins. Gene fragments encoding the transport signals of either **(a)** ATPS, **(b)** CoxIV or **(c)** SSU, were cloned upstream of the $GFP11_{x7}$ tag in the two PlaMiNGo vectors comprising either $mtRi_{1-100}$/GFP1-10 (*left panels*) or $FNR_{1-55}$/GFP1-10 (*right panels*). The resulting constructs were used to transform the lower epidermis of *Nicotiana benthamiana* leaves via *Agrobacterium* infiltration. For further details, see the legend of Figure 2. Scale bars correspond to 20 µm.

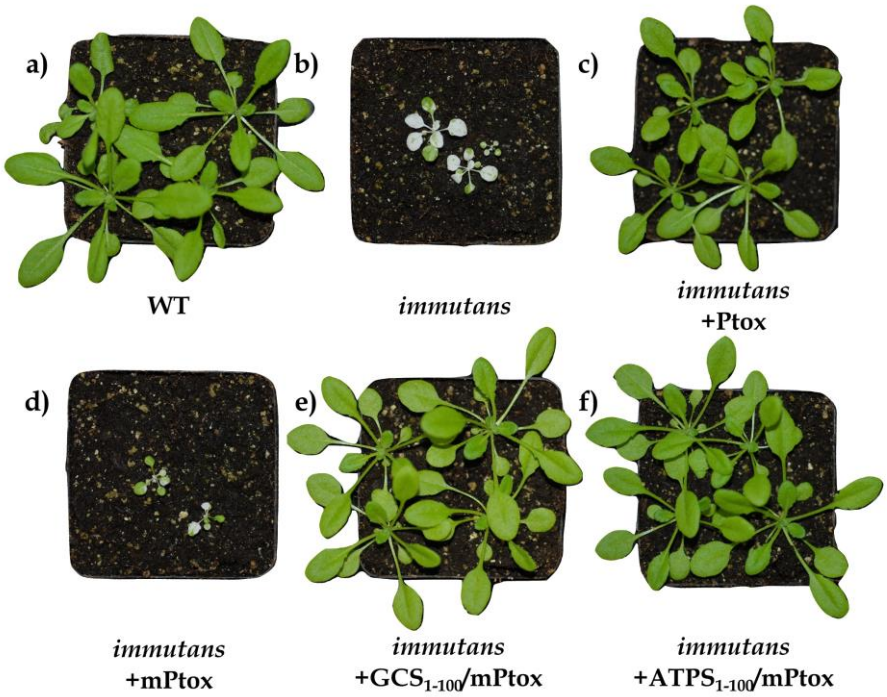


**Figure 9-** Complementation analysis of the *immutans* mutant of *Arabidopsis thaliana*. Phenotype of **(a)** wild type Col-0 and **(b)** *immutans* mutant plants. Gene constructions encoding either **(c)** the precursor PTOX (Ptox), **(d)** the mature PTOX (mPTOX) or PTOX fused to the transit peptides of the candidate proteins **(e)** GCS or **(f)** ATPS, were used to transform *immutans* mutant plants for generation of complementation lines. At least three independently transgenic lines were examined for variegated phenotype complementation of the variegated *immunats* phenotype. The figure shows representative 5 weeks old T2 generation plants.

**Supplementary File 1-**

Supplementary Figure 1-

Confocal laser scanning microscopy of the lower epidermis of *Nicotiana benthamiana* infiltrated with *Agrobacterium tumefaciens* strain GV3101 carrying constructs encoding (a) FNR1-55/GFP1-10 or (b) $mtRi_{1-55}$/GFP1-10 expressed under control of CaMV35S promoter. In both instances, no fluorescence signals were obtained. For further details, see the legend of Figure 2.

Supplementary Figure 2-

Quantification of fluorescence signal intensity. Comparative analysis of signal intensities obtained in **(a)** plastids and **(b)** mitochondria with the use of co-infiltration and PlaMiNGo constructs. Image acquisition was done with 20-x objective and 2% of full laser power in

several equally separated Z-stacks. Each bar represents the average pixel brightness from nine different digital images obtained from 2-3 independent experiments. Measure feature of Fiji was utilized for quantification of signals. (A.U. - Arbitrary units)

Supplementary Figure 3-
Assembly of *Golden Gate*-based vectors and construction of the PlaMiNGo toolkit. The Modular Cloning system (Weber et al., 2011) and DNA modules from the Plant Parts I and II toolkits (Engler et al., 2014; Gantner et al., 2018) were used for vector assembly. All generated modules and vectors are summarized here. Level 0 modules were generated using gene specific primers (summarized in supplementary table 3). Oligonucleotides for $GFP11_{x1}$ fragment were synthesized and hybridized before cloning. Five independent vector constructions (pTEI126, pTEI127, pTEI161, pTEI128, pTEI129) were generated in level I cloning steps. Two constructions namely pTEI176 and pTEI177 were generated additionally to extend the toolkit for the analysis of protein targeting to different sub-cellular compartments. The level I modules were then used to generate final vectors for analysis of protein targeting specificity to plastids (pLaNGo11-1, pLaNGo11-3, pLaNGo11-7) or mitochondria (pMiNGo11-1, pMiNGo11-3, pMiNGo11-7). These vectors carry a *ccd*B negative selection cassette flanked with *Bsa*I restriction site, which is replaced by cloning of the candidate transport signal.

Supplementary Figure 4-
Subcellular localization of the candidate proteins **(a)** $ATPS_{1-100}$/eYFP and **(b)** $GCS_{1-100}$/eYFP, as determined by confocal laser scanning microscopy in leaf tissue of transgenic *Arabidopsis thaliana* lines. All images are maximum intensity projections of several single images representing the complete cell in z-axis. Overlay pictures of both the eYFP channel (displayed in yellow) and the chlorophyll channel (displayed in red) are shown. Separate images of the two channels are displayed at higher magnification in the middle panel (chlorophyll) and left panel (eYFP). A few chloroplasts are encircled with a while line for better display of signals. The scale bars correspond to 10 µm.

Supplementary Table 1- Primer Sequences (for cloning in pRT100 mod vector)
Supplementary Table 2- DNA sequence of (GFP1-10, YFP1-10, GFP11 tags and organelle transit peptides)
Supplementary Table 3- PCR Primers for cloning of Level 0 modules