Proximity labeling of protein complexes and cell type-specific organellar proteomes in Arabidopsis enabled by TurboID

Andrea Mair 1,2, Shou-Ling Xu 3, Tess C. Branon 1,4,5,6, Alice Y. Ting 1,5,6,7, Dominique C. Bergmann 1,2

1 Department of Biology, Stanford University, Stanford
2 Howard Hughes Medical Institute
3 Carnegie Institution for Science - Department of Plant Biology, Stanford
4 Department of Chemistry, Massachusetts Institute of Technology, Cambridge
5 Department of Genetics, Stanford University, Stanford
6 Department of Chemistry, Stanford University, Stanford
7 Chan Zuckerberg Biohub, San Francisco

Abstract

Defining specific protein interactions and spatially or temporally restricted local proteomes improves our understanding of all cellular processes, but obtaining such data is challenging, especially for rare proteins, cell types, or events. Proximity labeling enables discovery of protein neighborhoods defining functional complexes and/or organellar protein compositions. Recent technological improvements, namely two highly active biotin ligase variants (TurboID and miniTurbo), allowed us to address two challenging questions in plants: (1) what are in vivo partners of a low abundant key developmental transcription factor and (2) what is the nuclear proteome of a rare cell type? Proteins identified with FAMA-TurboID include known interactors of this stomatal transcription factor and novel proteins that could facilitate its activator and repressor functions. Directing TurboID to stomatal nuclei enabled purification of cell type- and
subcellular compartment-specific proteins. Broad tests of TurboID and miniTurbo in Arabidopsis and *N. benthamiana* and versatile vectors enable customization by plant researchers.

**Key words:** proximity labeling, BioID, TurboID, miniTurbo, plant, protein complex composition, nuclear proteome, organellar proteome, cell type-specific proteome

**Recurring abbreviations:** AP, affinity purification; AP-MS, affinity purification-mass spectrometry; bHLH, basic helix-loop-helix; BirA*, promiscuous E. coli biotin ligase; FDR, false discovery rate; GC, guard cell; GO, gene ontology; HAT, histone acetyl transferase; HDAC, histone deacetylase; LC-MS/MS, liquid chromatography coupled to tandem mass spectrometry; MS, mass spectrometry; mTb, miniTurbo; NES, nuclear export signal; NLS, nuclear localization signal; PCA, principal component analysis; PL, proximity labeling; POI, protein of interest; SA, streptavidin; TF, transcription factor; TbID, TurboID; WT, wild-type; Y2H, yeast-2-hybrid
Introduction

All major processes of life, including growth and development and interactions among cells, organisms and the environment, rely on the activity and co-operation of hundreds of proteins. To fully understand these processes on a cellular level, we must know all players present in a cell or cell type at a specific location and time. This requires information about transcription and chromatin state, as well as about protein abundance and protein complex compositions. A large international effort, the ‘human cell atlas’ project, is taking a first step in this direction. It aims to characterize all cell types in the human body, using recent advancements in high-throughput single cell and multiplex techniques (Regev et al. 2017; Stuart and Satija 2019). Following this example, a call for a ‘plant cell atlas’ describing nucleic acid, protein and metabolite composition of different cell types in plants was issued at the start of this year (Rhee, Birnbaum, and Ehrhardt 2019). While several groups have produced single cell gene expression profiles (e.g. (Efroni et al. 2015; Ryu et al. 2019; Denyer et al. 2019; Nelms and Walbot 2019)) and tissue/cell type-specific profiles of active translation (e.g. (Vragovic et al. 2015; Tian et al. 2019)), we lack effective tools to obtain similarly precise information about protein distribution, abundance and the composition of protein complexes.

Today’s state-of-the-art for identification of in-planta protein interactors and complexes is affinity purification with protein-specific antibodies or single- and tandem affinity purification tags and subsequent mass spectrometry (MS) analysis (Struk et al. 2019; Xu et al. 2010). However, while affinity purification-mass spectrometry (AP-MS) strategies are undoubtedly useful in many instances, AP-MS is challenging for very low abundant proteins, those expressed only in rare cell types or developmental stages, and those with poor solubility like integral membrane proteins. Moreover, AP-MS tends to miss weak and transient interactions unless paired with
crosslinking (Qi and Katagiri 2009; Van Leene et al. 2007; Bontinck et al. 2018). For obtaining subcellular proteomes, typical traditional approaches rely on cell fractionation protocols that enrich organelles from whole tissues, followed by protein extraction and MS analysis. Besides cross-contamination with other organelles, cell fractionation has the issue that only compartments that can be purified are accessible (Agrawal et al. 2011). The usefulness of this strategy for studying local protein compositions in individual cell types is further limited by the requirement of prior enrichment of the cell type of interest. For rare and transient cell types, acquiring a sufficient amount of ‘pure’ organelle material for MS analysis would be a major challenge.

Recent technological innovations in the form of proximity labeling (PL) techniques provide the sensitivity and specificity needed to identify protein complexes and (local) proteomes on a cell type-specific level. These techniques employ engineered enzymes to covalently attach biotin to nearby proteins which can then be affinity purified from total protein extracts using streptavidin-coupled beads without the need for crosslinking to stabilize weak and transient protein interactions or cell sorting and fractionation to enrich organelles. Because proteins do not have to be isolated in their native state, harsher extraction and more stringent wash conditions can be applied, which can reduce false positives from post-lysis interactions or from non-specific binding of proteins to the beads, and can improve solubilization of membrane proteins (for recent review see (Gingras, Abe, and Raught 2019)). An ever growing number of applications for PL in animals, including the characterization of protein complexes (e.g. the nuclear pore complex (Kim et al. 2014)), of organellar proteomes (e.g. mitochondrial matrix and intermembrane space (Rhee et al. 2013; Hung et al. 2016)) as well as for local proteomes (e.g. inside cilia (Mick et al. 2015), at synaptic clefts (Loh et al. 2016) or at endoplasmic reticulum-mitochondria contact sites (Cho et al. 2017; Hung et al. 2017), demonstrate the usefulness and versatility of these techniques. Although the potential utility for PL in plants is equally
tremendous, its adaptation in plants has been slower, with the successes being limited primarily to systems employing transient and/or high expression. The first three of these (Lin et al. 2017) (Khan et al. 2018) (Conlan et al. 2018) utilized BioID, which is based on an *E. coli* biotin ligase (BirA) made promiscuous by a point mutation (R118G) to yield BirA*. BirA* is either fused to a protein of interest or targeted to a desired subcellular localization to mark protein interactors and local proteomes, respectively. When biotin is supplied in the presence of ATP, BirA* binds and activates biotin and releases reactive biotinyl-AMP which can covalently bind to close-by primary amines on lysine residues (Roux et al. 2012). Based on experiments with the nuclear pore complex (Kim et al. 2014), the labeling radius of BirA* was estimated to be approximately 10 nm, which is in the size-range of an average globular protein. The actual labeling range may vary between experiments and is dependent on characteristics of the BirA* fusion protein, such as linker length and bait mobility, and on the labeling time.

Structural features of plants, including cell walls and the cuticle, most plants’ growth temperature optima and the fact that plants produce and store biotin in their cells, provide major impediments to BioID-based PL systems (Bontinck et al. 2018). Recent engineering of improved versions of BirA, however, might improve PL efficiency in plants and provide the tools required to build a ‘plant cell atlas’. In a directed evolution approach, two new variants – TurboID and miniTurbo – with similar specificity as BirA* but greatly increased activity and lower temperature requirements were created (Branon et al. 2018). TurboID and miniTurbo were successfully utilized to generate different organellar proteomes in HEK cells and work *in vivo* in a broad range of species, including *Drosophila* and *C. elegans* which were previously inaccessible for BioID (Branon et al. 2018). While this work was in review, a new manuscript demonstrating the effectiveness of TurboID in tobacco plants (Zhang et al. 2019) was published. However, TurboID was again highly expressed and the activity of miniTurbo was not tested. Moreover, the
potential of TurboID to identify rare protein complexes or local proteomes in individual cell types of a complex multicellular organisms had not yet been addressed.

Here, we show that TurboID and miniTurbo enable effective PL in a wide variety of tissues and expression levels in plants. We use TurboID to identify partners of the stomatal guard cell transcription factor FAMA and to obtain the nuclear proteome of a rare cell type in Arabidopsis seedlings – young stomatal guard cells. Our work indicates high sensitivity of the new PL enzymes in plants and is complementary to recent work (Zhang et al. 2019) demonstrating high specificity in a plant-pathogen signaling context. To enable adoption by the plant research community, we provide reagents and a broadly applicable workflow for PL experiments under different experimental conditions in a variety of tissues in Arabidopsis and N. benthamiana and highlight critical steps in experimental design and execution.
Highlights box

- TurboID (TbID) and miniTurbo (mTb) work well in all tested tissues and growth stages of stably transformed Arabidopsis and in transiently transformed *N. benthamiana* leaves.
- Labeling times of under 10 minutes can give immunoblot-detectable signals, but longer incubation may be required for protein identification by mass spectrometry (MS).
- In Arabidopsis, TbID activity is higher than mTb activity, but ‘background’ labeling with endogenous biotin is also increased.
- Biotin concentrations in the range of 2.5-50 µM and 20-50 µM are suitable for enhanced labeling with TbID and mTb. For most Arabidopsis tissues, submergence in the biotin solution is sufficient but some tissues and other plants may require (vacuum) infiltration of the biotin solution for optimal labeling.
- TbID and mTb work at temperatures compatible with normal plant growth and at elevated temperatures, but are most likely not suitable for cold stress experiments.
- Proximity labeling (PL) with the FAMA-TbID fusion protein led to the identification of new putative co-activator and -repressor complex components for FAMA, a transcription factor in young guard cells.
- PL with nuclear TbID produced general and young guard cell-specific proteomes with high specificity for nuclear proteins and identified guard cell specific transcription factors.
- Important considerations for the experimental design of PL experiments: choice of biotin ligase (TbID vs. mTb); proper controls to distinguish specific labeling from background; optimization of labeling conditions (biotin concentration, treatment time); optimization of bead amount for affinity purification
- Important considerations for affinity purification and MS identification of biotinylated proteins: depletion of free biotin to reduce required bead amount (buffer exchange with gel filtration columns or centrifugal filters, dialysis); beads for affinity purification (avidin-, streptavidin-, or neutravidin beads vs. anti-biotin antibodies); MS sample prep (on-bead trypsin digest vs. elution and in-gel digest); MS and quantification method (label-free vs. isotopic labeling)
Results

TurboID and miniTurbo can biotinylate plant proteins under conditions appropriate for plant growth

With increased efficiency, BioID-based PL would be a valuable tool to study protein interactions and local proteomes on a cell type-specific level in plants. We therefore made an initial diagnosis of whether the improved BirA* variants TurboID and miniTurbo (hereafter called TbID and mTb) are appropriate for PL applications in plants, by testing their activity in the nucleus and cytosol of two plant model systems: transiently transformed N. benthamiana leaves and young seedlings of stably transformed Arabidopsis. To enable a comparison with previous experiments in the literature, we also included the original BirA* in our experiments and expressed all three versions under the ubiquitous UBIQUITIN10 (UBQ10) promoter. A YFP tag was added to confirm correct expression and localization of the TbID and mTb biotin ligases (Figure 1 – figure supplements 1 and 2).

To test biotin ligase activities, we treated leaf discs from N. benthamiana leaves or whole five day old transgenic Arabidopsis seedlings, each expressing TbID or mTb, with biotin and subsequently monitored biotinylation in total protein extracts using streptavidin immunoblots. For biotin treatment, we briefly vacuum infiltrated the plant tissue with the biotin solution and then incubated the tissue submerged in the solution for one hour at room temperature (approximately 22°C). Mock- or untreated plants were used as controls (Figure 1A). In TbID- and mTb-expressing N. benthamiana and Arabidopsis, biotin treatment induced strong labeling of proteins, demonstrating that the new biotin ligases work under our chosen conditions in plants. As was observed in other organisms, both TbID and mTb showed greatly increased activity compared to BirA*, which mainly achieved weak self-labeling within one hour of biotin
treatment (Figure 1B-C, Figure 1 – figure supplements 3 and 4). Since plants produce and store free biotin in their cells, we were concerned about 'background' labeling in the absence of exogenous biotin. Although it did appear, background labeling was in most cases negligible. Direct comparison of TbID and mTb in our plant systems revealed little difference in either activity or background labeling in N. benthamiana (Figure 1B, Figure 1 – figure supplement 3), possibly due to the high expression levels of the constructs. In Arabidopsis, however, TbID was clearly more active than mTb but also produced more background. Enhanced activity of in the absence of exogenous biotin was especially evident in lines that express TbID and mTb at low levels (Figure 1C, Figure 1 – figure supplement 4). Comparing nuclear and cytosolic constructs, we did not observe any significant differences in labeling efficiency at the resolution of immunoblots (Figure 1 – figure supplements 3 and 4).

From these experiments, we conclude that both TbID and mTb are well suited for use in plants. Which version is more suitable may depend on the individual question and whether high sensitivity (TbID) or tighter control over labeling time (mTb) is important. For this current study, we generated a versatile set of gateway-compatible entry and destination vectors that can be used to express TbID or mTb alone or as fusion with a protein of interest under a promoter of choice (Figure 1 – figure supplement 5). This ‘toolbox’ is accessible through Addgene (available vectors are listed in the materials and methods section).

Testing boundaries with TurboID – effects of labeling time, temperature, biotin concentration and application

Achieving an optimal enzyme efficiency by using the right experimental conditions, like labeling time, temperature, biotin concentration and mode of application, can be key for using PL with low-abundant proteins in plants. We therefore tested the effect of those parameters on biotin
labeling in four to five day old Arabidopsis seedlings expressing TbID and mTb under the UBQ10 promoter. In mammalian cell culture, 10 minutes of labeling with TbID were sufficient to visualize biotinylated proteins by immunoblot and to perform analysis of different organellar proteomes (Branon et al. 2018). Using immunoblots, we observed similarly fast labeling in plants. TbID induced labeling of proteins over background levels within 15-30 minutes of treatment with 250 or 50 µM biotin at room temperature (22°C) and labeling steadily increased over the next three to five hours (Figure 2A, Figure 2 – figure supplement 1). An increase in self-labeling of TbID was evident even earlier, after as little as five minutes (compare Figure 4 – figure supplement 3). Time course experiments in N. benthamiana suggest that mTb is equally fast, with clear labeling of proteins visible within 10 minutes of treatment with 50 µM biotin (Figure 1 – figure supplement 3). This is a significant improvement over BirA*, for which labeling times of 24 hours were applied in all three published plant experiments (Khan et al. 2018; Conlan et al. 2018; Lin et al. 2017).

We systematically tested the effect of different biotin treatment temperatures on TbID and mTb activity in Arabidopsis seedlings. Encouragingly, the activity of both variants was nearly as high at 22°C as at 30°C. Moreover, TbID showed only a moderate increase of activity at 37°C, while mTb activity was actually reduced at this temperature (Figure 2B). High activity at ambient temperatures was also observed in N. benthamiana (Figure 2 – figure supplement 2). Increasing temperatures above plant growth conditions to improve labeling is therefore not needed.

The biotin concentration used for PL is an important consideration. Endogenous levels of biotin in plants are sufficient for low-level labeling of proteins by TbID, and to some extent also by mTb. While this may be useful for some applications, most applications will require strongly enhanced and time-regulated labeling through the addition of exogenous biotin. Although using
excessive amounts of biotin is inconsequential for immunoblots, it poses a problem for downstream protein purification with streptavidin beads, as will be discussed later. We therefore tested biotin concentrations ranging from 0.5 to 250 µM to determine the optimal substrate concentration for TblID and mTb. We found that TblID has a larger dynamic range than mTb. Weak over-background labeling could already be seen with 0.5 µM biotin, which increased weakly through 20 µM, followed by a steeper increase with 30 µM and was more or less saturated at 50-75 µM. mTb required between 2.5 and 10 µM biotin for weak activity, showed a steep increase with 20 µM (comparable to TblID) and was also saturated at 50-75 µM (Figure 2C, Figure 2 – figure supplement 3). TblID and mTb are therefore comparable to BirA* in their biotin requirement and concentrations of 2.5-50 µM (TblID) and 20-50 µM (mTb) seem to be appropriate.

In initial experiments, we vacuum infiltrated the plant material with the biotin solution to maximize biotin uptake. At least in Arabidopsis seedlings, this is not necessary. Simply submerging the plantlets in the biotin solution resulted in the same amount of labeling as vacuum infiltration followed by incubation in the biotin solution did (Figure 2D). This finding is very important since it not only simplifies handling of the experiment, but also improves isolation of labeled proteins by reducing the amount of free biotin in the tissue.

TurboID works in a wide variety of developmental stages and tissues

For TblID to be widely applicable it must be able to biotinylate proteins in many developmental stages and plant tissues. One initial concern, especially with TblID, was that background labeling from endogenous biotin would accumulate over time, making timed experiments with older tissues unfeasible. This was, however, not the case. Labeling worked well in 4 to 14 day old plate-grown seedlings without significant increase of background (Figure 3 – figure
supplement 1). The same was true for separated roots and shoots of 6 to 14 day old seedlings and even for rosette leaves and flower buds of adult Arabidopsis plants grown on soil (Figure 3, Figure 3 – figure supplements 2 and 3). Background activity was low, especially in leaf tissue, and labeling worked well. Vacuum infiltration was not required for the tested plant sample types, except for unopened floral buds, where infiltration improved labeling relative to submergence in the biotin solution (Figure 3). This is likely because petals and reproductive tissues are not in direct contact with the biotin solution. Overall, our experiments suggest that TbID will be applicable in a wide range of developmental stages and tissues. Since TbID and mTb behaved similar in most experiments, it is likely that the same is true for mTb.

Testing TurboID’s potential to identify partners of a very low-abundant TF and to explore the nuclear proteome of a rare and transient cell type

After confirming general applicability of TbID for PL in plants, we wanted to test its performance for the identification of rare protein complexes and the characterization of cell type-specific organellar proteomes in a real experiment. For this purpose, we chose a cell type-specific transcription factor (FAMA) and a subcellular compartment of a rare cell type (nuclei of FAMA-expressing stomatal cells) for a case study. FAMA is a nuclear basic helix-loop-helix (bHLH) TF that is expressed in young stomatal guard cells (GCs) in the epidermis of developing aerial tissues (Ohashi-Ito and Bergmann 2006). The low abundance of FAMA and FAMA-expressing cells renders identification of interaction partners and cell type-specific nuclear proteins by traditional methods challenging and makes it well suited for a proof-of-concept experiment. Potential FAMA interaction partners were previously identified in yeast-2-hybrid (Y2H) and bimolecular fluorescence complementation studies (Chen et al. 2016; Kanaoka et al. 2008; Lee, Lucas, and Sack 2014; Li, Yang, and Chen 2018; Matos et al. 2014; Ohashi-Ito and Bergmann 2006), but only few have been confirmed by in vivo functional data.
For our study, we generated plants expressing TblID and a fluorescent tag for visualization under the FAMA promoter, either as a FAMA-protein fusion or alone with a nuclear localization signal (NLS) (Figure 4A, Figure 4 – figure supplement 1). By comparing proteins labeled in the FAMApro::FAMA-TblID-Venus (FAMA-TblID) and FAMApro::TblID-YFP_{NLS} (FAMA_{nuc}TblID) plants to each other and to wild-type (WT) plants and UBQ10pro::TblID-YFP_{NLS} (UBQ_{nuc}TblID) plants, we can test the ability of the system to identify (1) proteins in close proximity to FAMA (FAMA complexes), (2) the nuclear protein composition during the FAMA-cell stage, (3) the nuclear proteome in general, and (4) possible FAMA-stage specific nuclear proteins. The FAMA-TblID-Venus construct is functional, since it complements the seedling-lethal phenotype of the fama-1 mutant (Figure 4 – figure supplement 2).

Determining suitable incubation times is crucial since too short an incubation can yield insufficient protein amounts for identification but excessive incubation could label the whole subcellular compartment. We therefore performed labeling time-courses with the FAMA-TblID line, using immunoblots as a readout. FAMA auto-labeling could be observed after as little as five minutes but clear labeling of other proteins required approximately 15 to 30 minutes. Longer incubation led to further increase in labeling up to three hours, both in the form of stronger discrete bands and of diffuse labeling, but stayed more or less the same thereafter (Figure 4B, Figure 4 – figure supplement 3). Based on these observations, we chose 0.5 and 3 hour biotin treatments, after which we would expect abundant and most FAMA interactors to be labeled, respectively, for the ‘FAMA interactome’ experiment (Figure 4C). For the ‘nuclear proteome’ experiment (Figure 4D), only the longer 3 hour time point was used, since over-labeling of the compartment was not a concern. Accordingly, the FAMA-TblID, FAMA_{nuc}TblID and WT lines were treated for 0, 0.5 and 3 hours, and the UBQ_{nuc}TblID line only for 3 hours. As plant material, we chose seedlings five days post germination, which corresponds to a peak time in FAMA
promoter activity, as determined empirically by microscopy. We used three biological replicates per sample. To make the datasets as comparable as possible, all steps preceding data analysis were done together for the two experiments, as described in the next section.

**From labeling to identification of biotinylated proteins – identifying critical steps**

Through empirical testing of experimental conditions using the **UBQ**\textsubscript{nuc}**TbID** line, we identified steps and choices that have a big impact on success of protein purification and identification after PL with **TbID**. These include sample choice to maximize bait abundance, removal of free biotin, optimizing the amount of streptavidin beads for affinity purification (AP) and choosing among MS sample prep procedures. Below, we describe our experimental procedure (**Figure 4E**) and highlight key choices.

We first labeled five day old seedlings, by submerging them in a 50 µM biotin solution for 0, 0.5 or 3 hours, quickly washed them with ice cold water to stop the labeling reaction and to remove excess biotin and isolated total proteins for AP of biotinylated proteins. The protein extracts were then passed through PD-10 gel filtration columns to reduce the amount of free biotin in the sample before proceeding with AP using magnetic streptavidin beads. Successful labeling and purification was confirmed by immunoblots (**Figure 4 – figure supplements 4 and 5**). The inclusion of a biotin-depletion step was found to be critical as free biotin in the protein extracts competes with biotinylated proteins for binding of the streptavidin beads (**Figure 4 – figure supplement 6**). While for mammalian cell culture or rice protoplasts thorough washing of the cells seems to suffice for removal of free biotin, this is not the case for intact plant tissue (see also (Conlan et al. 2018; Khan et al. 2018)). Especially when large amounts of starting material and moderate amounts of biotin are used, little to none of the biotinylated proteins may be bound by the beads. To maximize the amount of purified proteins it is further advisable to
determine the appropriate amount of beads required for each experiment. We used 200 µl beads for approximately 16 mg total protein per sample. This amount was chosen based on tests with different bead-to-extract ratios (Figure 4 – figure supplement 7) and was sufficient to bind most biotinylated proteins in our protein extracts, although the beads were slightly oversaturated by the highly labeled UbQhucTbID samples (Figure 4 – figure supplement 5D).

Following AP, we performed liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) analysis to identify and quantify the captured proteins. Tryptic digest for LC-MS/MS analysis was done on-beads, since test experiments revealed that elution from the beads using two different methods (Cheah and Yamada 2017; Schopp and Bethune 2018) and subsequent in-gel digestion of biotinylated proteins yielded significantly lower protein amounts and less protein identifications (data not shown). This apparent sample loss is caused by the strong biotin-streptavidin interaction, which allows for stringent washing conditions but also prevents efficient elution of biotinylated proteins from the beads. Notably, highly biotinylated proteins, which likely comprise the most interesting candidates, will interact with more than one streptavidin molecule and will be especially hard to elute. After MS analysis, we identified and quantified the proteins by label-free quantification and filtered for significantly enriched proteins. This part was done separately for the ‘FAMA interactome’ and ‘nuclear proteome’ experiments and is described in the following sections.

Proximity labeling is superior to AP-MS for identification of candidate interactors of FAMA

FAMA acts as both an activator and repressor for hundreds of genes (Hachez et al. 2011), suggesting a need for coordinated action with other TFs, co-activators and -repressors (Matos et al. 2014). Identifying such proteins through classical affinity purification-mass spectrometry
approaches is hampered by the low overall abundance of FAMA. Apart from INDUCER OF CBF EXPRESSION 1 (ICE1), which is a known heterodimerization partner of FAMA (Kanaoka et al. 2008), we failed to identify any transcriptional (co-)regulators by AP-MS with FAMA-CFP, despite the use of crosslinking agents and large amounts of plant material (15 g of four day old seedlings per sample). Moreover, less than 20% of the AP-MS-derived ‘candidates’ were predicted to be nuclear, and one quarter were chloroplast proteins (Figure 5 – figure supplement 1, Supplementary file 2 – Table 5). We therefore wanted to see if PL would improve the identification of biologically relevant FAMA interactors.

For the ‘FAMA interactome’ experiment we compared proteins purified from plants expressing the FAMA-TbID fusion (FAMA-TbID) with proteins from WT and with proteins from plants expressing nuclear TbID (FAManucTbID) after 0, 0.5 and 3 hours of biotin treatment. In total, we identified 2,511 proteins with high confidence (quantified in all three replicates of at least one sample). Principal component analysis (PCA) showed a clear separation of the samples by genotype and time point (Figure 5 – figure supplement 2). Despite this clear separation, the majority of proteins were common to all samples, including the untreated WT control (Supplementary file 2 – Table 1), and moreover, were unchanged between samples (Figure 5 – figure supplements 2 and 3). This indicates that a large proportion of identified proteins bound to the beads non-specifically. This is not uncommon for affinity purification experiments, and underlines the importance of appropriate controls and data filtering pipelines.

To remove these ‘background proteins’ from our dataset and to narrow down the number of FAMA complex candidates, we applied three consecutive filtering steps (Figure 5, for details see methods section). First, we removed proteins that were not significantly enriched in the FAMA-TbID samples compared to WT. These comprise sticky proteins that bind the beads non-specifically and a handful of proteins that are biotinylated natively by Arabidopsis plant biotin
ligases (Alban 2011; Nikolau, Ohlrogge, and Wurtele 2003). This was done by pair-wise comparison of FAMA-TbID and WT samples at each time point, using only proteins that were found in all three replicates of the corresponding FAMA-TbID samples, and resulted in a list of 73, 85 and 239 significantly enriched proteins (including FAMA) at the 0, 0.5 and 3 hour time points, respectively (Figure 5, Supplementary file 2 – Table 2). Since TbID is highly active and endogenous levels of biotin are sufficient for low-level labeling, there is a risk that proteins are labeled stochastically and that, over time, the whole nuclear proteome would be labeled. Notably, more than half of the proteins enriched in the FAMA-TbID plants at any of the time points were also enriched in the FAMA_nucTbID plants (Figure 5, Figure 5 – figure supplement 4, Supplementary file 2 – Table 2). We therefore applied a second filtering step to remove proteins that were not significantly enriched in the FAMA-TbID versus the FAMA_nucTbID samples. Pair-wise comparison of FAMA-TbID and FAMA_nucTbID samples at the three time points, further reduced the dataset to 6, 15 and 57 proteins (including FAMA) (Figure 5, Supplementary file 2 – Table 2). Finally, we removed proteins that were not significantly enriched after biotin treatment compared to the untreated samples, since these proteins are likely genotype-specific contaminations. One protein, which was only enriched in the absence of exogenous biotin but not after biotin treatment, was removed as well.

This left us with 47 ‘high confidence’ candidates (Figure 5, Table 1), 35 of which were previously demonstrated to be in the nucleus using fluorescent protein fusions or were found in MS-based nuclear proteome studies (Figure 5, Table 1, Supplementary file 2 – Tables 2 and 4). Notably, more than half of the candidates have a role in regulation of transcription or are chaperones which could assist in FAMA’s role as a TF or in protein folding and stabilization, respectively (Figure 5, Table 1, Supplementary file 2 – Table 4). Moreover, several of these proteins have previously been shown to interact with each other, which suggests that they could be part of the same FAMA complexes. This is a huge improvement compared to our AP-MS
experiment, which could only confirm FAMA’s interaction with its obligate heterodimerization partner ICE1. The transcriptional regulators we found with PL can be roughly divided into two categories: TFs and transcriptional co-regulators. Among the TFs we again found ICE1 (as well as peptides shared between ICE1 and its orthologue SCRM2). We also found three other bHLH TFs (AIB/JAM1, JAM3, and BIM1) and the non-canonical bHLH-type TF BZR1. AIB and JAM3 play partially redundant roles in negative regulation of jasmonic acid (JA) signaling (Sasaki-Sekimoto et al. 2013; Fonseca et al. 2014), while BIM1 and BZR1 mediate brassinosteroid (BR) signaling (Yin et al. 2002; Wang et al. 2002). Both JA and BR signaling play roles in stomatal function or development (Acharya and Assmann 2009; Gudesblat et al. 2012; Kim et al. 2012).

Among the transcriptional co-regulators we found two significantly enriched transcriptional co-activators: MED16, which is part of the mediator complex that links TFs to RNA Pol II (Kidd et al. 2011), and HAC1, which is a histone acetyl transferase (HAT) (Deng et al. 2007). Combined with previous data showing a link between FAMA and RNA Pol II (Chen et al. 2016), this suggests that FAMA activates genes both directly by recruiting RNA Pol II and by opening up the chromatin for other transcriptional regulators. Among transcriptional co-repressors were TOPLESS (TPL)-related proteins TPR3 and TPR4 and LEUNIG (LUG) and LEUNIG HOMOLOG (LUH), which recruit histone deacetylases (HDACs) to TFs (Long et al. 2006). Additionally, we identified the linker protein SEUSS (SEU), which mediates interaction of LUG and LUH with TFs (Liu and Karmarkar 2008; Sitaraman, Bui, and Liu 2008). The identification of all three members of the SEU/LUG/LUH co-repressor complex is a strong indication of a functional complex with FAMA in the plant. Relaxing our filtering criteria to include proteins that are enriched in the FAMA-TbID vs FAMANucTbID samples but were not significant under our stringent cutoff, we find several more components of transcriptional co-regulator complexes, including three more MED proteins, another HAT, two more TPL-related proteins and TPL itself.
RBR1, a cell cycle regulator and a known interactor of FAMA (Lee, Lucas, and Sack 2014; Matos et al. 2014), is also among the FAMA-TbID enriched proteins but, due to a modest fold change, did not pass our last filters (Figure 5, Supplementary file 2 – Table 2). This suggests that by setting a stringent cutoff on enrichment between FAMA-TbID and FAMAnucTbID, we might lose some true interactors. This might be especially true of ubiquitously expressed proteins with many partners like RBR1, where FAMA-RBR1 interactions are likely to represent only a small fraction of all complexes.

For a small-scale validation of candidates biotinylated by FAMA-TbID, we tested four proteins (co-repressor complex components SEU and LUH and TFs BZR1 and BIM1) for direct interaction with FAMA in a Y2H system. Additionally, we co-expressed SEU and LUH with FAMA-TbID or nuclear TbID in N. benthamiana and tested for biotinylation of the candidates (Figure 5 – figure supplement 5). Our experiments suggest that BIM1 and SEU are direct interaction partners of FAMA, while LUH and BZR1 might be in indirect contact with FAMA or require a specific protein or DNA context to be present.

Overall, our ‘FAMA interactome’ experiment demonstrates the usefulness of PL to identify potential interaction partners of rare proteins. We identified several good FAMA-complex candidates which could support FAMA in its role as a key TF and provide a possible mechanism for FAMA to induce fate-determining and lasting transcriptional changes in developing GCs. Some of the FAMA-complex candidates identified through PL are also slightly enriched in FAMA AP-MS samples compared to their controls. However, the enrichment is not enough to call any of them, except ICE1, significant in the AP-MS experiment. PL therefore not only gave us higher specificity for nuclear proteins than the AP-MS did, but it is potentially more sensitive as well. It is worth noting, that most FAMA interaction candidates were identified at the 3 hour time point
and that longer biotin treatment greatly improved identification of biotinylated proteins (Figure 5, Figure 5 – figure supplements 4, Supplementary file 2 – Table 3).

Proximity labeling can be used to analyze the nuclear proteome in rare FAMA-expressing cells during GC development

The second question our PL experiment should answer was whether TblID could be used to take a snapshot of the nuclear proteome of FAMA-expressing cells. Traditional tools to study organellar proteomes are not well-suited for such an endeavor, since they require isolation of the cell type and organelle of interest and therefore lack the required sensitivity. Branon et al (Branon et al. 2018), showed that TblID can be used to efficiently and specifically purify proteins from different subcellular compartments without prior cell fractionation. Their work was done using a homogeneous population of cultured mammalian cells, however, so it remained to be shown whether it would be possible to isolate an organellar proteome from an individual cell type, especially a rare or transient one, in a complex multicellular organism.

To identify nuclear proteins in FAMA-expressing young GCs and compare them to the global nuclear proteome at this growth stage, we purified proteins from seedlings expressing nuclear TblID under the FAMA (FAMA\textsuperscript{nuc}TbID) and UBQ10 (UBQ\textsuperscript{nuc}TbID) promoter and from WT after three hours of biotin treatment. PCA, hierarchical clustering and multi scatterplots showed a clear separation of the three genotypes (Figure 6 – figure supplements 1 and 2). In total, we identified 3,176 proteins with high confidence (Supplementary file 3 – Table 1). 2,215 proteins were significantly enriched in UBQ\textsuperscript{nuc}TbID compared to WT (Figure 6, Figure 6 – figure supplement 3, Supplementary file 3 – Table 2). These proteins comprise our ‘global’ nuclear protein dataset. Despite the relative rareness of FAMA-expressing cells, the FAMA\textsuperscript{nuc}TbID dataset yielded 394 proteins that were enriched compared to WT (Figure 6, Figure 6 – figure
supplement 3, Supplementary file 3 – Table 3). Notably, most of them overlap with our global nuclear protein dataset (Figure 6), as would be expected since the \( \text{TbID} \) dataset also contains FAMA-stage cells.

To estimate how ‘pure’ our nuclear proteomes are, we curated published nuclear and subnuclear compartment proteomes (Bae et al. 2003; Bigeard et al. 2014; Calikowski, Meulia, and Meier 2003; Chaki et al. 2015; Palm et al. 2016; Pendle et al. 2005; Sakamoto and Takagi 2013; Goto et al. 2019) and searched the Arabidopsis protein subcellular localization database SUBA (version 4, (Hooper et al. 2017), http://suba.live/) for proteins that were observed in the nucleus as fluorescent-protein fusions. This resulted in a combined list of 4,681 ‘experimentally determined nuclear proteins’; 4,021 from MS and 975 from localization studies (Supplementary file 3 – Table 4). More than three quarters of the proteins enriched in our \( \text{UBQ} \) and \( \text{FAMA} \) datasets are either experimentally verified nuclear proteins or are predicted to be localized in the nucleus (Figure 6, Supplementary file 3 – Tables 2, 3 and 5). This suggests that most identified proteins are indeed nuclear proteins. Of the remaining proteins, most are predicted to be in the cytosol and could have been labeled by TbID right after translation and before nuclear import of the biotin ligase or by a small mis-localized fraction of TbID. Chloroplast proteins, which are a major source of contamination in plant MS experiments, make up only about 4 % of our identified proteins based on experimental and prediction data (Supplementary file 3 – Table 5). For a comparison, about 12 % and 6 % of the proteins identified in the two most recent Arabidopsis nuclear proteome studies (Palm et al. 2016; Goto et al. 2019), are predicted to be in the chloroplast (SUBAcon prediction, SUBA4). Gene ontology (GO) analysis is also consistent with nuclear enrichment in both nuclear TbID datasets (Figure 6 – figure supplement 4, Supplementary file 3 – Tables 2, 3 and 7). Importantly, our nuclear TbID successfully labeled all major sub-nuclear compartments and domains, including the nuclear pore complex, the nuclear envelope, the nuclear lamina, the nucleolus and other small
speckles, as well as DNA- and chromatin-associated proteins (subdomain markers from (Petrovska, Sebela, and Dolezel 2015; Tamura and Hara-Nishimura 2013), see Supplementary file 3 – Table 8 for examples).

After the general assessment of data quality and nuclear specificity, we asked whether we could identify known markers for FAMA-stage GC nuclei from our dataset by comparing the $\text{FAMA}_{\text{nucTbID}}$ to the $\text{UBQ}_{\text{nucTbID}}$ samples (Figure 6, Figure 6 – figure supplement 3, Supplementary file 3 – Table 3). Unlike ubiquitously expressed proteins, which should be enriched in the $\text{UBQ}_{\text{nucTbID}}$ samples, FAMA-cell specific or highly enriched proteins should be equally abundant in both sample groups. Indeed, looking at proteins that were equally abundant or slightly enriched in the $\text{FAMA}_{\text{nucTbID}}$ samples, we find several known stomatal lineage- and GC-associated TFs, namely FAMA itself, HOMEODOMAIN GLABROUS 2 (HDG2) and STOMATAL CARPENTER 1 (SCAP1) (Ohashi-Ito and Bergmann 2006; Peterson et al. 2013; Negi et al. 2013). Additionally, there were 10 proteins among the 44 highly $\text{FAMA}_{\text{nucTbID}}$ enriched proteins, that were previously identified as part of the GC proteome by Zhao and colleagues using MS analysis of 300 million GC protoplasts (Zhao et al. 2010; Zhao et al. 2008) (Supplementary file 3 – Tables 3 and 6). It is worth noting that neither FAMA nor SCAP1 are in the Zhao GC protoplast proteome, presumably due to their relatively low expression and the use of material from more mature plants. These results suggest that PL can find lowly expressed proteins, that our knowledge of the GC proteome is not complete, and that additional important regulators of GC development and function might be uncovered by looking at stage-specific proteomes.

Overall, this experiment demonstrates the usefulness of TbID as a tool for studying subcellular proteomes on a whole-plant as well as on a cell type-specific level. This will allow us to address
questions that were previously inaccessible and thus has the potential to greatly improve our understanding of cellular processes in a cell type-specific context.

Discussion

Suitability of TurboID and miniTurbo for application in plants and performance of TurboID in FAMA-complex identification and nuclear proteome analysis

Our experiments presented in this study demonstrate that the new biotin ligase versions TbID and mTb drastically improve the sensitivity of PL in plants, compared to previously used BirA*, and tolerate a range of experimental conditions. We observed rapid labeling of proteins by TbID and mTb in different species, tissues and at different growth stages from seedlings to mature plants, using a simple biotin treatment protocol at room temperature. This greatly broadens the range of possible PL applications for plants and will allow to address hitherto inaccessible or hard to address questions in the future. To test the usefulness of TbID for identification of protein interactors and organellar proteomes, we aimed to identify FAMA protein complexes and the nuclear proteome in FAMA-expressing cells. We deliberately picked a rare protein and cell type to observe its performance under conditions that make the use of traditional methods challenging or unfeasible.

In spite of FAMA’s low abundance, PL with the FAMA fusion protein in our ‘FAMA interactome’ experiment worked very well and outperformed our FAMA AP-MS experiments both in sensitivity and in specificity for nuclear proteins. Unlike in the AP-MS experiments, we identified a number of new proteins with gene regulatory functions that could support FAMA in its role as
a master regulator of GC differentiation. Beyond the known dimerization partner ICE1, which was found in both types of experiments, PL identified four additional bHLH TFs with the potential to form alternative FAMA heterodimers as well as three non-bHLH TFs. Strong, direct interaction with BIM1 could be confirmed by Y2H (Figure 5 – figure supplement 5A). We further identified several epigenetic regulators, which could fulfill roles predicted for FAMA complexes based on previous genetic and transcriptomic experiments (Adrian et al. 2015; Hachez et al. 2011; Matos et al. 2014) and fill a gap in our current model of gene regulation during GC formation. It was satisfying to see that when FAMA complex candidates were known to form functional complexes, we often identified multiple components of the complex. One example is the SEU-LUG/LUH co-repressor complex, of which SEU acts as an adapter protein, linking LUG/LUH to TFs (Liu and Karmarkar 2008). Our Y2H and PL experiments support a role of SEU as an adapter in FAMA repressive complexes (Figure 5 – figure supplement 5). MED14, which was found as potential FAMA complex component using relaxed criteria, could also be part of this interaction chain. LUG is known to act by recruiting HDACs to promote epigenetic gene repression, but can also interact with mediator complex subunits like MED14 to interfere with their interaction with transcriptional activators (Liu and Karmarkar 2008). MED14, if indeed a FAMA complex component, could therefore either interact with the FAMA heterodimer or with the repressors LUG or LUH. It is likely that our candidates are part of different activating and repressive FAMA complexes. Which proteins are in which complex and which are direct FAMA interactors will need to be further tested with independent methods.

Interestingly, we did not confirm all of FAMA’s previously postulated interaction partners in our PL experiment. This has both biological and technical causes. It is possible that some previously observed interactions are too rare or conditional (e.g. MED8 may require pathogen exposure) while others probably only happen under artificial conditions like in Y2H assays (e.g. bHLH71 and bHLH93, which seem not to be required for stomatal development (Ohashi-Ito and
Bergmann 2006)). Enrichment of NRPB3 in FAMA-TbID samples, on the other hand, could not be assessed because of non-specific binding of the protein to the beads. Other technical aspects that can complicate or prevent identification of interaction partners are removal of low-frequency interactions through stringent data filtering or a general lack of labeling. Because only proteins with exposed, deprotonated lysine residues can be labeled, sterically or chemically inaccessible proteins will not be detected.

PL of FAMA-expressing nuclei and comparison with a general nuclear proteome in our ‘nuclear proteome’ experiment showed that TbID is suitable to capture sub-cellular proteomes at a cell type-specific level, even when the cell type is rare. We identified 2,232 proteins in both nuclear datasets combined, which is 25% more than the most recent nuclear proteome obtained from cultured Arabidopsis cells (1,528 proteins, Goto et al. 2019). Judging from experimentally determined and predicted localization of the identified proteins and functional annotation with GO analysis, we obtained high specificity for nuclear proteins and identified proteins from all major nuclear subdomains. The biggest ‘contamination’ stems from the cytosol, which could be caused by a fraction of TbID in the cytosol (e.g. right after translation) or from activated biotin diffusing out of the nucleus. Chloroplast-predicted proteins made up only a small fraction. Among FAMA-nuclei enriched proteins, we identified several known nuclear markers of young GCs, confirming our ability to detect cell type-specific proteins, as well as proteins that have not yet been linked to GC development or function but could be interesting to investigate further. One of them is SHL (SHORT LIFE), which is a histone reader that can bind both H3K27me3 and H3K4me3 histone marks and has been implicated in seed dormancy and flower repression (Mussig et al. 2000; Qian et al. 2018). Interaction of FAMA with RBR1 and with newly identified co-repressors and co-activators from this study, strongly suggests that chromatin marks are important to lock GC in their terminally differentiated state (Matos et al. 2014; Lee, Lucas, and Sack 2014) and SHL could be involved in this process.
Considerations for a successful PL experiment

When designing a PL experiment, several things should be considered in order to achieve the best possible result. First, a choice has to be made which biotin ligase to use. Whether TbID or mTb is more suitable will depend on the research question. TbID is more active, which is an advantage for low abundant proteins, for cases where over-labeling is not a concern or when labeling times should be kept as short as possible. mTb is less active in the presence of endogenous levels of biotin and will give less background labeling, which could be beneficial in tissues with higher than average endogenous biotin levels (Shellhammer and Meinke 1990) or when a more restricted and controlled labeling time is desired. Additional factors that should be considered when choosing a biotin ligase are whether one version works better in a specific subcellular compartment (as was the case in human HEK cells (Branon et al. 2018)) and whether the larger TbID interferes with the activity or correct subcellular targeting of the protein of interest (POI). We added a fluorophore to all our biotin ligase constructs. This is very useful to confirm correct expression and subcellular localization of the TbID or mTb fusion protein, but may in some cases affect the activity of the ligase or the tagged protein. Interference with activity and targeting can depend on the position of the biotin ligase relative to the POI (N- or C-terminal tag). Another decision at the construct-design phase is the choice of linker length. For our experiments we added a short flexible linker to TbID. For identification of large protein complexes, increasing the linker length may improve labeling of more distal proteins.

Another crucial consideration are controls. Extensive and well-chosen controls are essential to distinguish between true candidates and proteins that either bind non-specifically to the beads or that are stochastically labeled because they are localized in the same subcellular compart as the POI. The former class of contaminants can be identified by including a non-transformed
control (e.g. WT). The best control for the latter will be situation-dependent. For identifying interaction partners of a POI, one could use free TbID or mTb targeted to the same subcellular localization as the POI, as we have done in our ‘FAMA interactome’ experiment. Alternatively, one or more unrelated proteins that are in the same (sub)compartment but do not interact with the POI can be used. This strategy might improve identification of proteins that are highly abundant or ubiquitously expressed but of which only a small fraction interacts with the bait. Such proteins may be lost during data filtering if a whole-compartment control is used, as we observed for RBR1 in our ‘FAMA interactome’ experiment. In either case, the control construct should have approximately the same expression level as the POI. For sub-organellar proteomes (e.g. a specific region at the plasma membrane) or for organellar proteomes in a specific cell type, it is useful to label the whole compartment or the compartment in all cell types for comparison.

Before doing a large-scale PL experiment, different experimental conditions should be tested to find a biotin concentration/treatment time/plant amount/bead amount combination that is suitable for the question and budget. Increasing the plant amount, labeling time or biotin concentration can improve protein coverage and can help to get more complete compartmental proteomes, but will also affect the amount of beads required. Excessive labeling is only advisable in closed compartments and when labeling of all proteins is the goal. Immunoblots are a useful tool to test different combinations of labeling concentration and time as well as for determining the correct bead amount for the pulldown so as not to over-saturate the beads. One should keep in mind, though, that the signal intensity on a western blot does not necessarily reflect the amount of labeled protein, because highly biotinylated and proteins have multiple binding sites for streptavidin, which leads to signal amplification. Moreover, labeling strength will also depend on the number of sterically/chemically available sites for biotinylation, and therefore also on the size and properties of a protein.
If biotinylation is induced by addition of exogenous biotin, a crucial step before AP of biotinylated target proteins is the depletion of free biotin in the sample to reduce the amount of beads required and therefore the per sample cost. We tested the effectiveness of two different approaches: gel filtration with PD-10 desalting columns (also used by (Conlan et al. 2018)) and repeated concentration and dilution with Amicon Ultra centrifugal filters (used by (Khan et al. 2018)). The latter method has the potential to remove more biotin and to be more suitable for large amounts of plant material. In our hands, though, using Amicon centrifugal filters led to considerable sample loss, presumably due to binding of the membrane, and was very slow. PD-10 columns, in contrast, did not lead to a notable loss of biotinylated proteins (Figure 4 – figure supplement 7). Surprisingly, consecutive filtering of protein extracts with two PD-10 columns did not improve the bead requirement. An alternative to these two methods is dialysis, which is suitable for larger volumes but is very time consuming. If the target is in an easy-to-isolate and sufficiently abundant organelle, cell fractionation prior to AP might also be considered to remove unbound biotin.

For AP, different kinds of avidin, streptavidin or neutravidin beads are available. Their strong interaction with biotin allows for efficient pulldowns and stringent wash conditions, but makes elution of the bound proteins difficult, especially if they are biotinylated at multiple sites. Should elution of the bound proteins be important for downstream processing or the identity of the biotinylated peptides be of major interest, the use of biotin antibodies might be preferable (Udeshi et al. 2017).

Finally, some consideration should also be given to the MS strategy. For example, digesting the proteins on the streptavidin beads instead of eluting them can increase the peptide yield. For data analysis, label-free quantification produces a more quantitative comparison of samples
than comparison of peptide counts. Care should be taken that an appropriate data normalization method is chosen, especially when samples are very different, as can be the case when different cellular compartments are compared. Isotopic labeling, which allows samples to be analyzed together, can further improve quantitative comparison.

Potential applications and challenges for PL in plants

We can see many potential applications for PL in plants, extending beyond the ones presented in this work. One that has the potential to be widely used is in vivo confirmation of suspected protein interactions or complex formation. The strategy is comparable to currently used co-immunoprecipitation (Co-IP) experiments but has the benefit that weak and transient, as well as other hard-to-purify interactions, are easier to detect. For this approach, proteins closely-related to the bait can be used as controls for interaction specificity.

One of the major applications we demonstrate in this study is de novo identification of protein interaction partners and complex components. Extending from that, PL can be used to observe changes in complex composition in response to internal or external cues (e.g. stress treatment). Currently used techniques like peptide arrays, two-hybrid screens in yeast or plant protoplasts and AP-MS have the disadvantage that they are either artificial or work poorly for low abundant and membrane proteins and tend to miss weak and transient interactions. PL could overcome some of their deficiencies. One should keep in mind, however, that rather than identifying proteins bound directly to a bait protein, PL will mark proteins in its vicinity. Labeling is generally strongest for direct interactors, but labeling radius will depend on properties of the bait such as size, mobility and linker length as well as on the duration of labeling. To define a protein interaction network, it will be useful to use several different baits (Gingras, Abe, and Raught 2019).
Another application of PL is characterization of subcellular proteomes, such as whole organellar proteomes, as we have demonstrated for the nucleus. Going forward, more detailed characterization of different organellar proteomes as well as sub-organellar proteomes and local protein composition, for example at membrane contact sites, can and should be addressed. This can be done on a whole plant level, but also at organ- and even cell type-specific level. Importantly, PL enables investigation of previously inaccessible compartments and of rare and transient cell types as we have demonstrated for FAMA-expressing GCs. Differences between individual cell types or treatments can be investigated as well. Labeling times will, among other things, depend on the 2D or 3D mobility and distribution of the bait. For whole-compartment labeling, a combination of several baits may increase efficiency and protein coverage. One drawback of PL compared to traditional biochemical methods is that it requires the generation of transgenic plants, which limits its use to plants that can currently be transformed.

PL can also be used in combination with microscopy, to visualize the subcellular localization of biotinylated proteins and reveal labeling patterns of individual bait proteins. This can be utilized to confirm that labeling is restricted to the desired compartment, but it can also be used to fine-map the subcellular localization of a protein of interest or to obtain information about its topology, as was demonstrated for an ER transmembrane protein in human cells (Lee et al. 2016).

Extended uses of PL techniques that will require some modification of TbID and mTb or the PL protocol before they can be applied include interaction-dependent labeling of protein complexes (split-BioID), identification of RNAs associated with biotinylated proteins and identification of proteins associated with specific DNA or RNA sequences (for a recent review of PL methods describing these applications see (Trinkle-Mulcahy 2019)).
While there are is a plethora of questions that can be addressed with PL, there are also limitations to what will be possible. For example, although TbID and mTb are much faster than BirA*, controlled short labeling pulses (as are possible with APEX-based PL techniques) will be hard to achieve. TbID and mTb are always active and use endogenous biotin to continuously label proteins, preventing a sharp labeling start. In addition, exogenous biotin needs time to enter the plant tissue to initiate labeling. In mammalian cell culture, 10 minutes of labeling might be sufficient, but in a whole multicellular organism it may take much longer, depending on the experimental setup and how complete labeling should be. In our ‘nuclear proteome’ experiment, for example, three hours of biotin treatment were not sufficient to reach labeling saturation and only very few proteins were enriched in FAMA\textsubscript{nuc}TbID samples by 30 minutes biotin treatment (Figure 5 – figure supplement 6, Supplementary file 2 – Table 3). Development of strategies to reduce background labeling, for example by (conditional) reduction of endogenous biotin levels, could improve labeling time control in the future. Another limitation stems from temperature sensitivity. Although TbID and mTb work well at room temperature and elevated temperatures, they are inactive at 4°C (Branon et al. 2018) and thus likely incompatible with cold treatment as might be done for cold adaptation and cold stress experiments. Further, it is likely that some compartments will be harder to work in than others. Insufficient ATP and biotin availability and adverse pH or redox conditions could reduce TbID and mTb activity. For example, although the final step of biotin synthesis happens in mitochondria (Alban 2011), free biotin in mitochondria is undetectable (Baldet et al. 1993). It is possible that active biotin export from mitochondria will be a challenge for PL.

Going forward, it will be interesting to see how TbID and mTb perform for different applications and which challenges arise from the use in other plants, tissues and organelles. Our
experiments in Arabidopsis and *N. benthamiana* suggest that PL will be widely applicable in plants and will provide a valuable tool for the plant community.
### Materials and methods

#### Key resource table

<table>
<thead>
<tr>
<th>Reagent type (species) or resource</th>
<th>Designation</th>
<th>Source or reference</th>
<th>Identifiers</th>
<th>Additional information</th>
</tr>
</thead>
<tbody>
<tr>
<td>gene (<em>E. coli</em> - modified)</td>
<td>BirA*</td>
<td>Branon et al. 2018; DOI: 10.1038/nbt.4201</td>
<td>TAIR:AT3G24140</td>
<td>R118G mutant of BirA; promiscuous bacterial biotin ligase</td>
</tr>
<tr>
<td>gene (<em>E. coli</em> - modified)</td>
<td>TurboID; TbID</td>
<td>Branon et al. 2018; DOI: 10.1038/nbt.4201</td>
<td>TAIR:AT5G53210</td>
<td>more active variant of BirA*</td>
</tr>
<tr>
<td>gene (<em>E. coli</em> - modified)</td>
<td>miniTurbo; mTb</td>
<td>Branon et al. 2018; DOI: 10.1038/nbt.4201</td>
<td>TAIR:AT3G06120</td>
<td>smaller and more active variant of BirA*</td>
</tr>
<tr>
<td>gene (<em>Arabidopsis thaliana</em>)</td>
<td>FAMA</td>
<td>NA</td>
<td>TAIR:AT1G43850</td>
<td>transcription factor involved in stomatal development</td>
</tr>
<tr>
<td>gene (<em>Arabidopsis thaliana</em>)</td>
<td>SPCH</td>
<td>NA</td>
<td>TAIR:AT2G32700</td>
<td>transcription factor involved in stomatal development</td>
</tr>
<tr>
<td>gene (<em>Arabidopsis thaliana</em>)</td>
<td>MUTE</td>
<td>NA</td>
<td>TAIR:AT1G75080</td>
<td>transcription factor involved in brassinosteroid signaling</td>
</tr>
<tr>
<td>gene (<em>Arabidopsis thaliana</em>)</td>
<td>SEU</td>
<td>NA</td>
<td>TAIR:AT5G08130</td>
<td>transcription factor involved in brassinosteroid signaling</td>
</tr>
<tr>
<td>gene (<em>Arabidopsis thaliana</em>)</td>
<td>LUH</td>
<td>NA</td>
<td>TAIR:AT3G26744</td>
<td>transcription factor involved in stomatal development and cold adaptation</td>
</tr>
<tr>
<td>gene (<em>Arabidopsis thaliana</em>)</td>
<td>BZR1</td>
<td>NA</td>
<td>TAIR:AT3G26744</td>
<td>transcription factor involved in stomatal development and cold adaptation</td>
</tr>
<tr>
<td>gene (<em>Arabidopsis thaliana</em>)</td>
<td>BIM1</td>
<td>NA</td>
<td>TAIR:AT5G08130</td>
<td>transcription factor involved in brassinosteroid signaling</td>
</tr>
<tr>
<td>gene (<em>Arabidopsis thaliana</em>)</td>
<td>ICE1</td>
<td>NA</td>
<td>TAIR:AT5G08130</td>
<td>transcription factor involved in brassinosteroid signaling</td>
</tr>
<tr>
<td>strain, strain background (<em>E. coli</em>)</td>
<td>TOP10</td>
<td>other</td>
<td></td>
<td>chemically competent <em>E. coli</em>, can be obtained from Invitrogen</td>
</tr>
<tr>
<td>strain, strain background (<em>Saccharomyces cerevisiae</em>)</td>
<td>AH109</td>
<td>Clontech</td>
<td></td>
<td></td>
</tr>
<tr>
<td>strain, strain background (<em>Agrobacterium thumefaciens</em>)</td>
<td>GV3101</td>
<td>other</td>
<td></td>
<td>electrocompetent <em>A. thumefaciens</em></td>
</tr>
<tr>
<td>strain, strain background (<em>Nicotiana benthamiana</em>)</td>
<td>NB-1</td>
<td>NA</td>
<td></td>
<td>standard lab strain</td>
</tr>
<tr>
<td>Genetic Reagent (Arabidopsis thaliana)</td>
<td>Strain, Strain Background (Arabidopsis thaliana)</td>
<td>Antibody</td>
<td>Source</td>
<td>Notes</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>--------------------------------------------------</td>
<td>----------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>fama-1</td>
<td>Columbia-0; Col-0</td>
<td>Streptavidin-HRP</td>
<td>Thermo Fisher Scientific: S911</td>
<td>0.2 µg/ml; 5% BSA in TBS-T</td>
</tr>
<tr>
<td>FAMApro::FAMA-TbID-mVenus</td>
<td>Fama-1/- -background; see materials and methods for line generation</td>
<td>Rat monoclonal anti-GFP antibody</td>
<td>Chromotek: 3H9</td>
<td>1:2000; 1-5% skim milk in TBS-T</td>
</tr>
<tr>
<td>SPCHpro::GFP</td>
<td>SPCHpro::GFP</td>
<td>Anti-HA High Affinity from rat IgG1</td>
<td>Roche: 11867423001</td>
<td>1:1000; 3-5% skim milk in TBS-T</td>
</tr>
<tr>
<td>MUTEpro::GFP</td>
<td>MUTEpro::GFP</td>
<td>Myc-Tag (71D10) Rabbit mAb</td>
<td>Cell Signaling: 2278S</td>
<td>1:1000; 5% BSA in TBS-T</td>
</tr>
<tr>
<td>FAMApro::FAMA-CFP</td>
<td>FAMApro::FAMA-CFP</td>
<td>AfinityPure Donkey Anti-Rat IgG-HRP</td>
<td>Jackson Immuno Research Laboratories: 712-035-153</td>
<td>1:10000; 1-5% skim milk in TBS-T</td>
</tr>
<tr>
<td>antibody</td>
<td>Streptavidin-HRP</td>
<td>Cell Signaling: 2278S</td>
<td></td>
<td>1:10000; 1-5% skim milk in TBS-T</td>
</tr>
<tr>
<td>antibody</td>
<td>Rabbit Anti-Rat IgG-HRP</td>
<td>Goat anti-Rabbit IgG (H&amp;L), HRP conjugated</td>
<td>Agrisera: AS09 602</td>
<td>1:20000; 3-5% skim milk in TBS-T</td>
</tr>
<tr>
<td>Recombinant DNA Reagent</td>
<td>Supplier and Details</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R4pGWB601</td>
<td>Nakamura et al. 2010; DOI: 10.1271/bbb.100184</td>
<td>RIKEN BRC:pdi00133 obtained from the Nakagawa lab (<a href="http://shimane-u.org/nakagawa/gbv.htm">http://shimane-u.org/nakagawa/gbv.htm</a>)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R4pGWB613</td>
<td>Nakamura et al. 2010; DOI: 10.1271/bbb.100184</td>
<td>RIKEN BRC:pdi00099 obtained from the Nakagawa lab (<a href="http://shimane-u.org/nakagawa/gbv.htm">http://shimane-u.org/nakagawa/gbv.htm</a>)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R4pGWB616</td>
<td>Nakamura et al. 2010; DOI: 10.1271/bbb.100184</td>
<td>RIKEN BRC:pdi00102 obtained from the Nakagawa lab (<a href="http://shimane-u.org/nakagawa/gbv.htm">http://shimane-u.org/nakagawa/gbv.htm</a>)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pB7m34GW,0</td>
<td>Karimi, De Meyer, and Hilson 2005; DOI: 10.1016/j.tplants.2005.01.008</td>
<td>Gateway entry vector for promoters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pK7m34GW,0</td>
<td>Karimi, De Meyer, and Hilson 2005; DOI: 10.1016/j.tplants.2005.01.008</td>
<td>Gateway entry vector for tags/genes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pENTR5'/TOPO</td>
<td>Invitrogen</td>
<td>Gateway entry vector for promoters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pENTR/D-TOPO</td>
<td>Invitrogen</td>
<td>Gateway entry vector for tags/genes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pDONR-P2R-P3</td>
<td>Invitrogen</td>
<td>Gateway entry vector for tags/genes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGADT7-GW</td>
<td>Lu et al. 2010; DOI: 10.1111/j.1365-313X.2009.04048.x</td>
<td>Addgene:61702 Gateway compatible Y2H prey vector (Gal4 activation domain)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pXDGAtcy86</td>
<td>Ding et al. 2007;DOI: 10.1385/1-59259-966-4:85</td>
<td>Gateway compatible Y2H bait vector (Gal4 DNA-binding domain)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V5-hBirA(R118G)-NES_pCDNA3</td>
<td>Branon et al. 2018; DOI: 10.1038/nbt.4201</td>
<td>Addgene:127363 UBQ10 promoter (2 kb), BirA* (cDNA) with nuclear export signal, YFP in Gateway vector R4pGWB601; see materials and methods for cloning and Addgene for vector map</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V5-hBirA-Turbo-NES_pCDNA3</td>
<td>Branon et al. 2018; DOI: 10.1038/nbt.4201</td>
<td>Addgene:107169 obtained from Ting lab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V5-hBirA-miniTurbo-NES_pCDNA3</td>
<td>Branon et al. 2018; DOI: 10.1038/nbt.4201</td>
<td>Addgene:107170 obtained from Ting lab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R4pGWB601_UBQ10_p_BirA(R118G)-NES-YFP</td>
<td>this paper</td>
<td>Addgene:127365 UBQ10 promoter (2 kb), BirA* (cDNA), YFP with nuclear import signal in Gateway vector R4pGWB601; see materials and methods for cloning and Addgene for vector map</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R4pGWB601_UBQ10_p_BirA(R118G)-NES-YFP</td>
<td>this paper</td>
<td>Addgene:127365 UBQ10 promoter (2 kb), BirA* (cDNA), YFP with nuclear import signal in Gateway vector R4pGWB601; see materials and methods for cloning and Addgene for vector map</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R4pGWB601_UBQ10_p_Turbo-NES-YFP-NLS</td>
<td>this paper</td>
<td>Addgene:127366 UBQ10 promoter (2 kb), TurboID (cDNA) with nuclear export signal, YFP in Gateway vector R4pGWB601; see materials and methods for cloning and Addgene for vector map</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R4pGWB601_UBQ10_p_Turbo-NES-YFP-NLS</td>
<td>this paper</td>
<td>Addgene:127366 UBQ10 promoter (2 kb), TurboID (cDNA) with nuclear export signal, YFP in Gateway vector R4pGWB601; see materials and methods for cloning and Addgene for vector map</td>
<td></td>
<td></td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>R4pGWB601_UBQ10 p_miniTurbo-NES-YFP (plasmid)</td>
<td>this paper</td>
<td>Addgene:127369</td>
<td>UBQ10 promoter (2 kb), miniTurbo (cDNA) with nuclear export signal, YFP in Gateway vector R4pGWB601; see materials and methods for cloning and Addgene for vector map</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------------------------------------</td>
<td>------------</td>
<td>----------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>R4pGWB601_UBQ10 p_miniTurbo-YFP-NLS (plasmid)</td>
<td>this paper</td>
<td>Addgene:127370</td>
<td>UBQ10 promoter (2 kb), miniTurbo (cDNA), YFP with nuclear import signal in Gateway vector R4pGWB601; see materials and methods for cloning and Addgene for vector map</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>pB7m34GW.0_FAMA p_gFAMA-Turbo-Venus (plasmid)</td>
<td>this paper</td>
<td></td>
<td>FAMA promoter (2.4 kb), FAMA (genomic DNA), TurboID, Venus in Gateway vector pB7m34GW.0; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>R4pGWB601_FAMA p_Turbo-YFP-NLS (plasmid)</td>
<td>this paper</td>
<td></td>
<td>FAMA promoter (2.4 kb), TurboID, YFP with nuclear import signal in Gateway vector R4pGWB601; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>pK7m34GW.0_UBQ10 p_cFAMA-TbID-Venus (plasmid)</td>
<td>this paper</td>
<td></td>
<td>UBQ10 promoter (2 kb), FAMA (cDNA), TurboID, Venus in Gateway vector pB7m34GW.0; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>R4pGWB613_UBQ10 p_ICE1-3xHA (plasmid)</td>
<td>this paper</td>
<td></td>
<td>UBQ10 promoter (2 kb), ICE1 (cDNA) in Gateway vector R4pGWB613; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>R4pGWB613_UBQ10 p_MUTE-3xHA (plasmid)</td>
<td>this paper</td>
<td></td>
<td>UBQ10 promoter (2 kb), MUTE (cDNA) in Gateway vector R4pGWB616; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>R4pGWB613_UBQ10 p_SEU-4xmyc (plasmid)</td>
<td>this paper</td>
<td></td>
<td>UBQ10 promoter (2 kb), SEU (cDNA) in Gateway vector R4pGWB616; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>R4pGWB613_UBQ10 p_LUH-3xHA (plasmid)</td>
<td>this paper</td>
<td></td>
<td>UBQ10 promoter (2 kb), LUH (cDNA) in Gateway vector R4pGWB613; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>pXDGATcy86-FAMA (plasmid)</td>
<td>this paper</td>
<td></td>
<td>FAMA (cDNA) in Gateway compatible Y2H bait vector pXDGATcy86; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>pGADT7-GW-ICE1 (plasmid)</td>
<td>this paper</td>
<td></td>
<td>ICE1 (cDNA) in Gateway compatible Y2H prey vector pGADT7-GW; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>pGADT7-GW-MUTE (plasmid)</td>
<td>this paper</td>
<td></td>
<td>MUTE (cDNA) in Gateway compatible Y2H prey vector pGADT7-GW; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>pGADT7-GW-SEU (plasmid)</td>
<td>this paper</td>
<td></td>
<td>SEU (cDNA) in Gateway compatible Y2H prey vector pGADT7-GW; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>pGADT7-GW-LUH (plasmid)</td>
<td>this paper</td>
<td></td>
<td>LUH (cDNA) in Gateway compatible Y2H prey vector pGADT7-GW; see materials and methods for cloning</td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>pGADT7-GW-BZR1 (plasmid)</td>
<td>this paper</td>
<td>BZR1 (cDNA) in Gateway compatible Y2H prey vector pGADT7-GW; see materials and methods for cloning</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------------------------</td>
<td>------------</td>
<td>------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>pGADT7-GW-BIM1 (plasmid)</td>
<td>this paper</td>
<td>BIM1 (cDNA) in Gateway compatible Y2H prey vector pGADT7-GW; see materials and methods for cloning</td>
<td></td>
</tr>
<tr>
<td>recombinant DNA reagent</td>
<td>additional plasmids</td>
<td></td>
<td>for list of Gateway compatible vectors to generate N- and C-terminal fusions with TbID or mTb (‘PL toolbox’) see table in materials and methods section</td>
<td></td>
</tr>
<tr>
<td>sequence-based reagent</td>
<td>primers</td>
<td></td>
<td>see primer table in materials and methods</td>
<td></td>
</tr>
<tr>
<td>peptide, recombinant protein</td>
<td>Biotin powder</td>
<td>Sigma</td>
<td>Sigma:B4639</td>
<td></td>
</tr>
<tr>
<td>commercial assay or kit</td>
<td>BioRad protein assay</td>
<td>BioRad</td>
<td>BioRad:5000006</td>
<td></td>
</tr>
<tr>
<td>commercial assay or kit</td>
<td>Novex Colloidal blue staining kit</td>
<td>Invitrogen</td>
<td>Invitrogen:LC6025</td>
<td></td>
</tr>
<tr>
<td>chemical compound, drug</td>
<td>Dynabeads MyOne Streptavidin C1</td>
<td>Invitrogen</td>
<td>Thermo Fisher:65002</td>
<td></td>
</tr>
<tr>
<td>chemical compound, drug</td>
<td>Dynabeads MyOne Streptavidin T1</td>
<td>Invitrogen</td>
<td>Thermo Fisher:65601</td>
<td></td>
</tr>
<tr>
<td>chemical compound, drug</td>
<td>GFP-Trap_MA beads</td>
<td>ChromoTek</td>
<td>ChromoTek:gtma-20</td>
<td></td>
</tr>
<tr>
<td>software, algorithm</td>
<td>MaxQuant</td>
<td>Tyanova, Temu, and Cox 2016; DOI: 10.1038/nprot.2016.136</td>
<td>version 1.6.2.6</td>
<td></td>
</tr>
<tr>
<td>software, algorithm</td>
<td>Perseus</td>
<td>Tyanova et al. 2016; DOI: 10.1038/nmeth.3901</td>
<td>version 1.6.2.3</td>
<td></td>
</tr>
<tr>
<td>software, algorithm</td>
<td>Normalyzer</td>
<td>Chawade et al. 2014; DOI: 10.1021/pr401264n</td>
<td>version 1.1.1.1 (web interface: <a href="http://normalyzer.immunoprot.lt.h.se/">http://normalyzer.immunoprot.lt.h.se/</a>)</td>
<td></td>
</tr>
<tr>
<td>software, algorithm</td>
<td>R studio</td>
<td>RStudio Team (2016)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>software, algorithm</td>
<td>AgriGO v2</td>
<td>Tian et al. 2017; DOI: 10.1093/nar/gkv382</td>
<td>web interface: <a href="http://systemsbiology.cau.edu.cn/agriGOv2/">http://systemsbiology.cau.edu.cn/agriGOv2/</a></td>
<td></td>
</tr>
<tr>
<td>software, algorithm</td>
<td>REVIGO</td>
<td>Supek et al. 2011; DOI: 10.1371/journal.pone.0021800</td>
<td>web interface: <a href="http://revigo.irb.hr/">http://revigo.irb.hr/</a></td>
<td></td>
</tr>
<tr>
<td>other</td>
<td>PD-10 Sesalting Column</td>
<td>GE-Healthcare</td>
<td>Fisher Scientific:45-000-148</td>
<td></td>
</tr>
</tbody>
</table>

**Generation of the ‘PL toolbox’ vectors**
Cloning of gateway-compatible entry vectors containing different BirA variants: BirA* (R118G), TurboID (TbID) or miniTurbo (mTb) were amplified from V5-hBirA(R118G)-NES_pCDNA3, V5-hBirA-Turbo-NES_pCDNA3 or V5-hBirA-miniTurbo-NES_pCDNA3 (Branon et al. 2018) using primers BirA-fw and BirA-rv or BirA-fw and BirA-NES-rv, thereby retaining the N-terminal V5 tag, either removing or retaining the C-terminal NES and adding a GGGGSGGG linker and an Ascl restriction site to both ends. The TbID and mTb PCR products were then cloned into a pENTR vector containing YFP between the attL1 and L2 sites to generate YFP-TbID and YFP-mTb fusions or into a pDONR-P2R-P3 vector containing mVenus-STOP between the attR2 and L3 sites to generate TbID-mVenus-STOP and mTb-mVenus-STOP fusions by restriction cloning with Ascl. The Ascl site in the pDONR-P2R-P3 vector was first introduced by mutagenesis PCR using primers pDONR-mut-Ascl-fw and pDONR-mut-Ascl-rv. These vectors can be used for 3-way gateway recombination with a promoter and gene of choice as described below. For 2-way recombination with a promoter of choice, TbID and mTb (with and without NES) were further amplified from the pDONR-P2R-P3 plasmids described above using primers BirA-YFPnls-fw and BirA-YFPnls-rv to remove the N-terminal linker and to add a NotI and Ncol site at the N-and C-terminus, respectively. BirA* was directly amplified from the PCR products with BirA-fw and BirA-rv or BirA-NES-rv in the same way. The resulting PCR products were either cloned into a pENTR vector containing YFP followed by a stop codon (versions with NES) or a YFP followed by an NLS and a stop codon (versions without NES) between the attL1 and L2 sites using NotI and Ncol to generate BirA-NES_YFP and BirA-YFP\textsubscript{NLS} fusions (BirA = BirA*, TbID or mTb). For a schematic overview over the cloning process and the composition of the vectors, see Figure 1 – figure supplement 5. For a list of vectors for 3- and 2-way recombination see the ‘PL toolbox’ vectors table at the end of the methods section.

Cloning of binary vectors for proximity labeling: The UBQ10pro::BirA-YFP\textsubscript{NLS} (BirA = BirA*, TbID or mTb) and FAMApro::TbID-YFP\textsubscript{NLS} constructs were generated by LR recombination of
the R4pGWB601 backbone (Nakamura et al. 2010) with either a pENTR5'/TOPO containing the 2 kb UBQ10 promoter or a pJET containing 2.4 kb of the FAMA (AT3G24140) upstream sequence (positions -2420 to -1, flanked by attL4 and R1, cloned from pENTR/D-TOPO (Ohashi-Ito and Bergmann 2006) by changing attL1 and L2 to L4 and R2) and the pENTR-BirA-YFP<sub>YFP</sub> plasmids described above. The UBQ10pro::NESYFP-BirA constructs were generated by LR recombination of the R4pGWB601 backbone with a pENTR5'/TOPO containing the 2 kb UBQ10 promoter and the pENTR-BirA-NESYFP plasmids described above. For a schematic overview over the cloning process and composition of the vectors, see Figure 1 – figure supplement 5. For a list of vectors, see ‘PL toolbox’ vectors table at the end of the methods section. The FAMAp::FAMA-TbID-Venus construct was generated by LR recombination of the pB7m34GW,0 backbone (Karimi, De Meyer, and Hilson 2005) with the pJET containing 2.4 kb of the FAMA upstream sequence, with pENTR containing the genomic sequence of FAMA without stop codon flanked by attL1 and L2 (amplified from genomic DNA with primers gFAMA-fw and gFAMA-rv and recombined with pENTR/D-TOPO) and with pDONR-P2R-P3-TbID-mVenus (described above).

Cloning of FAMA interaction partner candidates for Y2H and PL assays in <i>N. benthamiana</i>: SEU (AT1G43850), LUH (AT2G32700), BZR1 (AT1G75080), BIM1 (AT5G08130), ICE1 (AT3G26744) and MUTE (AT3G06120) were amplified from Col-0 cDNA with the primers listed in the primer table and cloned into pENTR/D-TOPO. FAMA in pENTR was cloned as described in Ohashi-Ito and Bergmann (2006). For Y2H assays, FAMA was recombined with Gateway compatible pXDGATcy86 (Ding, Zhang, and Song 2007), containing the Gal4 DNA-binding domain, and SEU, LUH, BZR1, BIM1, ICE1 and MUTE were recombined into Gateway compatible pGADT7-GW (Lu et al. 2010), containing the Gal4 activation domain. For PL assays in tobacco, pENTR5'/TOPO containing the 2 kb UBQ10 promoter, pENTR containing FAMA cDNA and pDONR-P2R-P3-TbID-mVenus were recombined with...
pK7m34GW,0 (Karimi, De Meyer, and Hilson 2005) to generate UBQ10pro::cFAMA-TbID-mVenus. Entry vectors carrying ICE1, MUTE and LUH and the entry vector carrying SEU were recombined with pENTR5'/TOPO containing the 2 kb UBQ10 promoter and R4pGWB613 (3x HA tag) and R4pGWB616 (4x myc tag), respectively (Nakamura et al. 2010). UBQ10pro::TbID-YFP<sub>NLS</sub> in R4pGWB601 is described above.

**Transformation and biotin activity assays in *N. benthamiana***

*N. benthamiana* ecotype NB-1 was transformed with UBQ10pro::BirA-YFP<sub>NLS</sub> and UBQ10pro::BirA<sup>nes</sup>YFP (BirA = BirA, TbID or mTb) by infiltrating young leaves with a suspensions of Agrobacteria (strain GV3101) carrying one of the binary vectors. Agrobacteria were grown from an overnight culture for two hours, supplemented with 150 µM Acetosyringone, grown for another four hours, pelleted and resuspended in 5 % sucrose to an OD<sub>600</sub> of 2. For more stable expression, Agrobacteria carrying a 35S::p19 plasmid (tomato bushy stunt virus (TBSV) protein p19) were co-infiltrated at a ratio of 1:1 for the temperature-dependency experiment. Two days after infiltration, expression was confirmed by epifluorescence microscopy and 5 mm wide leaf discs were harvested. Two to three discs were combined per sample. They were submerged in a 50 or 250 µM biotin solution, quickly vacuum infiltrated until the air spaces were filled with liquid and incubated at the indicated temperature for one hour. Control samples were not treated or were infiltrated with H<sub>2</sub>O. After biotin treatment, leaf discs were dried and flash-frozen for later immunoblotting. All experiments were done in duplicates with leaf discs for each of the two replicates taken from different plants if possible. Only one replicate is shown. Activity of different BirA variants was compared in four independent experiments, with similar results, temperature dependency of TbID and mTb was tested in two and one experiment, respectively.
Arabidopsis lines used in this study

Arabidopsis thaliana Col-0 was used as wild-type (WT). The fama-1 mutant line is SALK_100073 (Ohashi-Ito and Bergmann 2006). Plant lines for testing the activity of BirA*, TblD and mTb (UBQ10pro::BirA*-YFP<sub>NLS</sub>, UBQ10pro::TblD-YFP<sub>NLS</sub>, UBQ10pro::mTb-YFP<sub>NLS</sub>, UBQ10pro::BirA*-NES<sub>NES</sub>YFP, UBQ10pro::TblD-NES<sub>NES</sub>YFP, UBQ10pro::mTb-NES<sub>NES</sub>YFP) and for the ‘FAMA interactome’ and nuclear proteome’ experiments (FAMAprom::FAMA-TblD-mVenus in fama-1, FAMAprom::TblD-YFP<sub>NLS</sub>) were generated by floral dip of WT or fama-1 +/- plants with the plasmids described above using agrobacterium strain GV3101. Selection was done by genotyping PCR (fama-1) and segregation analysis. We did not observe any obvious decrease in viability or developmental delay in our transgenic Arabidopsis plants. All lines had a single insertion event and were either heterozygous T2 or homozygous T3 or T4 lines. While screening for biotin ligase lines with the UBQ10 promoter, we observed that most regenerants had very weak YFP signal, especially the nuclear constructs. This was not observed with any of the cell type-specific promoters we tested. Lines used for the FAMA-CFP AP-MS experiments were previously described in other studies: FAMAprom::FAMA-CFP (Weimer et al. 2018), SPCHprom::GFP<sub>NLS</sub> and MUTEprom::GFP<sub>NLS</sub> (Adrian et al. 2015).

Plant growth conditions and biotin assays in Arabidopsis

Seeds were surface sterilized with ethanol or bleach and stratified for two to three days. For biotin treatment in whole Arabidopsis seedlings or roots and shoots, seedlings were grown on ½ Murashige and Skoog (MS, Caisson labs) plates containing 0.5% sucrose for four to 14 days under long-day conditions (16 h light/8h dark, 22°C). For treatment of rosette leaves and flowers, seedlings were transferred to soil and grown in a long-day chamber (22°C) until the first
flowers emerged, at which point medium sized rosette leaves (growing but almost fully expanded) and inflorescences with unopened flower buds were harvested. All samples were pools from several individual plants. Biotin treatment was done by submerging the plant material in a biotin solution (0.5 – 250 μM biotin in water) and either vacuum infiltrating the tissue briefly until the air spaces were filled with liquid (approximately five minutes) or not, followed by incubation at room temperature (22°C), 30°C or 37°C for up to five hours. Controls were either treated with H₂O or not treated. Following treatment, the plant material was dried and flash-frozen for later immunoblotting. To confirm reproducibility of the experiment, most experiments were done in duplicates (only one replicate shown) or repeated more than once with similar results. Comparison of all different BirA variants in Arabidopsis was done once with two independent lines for each NLS and NES constructs. Difference in activity and background between TblD and mTb matches other comparison done with varying temperature and biotin concentration. Comparison of TblD and mTb at different temperatures was done three times. Comparison of the biotin ligase activity with different biotin concentrations was done twice for TblD and once for mTb. Three time courses with up to three or five hours of biotin treatment were done with the UBQ10pro::TblD-YFP<sub>NLS</sub> line and time courses with the FAMA-TblD line were done in duplicates. Experiments testing the effect of biotin application with and without vacuum infiltration in different tissues were done in duplicates. Activity of TblD in 4-14 day old seedlings was tested in two independent experiments. Activity of TblD in roots and shoots of 6-14 day old seedlings was tested once.

Optimization of streptavidin (SA) pulldown conditions

Saturation of SA beads by free biotin: To test the impact of free biotin from biotin treatment on the affinity purification (AP) efficiency with SA-coupled beads, WT and UBQ10pro::TblD-YFP<sub>NLS</sub> seedlings were submerged in H₂O or 50 μM biotin for one hour, washed twice and
frozen. AP was done in essence as described in (Schopp et al. 2017). Briefly, 1 ml of finely
ground plant material was resuspended in an equal volume of ice cold extraction buffer (50 mM
Tris pH 7.5, 500 mM NaCl, 0.4 % SDS, 5 mM EDTA, 1 mM DTT, 1 mM PMSF and 1x complete
proteasome inhibitor), sonicated in an ice bath four times for 30 seconds on high setting using a
Bioruptor UCD-200 (Diagenode) with 1.5-minute breaks on ice and supplemented with Triton-X-100 to reach a final concentration of 2 %. 2.3 ml of ice cold 50 mM Tris pH 7.5 were added to
dilute the extraction buffer to 150 mM NaCl and samples were centrifuged at top speed for 15
minutes. The supernatant was mixed with 50 μl Dynabeads MyOne Streptavidin T1 (Invitrogen)
that were pre-washed with equilibration buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.05 % Triton-
X-100, 1 mM DTT) and incubated over night at 4°C on a rotor wheel. The beads were washed
twice each with wash buffer 1 (2 % SDS), wash buffer 2 (50 mM Hepes pH 7.5, 500 mM NaCl, 1
mM EDTA, 1% Triton-X-100, 0.5% Na-deoxycholate), wash buffer 3 (10 mM Tris pH 8, 250 mM
LiCl, 1 mM EDTA, 0.5% NP-40, 0.5 % Na-deoxycholate) and wash buffer 4 (50 mM Tris pH 7.5, 50
mM NaCl, 0.1 % NP-40). Protein were eluted by boiling the beads for 15 minutes at 98°C in
elution buffer (10 mM Tris pH 7.5, 2% SDS, 5% beta mercaptoethanol, 2 mM biotin) and used
for immunoblotting and SDS-PAGE. Two experiments with similar results were done.

Testing of biotin depletion strategies and optimization of the SA bead amount: To compare different biotin depletion strategies and determine the required bead amount for the PL experiments, five day old UBQ10pro::TbID-YFP<sub>NLS</sub> seedlings were submerged in a 50 μM biotin solution for three hours, washed three times with ice cold water, frozen and later used for biotin
depletion and AP experiments. Proteins were extracted as described later in ‘Affinity purification
of biotinylated proteins’. For biotin depletion with one or two PD-10 gel filtration columns (GE-
Healthcare), the columns were equilibrated according to the manufacturer’s instructions with
extraction buffer without PMSF and complete protease inhibitor (50 mM Tris pH 7.5, 150 mM
NaCl, 0.1 % SDS, 1 % Triton-X-100, 0.5 % Na-deoxycholate, 1 mM EGTA, 1 mM DTT). 2.5 ml
protein extract were loaded onto the column. For the trial with one PD-10 column, the gravity protocol was used, eluting with 3.5 ml equilibration buffer. For the trial with two PD-10 columns, the proteins were eluted from the first column with 2.5 ml equilibration buffer using the spin protocol. The flow through was applied to a second PD-10 column and eluted with 3.5 ml equilibration buffer using the gravity protocol. For biotin depletion with an Amicon Ultra-4 Centrifugal filter (Millipore Sigma), 625 µl protein extract were concentrated three times to 10 – 20% of the starting volume by centrifuging in a swinging bucket rotor at 4,000g and 4°C and diluted with extraction buffer to the initial volume. To test bead requirements, a volume equivalent to 1/5 of the amount used in the ‘FAMA interactome’ and ‘nuclear proteome’ PL experiments was incubated with five to 30 µl Dynabeads MyOne Streptavidin C1 (Invitrogen) overnight at 4°C on a rotor wheel. The beads were either washed once with 1 M KCl and with 100 mM Na₂CO₃ and twice with extraction buffer (1x PD-10) or just twice with extraction buffer (Amicon Ultra, 2x PD-10) before elution of bound proteins by boiling for 10 minutes at 95°C in 4x Laemmli buffer supplemented with 20 mM DTT and 2 mM biotin. Each biotin depletion strategy was tested once.

‘FAMA interactome’ and ‘nuclear proteome’ PL experiments

Seedling growth and biotin treatment: Approximately 120 µl of seeds per sample and line were surface sterilized with ethanol and bleach and stratified for two days. The seeds were then spread on filter paper (Whatman Shark Skin Filter Paper, GE Healthcare 10347509) that was placed on ½ MS plates containing 0.5 % sucrose and grown in a growth chamber under long-day conditions for five days. For biotin treatment, seedlings were carefully removed from the filter paper, transferred into beakers, covered with 40 ml of a 50 µM biotin solution and incubated in the growth chamber for 0.5 or 3 hours. The biotin solution was removed and seedlings were quickly rinsed with ice cold water and washed three times for three minutes in
approximately 200 ml of ice cold water to stop the labeling reaction (Branon et al. 2018) and to remove excess biotin. The seedlings were then dried with paper towels, a sample was taken for immunoblots and the remaining seedlings were split into aliquots of about 1.5 g fresh weight for the three biological replicates, frozen in liquid nitrogen, ground to a fine powder and stored at -80°C until further use. Untreated 0 h samples were frozen directly after harvest. Treatment was timed to keep harvesting times as close together as manageable. Small samples for immunoblots were taken after treatment. The number of replicates was chosen to provide statistical power but to keep the experimental cost and sample handling in a feasible range.

**Affinity purification of biotinylated proteins:** For the affinity purification with streptavidin beads, 3 ml of densely packed ground plant material were resuspended in 2 ml extraction buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.1 % SDS, 1 % Triton-X-100, 0.5 % Na-deoxycholate, 1 mM EGTA, 1 mM DTT, 1x complete, 1 mM PMSF) and incubated on a rotor wheel at 4°C for 10 minutes. 1µl Lysonase (Millipore) was added to digest cell walls and DNA/RNA and the suspension was incubated on the rotor wheel at 4°C for another 15 minutes. The extracts were then distributed into 1.5 ml reaction tubes and sonicated in an ice bath four times for 30 seconds on high setting using a Bioruptor UCD-200 (Diagenode) with 1.5-minute breaks on ice. The suspension was centrifuged for 15 minutes at 4°C and 15,000 g to remove cell debris and the clear supernatant was applied to a PD-10 desalting column (GE healthcare) to remove excess free biotin using the gravity protocol according to the manufacturer’s instructions. Briefly, the column was equilibrated with five volumes of ice cold equilibration buffer (extraction buffer without complete and PMSF), 2.5 ml of the protein extract were loaded and proteins were eluted with 3.5 ml equilibration buffer. The protein concentration of the protein extract was then measured by Bradford (BioRad protein assay). The protein extract was diluted 1:5 to avoid interference of buffer components with the Bradford assay. A volume of each protein extract corresponding to 16 mg protein was transferred into a new 5 ml LoBind tube (Eppendorf)
containing Dynabeads MyOne Streptavidin C1 (Invitrogen) from 200 µl bead slurry that were
pre-washed with extraction buffer. Complete protease inhibitor and PMSF were added to reach
final concentrations of 1x complete and 1 mM PMSF, respectively, and the samples were
incubated on a rotor wheel at 4°C overnight (16 h). The next day, the beads were separated
from the protein extract on a magnetic rack and washed as described in (Branon et al. 2018)
with 1 ml each of the following solutions by incubating on the rotor wheel for 8 minutes and
removing the wash solution: 2x with cold extraction buffer (beads were transferred into a new
tube the first time), 1x with cold 1 M KCl, 1x with cold 100 mM Na₂CO₃, 1x with 2M Urea in 10
mM Tris pH 8 at room temperature and 2x with cold extraction buffer without complete and
PMSF. 2 % of the beads were boiled in 50 µl 4x Laemmli buffer supplemented with 20 mM DTT
and 2 mM biotin at 95°C for 5 minutes for immunoblots. The rest of the beads was spun down to
remove the remaining wash buffer and stored at -80°C until further processing. Sample prep
was done in three batches with one replicate of each sample in each batch. Samples from
replicate 2 (batch 2) were used for immunoblots to confirm the success of the procedure.

**MS sample preparation:** For on-beads tryptic digest, the frozen streptavidin beads were
thawed and washed twice each with 1 ml 50 mM Tris pH 7.5 (transferred to new tube with first
wash) and 1 ml 2 M Urea in 50 mM Tris pH 7.5. The buffer was removed and replaced by 80 µl
Trypsin buffer (50 mM Tris pH 7.5, 1 M Urea, 1m M DTT, 0.4 µg Trypsin). The beads were then
incubated for three hours at 25°C with shaking, and the supernatant was transferred into a fresh
tube. The beads were washed twice with 60 µl 1 M Urea in 50 mM Tris pH 7.5 and all
supernatants were combined (final volume 200 µl). The combined eluates were first reduced by
adding DTT to a final concentration of 4 mM and incubating at 25°C for 30 minutes with shaking
and then alkylated by adding iodoacetamide to a final concentration of 10 mM and incubating at
25°C for 45 minutes with shaking. Finally, another 0.5 µg Trypsin were added and the digest
completed by overnight (14.5 h) incubation at 25°C with shaking. The digest was acidified by
adding formic acid to a final concentration of ~ 1 % and desalted using OMIX C18 pipette tips 
(10 – 100 µL, Agilent). C18 desalting tips were first activated by twice aspiring and discarding 
200 µl buffer B2 (0.1 % formic acid, 50 % acetonitrile) and equilibrated by four times aspiring 
and discarding 200 µl buffer A2 (0.1 % formic acid). Peptides were bound by aspiring and 
dispensing the sample eight times. Then, the tip was washed by 10 times aspiring and 
discarding 200 µl buffer A2 and the peptides were eluted by aspiring and dispensing 200 µl 
buffer B2 in a new tube for eight times. The desalted peptides were dried in a speed vac and 
stored at -80°C until further processing.

**LC-MS/MS:** For LC-MS/MS analysis, peptides were resuspended in 0.1 % formic acid. Samples 
were analyzed on a Q-Exactive HF hybrid quadrupole-Orbitrap mass spectrometer (Thermo 
Fisher), equipped with an Easy LC 1200 UPLC liquid chromatography system (Thermo Fisher). 
Peptides were separated using analytical column ES803 (Thermo Fisher). The flow rate was 
300 nl/min and a 120-min gradient was used. Peptides were eluted by a gradient from 3 to 28 % 
solvent B (80 % acetonitrile, 0.1 % formic acid) over 100 minutes and from 28 to 44 % solvent B 
over 20 minutes, followed by short wash at 90 % solvent B. Precursor scan was from mass-to-
charge ratio (m/z) 375 to 1,600 and top 20 most intense multiply charged precursors were 
selection for fragmentation. Peptides were fragmented with higher-energy collision dissociation 
(HCD) with normalized collision energy (NCE) 27.

**MS data analysis – protein identification and label-free quantification:** Protein identification 
and label-free quantification (LFQ) were done in MaxQuant (version 1.6.2.6) (Tyanova, Temu, 
and Cox 2016) using default settings with minor modifications. The datasets for the ‘FAMA 
interactome’ and ‘nuclear proteome’ experiments were searched separately. For the ‘FAMA 
interactome’ PL experiment, all samples were sufficiently similar (majority of proteins expected 
to be equally abundant across samples). Therefore (global) data normalization was done in
MaxQuant as part of the LFQ algorithm (Cox et al., 2014). For the ‘nuclear proteome’ PL experiment, in which the UbQnuCTbID samples were markedly different from the other samples (majority of proteins cannot be assumed to be equally abundant across samples), normalization was skipped during LFQ and done separately as described later in the data analysis section.

MaxQuant settings for LFQ were as follows. Methionine oxidation and N-terminal acetylation were set as variable modifications and Carbamidomethylcysteine as fixed modification. Maximum number of modifications per peptide was five. Trypsin/P with a maximum of two missed cleavages was set as digest. Peptides were searched against the latest TAIR10 protein database containing a total of 35,386 entries (TAIR10_pep_20101214, updated 2011-08-23, www.arabidopsis.org) plus a list of likely contaminants containing Trypsin, human Keratin, streptavidin, YFP, TbID-mVenus and TbID-YFP and against the contaminants list of MaxQuant. Minimum allowed peptide length was seven. FTMS and TOF MS/MS tolerance were 20 ppm and 40 ppm, respectively, and the peptide FDR and protein FDR were 0.01. Unique and razor peptides were used for protein quantification. LFQ minimum ratio count was set to two and fast LFQ was active with a minimum and average number of neighbors of three and six, respectively. For the ‘nuclear proteome’ experiment ‘skip normalization’ was enabled to produce non-normalized LFQ values. Match between runs and second peptides were checked. The mass spectrometry proteomics data (raw data, MaxQuant analysis, normalization results (only ‘nuclear proteome’ experiment) and the following data analysis in Perseus) have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaíno et al. 2013) with the dataset identifiers PXD015161 (‘FAMA interactome’ experiment) and PXD015162 (‘nuclear proteome’ experiment).

**MS data analysis – identification of enriched proteins in the ‘FAMA interactome’ experiment:** Filtering and statistical analysis were done with Perseus (version 1.6.2.3) (Tyanova et al. 2016). The ‘proteinGroups.txt’ output file from MaxQuant was imported into
Perseus using the LFQ intensities as Main category. The data matrix was filtered to remove proteins marked as ‘only identified by site’, ‘reverse’ and ‘potential contaminant’. LFQ values were log2 transformed and proteins that were not identified/quantified in all three replicates of at least one time point of one genotype (low confidence) were removed. The clustering analysis shown in Figure 5 – figure supplement 2 and the multi scatterplot shown in Figure 5 – figure supplement 3, were done at this point, using the built-in ‘Hierarchical clustering’ function with the standard settings (Euclidean distance with average linkage, no constraints and preprocessing of k-means) and the ‘Multi scatter plot’ function (equal ranges, Pearson correlation values displayed), respectively. Next, missing values were imputed for statistical analysis using the ‘Replace missing values from normal distribution’ function (settings: width = 0.3, down shift = 1.8, mode = total matrix) and principal component analysis was done.

**Protein enrichment in FAMA-TbID samples:** To identify proteins enriched in TbID-expressing samples versus the WT control and to remove proteins that bind the beads non-specifically, unpaired 2-sided Students t-tests were performed comparing the FAMA-TbID or FAMA_nucTbID with the corresponding WT samples at each time point. An FDR of 0.05 (integrated modified permutation-based FDR (“Significance Analysis of Microarrays” (SAM) method)) and an S0 of 0.5 with 250 randomizations were chosen as cutoff for multiple sample correction. As an additional filter, only proteins that were quantified in all three replicates of the FAMA-TbID or FAMA_nucTbID samples at this time point were used for the tests (identified with high confidence; see Supplementary file 2 – Table 2 for number of proteins used in each test). Significantly enriched proteins were used for hierarchical clustering with standard settings, using Z-transformed non-imputed LFQ values. To further identify proteins that are significantly enriched in FAMA-TbID versus FAMA_nucTbID and to remove proteins that are stochastically labeled by FAMA-TbID, corresponding t-tests were performed on the reduced dataset using the same parameters as for the first t-tests. Finally, to filter out proteins that were not enriched in at least one of the two treatments compared to untreated FAMA-TbID, unpaired 2-sided Students t-tests
were performed on the remaining proteins comparing the samples at the 0.5 and 3 h time points to the 0 h time point using a p-value of 0.5 as cutoff.

**Protein enrichment over time:** To identify proteins enriched after short or long biotin treatment, unpaired 2-sided Students t-tests were performed comparing the 0.5 and 3 h treatment samples of WT, FAMA-TbID and FAMA\textsubscript{nuc}TbID to the corresponding 0 h samples. An FDR of 0.05 and an S0 of 0.5 with 250 randomizations were chosen as cutoff for multiple sample correction. Only proteins that were quantified in all three replicates of at least one of the two groups were tested (see Supplementary file 2 – Table 3 for number of proteins used in each test).

**Plots and heatmaps:** Scatterplots and PCA plots were made in RStudio (version 1.1.463 (RStudio Team (2016)), R version 3.5.1 (R Core Team (2018)) using t-test and PCA results exported from Perseus. Hierarchical clustering was done in Perseus and the plots were exported as pdf.

**MS data analysis – data normalization and identification of enriched proteins in the ‘nuclear proteome’ experiment:** Since the samples of the ‘nuclear proteome’ experiment were quite different in the number and abundance of identified proteins (majority of proteins cannot be assumed to be equally abundant across samples) and each sample was measured as one fraction, normalization was skipped in MaxQuant and done after peptide identification and quantification using the non-normalized LFQ values. Search for a suitable normalization method and normalization were done with Normalyzer (version 1.1.1) (Chawade et al., 2014), a tool to help select an optimal normalization method for a given dataset from a set commonly used global and local normalization methods. Subsequent data filtering and statistical analysis were done with Perseus (version 1.6.2.3) (Tyanova et al. 2016). First, the ‘proteinGroups.txt’ output file from MaxQuant was imported into Perseus using the LFQ intensities as Main category. The data matrix was filtered to remove proteins marked as ‘only identified by site’, ‘reverse’ and ‘potential contaminant’ (reduction of the dataset from 5,038 proteins to 4,964 proteins) and the
LFQ values of all remaining proteins were exported for normalization with Normalyzer (http://normalyzer.immunoprot.lth.se/normalize.php). Based on the assumption that the majority of proteins in our dataset is differentially abundant and on the performance of the supported normalization methods, we determined that LOESS-R (local normalization) is most suitable for the dataset. LOESS-R normalized LFQ values were then imported back into Perseus and log2 transformed. Proteins that were not identified/quantified in all three replicates of at least one genotype (low confidence) were removed. The clustering analysis shown in Figure 6 – figure supplement 1, as well as the multi scatterplot shown in Figure 6 – figure supplement 2, were done at this point, using the built-in ‘Hierarchical clustering’ and ‘Multi scatter plot’ functions as described for the ‘FAMA interactome’ experiment. Next, missing values were imputed for statistical analysis using the ‘Replace missing values from normal distribution’ function (settings: width = 0.3, down shift = 1.8, mode = total matrix) and principal component analysis was done.

**Protein enrichment in nuclear TbID samples:** To identify proteins enriched in nuclear TbID-expressing samples versus the WT control, unpaired 2-sided Students t-tests were performed comparing the UBQnucTbID or FAMA\textsuperscript{nuc}TbID with WT samples. An FDR of 0.05 (integrated modified permutation-based FDR (“Significance Analysis of Microarrays” (SAM) method)) and an S0 of 0.5 with 250 randomizations were chosen as cutoff for multiple sample correction. As an additional filter, only proteins that were quantified in all three replicates of the UBQ\textsuperscript{nuc}TbID or FAMA\textsuperscript{nuc}TbID sample were used for the tests (identified with high confidence; see Supplementary file 3 – Tables 2 and 3 for number of proteins used in each test). To further identify proteins that are specific for or highly enriched in FAMA nuclei compared to all nuclei, we looked for proteins that were equally or slightly more abundant in the FAMA\textsuperscript{nuc}TbID than the UBQ\textsuperscript{nuc}TbID samples, by performing unpaired 2-sided t-tests between FAMA\textsuperscript{nuc}TbID and UBQ\textsuperscript{nuc}TbID on the reduced dataset. As cutoff, an FDR of 0.01 and S0 of 0.1 were chosen. Protein not significantly enriched in the UBQ\textsuperscript{nuc}TbID samples were used for hierarchical clustering with standard settings, using Z-transformed non-imputed LFQ values.
**Plots and heatmaps:** Scatterplots and PCA plots were made as described for the ‘FAMA interactome’ experiment.

**Classification of the subcellular localization of nuclear proteins form PL:** Proteins identified by PL with nuclear TblID were divided into three broad categories: (1) previously identified in the nucleus experimentally, (2) not found experimentally but predicted to be in the nucleus and (3) predicted to be elsewhere. To determine proteins in the first category, a list of proteins identified in published nuclear- or subnuclear proteomics studies and of proteins detected in the nucleus in localization studies with fluorescent proteins (from SUBA4 (Hooper et al. 2017); http://suba.live/; date of retrieval: February 5 2019) was compiled (see Supplementary file 3 – Table 4). Localization predictions for proteins not identified experimentally in the nucleus was done using the SUBAcon prediction algorithm (Hooper et al. 2014) on SUBA4 (date of retrieval: February 5 2019). SUBAcon provides a consensus localization based on 22 different prediction algorithms and available experimental data.

**Gene ontology analysis:** Proteins that were significantly enriched in FAMA\textsubscript{nucTbID} and UBQ\textsubscript{nucTbID} compared to WT (FAMA- and global nuclear proteins) were used for GO analysis with AgriGO v2 (Tian et al. 2017) with the following settings: Singular enrichment analysis (SAE) with TAIR10\_2017 as background, statistical method: fisher, multi-test adjustment: Yekutieli (FDR), significance level: 0.05, minimum number of entries: five. Significantly enriched GO terms from AgriGO v2 (see Supplementary file 3 – Table 7) were visualized using REViGO (Supek et al. 2011) with the following settings: list size medium, associated numbers are p-values, GO term size: Arabidopsis thaliana, semantic similarity measure: SimRel. The R script provided by REViGO was used to draw the TreeMap plot in RStudio. Very small labels were removed and label size was adjusted in Adobe Illustrator.
Evaluation of selected FAMA interaction partner candidates by Y2H and PL assays in *N. benthamiana*:

**Y2H assays**: FAMA in pXDGATcy86 (Gal4 DNA binding domain) was tested for interaction with ICE1, MUTE, SEU, LUH, BZR1 and BIM1 in pGADT7-GW (Gal4 activation domain). Transformation of yeast strain AH109 with bait and prey plasmids (empty or containing one of the listed genes), selection for transformants and testing of pair-wise interactions by growth complementation assays on nutrient selective media and with X-α-Gal was done as described in the Matchmaker™ GAL4 Two-Hybrid System 3 manual (Clontech). To compensate for autoactivation of BD-FAMA, 3-AT was added to selective plates at different concentrations for growth complementation assays.

**PL assays in *N. benthamiana***: *N. benthamiana* ecotype NB-1 was co-transformed with three or four constructs as described before by infiltrating young leaves with suspensions of Agrobacteria carrying the following vectors: (1) one of UBQ10pro::cFAMA-TbID-Venus in pK7m34GW,0 or UBQpro::TbID-YFP<sub>NLS</sub> in R4pGWB601, (2) one or two of UBQ10pro::ICE1-3xHA, UBQ10pro::MUTE-3xHA, UBQ10pro::LUH-3xHA in R4pGWB613 and UBQ10pro::SEU-4xmyc in R4pGWB616 (Agrobacterium strain GV3101) and (3) an Agrobacterium strain carrying a 35S::p19 plasmid (tomato bushy stunt virus (TBSV) protein p19) for protein stabilization. Two days after infiltration, expression was confirmed by epifluorescence microscopy and 5 mm wide leaf discs were harvested and flash frozen. 2 x 15 leaf discs from two different leaves were used per sample for the pulldown experiments. Leaf discs were ground to a fine powder in a ball mill. Proteins were extracted by adding 600 µl ice cold extraction buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.1 % SDS, 0.5 % sodium deoxycholate, 1 % Triton-X-100, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 1 mM PMSF, 1x complete protease inhibitor). Extracts were sonicated in an ice bath four times for 30 seconds on high setting using a Bioruptor UCD-200 (Diagenode) with 1.5-
minute breaks. Suspensions were centrifuged twice for 10 minutes at 4°C and 15,000 g to remove cell debris. The cleared supernatants were mixed with pre-washed Dynabeads MyOne Streptavidin C1 (Invitrogen) from 20 µl bead slurry and incubated on a rotor at 4°C for two hours. Beads were washed three times with 300 µl cold extraction buffer, once with 2 M Urea in 10 mM Tris pH 8 (room temperature) and once more with cold extraction buffer. Proteins were eluted by boiling the beads in 4x Laemmli buffer supplemented with 20 mM DTT and 2 mM biotin for 10 minutes. Protein abundance in the extracts and bead elutions were tested by immune blotting.

Microscopy

Brightfield and epifluorescence images of N. benthamiana leaves and Arabidopsis seedlings, leaves and flowers were taken with a Leica DM6B microscope using a Leica CRT6 LED light source. Confocal microscopy images of Arabidopsis seedlings expressing different biotin ligase constructs were taken with a Leica SP5 microscope. For confocal microscopy, cell walls were stained with propidium iodide (Molecular Probes) by incubating in a 0.1 mg/ml solution for three to five minutes. Images were processed in FIJI (ImageJ) (Schindelin et al. 2012). Several independent Arabidopsis lines and transiently transformed N. benthamiana leaves were analyzed. Images shown in figures and figure supplements are representative.

Immunoblots

Samples for immunoblots were prepared by resuspending frozen and ground plant material from biotin treatment assays with 1x Laemmli buffer (60 mM Tris pH 6.8, 2 % SDS, 10 % glycerol, 2.5 % beta-mercaptoethanol, 0.025 % bromphenol blue) or mixing protein extracts 1:1
with 2x Laemmli buffer and boiling the samples for five minutes at 95°C. Proteins bound to SA or GFP-Trap beads were eluted from the beads as described in the respective methods sections. Proteins were separated by SDS-PAGE and blotted onto Immobilon-P PVDF membrane (0.45 µm, Millipore) using a Trans-Blot Semi-Dry transfer Cell (BioRad). The following antibodies were used: Streptavidin-HRP (S911, Thermo Fisher Scientific), Rat monoclonal anti-GFP antibody (3H9, Chromotek), Rat Anti-HA High Affinity (11867423001, Roche), Myc-Tag (71D10) Rabbit mAb (2278S Cell Signaling), AffiniPure Donkey Anti-Rat IgG-HRP (712-035-153, Jackson Immuno Research Laboratories), Rabbit Anti-Rat IgG-HRP (A5795, Sigma), Goat anti-Rabbit IgG-HRP (AS09 602, Agrisera). Blots were probed with primary antibodies overnight at 4°C or up to one hour at room temperature and with the secondary antibody for one hour at room temperature and incubated with ECL Western blotting substrates according to the manufacturer’s instructions. Signals were detected on X-ray films or on a ChemiDoc MP Imaging System (BioRad).

**FAMA-CFP AP-MS experiments**

**Seedling growth and crosslinking:** For the two FAMA-CFP AP-MS experiments, FAMApro::FAMA-CFP complementing the homozygous *fama-1* mutant and two lines expressing nuclear GFP under an early and late stomatal lineage specific promoter (SPCHpro::GFP<sub>NLS</sub> and MUTEpro::GFP<sub>NLS</sub>) were used. Large amounts of seedlings were grown for four days in liquid culture in ½ MS + 0.5 % sucrose (AP-MS experiment 1) or on filter paper that was placed on ½ MS + 0.5 % sucrose plates (AP-MS experiment 2). Seedlings were then treated with the proteasome inhibitor MG-132 by submerging them in liquid ½ MS + 0.5 % sucrose and 10 µM MG-132 for 2.5 (AP-MS 1) or 2 (AP-MS 2) hours to increase the abundance of FAMA and other unstable proteins and crosslinked with formaldehyde. Crosslinking was done by removing excess liquid, submerging the seedlings in crosslinking buffer (25 mM Hepes pH 7.5, 0.5 mM
EDTA, 1 mM PMSF, 10 mM MgCl$_2$, 75 mM NaCl) containing 0.25 % (AP-MS 1) or 0.125 % (AP-MS 2) formaldehyde and vacuum infiltrating for 15 minutes on ice. Crosslinking was stopped by adding glycine to a final concentration of 120 mM and vacuum infiltrating for another 10 minutes. Seedlings were washed with 800 ml ice cold H$_2$O and frozen in liquid nitrogen. Seedlings were treated in two batches per experiment.

**Affinity purification:** For AP, the frozen plant material was ground to a fine powder. Two biological replicates per genotype were used for AP-MS 1 and four replicates of the FAMA-CFP line and two of each of the control lines were used for AP-MS 2. AP-MS 1 was done to test overall performance of the experimental setup before doing a larger scale experiment. Therefore, only a small number of replicates was used. To increase statistical power in AP-MS 2, the replicate number was doubled. Approximately 15 g plant material (from one or two of the treatment batches) per sample were resuspended in 25 ml extraction buffer (25 mM Tris pH 7.5, 10 mM MgCl$_2$, 0.5 mM EGTA, 75 mM NaCl, 0.5/1 % Triton-X-100, 1 mM NaF, 0.5 mM Na$_3$VO$_4$, 15 mM beta glycerophosphate, 1 mM DTT, 1mM PMSF, 1x complete proteasome inhibitor) and incubated on a rotor wheel at 4°C for 30 to 45 minutes to thaw. The extracts were sonicated in 15 ml tubes in an ice bath using Bioruptor UCD-200 (Diagenode) on high setting. For experiment 1, samples were sonicated twice for five minutes with a 30 second on/off interval and a five-minute break on ice. For experiment 2, the sonication cycles were reduced to 2.5 minutes with a 30/60 seconds on/off interval. After sonication, 15 μl Lysonase (Millipore) were added and the extracts were incubated for another 30 minutes on a rotor wheel at 4°C before centrifuging at 12,000 g and 4°C for 10 minutes. The supernatants were filtered through two layers of Miracloth (Millipore) and incubated with 25 μl of GFP-Trap_MA beads (ChromoTek), pre-washed with extraction buffer, for five (AP-MS 1) or three (AP-MS 2) hours at 4°C on a rotor wheel. The beads were then washed six times with extraction buffer containing 100 mM NaCl.
(1x 15 ml, 5x 1ml), resuspended in 50 μl 2x Leammlí buffer and boiled for five minutes at 95°C. Samples were stored at -80°C until further processing.

**MS sample preparation:** Affinity purified proteins were separated by SDS-PAGE using Mini-PROTEAN TGX gels (BioRad) and stained with the Novex colloidal blue staining kit (Invitrogen). About 1 cm long sample areas were excised from the gel, and used for in-gel Tryptic digestion. Peptides were desalted using C18 ZipTips (Millipore). LC-MS/MS was done as described for the PL experiments.

**MS data analysis:** Protein identification and label-free quantification (LFQ) were done for both experiments together. RAW files from LC-MS/MS were searched in MaxQuant (version 1.6.2.6) (Tyanova, Temu, and Cox 2016) using the same settings as described for the PL experiments with one difference: LFQ minimum ratio count was set to 1 instead of 2. The mass spectrometry proteomics data (RAW files, MaxQuant search results and Perseus data analysis file) have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaino et al. 2013) with the dataset identifier PXD015212. Filtering and statistical analysis were done with Perseus (version 1.6.2.3) (Tyanova et al. 2016). The ‘proteinGroups.txt’ output file from MaxQuant was imported into Perseus using the LFQ intensities as Main category. The data matrix was filtered to remove proteins marked as ‘only identified by site’, ‘reverse’ and ‘potential contaminant’. LFQ values were log2 transformed and samples from AP-MS 1 and AP-MS 2 were separated for individual analysis. Proteins that were not identified/quantified in at least two (AP-MS 1) or three (AP-MS-2) FAMA-CFP or control samples (low confidence) were removed. Missing values were imputed for statistical analysis using the ‘Replace missing values from normal distribution’ function with the following settings: width = 0.3, down shift = 1.8 and mode = total matrix. Significantly enriched proteins were
identified by unpaired 2-sided Students t-tests comparing the FAMA-CFP with the control samples and using the integrated modified permutation-based FDR with 250 randomizations for multiple sample correction. Only proteins that were identified in two (AP-MS 1) or at least three (AP-MS 2) replicates of the FAMA-CFP samples were used for the test. FDR/S0 combinations were chosen to get a minimum number of ‘false negatives’ (statistically enriched proteins in controls). For AP-MS 1, the FDR was 0.2 and S0 was 0.5 and for AP-MS 2, the FDR was 0.01 and S0 was 0.5. Scatterplots with log2-fold changes and -log10 p-values form t-tests were done in RStudio.

**Primers**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>purpose</th>
<th>sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BirA-fw</td>
<td>cloning</td>
<td>CAGGCGCGCCGGTGGAGGCGGCTTCAGGAGGTGCGATGGGCAAGCCCATCCCGAC</td>
</tr>
<tr>
<td>BirA-NES-rv</td>
<td>cloning</td>
<td>CTGGCGCGCCCACCTCCGCACCTCCACCGCTCCGTCCAAGGCAGGCGCTCCAG</td>
</tr>
<tr>
<td>BirA-rv</td>
<td>cloning</td>
<td>TGCCGCGGCCACCTCCGCACCCTCCACCGCTCCGCCCTCCCCCGC</td>
</tr>
<tr>
<td>pDONR-mut-AscI-fw</td>
<td>cloning</td>
<td>GTACAAAGTGCTGGGCGCGCCTCCATGGTGAGCAAG</td>
</tr>
<tr>
<td>gFAMA-fw</td>
<td>cloning</td>
<td>CACCATGGATAAAGATTACTCGGTACGTACG</td>
</tr>
<tr>
<td>gFAMA-rv</td>
<td>cloning</td>
<td>AGTAAACACAATATTCCCAGGTAGGAAG</td>
</tr>
<tr>
<td>BirA-YFPnls-fw</td>
<td>cloning</td>
<td>GCATGCGGCGCATGGGCAAGCCCATC</td>
</tr>
<tr>
<td>BirA-YFPnls-rv</td>
<td>cloning</td>
<td>GATACCATGGACACCTCCGCCTTC</td>
</tr>
<tr>
<td>MUTE-fw</td>
<td>cloning</td>
<td>CACCATGTCTCATCGTGGTGAAGGAATCG</td>
</tr>
<tr>
<td>MUTE-rv</td>
<td>cloning</td>
<td>ATGGTAGAGACGATACCTTCATACG</td>
</tr>
<tr>
<td>ICE1-fw</td>
<td>cloning</td>
<td>CACCATGGGTCTTGACGGAAAACAATG</td>
</tr>
<tr>
<td>ICE1-rv</td>
<td>cloning</td>
<td>GATCATACCCATACCTTCG</td>
</tr>
<tr>
<td>SEU-fw</td>
<td>cloning</td>
<td>CACCATGGTACCAGAAGCCGCC</td>
</tr>
<tr>
<td>SEU-rv</td>
<td>cloning</td>
<td>CCGTTCCAATCAATTGTTGAAAC</td>
</tr>
<tr>
<td>LUH-fw</td>
<td>cloning</td>
<td>CACCATGGCTCAGGATATTGGGAAG</td>
</tr>
</tbody>
</table>
**LUH-rv** cloning  
CTTCCAAATCTTTACGGATTTGTCATG

**BZR1-fw** cloning  
CACCAGACCTCGGATGGAGCTACG

**BZR1-rv** cloning  
ACCACGAGCTTCCTCAACCTCGTC

**BIM1-fw** cloning  
CACCAGAGCCTTCCCATTTC

**BIM1-rv** cloning  
CTGTCCCGTCTTGAGCCGTT

**fama1-RP** genotyping  
CAATACAAAAAGCTCCCCTCAC

**fama1-LBb1.3** genotyping  
ATTTTGCCGATTTCGGAAC

---

**'PL toolbox’ vectors generated for common use**

<table>
<thead>
<tr>
<th>Plasmid name</th>
<th>Gateway sites</th>
<th>application</th>
<th>Addgene ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>pENTR_L1-YFP-Turbo-NES-L2</td>
<td>attL1-attL2</td>
<td>For N-terminal YFP-Tb1D fusion to non-nuclear proteins</td>
<td>127349</td>
</tr>
<tr>
<td>pENTR_L1-YFP-Turbo-L2</td>
<td>attL1-attL2</td>
<td>For N-terminal YFP-Tb1D fusion to nuclear proteins</td>
<td>127350</td>
</tr>
<tr>
<td>pENTR_L1-YFP-miniTurbo-NES-L2</td>
<td>attL1-attL2</td>
<td>For N-terminal YFP-mTb fusion to non-nuclear proteins</td>
<td>127351</td>
</tr>
<tr>
<td>pENTR_L1-YFP-miniTurbo-L2</td>
<td>attL1-attL2</td>
<td>For N-terminal YFP-mTb fusion to nuclear proteins</td>
<td>127352</td>
</tr>
<tr>
<td>pDONR_P2R-P3_R2-Turbo-NES-mVenus-STOP-L3</td>
<td>attR2-attL3</td>
<td>For C-terminal Tb1D-mVenus fusion to non-nuclear proteins</td>
<td>127353</td>
</tr>
<tr>
<td>pDONR_P2R-P3_R2-Turbo-mVenus-STOP-L3</td>
<td>attR2-attL3</td>
<td>For C-terminal Tb1D-mVenus fusion to nuclear proteins</td>
<td>127354</td>
</tr>
<tr>
<td>pDONR_P2R-P3_R2-miniTurbo-NES-mVenus-STOP-L3</td>
<td>attR2-attL3</td>
<td>For C-terminal mTb-mVenus fusion to non-nuclear proteins</td>
<td>127355</td>
</tr>
<tr>
<td>pDONR_P2R-P3_R2-miniTurbo-mVenus-STOP-L3</td>
<td>attR2-attL3</td>
<td>For C-terminal mTb-mVenus fusion to nuclear proteins</td>
<td>127356</td>
</tr>
</tbody>
</table>

---

**Entry vectors for expressing BirA*/TbID/mTb + YFP/mVenus under a promoter of choice**

<table>
<thead>
<tr>
<th>Plasmid name</th>
<th>Gateway sites</th>
<th>application</th>
<th>Addgene ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>pENTR_L1-BirA(R118G)-NES-YFP-STOP-L2</td>
<td>attL1-attL2</td>
<td>For expressing cytosolic BirA* under a promoter of choice</td>
<td>127357</td>
</tr>
<tr>
<td>pENTR_L1-Turbo-NES-YFP-STOP-</td>
<td>attL1-attL2</td>
<td>For expressing cytosolic Tb1D</td>
<td>127358</td>
</tr>
</tbody>
</table>
L2 under a promoter of choice

pENTR_L1-miniTurbo-NES-YFP-STOP-L2 attL1-attL2 For expressing cytosolic mTb under a promoter of choice 127359

pENTR_L1-BirA(R118G)-YFP-NLS-STOP-L2 attL1-attL2 For expressing nuclear BirA* under a promoter of choice 127360

pENTR_L1-Turbo-YFP-NLS-STOP-L2 attL1-attL2 For expressing nuclear TbID under a promoter of choice 127361

pENTR_L1-miniTurbo-YFP-NLS-STOP-L2 attL1-attL2 For expressing nuclear mTb under a promoter of choice 127362

Acknowledgements

We thank Annika Weimer for suggestions for proximity labeling construct design and testing of biotin ligase activity and other previous and current members of the Bergmann lab for creating tools that enabled this work and for critical discussion. Work in this publication has been funded with support from the following sources: AM was supported in part by a Schrödinger fellowship.
from the Austrian Science Fund FWF (grant number J4019-B29) and TCB was supported by Dow Graduate Research and Lester Wolfe Fellowships. DCB is an investigator of the Howard Hughes Medical institute. Further support came from the NIH grant R01-CA186568 (to AYT) and from the Carnegie endowment fund to the Carnegie mass spectrometry facility.

**Supplementary files**

**Supplementary File 1: non-cropped immunoblots**

**Supplementary File 2: XLS file containing the following additional protein lists for the ‘FAMA interactome’ experiment (Figure 5 and supplements):**

- **Table 1, ‘identifed proteins’:** Proteins identified in the ‘FAMA interactome’ experiment;
- **Table 2, ‘FAMA-TbID enriched’:** Proteins significantly enriched in FAMA-TbID vs. WT and FAMA-nucTbID samples and in FAMA-nucTbID vs. WT;
- **Table 3, ‘enrichment over time’:** Proteins significantly enriched in WT, FAMA-TbID, and FAMA-nucTbID after biotin treatment compared to the no-biotin control;
- **Table 4, ‘FAMA complex candidates’:** FAMA interaction candidates from Figure 5 and Table 1;
- **Table 5, ‘FAMA AP-MS’:** Proteins significantly enriched in AP-MS experiments with FAMA-CFP

**Supplementary File 3: XLS file containing the following additional protein lists for the ‘nuclear proteome’ experiment (Figure 6 and supplements):**

- **Table 1, ‘identified proteins’:** Proteins identified in the ‘nuclear proteome’ experiment;
- **Table 2, ‘enriched with UBQnucTbID’:** Proteins significantly enriched in UBQnucTbID vs. WT samples;
- **Table 3, ‘enriched with FAMA-nucTbID’:** Proteins significantly enriched in FAMA-nucTbID vs. WT and FAMA-nucTbID vs. UBQnucTbID samples;
- **Table 4, ‘published nuclear proteomes’:** Proteins found in proteomics studies of purified nuclei or sub-nuclear compartments by mass spectrometry or in localization
studies with fluorescent protein fusions; Table 5, ‘localization prediction’: Protein localization data used for pie charts in Figure 6; Table 6, ‘published GC proteome’: Proteins found in published guard cell proteomics experiments; Table 7, ‘GO term enrichment’: Enriched GO terms of nuclear proteins in Figure 6 and supplements; Table 8, ‘nuclear compartments’: Selected marker proteins and protein classes for different nuclear compartments and domains identified in the UBQnucTbID and FAMA nucTbID dataset.

References


Khan, M., J. Y. Youn, A. C. Gingras, R. Subramaniam, and D. Desveaux. 2018. 'In planta proximity dependent biotin identification (BioID)', *Sci Rep*, 8: 9212. DOI: 10.1038/s41598-018-27500-3

64


Li, M., and F. D. Sack. 2014. 'Myrosin idioblast cell fate and development are regulated by the Arabidopsis transcription factor FAMA, the auxin pathway, and vesicular trafficking', *Plant Cell*, 26: 4053-66. DOI: 10.1105/tpc.114.129726

Li, X., R. Yang, and H. Chen. 2018. 'The Arabidopsis thaliana Mediator subunit MED8 regulates plant immunity to Botrytis Cinerea through interacting with the basic helix-loop-helix (bHLH) transcription factor FAMA', *PLoS One*, 13: e0193458. DOI: 10.1371/journal.pone.0193458


Matos, J. L., O. S. Lau, C. Hachez, A. Cruz-Ramirez, B. Scheres, and D. C. Bergmann. 2014. 'Irreversible fate commitment in the Arabidopsis stomatal lineage requires a FAMA and RETINOBLASTOMA-RELATED module', *Elife*, 3. DOI: 10.7554/eLife.03271


Nelms, Brad, and Virginia Walbot. 2019. 'Defining the developmental program leading to meiosis in maize', *Science*, 364: 52-56. DOI: 10.1126/science.aav6428


Regev, A., S. A. Teichmann, E. S. Lander, I. Amit, C. Benoist, E. Birney, B. Bodenmiller, P.
Campbell, P. Carninci, M. Clatworthy, H. Clevers, B. Deplancke, I. Dunham, J.
Eberwine, R. Eils, W. Enard, A. Farmer, L. Fugger, B. Gottgens, N. Hacohen, M. Haniffa,
M. Hemberg, S. Kim, P. Kleenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,
M. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg,
J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea,


Stuart, T., and R. Satija. 2019. 'Integrative single-cell analysis', *Nat Rev Genet*. DOI: 10.1038/s41576-019-0093-7


### Table 1: FAMA complex candidates from Figure 5

<table>
<thead>
<tr>
<th>AGI</th>
<th>Gene name</th>
<th>Functional annotation</th>
<th>Subcellular localization</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G26744</td>
<td>ICE1, SCR2</td>
<td>bHLH transcription factor</td>
<td>N</td>
</tr>
<tr>
<td>AT2G46510</td>
<td>AIB, JAM1</td>
<td>bHLH transcription factor</td>
<td>N</td>
</tr>
<tr>
<td>AT4G16430</td>
<td>JAM3</td>
<td>bHLH transcription factor</td>
<td>N, C</td>
</tr>
<tr>
<td>AT5G08130</td>
<td>BIM1</td>
<td>bHLH transcription factor</td>
<td>N</td>
</tr>
<tr>
<td>AT1G75080</td>
<td>BZR1</td>
<td>transcription factor</td>
<td>N, C</td>
</tr>
<tr>
<td>AT5G11060</td>
<td>KNAT4</td>
<td>homeobox transcription factor</td>
<td>N, C</td>
</tr>
<tr>
<td>AT2G41900</td>
<td>OX22, TCF7</td>
<td>zinc finger transcription factor</td>
<td>(N), C</td>
</tr>
<tr>
<td>AT1G79000</td>
<td>HAC1, PCA12</td>
<td>transcriptional co-activator (histone acetyltransferase)</td>
<td>N, C</td>
</tr>
<tr>
<td>AT4G04920</td>
<td>MED16, SFR6</td>
<td>transcriptional co-activator (mediator complex)</td>
<td>N</td>
</tr>
<tr>
<td>AT1G43850</td>
<td>SEU</td>
<td>transcriptional co-repressor adapter</td>
<td>N</td>
</tr>
<tr>
<td>AT4G32551</td>
<td>LUG, RON2</td>
<td>transcriptional co-repressor</td>
<td>N</td>
</tr>
<tr>
<td>AT2G32700</td>
<td>LUIH, MUM1</td>
<td>transcriptional co-repressor</td>
<td>N</td>
</tr>
<tr>
<td>AT3G15880</td>
<td>TPR4, WSIP2</td>
<td>transcriptional co-repressor</td>
<td>N</td>
</tr>
<tr>
<td>AT5G27030</td>
<td>TPR3</td>
<td>transcriptional co-repressor</td>
<td>N</td>
</tr>
<tr>
<td>AT5G02500</td>
<td>HSP70-1</td>
<td>HSP70 chaperone</td>
<td>N, C</td>
</tr>
<tr>
<td>AT5G02490</td>
<td>HSP70-2</td>
<td>HSP70 chaperone</td>
<td>N, C</td>
</tr>
<tr>
<td>AT3G09440</td>
<td>HSP70-3</td>
<td>HSP70 chaperone</td>
<td>N, C</td>
</tr>
<tr>
<td>AT3G12580</td>
<td>HSP70-4</td>
<td>HSP70 chaperone</td>
<td>N, C</td>
</tr>
<tr>
<td>AT5G22060</td>
<td>J2</td>
<td>HSP70 co-chaperone</td>
<td>N</td>
</tr>
<tr>
<td>AT3G44110</td>
<td>J3</td>
<td>HSP70 co-chaperone</td>
<td>N, C, MA</td>
</tr>
<tr>
<td>AT1G62740</td>
<td>HOP2</td>
<td>HSP70/70 co-chaperone</td>
<td>(N), C</td>
</tr>
<tr>
<td>AT3G25230</td>
<td>FKBP62, ROF1</td>
<td>HSP70/70 co-chaperone</td>
<td>(N), C</td>
</tr>
<tr>
<td>AT4G22670</td>
<td>HIP1, TPR11</td>
<td>HSP70/70 co-chaperone</td>
<td>N, C</td>
</tr>
<tr>
<td>AT4G02450</td>
<td>P23-1</td>
<td>HSP90 co-chaperone</td>
<td>N, C</td>
</tr>
<tr>
<td>AT5G56460</td>
<td></td>
<td>putative protein kinase</td>
<td>PM</td>
</tr>
<tr>
<td>AT5G35410</td>
<td>SOSS2, CIPK24</td>
<td>protein kinase</td>
<td>N, C, PM</td>
</tr>
<tr>
<td>AT3G54170</td>
<td>FIP3</td>
<td>m6A methyltransferase complex component</td>
<td>N</td>
</tr>
<tr>
<td>AT1G02140</td>
<td>HAP1, MAGO</td>
<td>exon-junction complex component</td>
<td>N, C</td>
</tr>
<tr>
<td>AT5G41880</td>
<td>POLA3, POLA4</td>
<td>putative DNA polymerase alpha subunit</td>
<td>N</td>
</tr>
<tr>
<td>AT1G72390</td>
<td>TIC</td>
<td>nuclear clock regulation factor</td>
<td>N</td>
</tr>
<tr>
<td>AT2G41100</td>
<td>TCH3, CAL12</td>
<td>calcium-binding protein</td>
<td>N</td>
</tr>
<tr>
<td>AT1G72290</td>
<td>PHL</td>
<td>nuclear receptor/co-activator</td>
<td>N, C</td>
</tr>
<tr>
<td>AT1G20110</td>
<td>FREE1, FYVE1</td>
<td>ESCRT-I complex component</td>
<td>C, ES, N</td>
</tr>
<tr>
<td>AT1G51860</td>
<td>IAP1</td>
<td>C3HC4-type RING-finger domain protein</td>
<td>MA, N</td>
</tr>
<tr>
<td>AT1G12200</td>
<td>FMO</td>
<td>putative flavin monoxygenase</td>
<td>N/A</td>
</tr>
<tr>
<td>AT3G353200</td>
<td>PAL2</td>
<td>phenylalanine ammonia-lyase</td>
<td>N, C, EX</td>
</tr>
<tr>
<td>AT3G23840</td>
<td>CER28-LIKE</td>
<td>acyl-CoA-dependent acyltransferase</td>
<td>N/A</td>
</tr>
<tr>
<td>AT5G13710</td>
<td>CPH, SMT1</td>
<td>C-24 sterol methyl transferase</td>
<td>N</td>
</tr>
<tr>
<td>AT1G63180</td>
<td>UGE3</td>
<td>UDP-Glucose 4-E-pimerase</td>
<td>C*</td>
</tr>
<tr>
<td>AT5G17990</td>
<td>PAT1, TRP1</td>
<td>phosphorobiosylnlantrile transferase</td>
<td>CP*</td>
</tr>
<tr>
<td>AT1G15980</td>
<td>NDH48, NDF1</td>
<td>chloroplast NAD(P)H dehydrogenase complex subunit</td>
<td>CP</td>
</tr>
<tr>
<td>AT4G30720</td>
<td>PDE327</td>
<td>putative oxidoreductase/electron carrier</td>
<td>CP</td>
</tr>
<tr>
<td>AT1G50570</td>
<td></td>
<td>undescribed protein</td>
<td>N</td>
</tr>
<tr>
<td>AT1G30070</td>
<td></td>
<td>undescribed protein</td>
<td>N</td>
</tr>
<tr>
<td>AT5G15680</td>
<td></td>
<td>undescribed protein</td>
<td>N/A</td>
</tr>
<tr>
<td>AT5G53230</td>
<td></td>
<td>undescribed protein</td>
<td>N</td>
</tr>
<tr>
<td>AT4G25290</td>
<td></td>
<td>undescribed protein</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Column labels:** Enriched at time point: time points at which a protein was significantly enriched are marked with Y. AGI: Arabidopsis gene identifier. Subcellular localization: as described for fluorescent protein fusions in literature unless marked with * (localization inferred from functional annotation): N, nucleus; (N), nucleus under heat or other stress; C, cytosol; EX, extracellular; PM, plasma membrane; ES, endosomes; MA, membrane (associated); CP, chloroplast; N/A, localization unknown (no experimental evidence found and localization cannot be clearly inferred from function).

For further information on the candidate proteins and selected references see Supplementary file 2 – Table 4.
Figure legends

Figure 1 and supplements:

**Figure 1:** TbdID and mTb exhibit robust biotinylation activity in *N. benthamiana* and Arabidopsis

**(A) Overview of the experimental setup.** UBQ10pro::BirA*/TbdID/mTb-YFP constructs with an NLS or NES for nuclear or cytosolic localization were used for transient and stable transformation of *N. benthamiana* and *A. thaliana*, respectively. Tobacco leaf discs or whole Arabidopsis seedlings were submerged in a 250 µM biotin solution, briefly vacuum infiltrated, incubated for one hour at room temperature (RT, ~22°C) and frozen. Untreated controls were infiltrated with H2O or frozen directly. Expression and activity of the BirA versions were analyzed by immunoblotting.

**(B-C) Biotin ligase activity in *N. benthamiana* (B) and Arabidopsis (C)** Streptavidin (SA) and anti-GFP immunoblots (IB) of protein extracts from tobacco leaf discs and Arabidopsis expressing the cytosolic BirA variants without (-) and with (+) biotin treatment. Untransformed tobacco leaves and Col-0 wild-type (WT) seedlings were used as controls. Each sample is a pool of 3 leaf discs or ~ 30 seedlings. Coomassie Brilliant Blue-stained membranes (CBB) are shown as a loading controls. Asterisks mark the positions of naturally biotinylated proteins.

For microscopy images showing the subcellular localization of the BirA variants in *N. benthamiana* and Arabidopsis see Figure 1 – figure supplements 1 and 2. For immunoblots showing the activity and expression of both cytosolic and nuclear BirA versions in *N. benthamiana* and Arabidopsis see Figure 1 – figure supplements 3 and 4. For a schematic overview over the generation and composition of the available vectors in the ‘PL toolbox’ see Figure 1 – figure supplement 5.

**Figure 1 – figure supplement 1: Subcellular localization of biotin ligase constructs in transiently transformed *N. benthamiana* leaves**

Epifluorescence images of *N. benthamiana* leaves transformed with UBQ10pro::BirA-YFP	extsubscript{NLS} or UBQ10pro::BirA-

**Figure 1 – figure supplement 2: Subcellular localization of biotin ligase constructs in stable Arabidopsis lines**
Confocal microscopy images of the cotyledon epidermis of five day old Arabidopsis seedlings transformed with UBQ10pro::BirA-YFP<sub>NLS</sub> or UBQ10pro::BirA<sup>−NES</sup>YFP expression vectors (BirA = BirA*, mTb or TblD). Shown are the YFP channel (yellow), the cell walls stained with propidium iodide (purple) and an overlay. Scale bar = 50 µm. Constructs and lines used for images are indicated on the left. Nuclear targeted BirA-YFP (NLS) constructs were mostly nuclear, and cytosol targeted (NES) constructs were excluded from the nucleus.

**Figure 1 – figure supplement 3: TblD and mTb are highly active in the cytosol and nucleus of transiently transformed *N. benthamiana* leaves**

*N. benthamiana* leaves were transformed with UBQ10pro::BirA-YFP<sub>NLS</sub> or UBQ10pro::BirA<sup>−NES</sup>YFP expression vectors (BirA = BirA*, mTb or TblD). Two days after infiltration, leaf discs expressing either of the six constructs or non-transformed leaves (not transf.) were briefly vacuum infiltrated with 50 µM biotin and incubated for 10, 30 or 60 minutes at room temperature (~22°C). Untreated samples were used as controls to visualize the background activity of the biotin ligases with endogenous biotin. Each sample is a pool of 2 leaf discs. Activity and expression of the BirA-YFP variants were analyzed by immunoblots (IB) with streptavidin-HRP (SA) and anti-GFP antibodies. Coomassie Brilliant Blue-stained membranes (CBB) are shown as loading controls. Asterisks mark the positions of naturally biotinylated proteins.

**Figure 1 – figure supplement 4: TblD is more active than mTb, but also produces more background labeling in Arabidopsis**

Five day old Arabidopsis seedlings expressing UBQ10pro::BirA-YFP<sub>NLS</sub> or UBQ10pro::BirA<sup>−NES</sup>YFP constructs (BirA = BirA*, mTb or TblD) were submerged in 250 µM biotin, briefly vacuum infiltrated and incubated at room temperature (22°C) for one hour (+). Non-treated seedlings were used as controls (-) to visualize the background activity of the biotin ligases with endogenous biotin. Activity and expression of the BirA-YFP variants were analyzed by immunoblots (IB) with streptavidin-HRP (SA) and anti-GFP antibodies. Coomassie Brilliant Blue-stained membranes (CBB) are shown as loading controls. Asterisks mark the positions of naturally biotinylated proteins. Each sample is a pool of ~ 30 seedlings. At least two lines were tested per construct; strong cytosolic lines and a strong nuclear TblD line are on the left, weak nuclear lines and WT are on the right. Blots were made and exposed in parallel. Different exposure times are shown to enable comparison of weak and strong lines. Short exposure times were five seconds, long exposures times were 30 seconds (SA blot) and 70 seconds (anti-GFP blot). mTb is less active than TblD, but has less background labeling with endogenous biotin, especially in low-expressing lines. WT samples and samples from lines 3, 10 and 3 of the BirA*<sup>−NES</sup>YFP, mTb<sup>−NES</sup>YFP and TblD<sup>−NES</sup>YFP plants were also used for the blots shown in Figure 1C.

**Figure 1 – figure supplement 5: Generating a toolbox of gateway-compatible vectors for PL in plants**
Schematic overview over cloning steps involved in the generation of the PL toolbox and over the available vectors (highlighted as orange boxes). Analogous constructs for different biotin ligase variants are summarized (indicated by BirA*/TbID/mTb or TblID/mTb). Primers and restriction sites used for PCR and restriction cloning are indicated on the top or bottom of the constructs. V5 = V5 tag, L = 8 amino acid linker (GGGSGGGG), NES = nuclear export signal, NLS = nuclear localization signal, POI = protein of interest. Available plasmids include: ❶ 8 entry vectors for N- and C-terminal tagging of proteins of interest with TbID or mTb plus a fluorophore, ❷ 6 entry vectors for expressing any of the three BirA variants under a promoter of interest with a NES or NLS and a fluorophore, and ❸ 6 binary plant transformation vectors for expressing any of the three BirA variants with an NES or NLS under the UBQ10 promoter. A detailed description and primer sequences are included in the materials and methods section. A list of available plasmids with Addgene catalogue number can also be found in the methods section.

Figure 2 and supplements:

Figure 2: TbID and mTb work quickly and tolerate a range of experimental conditions in Arabidopsis seedlings

(A-D) Dependency of TbID and mTb activity on labeling time, temperature, biotin concentration and biotin application. Four to five day old seedlings were treated with biotin as described below. Activity and expression of the TblID/mTb-YFP constructs were analyzed by immunoblots (IB) with streptavidin-HRP (SA) and anti-GFP antibodies. Coomassie Brilliant Blue-stained membranes (CBB) are shown as a loading controls. Asterisks mark the positions of naturally biotinylated proteins. Each sample is a pool of ~30-50 seedlings. (A) Labeling time. Wild-type (WT) and UBQ10pro::TbID-YFP<sub>NLS</sub> (TbID-YFP) seedlings were submerged in 250 µM biotin, briefly vacuum infiltrated and incubated for the indicated time at room temperature (22°C). A control sample was taken before treatment (0'). (B) Temperature-dependency. WT, UBQ10pro::TbID-YFP<sub>NLS</sub> (TbID-YFP) and UBQ10pro::mTb<sub>NES</sub>-YFP (mTb-YFP) seedlings were submerged in 250 µM biotin and incubated for one hour at the indicated temperature. Control samples (-) were incubated in H₂O at 22°C. (C) Biotin concentration. UBQ10pro::TbID<sub>NES</sub>-YFP (TbID-YFP) and UBQ10pro::mTb<sub>NES</sub>-YFP (mTb-YFP) seedlings were submerged in 0.5 to 250 µM biotin and incubated for one hour at room temperature. A control sample was taken before treatment (0 µM). (D) Biotin application. WT and UBQ10pro::TbID-YFP<sub>NLS</sub> (TbID-YFP) seedlings were submerged in 250 µM biotin, briefly vacuum infiltrated (vac. inf.) or not and incubated for one hour at room temperature. A control sample was taken before treatment.

For a longer time course in Arabidopsis and quantification of the immunoblots shown in (C) see Figure 2 – figure supplements 1 and 3. For a short time course and temperature dependency of TbID and mTb in <i>N. benthamiana</i> see Figure 1 – figure supplement 3 and Figure 2 – figure supplement 2.
Figure 2 – figure supplement 1: Biotinylation by TblID in Arabidopsis increases over time

Labeling time course with TblID in four day old Arabidopsis seedlings expressing the UBQ10pro::TbID-YFP<sub>NLS</sub> construct (TbID-YFP<sub>NLS</sub>). Seedlings were submerged in 50 μM biotin, briefly vacuum infiltrated and incubated for the indicated time at room temperature (22°C). A control sample and a wild-type (WT) sample were taken before biotin treatment. Activity of TblID over time was analyzed by immunoblot (IB) with streptavidin-HRP (SA). The Coomassie Brilliant Blue-stained membrane (CBB) is shown as loading control. Asterisks mark the positions of naturally biotinylated proteins. Each sample is a pool of seedlings. The increase of labeling between three and five hours of biotin treatment demonstrates that labeling of nuclear proteins is not saturated after three hours.

Figure 2 – figure supplement 2: TblID and mTb are active from 22°C to 37°C in N. benthamiana

Streptavidin-HRP (SA) and anti-GFP immunoblots (IB) showing the activity and expression of TblID and mTb at room temperature (22°C), 30°C and 37°C. N. benthamiana leaves were transformed with UBQ10pro::TbID<sub>NES</sub>-YFP (TbID-NES-YFP) or UBQ10pro::mTb<sub>NES</sub>-YFP (mTb-NES-YFP) expression vectors. Two days after infiltration, leaf discs expressing the TblID and mTb constructs and non-transformed leaves were briefly vacuum infiltrated with H<sub>2</sub>O (-) or 250 μM biotin (+) and incubated at the indicated temperature for one hour. Each sample is a pool of tree leaf discs. The Coomassie Brilliant Blue-stained membrane (CBB) is shown as loading control. Asterisks mark the positions of naturally biotinylated proteins.

Figure 2 – figure supplement 3: Quantification of TblID and mTb activity in Arabidopsis at different biotin concentrations

Quantification of the streptavidin-HRP (SA) immunoblots (IB) shown in Figure 2C. Five day old seedlings transformed with UBQ10pro::TbID<sub>NES</sub>-YFP (TbID-NES-YFP) and UBQ10pro::mTb<sub>NES</sub>-YFP (mTb-NES-YFP) were submerged in 0.5 to 250 μM biotin and incubated for one hour at room temperature (22°C). A control sample was taken before treatment (0 μM). The anti-GFP blot shows expression of the YFP fusion proteins. Quantification was done in FIJI (ImageJ) by drawing five parallel vertical lines through each of the lanes at different positions and reading out the signal intensity along the line. The graphs on the right show the average signal intensity of the five lines along the length of the gel. The peaks associated with TblID-YFP and mTb-YFP are indicated and the positions of naturally biotinylated proteins are marked by asterisks.

Figure 3 and supplements:

Figure 3: TblID works in different developmental stages and organs of Arabidopsis and does not require vacuum infiltration of biotin
TbID activity in shoots and roots of 10 day old plate-grown UBQ10pro::TbID-YFP<sub>NLS</sub> (TbID-YFP<sub>NLS</sub>) seedlings, and in rosette leaves and unopened flower buds of mature soil-grown plants. Col-0 wild-type (WT) was used as control. The plant material was submerged in a 250 μM biotin solution, briefly vacuum infiltrated until air spaces were filled with liquid or not vacuum infiltrated and incubated at room temperature (22°C) for one hour. Control samples were taken before biotin treatment. Samples are pools of three shoots or roots, two rosette leaves or four inflorescences. Activity and expression of TbID-YFP were analyzed by immunoblots (IB) with streptavidin-HRP (SA) and anti-GFP antibodies. Coomassie Brilliant Blue-stained membranes (CBB) are shown as a loading controls. Asterisks mark the positions of naturally biotinylated proteins. Epifluorescence images of seedlings and mature tissues of the TbID-YFP line are shown on top.

For immunoblots showing TbID activity and background in four to 14 day old whole seedlings and shoots and roots of six to 14 day old seedlings see Figure 3 – figure supplements 1 and 2. For further microscopy images of the TbID-YFP line see Figure 3 – figure supplement 3.

**Figure 3 – figure supplement 1: Activity and background labeling of TbID are similar in seedlings ranging from 4 to 14 days of age**

Plate-grown Arabidopsis wild-type (WT) and UBQ10pro::TbID-YFP<sub>NLS</sub> (TbID-YFP<sub>NLS</sub>) seedlings of the indicated age (dp = days post germination) were submerged in a 250 μM biotin solution, briefly vacuum infiltrated and incubated for one hour at room temperature (22°C; + biotin). Non-treated control samples (- biotin) were taken at the same time. Each sample is a pool of seedlings. Activity and expression of TbID-YFP<sub>NLS</sub> were analyzed by immunoblots (IB) with streptavidin-HRP (SA) and anti-GFP antibodies. Coomassie Brilliant Blue-stained membranes (CBB) are shown as a loading controls. Asterisks mark the positions of naturally biotinylated proteins.

**Figure 3 – figure supplement 2: Activity and background labeling of TbID are similar in roots and shoots of 6 to 14 day old seedlings**

Six, 10 and 14 days post germination (dp = days post germination), plate-grown Arabidopsis wild-type (WT) and UBQ10pro::TbID-YFP<sub>NLS</sub> (TbID-YFP<sub>NLS</sub>) seedlings were divided into a shoot and root section. Shoots and roots were submerged separately in a 250 μM biotin solution, briefly vacuum infiltrated and incubated for one hour at room temperature (22°C; +). Non-treated control samples (-) were taken at the same time. Each sample is a pool of seedlings. Activity and expression of TbID-YFP<sub>NLS</sub> were analyzed by immunoblots (IB) with streptavidin-HRP (SA) and anti-GFP antibodies, respectively. Coomassie Brilliant Blue-stained membranes (CBB) are shown as a loading control. Asterisks mark the positions of naturally biotinylated proteins.

**Figure 3 – figure supplement 3: UBQ10pro::TbID-YFP<sub>NLS</sub> is expressed throughout the whole plant**
Epifluorescence microscopy images of plate-grown Arabidopsis seedlings and soil-grown mature tissues from the strong TbID-YFP<sub>NLS</sub> line used in Figures 2 and 3 and the PL experiments in Figures 4 to 6. Shown are the YFP and brightfield (BF) channel and an overlay. Scale bar = 100 μm. Abbreviations: P, petal; S, sepal; A, anther; St, stigma; O, ovary. The overlay images of the 5 day old cotyledon and root and of mature flowers are also used in Figure 3.

Figure 4 and supplements:

Figure 4: Testing TbID's potential to label protein interactors and subcellular proteomes in a rare cell type in Arabidopsis

(A) Plant lines generated for the ‘FAMA interactome’ and ‘nuclear proteome’ experiments. Line names and genotypes are given on the top, schematic expression of the TbID fusion proteins (yellow dots) in a leaf and confocal microscopy images of the epidermis of five day old seedlings are shown below (TbID fusion protein = yellow; propidium iodide-stained cell wall = purple; scale bar = 50 μM). FAMA-TbID and FAMA<sub>nuc</sub>TbID constructs are expressed in young guard cells, while the <sub>ubQ</sub>nucTbID construct is expressed ubiquitously. (B) Time course to optimize time points for the experiments. Five day old wild-type (WT) and FAMA-TbID seedlings were submerged in 250 μM biotin, briefly vacuum infiltrated and incubated for the indicated time at room temperature 22°C. Control samples were taken before treatment (0'). Samples are pools of ~ 30 seedlings. Activity and expression of FAMA-TbID were analyzed by immunoblots (IB) with streptavidin-HRP (SA) and anti-GFP antibodies. The Coomassie Brilliant Blue-stained membrane (CBB) is shown as loading control. Asterisks and circles mark the positions of naturally biotinylated proteins and putative FAMA-TbID targets, respectively. (C-D) Scheme of samples and comparisons used in the ‘FAMA interactome’ (C) and ‘nuclear proteome’ (D) experiments. (E) Simplified workflow of the experimental procedure from biotin labeling to protein identification by liquid chromatography coupled to mass spectrometry (LC-MS/MS). Three biological replicates were used. Abbreviations: SA, streptavidin; LFQ, label free quantification.

For larger extracts of the confocal microscopy images of the plant lines used in the PL experiments shown in (A) see Figure 4 – figure supplement 1. For complementation of the <i>fama</i>-1 phenotype by the FAMApro::FAMA-TbID-Venus construct see Figure 4 – figure supplement 2. For another labeling time course with the FAMA-TbID line using shorter labeling times with and without vacuum infiltration of biotin see Figure 4 – figure supplement 3. For immunoblots showing successful labeling and purification of proteins for the ‘FAMA interactome’ and ‘nuclear proteome’ experiments see Figure 4 – figure supplements 4 and 5. For immunoblots demonstrating the importance of the biotin depletion step and a comparison of different biotin depletion strategies see Figure 4 – figure supplements 6 and 7.
Figure 4 – figure supplement 1: Expression of the TbID constructs in lines used for the ‘FAMA interactome’ and ‘nuclear proteome’ PL experiments

Confocal microscopy images of the cotyledon epidermis of five day old Arabidopsis seedlings. Shown are the YFP channel (yellow), the cell walls stained with propidium iodide (purple) and an overlay. Scale bar = 50 µm. Excerpts of the overlay images were used in Figure 4A. FAMA-TbID-Venus and TbID-YFP\textsubscript{NLS} controlled by the FAMA promoter are expressed in developing guard cells (GCs). Expression is highest in dividing and young GCs and decreases as GCs mature. We did not see expression in developing myrosin cells (Li and Sack 2014), but cannot rule out expression at growth stages that were not used in our experiments. TbID-YFP\textsubscript{NLS} controlled by the UBQ10 promoter is expressed ubiquitously (see also Figure 3 – figure supplement 3).

Figure 4 – figure supplement 2: The FAMApro::FAMA-TbID-Venus construct rescues the \textit{fama-1} mutant phenotype

(A) FAMA-TbID-Venus rescues the seedling-lethal growth phenotype of \textit{fama-1}. Col-0 wild-type (WT), \textit{fama-1}-/- and the FAMA-TbID rescue line (FAMApro::FAMA-TbID-Venus in \textit{fama-1}-/- background) were grown on \(\frac{1}{2}\) MS plates supplemented with 0.5% sucrose. Pictures of the seedlings were taken after 9 and 12 days. (B) FAMA-TbID-Venus rescues the stomatal development phenotype of \textit{fama-1}. Brightfield images of the cotyledon epidermis of seedlings five and 9 days post germination (dpg). Scale bar = 50 µm. In the \textit{fama-1} mutant, guard cell precursors continue dividing and fail to differentiate into stomata with a central pore, resulting in clusters of guard cell precursors (black arrowheads).

Figure 4 – figure supplement 3: FAMA-TbID self-labeling increases within five minutes of biotin treatment regardless of vacuum infiltration

Time course to determine time points for the PL experiments with FAMA-TbID. Five day old wild-type (WT) and FAMApro::FAMA-TbID-Venus (FAMA-TbID-Venus) seedlings were submerged in 250 µM biotin, briefly vacuum infiltrated or not and incubated for the indicated time at room temperature (22°C). Control samples were taken before treatment (0 minutes). Activity and expression of TbID were analyzed by immunoblots (IB) with streptavidin-HRP (SA) and anti-GFP antibodies. The Coomassie Brilliant Blue-stained membrane (CBB) is shown as loading control. Asterisks and circles mark the positions of naturally biotinylated proteins and putative FAMA-TbID targets, respectively. Each sample is a pool of ~ 20 seedlings.

Figure 4 – figure supplement 4: Confirming successful labeling of proteins in the PL experiment

For the PL experiments shown in Figures 4-6, five day old wild-type (WT), FAMA-TbID, FAMA\textsubscript{nuc}TbID and UBQ\textsubscript{nuc}TbID seedlings were submerged in a 50 µM biotin solution for 0, 0.5 and 3 hours and washed repeatedly with ice cold water. Activity and expression of the TbID constructs at the tree time points were analyzed by immunoblots (IB) with
streptavidin-HRP (SA) and anti-GFP antibodies. The Coomassie Brilliant Blue-stained membrane (CBB) is shown as loading control. Asterisks mark the positions of naturally biotinylated proteins. See Figure 4 – figure supplement 5 for affinity purification of the biotinylated proteins.

**Figure 4 - figure supplement 5: Affinity purification of biotinylated proteins in the PL experiment**

Five day old wild-type (WT), FAMA-TbID, FAManucTbID and UBQnucTbID seedlings were treated with biotin for 0, 0.5, and 3 hours (see Figure 4 – figure supplement 4) and used for affinity purification (AP) of biotinylated proteins. Samples were taken at different steps of the AP procedure and used for immunoblots (IB) with streptavidin-HRP (SA) to confirm successful biotin depletion and binding of biotinylated proteins to the beads (A-D). One of three biological replicates is shown for each sample. The positions of the TbID fusion proteins are indicated on the right and naturally biotinylated proteins are marked by asterisks. Protein extracts before (extract) and after biotin depletion (after PD-10) confirm that no significant sample loss occurred at this step. The protein extracts after AP (supern.) were depleted of biotinylated proteins, indicating binding of these proteins to the beads. Small aliquots of the beads were used for elution of bound proteins. In most cases, only the naturally biotinylated proteins could be eluted, presumably because TbID targets are biotinylated on multiple sites and bind the beads more strongly.

**Figure 4 – figure supplement 6: Free biotin from biotin treatment out-competes biotinylated proteins for streptavidin bead binding**

Four day old wild-type (WT) and UBQ10pro::TbID-YFP<sub>NLS</sub> (TbID-YFP<sub>NLS</sub>) seedlings were submerged in H<sub>2</sub>O (-) or 50 μM biotin for one hour and used for affinity purification of biotinylated proteins with streptavidin (SA) beads. Protein extracts before and after overnight incubation with the beads and proteins eluted from the beads were used for immunoblotting (IB) with streptavidin-HRP (SA) (A) and for SDS-PAGE (B) to analyze binding of biotinylated proteins to the beads. Biotin treatment led to an increase of labeling in the TbID-YFP<sub>NLS</sub> line, but also inhibited AP of biotinylated proteins as can be seen both in the protein extracts after incubation with the beads and in the eluates. Coomassie Brilliant Blue-stained membranes (CBB) are shown as loading controls for the IB. Asterisks mark the positions of naturally biotinylated proteins. The positions of TbID-YFP<sub>NLS</sub>, as well as of BSA (beads used in this experiment are blocked with BSA) and SA monomers that are eluted from the beads, are indicated.

**Figure 4 – figure supplement 7: Comparison of different biotin depletion methods and bead concentrations for an effective pulldown of biotinylated proteins**

Five day old seedlings expressing the UBQ10pro::TbID-YFP<sub>NLS</sub> construct were submerged in a 50 μM biotin solution for three hours and used for biotin depletion-affinity purification (AP) experiments. For biotin depletion, protein extracts were either filtered through one (1x) or two consecutive (2x) PD-10 desalting columns or concentrated repeatedly using an Amicon Ultra centrifugal filter. Samples were taken before and after this step to check for protein loss. From the biotin-depleted extracts, a volume corresponding to 1/5 of the amount used for the PL experiments...
shown in Figures 4-6 was used for AP with the indicated amount of streptavidin (SA) beads. Samples of the supernatants after AP and of the proteins eluted from the beads were taken to determine the ratio of biotinylated proteins remaining in the protein extract and bound to the beads. All samples were analyzed by immunoblotting (IB) with SA-HRP. The amount of protein extracts and bead eluates loaded correspond to 0.05% and 0.5% of the total volumes, respectively. The Coomassie Brilliant Blue-stained membranes (CBB) are shown as loading controls for the IBs. Asterisks mark the positions of naturally biotinylated proteins. Biotin depletion with the Amicon Ultra Centrifugal filter, but not with PD-10 columns resulted in considerable protein loss. Increasing the amount of beads increased the AP efficiency.

**Figure 5 and supplements:**

**Figure 5: ‘FAMA interactome’ experiment – PL with TbID reveals potential FAMA interactors involved in transcriptional regulation**

**Workflow (left) and results (right) of the experimental setup and data filtering process.** Biotinylated proteins from seedlings expressing FAMA-TbID or nuclear TbID in FAMA-stage cells (FAMA\textsubscript{nuc}TbID) and from wild-type (WT) after 0, 0.5 and 3 hours of biotin treatment were affinity purified (AP) with streptavidin beads and analyzed by mass spectrometry (MS). Proteins that were identified in all three biological replicates of at least one genotype and time point (2,511 proteins) were used to filter for FAMA-complex candidates by three consecutive filtering steps: First, proteins enriched in FAMA-TbID compared to WT were determined for each time point using unpaired 2-sided t-tests with a permutation-based FDR for multiple sample correction to remove background from non-specific binding to the beads. Only proteins that were identified in all three replicates of FAMA-TbID at this time point (high confidence identifications) were used for the test (cutoff: FDR = 0.05, S0 = 0.5; ❶). Significantly enriched proteins were then filtered for enrichment compared to FAMA\textsubscript{nuc}TbID (same t-test parameters; ❷) to remove stochastically labeled proteins. Finally, the remaining proteins were filtered for enrichment in biotin-treated versus untreated FAMA-TbID samples (t-test: p < 0.05; ❸) to remove proteins that were not labeled in response to biotin treatment. The Venn diagram on the bottom left shows the distribution of the 47 candidates between time points (see Table 1 for candidate list). Scatter plots on the right show log2-fold changes and -log10 p-values from t-test comparisons between FAMA-TbID and WT (top) and FAMA-TbID and FAMA\textsubscript{nuc}TbID (center). Proteins significantly enriched in FAMA-TbID are shown in yellow, orange and red. All filtering steps and statistical analyses were done in Perseus. Subcellular localization and functional distribution of the candidate proteins is shown as pie charts on the bottom right. Data were manually curated from literature. Abbreviations: Nuc, nucleus; Cyt, cytosol, CP, chloroplast, PM, plasma membrane, +, and other.

For hierarchical clustering and PCA of samples used in this experiment, see Figure 5 – figure supplement 2. For a multi scatterplot of all samples, scatterplots and heatmaps of proteins enriched in FAMA-TbID and FAMA\textsubscript{nuc}TbID
compared to WT and for enrichment of proteins over time in all three genotypes, see Figure 5 – figure supplements 3, 4 and 6, respectively. For tables summarizing all identified and enriched proteins, see Supplementary file 2 – Tables 1-4. For an overview over the workflow and proteins identified by affinity purification of FAMA-interacting proteins using a classical AP-MS strategy with GFP-Trap beads, see Figure 5 – figure supplement 1 and Supplementary file 2 – Table 5. For small-scale validation of selected candidates by Y2H and pairwise PL in *N. benthamiana*, see Figure 5 – figure supplement 5.

**Figure 5 – figure supplement 1: FAMA-CFP AP-MS experiments identified ICE1, but no novel transcriptional regulators as putative FAMA partners**

**(A) Simplified workflow of the AP-MS experiments.** Four day old seedlings of a FAMA-CFP line (FAMApro::FAMA-CFP in *fama-1*) and two lines expressing nuclear GFP under stomatal lineage-specific promoters (SPCHpro::GFP<sub>NLS</sub> and MUTEpro::GFP<sub>NLS</sub>) were treated with MG-132 to stabilize proteins and with formaldehyde (FA) to cross-link (X-link) proteins. X-linking was done with 0.25 % FA in experiment 1 and with 0.125 % in experiment 2. Proteins extracted from 15 g of plant material per sample were used for AP with GFP-Trap beads and purified proteins were identified by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). Two biological replicates of each line were used in experiment 1, and four of the FAMA-CFP line and two of each control in experiment 2. **(B) Significantly enriched proteins.** Scatterplots show the log2-fold change and -log10 p-value from unpaired 2-sided t-tests between FAMA-CFP samples and the controls with a permutation-based FDR for multiple sample correction (cutoff: AP-MS 1: FDR = 0.2, S0 = 0.5; AP-MS 2: FDR = 0.01, S0 = 0.5; N = number of proteins used in each test). Significantly enriched proteins are shown in orange and FAMA and its heterodimerization partner ICE1 are indicated. The Venn diagram shows the overlap between the candidate proteins identified in the two experiments (including FAMA). Only FAMA and ICE1 were enriched in the PL experiment shown in Figure 5. Filtering and statistical analysis was done in Perseus. **(C) Subcellular and functional distribution of candidates.** Candidates were classified into nuclear and non-nuclear proteins based on localization predictions (SUBA4 consensus prediction) and experimentally determined localization data and nuclear localized proteins were further divided into such that are involved in transcriptional regulation (2 proteins, including FAMA) or have a different function (6 proteins). For a list of identified proteins and the localization predictions see Supplementary file 2 – Table 5.

**Figure 5 – figure supplement 2: Clustering and PCA of samples for the ‘FAMA interactome’ PL experiment**

**(A) Hierarchical clustering.** Proteins that were identified in all three replicates of at least one time point of one genotype were used for average linkage clustering with Euclidian distance. Rows and columns represent individual proteins and samples. In the sample dendrogram, genotypes, replicate numbers and time points are labeled with differently colored and numbered triangles, circles and square (same designations as in (B)). In the heatmap, proteins are colored according to their abundance (label free quantification value). White areas represent proteins not identified in a sample. **(B) Principal component analysis (PCA).** Missing value were imputed (from normal
distribution) and the data matrix was used for PCA. Principal component 1 (PC1) was plotted against PC2 or 3. The PC1-PC3 plot (top) shows separation of the samples by genotype and time. Areas where samples cluster by genotypes and time point are shaded. A batch effect separating replicate 1 from replicates 2 and 3 is visible in PC2 (bottom) as well as in (A). Clustering and PCA analysis were done in Perseus using built-in functions.

**Figure 5 – figure supplement 3: Multi scatter plot of samples for the ‘FAMA interactome’ PL experiment**

Proteins that were identified in all three replicates of at least one time point of one genotype were used for a multi scatter plot. For each pair-wise comparison, LFQ values of individual proteins in sample A (labeled on top, x-axis) and sample B (labeled on the left, y-axis) are plotted against each other. Plots are scaled equally. Pearson correlation coefficients for each comparison are given in the upper left corner of the plots in blue.

**Figure 5 – figure supplement 4: Significantly enriched proteins in the FAMA-TbID and FAMA\textsubscript{nucTbID} lines**

For each time point, proteins that were significantly enriched in FAMA-TbID or FAMA\textsubscript{nucTbID} compared to wild-type (WT) were determined by 2-sided t-tests with a permutation-based FDR for multiple sample correction (cutoff: FDR = 0.05, S0 = 0.5). Only proteins identified in all three biological replicates of either FAMA-TbID (A) or FAMA\textsubscript{nucTbID} (B) at the indicated time point were used for these tests (N = number of proteins used in each test). Scatterplots on the left show log2-fold changes and -log10 p-values from t-tests (plots in (A) are also shown in Figure 5). Significantly enriched proteins are shown in orange/yellow (A) or dark/light blue (B) and increase with the duration of biotin treatment. Known FAMA interactors (ICE1 and RBR1) and nuclear proteins in guard cells (FAMA and SCAP1) are indicated. The heatmaps on the right of each scatter plot show average linkage clustering of significantly enriched proteins in this plot. Proteins are colored according to their abundance (Z-transformed label free quantification values). White areas represent proteins not identified in a sample.

**Figure 5 – figure supplement 5: Validation of FAMA complex candidates**

Y2H (A-B) and pairwise proximity labeling assays (C) to test direct interaction and relative proximity of four putative FAMA interaction candidates: co-repressor complex components SEU and LUH and BR signaling TFs BZR1 and BIM1. ICE1 and MUTE were used as positive and negative controls, respectively. (A-B) Y2H assays: Yeast strain AH109 was co-transformed with pairwise combinations of BD (Gal4 DNA binding domain) or AD-FAMA and AD-BZ1 (Gal4 activation domain) or AD-bait protein fusions. Successful transformation and bait-prey interactions were tested for by growth on synthetic defined media (-LT: selection for transformants, -LTHA: selection for interaction) (A). To overcome the autoactivation activity of BD-FAMA (see -LTHA plates), 3-AT was added to the plates at two different concentrations to increase selection stringency. For each combination, three dilutions were spotted on the plates and pictures were taken after six days. Three clones per combination were tested with the same result. One replicate is shown. We noticed that both ICE1 and SEU cause a delay in yeast growth. Therefore, the clones used for the growth-based interaction assay shown in (A), were also streaked on a selective plate containing X-α-Gal for
blue/white selection (B). Blue color indicates positive interaction. The picture was taken after two days. (C) Pairwise proximity labeling assays: FAMA-TbID-Venus or TbID-YFP<sub>NLS</sub> were co-transformed with HA- or myc-tagged versions of ICE1, MUTE, SEU and LUH (all under the control of the UBQ10 promoter) into N. benthamiana leaves. After 2 days, expression of the TbID construct was verified by fluorescent microscopy (see images on the right) and samples were frozen for pulldowns of biotinylated proteins with streptavidin (SA) beads. To avoid over-labeling due to high expression of bait and prey proteins and to label only close interactors of FAMA, no exogenous biotin was added to the leaves. Immunoblots on the right show total amounts of FAMA-TbID-Venus or TbID-YFP<sub>NLS</sub> (IB: GFP) and the prey proteins (IB: myc, IB: HA) in the protein extracts before pulldowns (input) and as biotinylated proteins eluted from the SA beads (IP). Coomassie stained membranes (CBB) are shown as a loading control. Red arrowheads mark evidence of biotinylation of the bait protein by FAMA-TbID. (D) Summary of observed interactions: Positive interactions with FAMA in (A-C) are indicated by arrows. Line thickness indicates interaction strength. LUH shows weak interaction in the Y2H but was not biotinylated in the PL assay, which could indicate that it is not a direct interactor of FAMA.

Figure 5 – figure supplement 6: Increase of biotinylation in the PL samples over time

Proteins that were significantly enriched in wild-type (WT), FAMA-TbID or FAMA<sub>nuc</sub>TbID samples after 0.5 and after 3 hours of biotin treatment compared to untreated samples were determined by 2-sided t-tests with a permutation-based FDR for multiple sample correction (cutoff: FDR = 0.05, S0 = 0.5). Only proteins identified in all three biological replicates of wild-type (WT) (A), FAMA-TbID (B) or FAMA<sub>nuc</sub>TbID (C) at any of the two time points were used for the tests (N = number of proteins used in each test). Scatterplots show log2-fold changes and -log10 p-values from t-tests. Significantly up-and down-regulated proteins are shown in red and blue and their number and percentage of all analyzed proteins are given. Only few proteins are enriched in any of the genotypes after 0.5 hours of biotin treatment.

Figure 6 and supplements:

Figure 6: ‘Nuclear proteome’ experiment – PL with nuclear TbID results in identification of FAMA- and global nuclear proteomes with high organellar specificity

Workflow (left) and results (right) of the experimental setup and data filtering process. Biotinylated proteins from seedlings expressing nuclear TbID under the FAMA (FAMA<sub>nuc</sub>TbID) and UBQ10 (UBQ<sub>nuc</sub>TbID) promoters and from wild-type (WT) after 3 hours of biotin treatment were affinity purified (AP) with streptavidin beads and analyzed by mass spectrometry (MS). Proteins that were identified in all three replicates of at least one genotype (3,176) were used to filter for the FAMA cell- and global nuclear proteomes: proteins enriched in FAMA<sub>nuc</sub>TbID or UBQ<sub>nuc</sub>TbID compared to WT were determined by unpaired 2-sided t-tests with a permutation-based FDR for multiple sample correction. Only proteins that were identified in all three replicates of the respective nucTbID line (high confidence
identifications) were used for the test (cutoff: FDR = 0.05, S0 = 0.5; ❶). Scatter plots of the log2-fold changes and log10 p-values from the t-test comparisons are on the top right. Significantly enriched proteins are shown in blue and green. Overlap of the enriched proteins and subcellular distribution are shown as Venn diagrams and pie charts on the center right. Proteins previously found in the nucleus in MS- or fluorescent protein (FP) fusion-based experiments, and proteins predicted to be in the nucleus (SUBA4 consensus prediction) are depicted in different shades of blue (N, nucleus; C, cytosol; CP, plastid; MT, mitochondrion; ER, endoplasmic reticulum; G, golgi; P, peroxisome; V, vacuole; PM, plasma membrane; EX, extracellular; #N/A, not available). Proteins that are highly enriched in or even specific to FAMA nuclei were identified through further filtering by comparing the abundance of the 394 FAMA nuclear proteins between FAMA<sup>nuc</sup>TbID and UBO<sup>nuc</sup>TbID using a 2-sided t-test with a permutation-based FDR for multiple sample correction (FDR = 0.01, S0 = 0.1; ❷). Proteins that were not enriched in the UBO<sup>nuc</sup>TbID samples (higher in FAMA<sup>nuc</sup>TbID or equally abundant) were considered ‘FAMA nuclei enriched/specific’.

A scatterplot of the t-test results with ‘FAMA nuclei enriched/specific’ proteins in purple is shown at the center bottom. For hierarchical clustering and PCA of samples used in this experiment, see Figure 6 – figure supplement 1. For multi scatterplots, scatterplots and heatmaps of proteins enriched in FAMA<sup>nuc</sup>TbID and UBO<sup>nuc</sup>TbID, see Figure 6 – figure supplements 2 and 3. For a visualization of enriched GO terms, see Figure 6 – figure supplement 4. For tables summarizing all identified and enriched proteins, published nuclear proteins and guard cell proteins, the localization data used for the pie charts in Figure 6, enriched GO terms of the nuclear proteomes and examples of identified proteins from different nuclear compartments, see Supplementary file 3 – Tables 1-8.

**Figure 6 – figure supplement 1: Clustering and PCA of samples for the ‘nuclear proteome’ PL experiment**

**(A) Hierarchical clustering.** Proteins that were identified in all three replicates of at least one genotype were used for average linkage clustering with Euclidian distance. Rows and columns represent individual proteins and samples. In the sample dendrogram, genotypes and replicate number are labeled with differently colored and numbered circles. In the heatmap, proteins are colored according to their abundance (label free quantification value). White areas represent proteins not identified in a sample. **(B) Principal component analysis (PCA).** Missing value were imputed (from normal distribution) and the data matrix was used for PCA. Principal component 1 (PC1) was plotted against PC2 or 3. The PC1-PC2 plot (top) shows clear separation of the samples by genotype. Areas where samples cluster are shaded. A batch effect separating replicate 1 from replicates 2 and 3 is visible in PC3 (bottom). Clustering and PCA analysis were done in Perseus using built-in functions.

**Figure 6 – figure supplement 2: Multi scatter plot of samples for the ‘nuclear proteome’ PL experiment**

Proteins that were identified in all three replicates of at least one genotype were used for a multi scatter plot. For each pair-wise comparison, LFQ values of individual proteins in sample A (labeled on top, x-axis) and sample B (labeled on the left, y-axis) are plotted against each other. Plots are scaled equally. Pearson correlation coefficients for each comparison are given in the upper left corner of the plots in blue.
Figure 6 – figure supplement 3: Significantly enriched proteins in FAMA-expressing and all nuclei

Proteins that were significantly enriched in \textit{\textsuperscript{FAMA}}\textsuperscript{nuc}TbID or \textit{\textsuperscript{UBQ}}\textsuperscript{nuc}TbID compared to wild-type (WT) were determined by 2-sided t-tests with a permutation-based FDR for multiple sample correction (cutoff: FDR = 0.05, S0 = 0.5). Only proteins identified in all three biological replicates of \textit{\textsuperscript{FAMA}}\textsuperscript{nuc}TbID (A) or \textit{\textsuperscript{UBQ}}\textsuperscript{nuc}TbID (B) were used for these tests (N = number of proteins used in each test). Proteins that were enriched in FAMA nuclei compared to all nuclei (C) were determined by a 2-sided t-tests with an FDR of 0.01 (S0 = 0.1) using only proteins significantly enriched in \textit{\textsuperscript{FAMA}}\textsuperscript{nuc}TbID compared to WT. Scatterplots on the left are from Figure 6 and show log2-fold changes and -log10 p-values from t-tests. Significantly enriched proteins are shown in blue (A) and green (B) and FAMA nuclei-enriched proteins in purple (C). Known FAMA-nuclei enriched proteins (FAMA, SCAP1, HDG2) are indicated. The heatmaps on the right of each scatter plot show average linkage clustering of significantly enriched proteins in this plot (A-B) or FAMA nuclei-enriched proteins (C). Proteins are colored according to their abundance (Z-transformed label free quantification values). White areas represent proteins not identified in a sample.

Figure 6 – figure supplement 4: GO terms enriched in global and FAMA nuclear proteomes

GO terms enriched for global and FAMA nuclear proteins (Figure 6) were determined with AgriGO v2 and visualized with REViGO. Shown is the TreeMap plot of enriched ‘biological processes’ in global (UBQ10 promoter, top) and FAMA (bottom) nuclei. Related GO-terms are clustered into same colored boxes and given summary titles.
A

3 versions of BirA

BirA⁺

TbID

mTb

Transient expression in N. benthamiana

Stable expression in A. thaliana

infiltration

floral dip

2 days

leaf discs punched out

• submerged in biotin sol. / H₂O
• briefly vac. infiltrated
• 1 h @ RT

leaf discs washed, frozen and used for immunoblotting

stable lines (T2/T3)

5 day old seedlings

• submerged in biotin solution
• 5' vac. infiltrated
• 1 h @ RT

seedlings washed, frozen and used for immunoblotting

B

<table>
<thead>
<tr>
<th></th>
<th>not transf.</th>
<th>BirA⁺-YFP</th>
<th>mTb-YFP</th>
<th>TbID-YFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>biotin</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

C

<table>
<thead>
<tr>
<th></th>
<th>WT</th>
<th>BirA⁺-YFP</th>
<th>mTb-YFP</th>
<th>TbID-YFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>biotin</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
A FAMA-TbID
FAMApro::FAMA-TbID-
Venus; fama-1

FAMAnucTbID
FAMApro::TbID-
YFPNLS

UBQnucTbID
UBQ10pro::TbID-
YFPNLS

WT
Col-0

B

<table>
<thead>
<tr>
<th></th>
<th>WT</th>
<th>FAMA-TbID</th>
<th>WT</th>
<th>FAMA-TbID</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 h</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>0 h</td>
<td>15</td>
<td>15</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>0 h</td>
<td>30</td>
<td>30</td>
<td>1 h</td>
<td>1 h</td>
</tr>
<tr>
<td>0 h</td>
<td>1 h</td>
<td>1 h</td>
<td>3 h</td>
<td>3 h</td>
</tr>
<tr>
<td>0 h</td>
<td>3 h</td>
<td>3 h</td>
<td>5 h</td>
<td>5 h</td>
</tr>
</tbody>
</table>

Biotin treatment: 50 µM, 0.5/3 h

C

<table>
<thead>
<tr>
<th></th>
<th>FAMA-TbID</th>
<th>WT</th>
<th>FAMAnucTbID</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 h</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.5 h</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>3 h</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Optimal labeling time?

D

<table>
<thead>
<tr>
<th></th>
<th>FAMAnucTbID</th>
<th>UBQnucTbID</th>
<th>WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 h</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>0.5 h</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3 h</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Nuclear proteomes?

FAMA nuclei enriched proteins?

E

STOP reaction & WASH (ice cold water)

protein extraction

free biotin depletion (PD-10 gel filtration)

SA bead pulldown (optimized bead amount)

tryptic digest (on-bead/in-solution)

LC-MS/MS

LFQ data analysis (MaxQuant/Perseus)
PL with FAMA-TbID and FAMAnucTbID

AP-MS of biotinylated proteins

2,511 proteins

FILTER: enriched in FAMA-TbID vs. WT

73/85/239 proteins

FILTER: enriched in FAMA-TbID vs. FAMAnucTbID

6/15/57 proteins

FILTER: enrichment over time

47 FAMA-complex candidates

0 h biotin

0.5 h biotin

3 h biotin

FAMA-TbID vs. WT (enriched vs. wt in FAMA-TbID / FAMA-TbID & FAMAnucTbID)

FAMA-TbID vs. FAMAnucTbID (enriched in FAMA-TbID vs. FAMAnucTbID)

Curated subcellular localization

Functional classification

- transcription factor
- transcriptional co-regulator
- photosynthesis
- metabolic enzyme
- light sign
- Ca^{2+} sign
- clock
- replication
- mRNA reg.
- kinase
- chaperone
PL of FAMA-stage and global nuclear proteins

AP-MS of biotinylated proteins

3,176 proteins

FILTER: enriched in FAMA nuclei

FAMA nuclei proteins

Nuclear proteins: 394 in FAMA nuclei
2,215 in global nuclei

‘FAMA nuclei proteins’

FILTER: equal or higher in FAMA nuclei than UBQ nuclei

FAMA nuclei enriched specific proteins

(include known marker genes for young GCs and regulators of stomatal development: FAMA, HDG2, SCAP1)