

**Analysis of GET pathway receptors**  
**in *Arabidopsis thaliana***

**Dissertation**

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Lisa Yasmin Asseck

aus Göppingen

Tübingen

2020

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 21.10.2020

Stellvertretender Dekan: Prof. Dr. József Fortágh

1. Berichterstatter: Prof. Dr. Christopher Grefen

2. Berichterstatter: Prof. Dr. Gerd Jürgens

# CONTENTS

|   |        |
|---|--------|
| Abbreviations.....  | - 1 -  |
| Summary.....  | - 3 -  |
| Zusammenfassung.....  | - 4 -  |
| List of publications in the thesis .....  | - 6 -  |
| Personal contribution .....   | - 7 -  |
| Introduction.....   | - 9 -  |
| Types of membrane proteins .....  | - 9 -  |
| TA proteins.....  | - 10 - |
| SNARE proteins.....   | - 11 - |
| Co-translational targeting via the SRP pathway .....  | - 12 - |
| Post-translational targeting via the GET pathway in yeast .....   | - 13 - |
| The GET1-GET2 transmembrane complex in yeast.....   | - 15 - |
| Post-translational targeting via the TRC pathway in mammals.....  | - 16 - |
| The WRB-CAML transmembrane complex in mammals.....  | - 17 - |
| Objectives and expected output of the thesis.....   | - 19 - |
| Results.....  | - 20 - |
| Chapter one: Loss of GET pathway orthologs in <i>Arabidopsis thaliana</i> causes root hair growth defects and affects SNARE abundance ..... | - 20 - |
| Chapter two: ER membrane receptors of the GET pathway are conserved throughout eukaryotes.....  | - 22 - |
| Chapter three: Detecting interactions of membrane proteins: the split-ubiquitin system.....   | - 24 - |
| Discussion.....   | - 25 - |
| Conservation and divergence of the GET pathway in eukaryotes .....  | - 25 - |
| The GET receptor complex in <i>Arabidopsis thaliana</i> .....   | - 28 - |
| GET pathway dependence of TA proteins .....   | - 30 - |
| Alternative insertion pathways for TA proteins.....   | - 31 - |

|                                    |         |
|------------------------------------|---------|
| Conclusions and perspectives ..... | - 34 -  |
| References.....                    | - 35 -  |
| Appendices.....                    | - 44 -  |
| I. Xing et al., 2017.....          | - 44 -  |
| II. Asseck et al., submitted ..... | - 68 -  |
| III. Asseck and Grefen, 2018 ..... | - 97 -  |
| Curriculum vitae .....             | - 110 - |
| Acknowledgements.....              | - 112 - |

## ABBREVIATIONS

|        |   |
|--------|---|
| ALB3   | Albino 3  |
| BAG6   | BCL-2-associated athanogene 6   |
| CAML   | Calcium-modulating cyclophilin ligand   |
| EMC    | ER membrane complex   |
| ER     | Endoplasmic reticulum   |
| GET    | Guided entry of tail-anchored proteins  |
| G1IP   | <i>At</i> GET1-interacting protein  |
| HSC70  | Heat shock cognate 70   |
| HSP40  | Heat shock protein 40   |
| IP-MS  | Immunoprecipitation-mass spectrometry   |
| mbSUS  | Mating-based split-ubiquitin system   |
| MSP1   | Mitochondrial sorting of proteins 1   |
| OXA1   | Oxidase assembly protein 1  |
| PPIs   | Protein-protein interactions  |
| rBiFC  | Ratiometric bimolecular fluorescence complementation                          |
| RNC    | Ribosome-nascent chain  |
| RNF126 | Ring finger protein 126   |
| SEC61  | Secretory 61  |
| SGT2   | Small glutamine-rich tetratricopeptide repeat-containing protein 2            |
| SGTA   | Small glutamine-rich tetratricopeptide repeat-containing protein alpha        |
| SNARE  | Soluble <i>N</i> -ethylmaleimide-sensitive factor attachment protein receptor |
| SND    | SRP-independent   |
| SR     | SRP receptor  |
| SRP    | Signal recognition particle   |
| SYP72  | Syntaxin of plants 72   |
| SYP123 | Syntaxin of plants 123  |
| TA     | Tail-anchored   |
| TMD    | Transmembrane domain  |
| TRC    | Transmembrane domain recognition complex                                      |

|         |   |
|---------|---|
| TRC35   | Transmembrane domain recognition complex 35 kDa subunit |
| TRC40   | Transmembrane domain recognition complex 40 kDa subunit |
| UBL4A   | Ubiquitin-like protein 4a                               |
| UBQ12   | Ubiquitin 12  |
| VAMP721 | Vesicle-associated membrane protein 721                 |
| WRB     | Tryptophan-rich basic protein                           |

## SUMMARY

Proper targeting and insertion of proteins into membranes is essential for the structure and function of all cells. The Guided Entry of Tail-anchored proteins (GET) pathway has been described as the major targeting route for tail-anchored (TA) membrane proteins to the endoplasmic reticulum (ER) in yeast and mammals (Stefanovic and Hegde, 2007; Schuldiner et al., 2008). Here, the cytosolic targeting factor GET3 (in yeast; TRC40 in mammals) chaperones newly synthesized TA proteins from the ribosome to the ER where the GET1-GET2 (in yeast; WRB-CAML in mammals) receptor complex facilitates insertion (Yamamoto and Sakisaka, 2012; Wang et al., 2014). Despite the importance of this pathway for regulating membrane protein insertion at the ER, hardly anything is known about its conservation and function in higher plants.

In this work, we identified and functionally characterized several GET pathway components in the model plant *Arabidopsis thaliana*, including orthologues of GET1, GET3 and GET4. Detailed phenotypic analyses of corresponding *Arabidopsis* T-DNA insertion lines uncovered a role of the GET pathway in root hair elongation. Homozygous *Atget* mutants have shorter root hairs compared to wild type, but otherwise develop normally. This phenotype coincides with reduced protein levels of the TA SNARE SYP123 in root hairs, indicating a conserved function of the GET pathway in regulating TA protein insertion. However, the mild phenotype in the *Atget* mutants suggests the existence of an alternative targeting route to the ER. We further investigated the physiological function of the GET pathway by characterizing *Arabidopsis* plants overexpressing *AtGET3a* constitutively in the absence of *AtGET1*. Homozygous lines display severe growth defects, emphasising the functional conservation of the GET pathway in plants.

While orthologues of GET1, GET3 and GET4 could be identified through *in silico* sequence comparison, there is no obvious GET2 orthologue in the *Arabidopsis* genome. We therefore performed immunoprecipitation-mass spectrometry (IP-MS) of *AtGET1*-GFP expressing lines and found an unknown protein containing structural characteristics of GET2 and CAML, that we named G1IP (*AtGET1*-interacting protein). Subcellular localisation and functional analyses strongly suggest G1IP as the elusive GET co-receptor in *Arabidopsis*.

Additionally, this work details the application of the yeast mating-based split-ubiquitin system (mbSUS) and its modification cytoSUS to analyse protein-protein interactions *in vivo*, using *AtGET1* and *AtGET3a*, respectively, as examples.

## ZUSAMMENFASSUNG

Die korrekte Zielführung und Insertion von Membranproteinen ist für die Struktur und Funktion aller Zellen von essenzieller Bedeutung. Der Guided Entry of Tail-anchored proteins (GET) Pfad gilt in Hefe und Säugetieren als Hauptweg für den Transport von tail-anchored (TA) Membranproteinen in das endoplasmatische Retikulum (ER) (Stefanovic and Hegde, 2007; Schuldiner et al., 2008). Hier überführt das zytosolische GET3-Protein (in Hefe; TRC40 in Säugetieren) neu synthetisierte TA-Proteine vom Ribosom zum ER, wo der GET1-GET2 (in Hefe; WRB-CAML in Säugetieren) Rezeptorkomplex die Insertion vermittelt (Yamamoto and Sakisaka, 2012; Wang et al., 2014). Trotz der Wichtigkeit dieses Weges für die Regulation der Membranproteininsertion ins ER ist kaum etwas über seine Konservierung und Funktion in höheren Pflanzen bekannt.

In dieser Arbeit haben wir mehrere Komponenten des GET-Transportwegs in der Modellpflanze *Arabidopsis thaliana* identifiziert und funktionell charakterisiert, darunter Orthologe von GET1, GET3 und GET4. Detaillierte phänotypische Analysen der entsprechenden *Arabidopsis* T-DNA Insertionslinien deckten eine Rolle des GET-Pfads bei der Wurzelhaarverlängerung auf. Homozygote *Atget* Mutanten haben im Vergleich zum Wildtyp kürzere Wurzelhaare, entwickeln sich jedoch ansonsten normal. Dieser Phänotyp geht mit einer reduzierten Proteinkonzentration des TA SNAREs SYP123 in Wurzelhaaren einher, was auf eine konservierte Funktion des GET-Pfads bei der Regulierung der Insertion von TA-Proteinen schließen lässt. Der milde Phänotyp der *Atget* Mutanten deutet jedoch auf das Vorhandensein eines alternativen Transportwegs zum ER hin. Um weitere Erkenntnisse über die physiologische Funktion des GET-Pfads zu gewinnen, haben wir *Arabidopsis* Pflanzen charakterisiert, die *AtGET3a* in Abwesenheit von *AtGET1* konstitutiv überexprimieren. Homozygote Linien zeigen schwere Wachstumsdefekte, was die funktionelle Konservierung des GET-Pfads in Pflanzen unterstreicht.

Während Orthologe von GET1, GET3 und GET4 durch *in silico* Sequenzvergleiche identifiziert werden konnten, gibt es im *Arabidopsis* Genom kein offensichtliches GET2 Ortholog. Daher führten wir Immunpräzipitations-Massenspektrometrie (IP-MS) mit *AtGET1*-GFP exprimierenden Linien durch und fanden ein unbekanntes Protein mit strukturellen Ähnlichkeiten zu GET2 und CAML, das wir G1IP (*AtGET1*-interagierendes Protein) nannten. Subzelluläre Lokalisations- und Funktionsanalysen deuten stark auf G1IP als den fehlenden GET-Korezeptor in *Arabidopsis* hin.

Zusätzlich wird in dieser Arbeit die Anwendung des mating-based Split-Ubiquitin-Systems (mbSUS) in Hefe und des modifizierten cytoSUS zur *in vivo* Analyse von Protein-Protein-Interaktionen am Beispiel von *AtGET1* und *AtGET3a* beschrieben.

## LIST OF PUBLICATIONS IN THE THESIS

### Accepted papers:

Xing, S., Mehlhorn, D.G., Wallmeroth, N., **Asseck, L.Y.**, Kar, R., Voss, A., Denninger, P., Schmidt, V.A., Schwarzländer, M., Stierhof, Y.D., Grossmann, G., Grefen, C. (2017). "Loss of GET pathway orthologs in Arabidopsis thaliana causes root hair growth defects and affects SNARE abundance." *Proceedings of the National Academy of Sciences* 114(8): E1544-E1553.

**Asseck, L. Y.** and Grefen, C. (2018). "Detecting interactions of membrane proteins: the split-ubiquitin system." *Two-Hybrid Systems*, Springer: 49-60.

### Submitted manuscripts:

**Asseck, L.Y.**, Mehlhorn, D.G., Rivera Monroy, J., Ricardi, M.M., Breuninger, H., Wallmeroth, N., Berendzen, K.W., Nouwrosian, M., Xing, S., Schwappach, B., Bayer, M., Grefen, C. "ER membrane receptors of the GET pathway are conserved throughout eukaryotes." Submitted.

## PERSONAL CONTRIBUTION

### Chapter one:

Xing, S., Mehlhorn, D.G., Wallmeroth, N., **Asseck, L.Y.**, Kar, R., Voss, A., Denninger, P., Schmidt, V.A., Schwarzländer, M., Stierhof, Y.D., Grossmann, G., Grefen, C. (2017). "Loss of GET pathway orthologs in *Arabidopsis thaliana* causes root hair growth defects and affects SNARE abundance." *Proceedings of the National Academy of Sciences* 114(8): E1544-E1553.

- Figure 4D qPCR analysis of SYP123 transcript levels in Col-0, *Atget1* and *Atget3a*
- Figure S5A cloning of pZU-LC-*ScGET1p* and complementation assay of the yeast *get1* mutant, figure preparation
- Figure S5B cloning of vector pYOX1-Dest

### Chapter two:

**Asseck, L.Y.**, Mehlhorn, D.G., Rivera Monroy, J., Ricardi, M.M., Breuninger, H., Wallmeroth, N., Berendzen, K.W., Nouwrosian, M., Xing, S., Schwappach, B., Bayer, M., Grefen, C. "ER membrane receptors of the GET pathway are conserved throughout eukaryotes." Submitted.

- Figure 1B qPCR analysis of *AtGET1*, G1IP and G1IP-*like* transcript levels in Col-0 organs
- Figure 1C-J cloning and subcellular localisation study of GFP-G1IP and GFP-G1IP-*like*
- Figure 1K-M cloning and rBiFC analysis of G1IP and G1IP-*like* with *Arabidopsis* GET pathway components
- Figure 1N-P cloning and co-immunoprecipitation of *AtGET3a*-mVenus with G1IP-3xHA and G1IP-*like*-3xHA
- Figure 2A cloning and generation of the CRISPR line *glip-4*
- Figure 2B root hair imaging and measurements
- Figure 2C,D cloning and complementation assays of the yeast *get1/get2* double-deletion mutant
- Figure 3G cloning of pGEX-G1IP<sup>4E</sup>cyt

- Figure S1B cloning and rBiFC analysis of G1IP with *AtGET1* to test protein topology
- figure preparation, writing of the manuscript with Prof. Dr. Christopher Grefen

### **Chapter three:**

**Asseck, L. Y.** and Grefen, C. (2018). "Detecting interactions of membrane proteins: the split-ubiquitin system." *Two-Hybrid Systems*, Springer: 49-60.

- Figure 2 mating-based SUS and cytoSUS analysis of *AtGET1* and *AtGET3a*, respectively, Western Blot analysis of Cub- and Nub-fusions
- figure preparation, writing of the chapter with Prof. Dr. Christopher Grefen

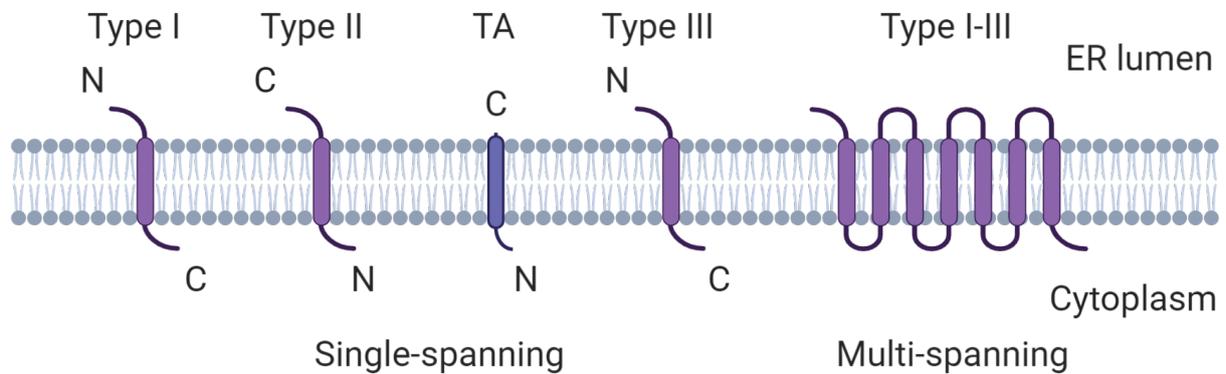
## INTRODUCTION

### Types of membrane proteins

Membrane proteins play a central role in many biological processes by functioning for example as receptors, ion channels, enzymes or transport proteins and can be classified into two groups depending on how they associate with the membrane: Peripheral membrane proteins are only transiently attached, either by direct interaction with membrane lipids or indirectly by interaction with other membrane binding proteins. In contrast, integral membrane proteins are permanently anchored in the lipid bilayer and can be further subdivided according to the number of transmembrane segments (single- and multi-spanning membrane proteins), their transmembrane topology (type I-III) and/or their mechanism of membrane insertion (co- or post-translational) (Figure 1, Lodish et al., 2000; Cournia et al., 2015). Type I proteins are anchored within the lipid bilayer with the N-terminal end in the extracellular space or the lumen of an organelle and the C-terminus in the cytoplasm. Type II membrane proteins are integrated in the opposite orientation (Figure 1, Goder and Spiess, 2001). Targeting and translocation of type I proteins to the endoplasmic reticulum (ER) is directed by a cleavable N-terminal signal sequence and occurs co-translationally via the signal recognition particle (SRP) pathway. Proteins of type II contain a non-cleavable transmembrane signal-anchor sequence, that functions as both, targeting signal and membrane anchor, and are inserted by the same machinery (Goder and Spiess, 2001; Higy et al., 2004; Shao and Hegde, 2011; Park and Rapoport, 2012). Tail-anchored (TA) proteins are a subclass of single-spanning type II membrane proteins that share the same topology but differ in their insertion pathway due to an extreme C-terminal transmembrane domain (TMD) (Figure 1). Unlike the classical type II protein family, TA proteins cannot access the co-translational route but instead are delivered to the ER via the post-translational Guided Entry of Tail-anchored proteins (GET)/Transmembrane domain Recognition Complex (TRC) pathway (Borgese et al., 2003; Borgese and Fasana, 2011; F Colombo and Fasana, 2011; Park and Rapoport, 2012).

Type III membrane proteins contain a reverse signal anchor-sequence to direct SRP-dependent translocation of the N-terminus across the membrane, thus adopting the same orientation within the membrane as type I proteins (Figure 1, Goder and Spiess, 2001; Higy et al., 2004). Similarly, multi-spanning membrane proteins can be classified as type I-III based on the nature of their first hydrophobic element (cleavable signal sequence or TMD), and become co-

translationally inserted into the lipid bilayer via the SRP/translocon machinery (Spiess, 1995; Goder and Spiess, 2001).



**Figure 1: Classification and topology of integral membrane proteins.** Integral membrane proteins can be distinguished according to the number of TMDs (single- and multi-spanning transmembrane proteins), their topology (type I/III: luminal N-terminus, cytosolic C-terminus; type II/TA (tail-anchored): cytosolic N-terminus, luminal C-terminus) and/or their membrane targeting (co- or post-translational; not shown here) (created with Biorender.com).

## TA proteins

Tail-anchored (TA) proteins are a class of integral membrane proteins involved in various cellular processes such as protein translocation (SEC61), vesicle trafficking (SNAREs) and apoptosis (BCL-2 family) (Wattenberg and Lithgow, 2001; Borgese et al., 2003). They are characterized by an N-terminal cytosolic region that lacks a signal sequence, and a single hydrophobic TMD close to the C-terminus (Figure 1). The N-terminus is usually the functional domain, while the TMD is important for targeting and anchoring of the TA protein to the membranes of different organelles such as the ER, mitochondria and chloroplasts (Kutay et al., 1993; Borgese et al., 2007). Since the C-terminal targeting motif only emerges from the ribosome after translation has terminated, co-translational insertion cannot function properly. Hence, targeting of TA proteins to their destined membrane occurs post-translationally and depends on multiple features of the TMD such as length, hydrophobicity, and charge of the flanking sequences (Kanaji et al., 2000; Hwang et al., 2004; Wattenberg et al., 2007). For example, mitochondrial TA proteins usually have short TMDs of moderate hydrophobicity with positive flanking charges whereas the TMDs of ER-directed TA proteins tend to be longer and more hydrophobic, and their C-terminal sequence downstream of the TMD is often less positively charged (Borgese et al., 2007; Moog, 2019).

Post-translational targeting and insertion of TA proteins into the ER membrane was identified to be mediated by the GET (Figure 3) and TRC pathway, respectively, whereas targeting to either mitochondria or chloroplasts is likely GET/TRC independent (Borgese et al., 2001; Stefanovic and Hegde, 2007; Schuldiner et al., 2008; Dhanoa et al., 2010). However, the molecular mechanisms underlying the import of mitochondrial and chloroplast TA proteins are not yet fully understood. TA proteins destined to organelles of the secretory pathway are initially inserted into the ER via the GET/TRC machinery and subsequently delivered to their target membrane by vesicular transport (Jantti et al., 1994; Kutay et al., 1995; Linstedt et al., 1995).

### **SNARE proteins**

A bioinformatics approach identified 454 gene loci encoding TA proteins in *Arabidopsis thaliana*, including 52 out of 64 known soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins. Hence, SNAREs constitute the largest group of TA proteins in *Arabidopsis* (Kriechbaumer et al., 2009).

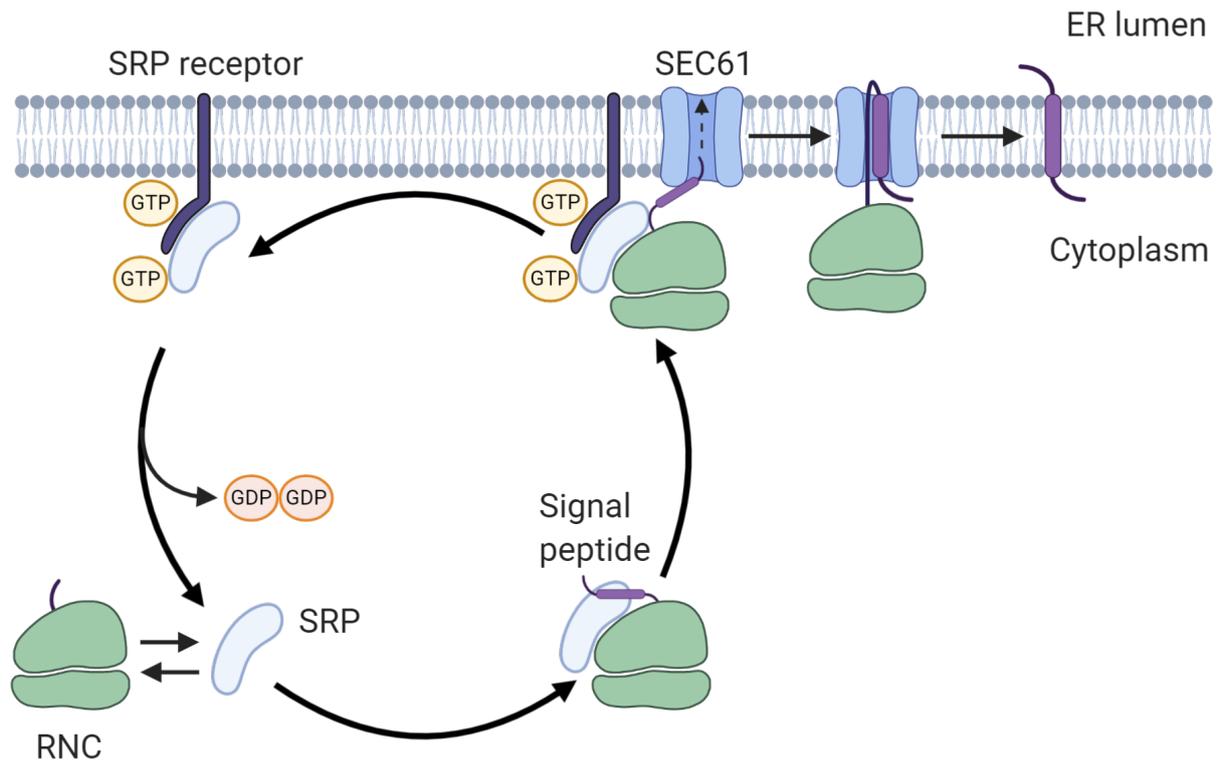
The SNARE family plays a key role in membrane fusion along the secretory and endocytic pathways and is highly conserved among eukaryotes. Gene duplication events increased the number of SNARE genes in plants compared to yeast and humans by three- and two-fold, respectively, likely due to the necessity of some SNAREs for plant-specific processes such as plant cytokinesis, gravitropism, plant-microbe interactions and transport of phytohormones (Lipka et al., 2007). Thus, SNARE proteins contribute essentially to plant development and physiology, and loss of SNARE function can cause severe growth defects or lethality in *Arabidopsis*.

A characteristic feature of SNARE proteins is the SNARE motif, an evolutionary conserved alpha-helical coiled-coil domain formed by heptad repeats that is exposed to the cytosol (Weimbs et al., 1997). Most SNAREs are localised to specific compartments of the endomembrane system or the plasma membrane. However, some SNAREs exhibit dual or multiple localisation patterns likely due to shuttling between organelles (Uemura et al., 2004). SNARE proteins were originally classified as v- and t-SNAREs based on their localisation on the vesicle or target membrane (Söllner et al., 1993). This categorisation, however, is confusing in case of homotypic membrane fusion events. Therefore, SNARE proteins were reclassified as Q- and R-SNAREs according to the presence of either a conserved glutamine (Q) or arginine

(R) residue in the centre of the SNARE motif. One R-SNARE and three Q-SNAREs (Qa, Qb, and Qc) on opposing membranes interact with each other via their SNARE motifs to assemble into a highly stable complex. The formation of this four-helix bundle forces the two membranes into close proximity and initiates membrane fusion (Fasshauer et al., 1998).

### **Co-translational targeting via the SRP pathway**

The signal recognition particle (SRP) pathway mediates the co-translational targeting and translocation of secretory and membrane proteins across or into the ER membrane and is evolutionary conserved (Akopian et al., 2013; Nyathi et al., 2013). SRP is a cytosolic ribonucleoprotein complex that binds the hydrophobic N-terminal signal sequence or first TMD of nascent protein chains emerging from the ribosome and stops translation temporarily (Walter and Blobel, 1981; Halic et al., 2004). The SRP/ribosome-nascent chain (RNC) complex is then targeted to the ER membrane where SRP interacts with its receptor (SR) via their GTPase domains to form a GTP-dependent heterodimer (Gilmore et al., 1982; Gilmore et al., 1982). Conformational changes upon formation of the SRP-SR complex drive the unloading of RNC from SRP to the SEC61 (secretory 61) translocon channel where translation proceeds directly into the ER membrane or lumen (Connolly and Gilmore, 1989; Shan et al., 2007). Further, reciprocal activation of GTP hydrolysis between SRP and SR triggers disassembly of the complex and recycling of the components for additional rounds of protein targeting (Figure 2, Connolly et al., 1991). Loss of SRP can lead to the mistargeting of ER proteins to mitochondria and cause mitochondrial fragmentation (Costa et al., 2018).



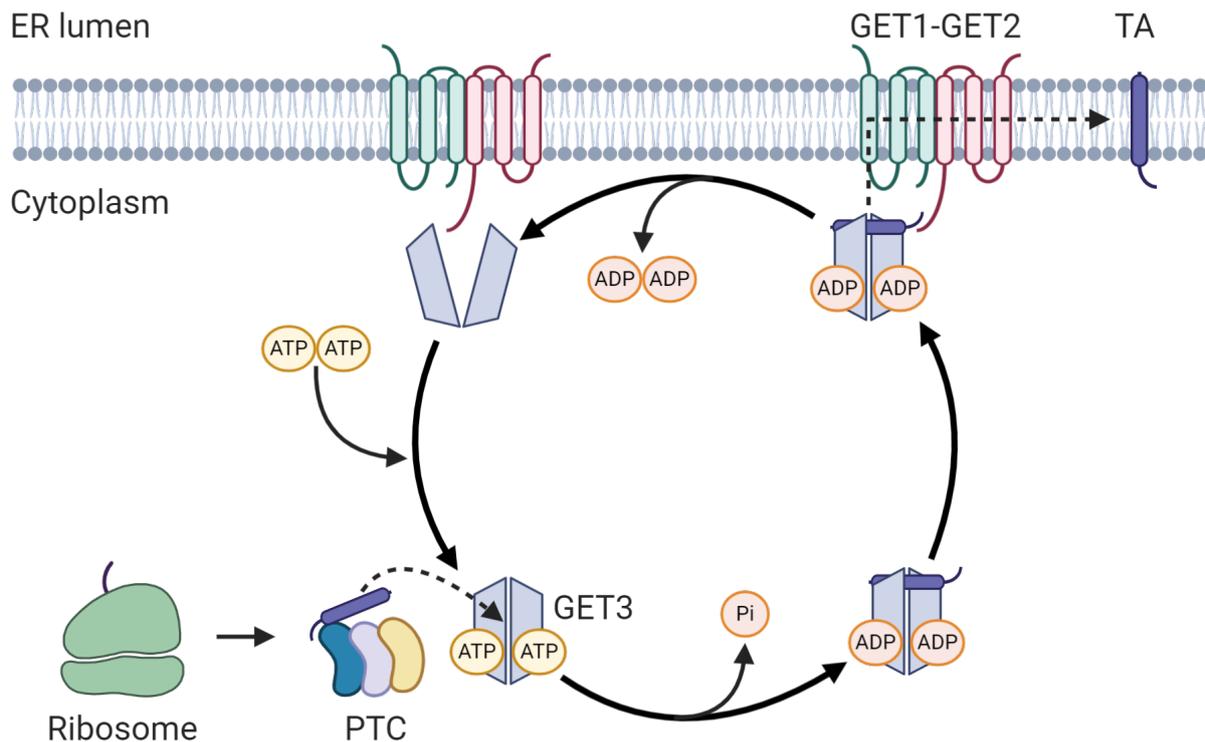
**Figure 2: The SRP pathway.** The SRP recognizes and binds to the signal peptide of nascent proteins on the ribosome and pauses the translation. The SRP-RNC complex is then recruited to the ER membrane via a GTP-dependent interaction with the SRP receptor. The RNC is transferred to the SEC61 translocon and translation resumes through the membrane pore. After GTP hydrolysis, SRP is released from its receptor and returns to the cytosol (created with Biorender.com).

### Post-translational targeting via the GET pathway in yeast

In the yeast Guided Entry of Tail-anchored proteins (GET) pathway, newly synthesized TA proteins are initially captured on the ribosome by a cytosolic pre-targeting complex comprising SGT2 (small glutamine-rich tetratricopeptide repeat-containing protein 2), GET5 and GET4 (Figure 3, Chang et al., 2010; Wang et al., 2010). SGT2 is a co-chaperone that binds to the hydrophobic TMD of nascent TA proteins and interacts with GET5 (Chang et al., 2010; Wang et al., 2010; Kohl et al., 2011). GET4 forms a stable heterotetrameric complex with GET5 and interacts directly with an ATP-bound, closed dimer of GET3, thus enabling loading of the TA protein from SGT2 onto GET3 (Jonikas et al., 2009; Chartron et al., 2010; Gristick et al., 2014). The cytosolic GET3 ATPase is key component of the GET pathway and undergoes conformational changes depending on its nucleotide-binding status (Wereszczynski and McCammon, 2012). Nucleotide-free GET3 is in an open conformation while ATP-binding

drives closure of the dimer, thereby creating a hydrophobic groove, which binds and shields the TMD of the TA protein (Bozkurt et al., 2009; Mateja et al., 2009; Wereszczynski and McCammon, 2012). The transfer of the TA protein to GET3 requires hydrolysis of ATP giving a stable, fully closed GET3-TA protein targeting complex loaded with ADP (Figure 3, Rome et al., 2013). Upon substrate binding, GET3 dissociates from the pre-targeting complex and chaperones the TA protein from the cytosol to the GET1-GET2 transmembrane complex at the ER (Figure 3, Schuldiner et al., 2008). The targeting complex is first captured by the long cytosolic N-terminal domain of GET2, thereby bringing GET3 into proximity with GET1 (Wang et al., 2011). Interaction of GET3 with the GET1 coiled-coil domain induces transition of GET3 from a closed to an open state with release of ADP (Figure 3). Moreover, separation of the GET3 monomers in the open dimer conformation disrupts the hydrophobic groove, leading to release of the bound TA protein for ER membrane insertion (Mariappan et al., 2011; Kubota et al., 2012). Finally, GET3 dissociates from the membrane by ATP binding and recycles back to the cytosol for further rounds of targeting (Figure 3, Mariappan et al., 2011; Wang et al., 2011).

In yeast, the GET pathway is not essential for viability, but deletion of GET components results in a wide range of phenotypic defects, such as increased sensitivity to heat, antifungal drugs or oxidative stress, and secretion of ER-resident proteins (Schuldiner et al., 2008; Yeh et al., 2014). Furthermore, impairment of the GET pathway can lead to aggregation of GET-dependent TA proteins in the cytosol or mistargeting and -insertion into the outer mitochondrial membrane (Schuldiner et al., 2008). Mislocalised TA proteins are recognized by the conserved AAA ATPase MSP1 (mitochondrial sorting of proteins 1) through exposed hydrophobic surfaces and are degraded (Okreglak and Walter, 2014; Opaliński et al., 2014; Wohlever et al., 2017; Li et al., 2019).



**Figure 3: The GET pathway in yeast.** The soluble pre-targeting complex (SGT2, GET5 and GET4) binds to and transfers the newly synthesized TA protein to the dimeric ATPase GET3 in an ATP hydrolysis-dependent manner. GET3 then delivers the TA protein to the ER-resident GET1-GET2 receptor complex for membrane insertion. Following release of ADP and binding of ATP, GET3 dissociates from the membrane and returns to the cytosol (created with Biorender.com).

### The GET1-GET2 transmembrane complex in yeast

A single heterodimer of GET1 and GET2 forms the receptor complex for GET3-mediated TA protein insertion into the ER (Schuldiner et al., 2008; Mariappan et al., 2011; Wang et al., 2011; Zalisko et al., 2017). GET1 and GET2 are integral membrane proteins containing three predicted TMDs for anchoring of the proteins in the bilayer and complex formation through direct interaction (Mariappan et al., 2011). In contrast, the cytoplasmic domains of GET1 and GET2 do not interact with each other but bind to the GET3 dimer, thus recruiting GET3 from the cytosol to the ER (Schuldiner et al., 2008; Wang et al., 2011). In the absence of either GET1 or GET2, ER recruitment fails, leading to cytosolic aggregation of the GET3-TA protein complexes (Schuldiner et al., 2005; Auld et al., 2006; Schuldiner et al., 2008). Furthermore, GET1 deficiency leads to a reduced protein level of GET2 and vice versa, demonstrating reciprocal regulation between the two subunits of the GET receptor complex (Schuldiner et al., 2008; Stefer et al., 2011).

The long cytosolic N-terminus of GET2 (residues 1-151) is mostly unstructured but comprises two alpha-helices that are connected by a flexible glycine linker (Stefer et al., 2011). The first alpha-helix contains a highly conserved motif (14-RERR) that is critical for GET3 binding (Stefer et al., 2011; Wang et al., 2011). The residues that participate in the molecular interplay between GET2 and GET3 are not located in the GET3 dimer interface so that GET2 binds only one subunit of each homodimer (Stefer et al., 2011). The long N-terminal tail structure implies that GET2 is responsible for initial capture of the closed GET3-TA protein targeting complex in the cytosol, thereby increasing its local concentration at the ER (Wang et al., 2011).

GET1 contains a conserved cytoplasmic coiled-coil domain (residues 19-103) between the first and the second TMD that was found to be the GET3 binding site (Stefer et al., 2011; Wang et al., 2011). The GET1 coiled-coil domain wedges into the GET3 dimer interface, thus stabilizing an open dimer conformation of GET3 in which the hydrophobic groove is disrupted (Mariappan et al., 2011; Stefer et al., 2011; Wang et al., 2011; Kubota et al., 2012). The binding sites of GET1 partially overlap with the interface for the GET2-GET3 complex, suggesting that GET1 displaces GET2, at least to a certain extent, or binds simultaneously with GET2 to opposite sides of the GET3 homodimer (Stefer et al., 2011; Zalisko et al., 2017). Interaction between GET3 and the GET1-GET2 heterodimer leads to release of the bound TA substrate and its integration into the ER membrane via the insertase function of the GET1-GET2 TMDs (Wang et al., 2014). Rebinding of ATP to GET3 weakens the GET3-GET1 interaction by transition into the closed conformation, thereby triggering dissociation of GET3 from the membrane back to the cytosol where it forms a complex with GET4/5 (Chartron et al., 2010; Stefer et al., 2011). The binding sites of GET4 on GET3 overlap with the receptor binding sites, thus preventing rebinding of GET3 to the GET1-GET2 transmembrane complex (Gristick et al., 2014; Rome et al., 2014).

Recently, a novel, GET3-independent function of GET1 and GET2 in mitophagy has been identified, however, their contribution remains unknown (Onishi et al., 2018).

### **Post-translational targeting via the TRC pathway in mammals**

The GET pathway is conserved in metazoans, where it has maintained its function in regulating TA protein insertion and is called Transmembrane domain Recognition Complex (TRC) pathway (Stefanovic and Hegde, 2007). Almost all yeast GET proteins share sequence similarity with their mammalian counterparts except for GET2 which has a functional

homologue called CAML (calcium-modulating cyclophilin ligand) (Yamamoto and Sakisaka, 2012).

The mammalian pre-targeting complex is composed of the GET4 and GET5 orthologues, TRC35 (transmembrane domain recognition complex 35 kDa subunit) and UBL4A (ubiquitin-like protein 4a), respectively, and the ubiquitin-like protein BAG6 (BCL-2-associated athanogene 6, also known as BAT3/Scythe), that is not present in yeast (Mariappan et al., 2010). Unlike their counterparts in yeast, TRC35 and UBL4A do not directly interact with each other, but do both interact with BAG6, which functions as scaffold to link TRC35 to UBL4A (Mock et al., 2015). This heterotrimeric complex binds the mammalian orthologue of SGT2, SGTA (small glutamine-rich tetratricopeptide repeat-containing protein alpha), either via the BAG6 or UBL4A subunit (Leznicki et al., 2013). TRC35 recruits TRC40 (transmembrane domain recognition complex 40 kDa subunit, also known as ASNA1), the mammalian counterpart of yeast GET3, to the complex and thereby facilitates substrate handover from SGTA to the targeting factor. TRC40 delivers the TA protein to the ER, where its receptor complex composed of CAML and the GET1 orthologue WRB (tryptophan-rich basic protein, also known as CHD5) promotes substrate release and insertion into the membrane (Yamamoto and Sakisaka, 2012; Vilardi et al., 2014).

In mammals, loss of GET/TRC pathway components results in more severe phenotypes than in yeast. Homozygous deletion of TRC40 or the ER membrane receptor component CAML leads to early embryonic lethality in mice while conditional knockout of CAML causes deafness (Tran et al., 2003; Mukhopadhyay et al., 2006; Bryda et al., 2012). Likewise, deletion of WRB in mouse inner hair cells causes synaptic hearing impairment due to reduced ER-insertion of the TA protein otoferlin (Vogl et al., 2016). Interestingly, knockout of TRC40 in HeLa cells affects the biogenesis of only a few TA proteins as was previously reported for tissue-specific knockout of WRB in mouse cardiomyocytes and hepatocytes. Most TA proteins, however, are not severely affected by disruption of the TRC pathway, suggesting the existence of alternative ER-targeting routes (Rivera-Monroy et al., 2016; Casson et al., 2017).

### **The WRB-CAML transmembrane complex in mammals**

WRB, also called CHD5 (congenital heart disease 5 protein), has been identified as the mammalian GET1 orthologue and functions as a subunit of the TRC40 receptor complex (Vilardi et al., 2011). The highest sequence conservation between GET1 orthologues occurs in

the cytosolic coiled-coil domain which contains the GET3/TRC40 binding site (Stefer et al., 2011; Vilardi et al., 2011).

Interestingly, mammals have no obvious sequence orthologue for yeast GET2 but instead acquired CAML as a functional homologue (Yamamoto and Sakisaka, 2012; Vilardi et al., 2014). GET2 and CAML share very similar structures, with a long cytosolic N-terminus, three transmembrane segments and a luminal C-terminal region (Bram and Crabtree, 1994; Schuldiner et al., 2005). A cluster of positively charged residues (32-RRRK) within the N-terminal domain of CAML is a major determinant for interaction with TRC40, comparable to the 14-RERR motif of GET2 (Stefer et al., 2011; Yamamoto and Sakisaka, 2012). CAML binding to TRC40 is competed by WRB, suggesting that the binding sites on TRC40 may partially overlap (Yamamoto and Sakisaka, 2012).

CAML and WRB interact via their TMDs to form a heterodimeric receptor complex with probable TMD-insertase activity and can functionally replace GET1 and GET2 in yeast (Yamamoto and Sakisaka, 2012; Vilardi et al., 2014). It has been shown that WRB and CAML regulate each other at the transcriptional level, however, latest reports suggest that the regulation may rather occur post-translationally (Colombo et al., 2016; Rivera-Monroy et al., 2016; Haßdenteufel et al., 2017).

## OBJECTIVES AND EXPECTED OUTPUT OF THE THESIS

The GET pathway is considered as major route for the targeting of TA proteins to the ER in opisthokonts (fungi and metazoa). However, its conservation in plants has not yet been demonstrated. This thesis aims at identifying and analysing the components involved in a putative GET pathway in *Arabidopsis thaliana*, with the main focus on studying the two membrane receptors. Potential orthologues of GET candidates will be identified by sequence comparison and validated using *in vivo* localisation and interaction studies. Detailed phenotypic and biochemical analyses of loss-of-function lines will provide new insight into the physiological role of the GET pathway in plants.

The *Arabidopsis* genome does not encode an obvious orthologue of the co-receptor GET2. However, direct *in planta* interaction analysis using immunoprecipitation-mass spectrometry (IP-MS) of AtGET1-GFP identified a promising candidate. As part of this work, structural and functional analyses including gene expression, subcellular localisation and protein-protein interaction studies will be performed to investigate a putative role of this candidate as the functional orthologue of GET2 in *Arabidopsis*. In addition, loss-of-function lines will be generated and analysed phenotypically.

Furthermore, this work will highlight an *in vivo* method to analyse binary protein-protein interactions, with *Arabidopsis* GET orthologues as examples.

## RESULTS

### Chapter one: Loss of GET pathway orthologs in *Arabidopsis thaliana* causes root hair growth defects and affects SNARE abundance

Xing, S., Mehlhorn, D.G., Wallmeroth, N., Asseck, L.Y., Kar, R., Voss, A., Denninger, P., Schmidt, V.A., Schwarzländer, M., Stierhof, Y.D., Grossmann, G., Grefen, C., 2017, PNAS.

Targeting and insertion of TA proteins into the ER membrane is a challenging task for eukaryotic cells. The C-terminal TMD of TA proteins not only prohibits co-translational membrane insertion but also requires constant chaperoning to prevent aggregation in the cytosol. Such TA proteins are post-translationally targeted into the ER membrane via the GET pathway that was previously described in mammals and yeast (Stefanovic and Hegde, 2007; Schuldiner et al., 2008). However, nothing is known about the underlying mechanism for TA protein insertion in plants.

To address this issue, we performed *in silico* sequence comparison and identified a single *Arabidopsis* orthologue for GET1 and GET4, and three putative orthologues of GET3 (*AtGET3a-c*) that show differences in conserved sequence motifs and subcellular localisation. *AtGET3a* localises to the cytosol whereas *AtGET3b* and *AtGET3c* are targeted to the chloroplast stroma and mitochondrial matrix, respectively. Confocal imaging also confirmed an ER localisation for *AtGET1* and cytosolic localisation of *AtGET4*. We then examined the interaction among the *Arabidopsis* orthologues using the mating-based split-ubiquitin system (mbSUS) and ratiometric bimolecular fluorescence complementation (rBiFC) and found that *AtGET3a*, but neither *AtGET3b* nor *AtGET3c*, interacts with the other *Arabidopsis* GET pathway components. Additionally, we identified several TA proteins as interacting with *AtGET3a* and *AtGET1*, including the Qa-SNARE SYP123 (syntaxin of plants 123) and its R-SNARE partner VAMP721 (vesicle-associated membrane protein 721).

To test whether the *Arabidopsis* genes can functionally substitute for the loss of their yeast orthologues, we analysed their ability to complement the temperature-sensitive growth defect in the yeast *get1* and *get3* mutant, respectively. Overexpression of *AtGET1* and *AtGET3a* slightly rescues the growth defect of the corresponding yeast deletion strain, whereas the chloroplast and mitochondrial *AtGET3* paralogues fail to complement.

In order to reveal the physiological function of GET in plants, we phenotypically analysed appropriate *Arabidopsis* T-DNA insertion lines. Loss of *AtGET1*, *AtGET3a*, and *AtGET4* but

not *AtGET3b* and *AtGET3c* leads to reduced root hair growth in otherwise normally developing plants. We further demonstrate that this phenotype can be rescued by introducing genomic fragments of the corresponding genes and is not enhanced in double or triple mutants. We next introduced the marker Qa-SNARE SYP123 into the *Atget1* and *Atget3a* backgrounds and assessed its expression by immunoblot analysis of membrane fractions and quantitative PCR. Our data show a significant decrease in both transcript and protein abundance of SYP123, whereas its subcellular distribution is unaffected. To gain further insight into the physiological relevance of the GET pathway in plants, we investigated the effect of *AtGET3a* overexpression. Overexpression of *AtGET3a* in the absence of *AtGET1* results in severe dwarfism including reduced root growth and fertility, and confocal imaging of these plants revealed abnormal clustering of *AtGET3a* into aggregate-like structures.

For details see appendix I.

## Chapter two: ER membrane receptors of the GET pathway are conserved throughout eukaryotes

Asseck, L.Y., Mehlhorn, D.G., Rivera Monroy, J., Ricardi, M.M., Breuninger, H., Wallmeroth, N., Berendzen, K.W., Nouwrosian, M., Xing, S., Schwappach, B., Bayer, M., Grefen, C., submitted.

The GET pathway for TA protein insertion has been previously shown to be partially conserved in plants. GET1, GET3 and GET4 were identified in *Arabidopsis thaliana* based on their sequence similarity to yeast and human GET proteins (Xing et al., 2017). However, no orthologue of the receptor GET2/CAML has been found in *Arabidopsis* or other plant genomes.

To find a potential protein that takes over the function of GET2/CAML in plants, we performed immunoprecipitation of *At*GET1-GFP followed by mass spectrometry. Among the interacting partners, we found a protein of unknown function, G1IP (*At*GET1-interacting protein), which resembles a similar TMD architecture as GET2/CAML and carries a conserved positively charged motif in its N-terminus. Additionally, our IP-MS identified a close homologue of G1IP, named G1IP-like, as an interactor of *At*GET1. G1IP shows a broad expression pattern in different developmental stages and tissues, whereas G1IP-like is expressed only in inflorescence. The subcellular localisation of GFP-labelled G1IP and G1IP-like in *Arabidopsis* leaf cells has shown that both proteins are localised at the ER membrane. To examine whether G1IP and G1IP-like interact with the previously identified *Arabidopsis* GET pathway components, rBiFC assays in *Nicotiana benthamiana* leaves were performed. Both homologues interact with *At*GET1 but not *At*GET3a and *At*GET4. We next carried out co-immunoprecipitation experiments in *Arabidopsis* and found that *At*GET3a binds to G1IP in wild type but not in an *Atget1* mutant background, whereas G1IP-like was not detected in the co-immunoprecipitate.

To further investigate whether the identified proteins are part of the *Arabidopsis* GET pathway, we phenotypically analysed appropriate T-DNA insertion lines as well as a CRISPR-based gene knockout of G1IP. We found that loss of G1IP leads to reduced root hair growth phenocopying the *Atget1* mutant, whereas the root hairs of *gilp-like* mutant seedlings are similar in length to wild type. Moreover, we demonstrate that this phenotype can be rescued by introducing the genomic sequence of G1IP and is not enhanced further in the *Atget1gilp* double mutant. To test whether *At*GET1 and G1IP or G1IP-like can functionally replace GET1 and GET2 in yeast, we performed complementation assays using a GET-receptor deficient

strain. Co-overexpression of *AtGET1* and G1IP partially rescues the temperature-sensitive growth phenotype of the yeast mutant, whereas *AtGET1* and G1IP-like fail to complement. Moreover, we observed that a mixed expression of *Arabidopsis* and yeast GET receptor proteins does not rescue as efficiently as the homologous combinations.

We next studied the function of the different domains of G1IP to define their putative roles. Using rBiFC and co-immunoprecipitation experiments we found that the TMDs of G1IP but not its cytoplasmic region interact with *AtGET1*. *In vitro* insertion assays have shown that the G1IP cytoplasmic domain can block the insertion of TA proteins into mammalian microsomes whereas the coiled-coil domain of *AtGET1* has no effect. Mutation of the N-terminal conserved cluster in G1IP, however, affects this function.

For details see appendix II.

### **Chapter three: Detecting interactions of membrane proteins: the split-ubiquitin system**

**Asseck, L. Y.** and Grefen, C., 2018, Two-Hybrid Systems, Springer.

Protein-protein interactions (PPIs) play a crucial role in almost all biological processes and pathways. Also, the GET pathway consists of a cascade of PPIs chaperoning newly synthesized TA proteins to the ER for membrane insertion. Here, we describe a protocol for the mating-based split-ubiquitin system (mbSUS) and its modification cytoSUS to analyse binary PPIs *in vivo*. The system is based on the reassembly of ubiquitin from complementary N- and C-terminal fragments fused to proteins of interest and subsequent cleavage of an artificial transcription factor, which activates reporter genes for visualisation. By using the cytosolic ATPase *AtGET3a* and its membrane receptor *AtGET1*, we demonstrate exemplarily that the SUS approach can be used to study interactions of both membrane and soluble proteins. The system allows for simple, fast and inexpensive detection of PPIs and can be also used for large-scale interaction screens (Asseck et al., 2018).

For details see appendix III.

## DISCUSSION

The GET pathway for TA protein insertion into the ER has been intensively studied in yeast and mammals but has not yet been described in plants. We have identified components of a putative GET pathway in *Arabidopsis thaliana* and investigated their role in membrane insertion of SNARE proteins. Our results demonstrate functional conservation of the GET pathway in plants but also question its monopoly as the sole ER-targeting route for TA proteins.

### Conservation and divergence of the GET pathway in eukaryotes

Although TA protein insertion into the ER membrane is an important housekeeping function in all eukaryotic cells, it was not yet known whether the mechanism in plants diverges from that described in yeast and mammals. Homology analysis of protein sequences with yeast and human GET proteins revealed conservation of putative orthologues for almost all GET pathway components in *Arabidopsis thaliana*, except for the receptor GET2 (Srivastava et al., 2017; Xing et al., 2017). Similarly, orthologues of GET1, GET3 and GET4 were recently identified in *Oryza sativa* subsp. *indica* and *Solanum tuberosum*, while GET2 and GET5 were not observed in both the plant species (Manu et al., 2018). The absence of a putative plant orthologue of GET2 may be attributed to low evolutionary selection pressure compared to other GET proteins. Notably, GET2 has also no sequence orthologue in metazoa but shares structural and functional similarities with a calmodulin-binding protein named CAML (Yamamoto and Sakisaka, 2012). Together with the GET1 orthologue WRB, CAML has been demonstrated to act as membrane receptor for TA protein insertion in mammalian cells (Vilardi et al., 2014). However, also no gene homologous to CAML can be found in plants, highlighting the evolutionary divergence of the GET2 sequences among eukaryotes. We have now identified an unknown transmembrane protein with high overall structural (but low sequence) homology to GET2/CAML, that takes over its function in plants (Asseck et al., submitted) (discussed below).

Sequence comparison and phylogenetic analysis have revealed that plants possess a higher, although variable, number of GET3 orthologues compared to other eukaryotes (Xing et al., 2017; Manu et al., 2018; Bodensohn et al., 2019). The *Arabidopsis* genome encodes three GET3 paralogues that belong to two distinct clades, termed GET3a and GET3bc. Clade a comprises cytosolic GET3 proteins, whereas members of clade bc are localised to organelles

such as chloroplasts and mitochondria and are not present in opisthokonts. Among the three GET3 paralogues in *Arabidopsis*, only the cytosolic *AtGET3a* has been shown to be part of the GET pathway. Interaction studies found that *AtGET3a* homodimerizes and binds to *AtGET1* and *AtGET4* *in vivo*, demonstrating that the *AtGET3a* interactome resembles the orthologous yeast network (Xing et al., 2017). Interestingly, some plant species possess several cytosolic GET3 proteins, however, whether they act redundantly in the GET pathway, remains to be elucidated (Manu et al., 2018; Bodensohn et al., 2019).

The biological function of the plastidic and mitochondrial GET3 orthologues is still unknown. Despite their lack of the GET1 and TA protein binding sites, it is speculated that the GET3bc proteins are involved in TA protein biogenesis in the respective organelle (Zhuang et al., 2017; Anderson et al., 2019; Bodensohn et al., 2019). Interestingly, orthologues of GET1/WRB can be also found in the chloroplast thylakoid and mitochondrial inner membrane and are named ALB3 (albino 3) and OXA1 (oxidase assembly protein 1), respectively. Both proteins can function as insertases and are important for co- and post-translational insertion of transmembrane proteins (Anghel et al., 2017). Based on its localisation in the chloroplast stroma, *AtGET3b* could act upstream of ALB3 in targeting of chloroplast-encoded or imported TA proteins to the thylakoid membrane for insertion. However, since GET3bc proteins do not contain the binding motifs for either GET1 or TA proteins, interactions have to be confirmed first.

*AtGET3c* was previously shown to localise to the outer mitochondrial membrane, where it was thought to be involved in the biogenesis of mitochondrial TA proteins (Duncan et al., 2013; Zhuang et al., 2017). However, this localisation was only based on transient expression studies in *Arabidopsis* cell culture protoplasts. Confocal laser scanning, and transmission electron microscopy of stable lines clearly showed that *AtGET3c* localises to the mitochondrial matrix but not to the mitochondrial outer membrane (Xing et al., 2017). Therefore, *AtGET3c* might rather act in the biogenesis of inner membrane proteins, possibly upstream of OXA1. However, mitochondrial GET3 proteins are not present in all plant species (and not in opisthokonts), challenging its role as a critical targeting factor for mitochondrial inner membrane proteins (Bodensohn et al., 2019).

The pre-targeting steps of the GET pathway in plants are mostly unresolved. Since only an orthologue of GET4/TRC35 has yet been annotated in plants, it is unclear whether it functions in a complex similar to that in other eukaryotes or independently. Recently, it has been

proposed that UBQ12 (ubiquitin 12, AT1G55060) might be the *Arabidopsis* orthologue of GET5/UBL4A, however, it has not been explained how exactly they identified this candidate, and they did not characterize it further (Srivastava et al., 2017). The *UBQ12* gene, however, contains three premature in-frame stop codons within the first two ubiquitin repeats and is therefore assumed to be a pseudogene (Bachmair et al., 2001). Thus, a role for UBQ12 as a pre-targeting factor of the GET pathway seems unlikely. In contrast, our BLASTp analysis of GET5/UBL4A revealed multiple ubiquitin family proteins as potential orthologues, as for BAG6. Interestingly, BAG6 has been shown to have dual functions in TA protein biogenesis and quality control. Substrates bound to BAG6 can be either handed off to the targeting factor TRC40 for ER insertion or are ubiquitinated by the BAG6-associated E3 ubiquitin ligase RNF126 (ring finger protein 126), which is recruited via the N-terminal ubiquitin-like domain of BAG6, and are targeted for proteasomal degradation (Hessa et al., 2011; Rodrigo-Brenni et al., 2014). This process can be reversed by the co-chaperone SGTA, which participates not only in the handoff of substrate to TRC40 but also actively promotes the deubiquitination and stabilisation of target proteins (Leznicki and High, 2012; Leznicki et al., 2013). However, also yeast lacks an obvious BAG6 orthologue, questioning how quality control is regulated in organisms other than mammals. It is conceivable that the processes in plants require multiple ubiquitin family proteins that act as specific components in either (pre-)targeting or degradation of membrane proteins or operate redundantly in both pathways. As for GET5/UBL4A and BAG6, multiple potential orthologues (tetratricopeptide proteins) were predicted for SGT2/SGTA, further supporting the notion that the physiological network related to protein (pre-)targeting and quality control in plants might be more complex.

Moreover, our data provide experimental evidence that not only the components of the GET pathway but also its function in TA protein biogenesis is common to all eukaryotes. We show that the insertion of the *Arabidopsis* SNARE SYP123 primarily depends on the GET pathway, as the absence of either *AtGET1* or *AtGET3a* leads to reduced protein levels of SYP123 at the plasma membrane (Xing et al., 2017). Similarly, another study demonstrated that the *Arabidopsis* GET system is required for proper insertion of SYP72 into the ER membrane. They further showed *in vitro* that also yeast GET proteins can mediate the insertion of plant-specific SYP72, underscoring functional conservation of the pathway in eukaryotes (Srivastava et al., 2017). However, although SYP72 is a pollen-specific SNARE protein, we were unable to observe a pollen-related phenotype in the *Atget* mutants (Xing et al., 2017). The segregation ratio of the lines is normal and pollen tube growth is not affected, suggesting gene redundancy

or compensation by alternative targeting pathways. Loss of GET components, however, leads to impaired root hair elongation, at least partly due to reduced biogenesis of SYP123, whose absence has been shown to result in shortened root hairs (Ichikawa et al., 2014; Xing et al., 2017). Interestingly, the phenotype in the *Atget* mutants is more pronounced than in the *syp123* mutant, indicating additive effects.

However, there are also findings suggesting that the functions of the GET components may have diverged between plants and opisthokonts. Using a complementation approach in yeast, it has been found that the mammalian orthologues can functionally replace their counterparts in yeast (Vilardi et al., 2014). In contrast, *Arabidopsis* GET proteins cannot fully rescue the temperature-sensitive growth defect of the corresponding yeast mutants, indicating functional dissimilarities between the proteins (Xing et al., 2017; Asseck et al., submitted).

### **The GET receptor complex in *Arabidopsis thaliana***

In yeast, the final insertion step of TA proteins into the ER membrane is mediated by the GET1-GET2 transmembrane complex (Wang et al., 2014). While GET1 is evolutionary conserved, GET2 seems to be specific to yeast. However, we identified a previously uncharacterized protein, G1IP, which takes over the function of GET2 in *Arabidopsis thaliana* (Asseck et al., submitted). Although the overall similarity at the amino acid level is quite low, the predicted structures are highly similar. Both proteins possess a long cytosolic N-terminus, three TMDs and a luminal C-terminal region (Schuldiner et al., 2005; Asseck et al., submitted). A similar domain structure was also reported for CAML, the functional equivalent of GET2 in mammalian cells (Bram and Crabtree, 1994; Yamamoto and Sakisaka, 2012). Besides structural homology, another key feature of these proteins is the presence of a positively charged motif at the N-terminus where the GET3/TRC40 binding site is located (Stefer et al., 2011; Yamamoto and Sakisaka, 2012). This strongly implies evolutionary pressure on functional sites and domains for maintaining a common function and suggests shared ancestry. Consistent with this, phylogenetic analysis of the G1IP amino acid sequence points to an early evolutionary divergence of the GET2 proteins rather than convergence (Asseck et al., submitted).

G1IP is constitutively co-expressed with *AtGET1* throughout development and localises at the ER membrane, consistent with a putative role as a receptor component of the GET pathway in plants. Disruption of the *G1IP* gene reduces root hair growth, as has been shown for other *Atget*

mutants (Xing et al., 2017; Asseck et al., submitted). Moreover, protein interaction analyses demonstrated that G1IP interacts with *AtGET1* via its TMDs, whereas the cytosolic N-terminal domain seems to be dispensable for complex formation (Asseck et al., submitted). Due to its length, however, the N-terminus might function as the first docking point for *AtGET3a*, bringing it into proximity to the membrane and *AtGET1*. The precise order of binding, however, remains to be experimentally determined. *AtGET1* interacts with *AtGET3a* and G1IP but, significantly, *AtGET3a* and G1IP do not interact in the absence of *AtGET1* (Asseck et al., submitted). We therefore infer that the association between the TMDs of the two receptors causes G1IP to adopt a conformation favourable for binding to *AtGET3a*. Conversely, whether the interaction of *AtGET1* with *AtGET3a* also depends on the presence of G1IP is still unclear. Recently, it has been shown that the functional mammalian homologue CAML requires its partner WRB to be correctly inserted in the ER membrane. In the absence of WRB, CAML adopts an aberrant topology with its TMDs (TMD-2 and/or TMD-3) exposed to the ER lumen or cytosol and gets degraded (Carvalho et al., 2019; Inglis et al., 2020). However, we did not observe instability of G1IP in the absence of *AtGET1*, indicating that this mode of regulation is rather specific for the mammalian receptor components (Asseck et al., submitted).

As for GET2 and CAML, a cluster of basic residues is present in the N-terminal domain of G1IP which is composed of three arginines and one lysine (Asseck et al., submitted). This positively charged region has been reported to participate in the binding site for GET3 and TRC40, respectively (Stefer et al., 2011; Yamamoto and Sakisaka, 2012). Charge inversion at the CAML N-terminus is sufficient to completely abolish the interaction with TRC40, resulting in reduced membrane insertion of TA proteins (Yamamoto and Sakisaka, 2012). Similarly, substitution of the four residues in G1IP eliminates its ability to interfere with the mammalian insertion system, highlighting the importance of the positive charges at the N-terminus (Asseck et al., submitted). It is conceivable that the cluster of basic amino acids in G1IP forms ionic contacts to negatively charged residues in *AtGET3a*, similar as proposed for GET2 and CAML (Mariappan et al., 2011; Stefer et al., 2011; Yamamoto and Sakisaka, 2012). The interference of the native G1IP N-terminus with the mammalian machinery for TA protein insertion suggests a conserved role for this domain in binding of TRC40/GET3. The coiled-coil motif of *AtGET1*, however, does not inhibit membrane insertion, indicating that the binding sites or functional residues may have diverged from those of its orthologue in mammals (Asseck et al., submitted). Functional differences are also evident from complementation assays in yeast. WRB and CAML have been shown to fully replace the function of GET1 and GET2 *in vivo*,

whereas *AtGET1* and G1IP only partially compensate the lack of the yeast receptor components (Vilardi et al., 2014; Asseck et al., submitted). Interestingly, a combination of yeast and *Arabidopsis* proteins is even less efficient in rescuing the mutant phenotype, suggesting that they are probably unable to form a fully functional receptor complex. Co-expression of *AtGET1* and *GET2*, however, seems to complement the defect slightly better than *GET1* with G1IP (Asseck et al., submitted). A similar observation has been made for the combinations of yeast and mammalian proteins. While WRB and *GET2* can rescue the growth defect of the *get1/get2* deletion mutant, *GET1* and CAML fail to complement under most conditions (Vilardi et al., 2014). These results indicate that *GET2* and its orthologues in plants and mammals have become more specialized in their function than the *GET1* orthologues and are therefore not interchangeable across kingdoms.

There is a close homologue of G1IP, called *G1IP-like*, in *Arabidopsis*. Although the two proteins share structural features and ER localisation, we provide evidence that they do not act redundantly in the GET pathway. For example, *G1IP-like* is specifically expressed in flowers, but not in roots, stems, and leaves. Thus, in contrast to disruption of *GET* genes, loss of *G1IP-like* in *Arabidopsis* has no impact on root hair growth. Moreover, *G1IP-like* cannot interact with *AtGET3a* even in the presence of *AtGET1*, although its N-terminal domain also carries a cluster of positively charged residues. Furthermore, *G1IP-like* in combination with *AtGET1* is not able to complement the loss of *GET1* and *GET2* in yeast (Asseck et al., submitted). The function of *G1IP-like* *in vivo* is not yet known. However, due to its expression profile, we assume that it may be involved in flower-specific processes. Thus, future studies should concentrate on a more detailed analysis of flower development in *glip-like* mutant plants.

## **GET pathway dependence of TA proteins**

The GET pathway is considered as the dominant cellular mechanism for post-translational membrane insertion of TA proteins. However, loss of GET components in *Arabidopsis thaliana* is not lethal but only impacts on root hair length, indicating that disruption of the GET pathway does not globally impair TA protein insertion. Also, IP-MS analysis of *AtGET3a*-GFP expressing lines identified only 23 of the 454 TA proteins present in *Arabidopsis thaliana*, further suggesting that not all TA proteins strictly depend on the GET pathway for membrane insertion. Consistent with this, validation of a subset of putative candidates revealed two TA proteins (SYP43 and AT5G40510) that fail to interact with both *AtGET3a* and *AtGET1* (Xing

et al., 2017). We assume that there are likely to be different classes of TA proteins with varying degrees of dependence on the GET pathway for their biogenesis. While some TA proteins seem to be strictly GET-dependent or -independent in their targeting, we presume that most TA proteins favour the GET pathway under normal conditions but can also enter alternative, yet undefined targeting routes when the GET pathway is disrupted. The exact protein features that define the differential dependence on the GET system are still unknown. However, analysis of conditional WRB knockout mice has implied that there might be no correlation between hydrophobicity of the TMD and GET pathway dependency *in vivo*, as has been proposed in previous *in vitro* studies (Rabu et al., 2008; Rivera-Monroy et al., 2016).

In *Arabidopsis*, two TA proteins have so far been experimentally demonstrated to depend on the GET pathway *in vivo*. Our results show that insertion of SYP123 is impaired, although not completely prevented, in the absence of a functional GET pathway, suggesting that a backup targeting system for this TA protein exists (Xing et al., 2017). Additionally, another recent study identified the TA SNARE SYP72 as a substrate of the *Arabidopsis* GET pathway (Srivastava et al., 2017).

### **Alternative insertion pathways for TA proteins**

Studies on the GET pathway in yeast and mammals have established its functional importance in post-translational membrane insertion of TA proteins (Stefanovic and Hegde, 2007; Schuldiner et al., 2008). Orthologues of GET components can be also found in plants and their role in TA protein biogenesis seems to be conserved (Srivastava et al., 2017; Xing et al., 2017; Manu et al., 2018). However, there is growing evidence that additional targeting routes to the ER exist to ensure efficient membrane insertion. For example, the GET pathway is not essential for viability in yeast and tissue-specific knockout of GET components in mice has been shown to affect the biogenesis of only a few TA proteins (Schuldiner et al., 2008; Rivera-Monroy et al., 2016). Similarly, loss of GET function in *Arabidopsis thaliana* is not lethal and instead causes only mild phenotypes (Srivastava et al., 2017; Xing et al., 2017). Moreover, only a limited number of TA proteins has yet been identified to depend on the GET system *in vivo*, further suggesting the existence of an alternative insertion route to the ER (Rivera-Monroy et al., 2016; Xing et al., 2017). Indeed, in the last few years multiple such pathways have been identified in yeast and mammals that may also be present in plants.

Several studies report a novel post-translational function for SRP in the targeting of TA proteins to the mammalian ER. They show that SRP can associate post-translationally with the C-terminal hydrophobic domains of TA proteins and facilitate their SR-dependent membrane insertion, although the exact mechanism of integration remains to be elucidated (Abell et al., 2004; Casson et al., 2017). Previous research in plants, however, has mostly focused on the mechanism and role of SRP in chloroplasts, whereas its function in protein targeting to the ER has not yet been investigated. It is therefore difficult to conclude whether SRP provides an alternative pathway for TA proteins to the ER in plants as well.

Moreover, analyses *in vitro* and in cultured cells suggest that a minimal combination of HSP40 (heat shock protein 40) and HSC70 (heat shock cognate 70, also known as HSP70-8) is sufficient to promote the ATP-dependent membrane insertion of mammalian TA proteins (Abell et al., 2007; Rabu et al., 2008). HSP40 is a co-chaperone for HSC70 and regulates ATP-dependent substrate binding and ATPase activity of HSC70 (Meimaridou et al., 2009). Both chaperones have been demonstrated to bind directly to TA protein substrates, however, the exact mechanism by which proteins utilising these chaperones are targeted and inserted is still unknown. The HSP40/HSC70-mediated pathway has been found to be essential for the insertion of TA proteins with TMDs of low hydrophobicity. TA proteins with more hydrophobic TMDs can also exploit the chaperone-mediated route when no alternative pathways are available (Abell et al., 2007; Rabu et al., 2008). However, a relevance of HSP40/HSC70-mediated membrane insertion *in vivo* remains to be demonstrated. Since heat shock proteins of the HSP70 family and their HSP40 co-chaperones can be found in all organisms, it seems conceivable that also their role in TA protein targeting is conserved across kingdoms. The HSP40/HSC70 complex might work in parallel with GET3/TRC40 in chaperoning TA proteins to the insertion machinery at the ER, possibly explaining why only about 5% of all TA proteins have been identified as interacting with *At*GET3a by IP-MS. Analysing the interaction between HSP40/HSC70 and the GET1-GET2 receptors would provide first evidence whether they actually facilitate TA protein membrane insertion via binding to the GET complex.

More recently, another alternate targeting mechanism has been discovered in yeast, termed the SRP-independent (SND) pathway. This pathway consists of (at least) three components, SND1-3 (Aviram et al., 2016). SND1 is localised in the cytosol and predicted to interact with ribosomes, where it may be involved in co-translational capturing of nascent proteins (Fleischer et al., 2006; Aviram et al., 2016). SND2 and SND3 are both membrane-bound proteins that

form a complex with the SEC61 translocon and could act as receptors. However, the exact molecular mechanisms remain to be uncovered. The SND pathway has been shown to be particularly important for targeting of proteins with central TMDs but can also function as a backup system for substrates that commonly depend on either the SRP or GET pathway for translocation. Thus, double deletions of SND and GET components are synthetically lethal in yeast (Aviram et al., 2016). The SND pathway seems to be functionally conserved in mammals, although only a human orthologue of SND2 (TMEM208, named hSND2) has yet been identified (Casson et al., 2017; Haßdenteufel et al., 2017). Similarly, the *Arabidopsis* genome encodes two putative SND2 orthologues (AT4G30500, SND2a; AT2G23940, SND2b) whereas SND1 and SND3 are missing (C. Grefen, personal communication). Thus, SND-mediated targeting might also exist in plants and compensate for the absence of a functional GET pathway, possibly explaining the mild phenotype observed in *Atget* mutants. The phenotypic analysis of double mutants for SND and GET proteins would provide first evidence whether there is redundancy between these two targeting routes in plants as well.

Recent studies also suggest a novel role for the ER membrane complex (EMC) in TA protein insertion (Guna et al., 2018). The EMC is a large multifunctional, multi-subunit protein complex and has been shown to serve as a post-translational insertase for ER-destined TA proteins with TMDs of moderate to low hydrophobicity in mammalian cells (Jonikas et al., 2009; Wideman, 2015; Guna et al., 2018). However, some TA proteins showed partial dependence on both EMC and TRC40, suggesting overlapping substrate specificity and functional redundancy between these two pathways (Guna et al., 2018; Volkmar et al., 2019). Furthermore, the EMC has been implicated in the co-translational insertion of multi-pass transmembrane proteins by guiding insertion of the first TMD, whereas downstream TMDs are inserted by the SEC61 translocon (Satoh et al., 2015; Chitwood et al., 2018; Shurtleff et al., 2018). Although the EMC is highly conserved throughout eukaryotes, it is not clear whether its functions have diverged. However, it is possible that the EMC is involved in TA protein biogenesis in plants as well and acts redundantly with the GET pathway to ensure robust targeting to the ER.

## CONCLUSIONS AND PERSPECTIVES

In this work, we have shown that the GET pathway is conserved in plants where it also regulates membrane insertion of TA proteins. However, not all TA proteins seem to depend on this pathway in their targeting and its disruption is not lethal, indicating the existence of alternative targeting routes to the ER to maintain sufficient insertion efficiency and thus ensure survival. Several potential alternative pathways have been recently identified in yeast and mammals. Future studies should investigate which of these or other pathways exist in plants, their relative physiological contribution, and their potential inter-relationship to improve our understanding of plant TA protein biogenesis. In this context, it is also important to identify the protein features that determine either their dependence on a specific pathway or their ability to use multiple routes.

The identification of the functional *Arabidopsis* orthologue of the receptor GET2/CAML gives further insight into the insertion mechanism at the membrane in plants and its protein sequence now allows us to identify and to characterize GET2 orthologues in other plant (and algae) species. Clearly, the interaction between the two membrane receptors and *At*GET3a needs to be analysed in more detail. For example, it is not known at present whether *At*GET1 and G1IP bind sequentially or simultaneously and how they depend on each other for the interaction with *At*GET3a. Therefore, structural insights into the *At*GET3a/receptor complex will be required.

Although with G1IP most of the GET pathway components are now identified in plants, further research is required to elucidate the existence and composition of a pre-targeting complex. So far, only one orthologue of GET4/TRC35 has been identified and functionally characterized, whereas the other subunits are still speculative as several potential orthologues were predicted.

Moreover, it is not yet known why plants have evolved multiple, organelle-specific GET3 paralogues. While GET3a proteins seem to be involved in the targeting of TA proteins to the ER in a manner similar to GET3 in yeast, the function of GET3bc proteins remains to be identified. Despite their lack of residues related to the targeting function, it is speculated that GET3bc proteins may play an analogous role in the biogenesis of chloroplast and mitochondrial TA proteins. Hence, interaction studies or screens are required to assess their potential binding ability for TA proteins and to identify up- and downstream interaction partners to reveal the molecular network of these proteins.

## REFERENCES

- Abell, B. M., et al. (2004). "Signal recognition particle mediates post-translational targeting in eukaryotes." The EMBO journal **23**(14): 2755-2764.
- Abell, B. M., et al. (2007). "Post-translational integration of tail-anchored proteins is facilitated by defined molecular chaperones." Journal of Cell Science **120**(10): 1743-1751.
- Akopian, D., et al. (2013). "Signal recognition particle: an essential protein-targeting machine." Annual review of biochemistry **82**: 693-721.
- Anderson, S. A., et al. (2019). "Membrane-Specific Targeting of Tail-Anchored Proteins SECE1 and SECE2 Within Chloroplasts." Front Plant Sci **10**: 1401.
- Anghel, S. A., et al. (2017). "Identification of Oxal homologs operating in the eukaryotic endoplasmic reticulum." Cell reports **21**(13): 3708-3716.
- Asseck, L. Y., et al. (2018). ER membrane protein interactions using the split-ubiquitin system (SUS). The Plant Endoplasmic Reticulum, Springer: 191-203.
- Auld, K. L., et al. (2006). "The conserved ATPase Get3/Arr4 modulates the activity of membrane-associated proteins in *Saccharomyces cerevisiae*." Genetics **174**(1): 215-227.
- Aviram, N., et al. (2016). "The SND proteins constitute an alternative targeting route to the endoplasmic reticulum." Nature **540**(7631): 134.
- Bachmair, A., et al. (2001). "Ubiquitylation in plants: a post-genomic look at a post-translational modification." Trends in Plant Science **6**(10): 463-470.
- Bodensohn, U. S., et al. (2019). "The intracellular distribution of the components of the GET system in vascular plants." Biochimica et Biophysica Acta (BBA)-Molecular Cell Research **1866**(10): 1650-1662.
- Borgese, N., et al. (2007). "How tails guide tail-anchored proteins to their destinations." Curr Opin Cell Biol **19**(4): 368-375.
- Borgese, N., et al. (2003). "The tale of tail-anchored proteins: coming from the cytosol and looking for a membrane." The Journal of cell biology **161**(6): 1013-1019.
- Borgese, N. and E. Fasana (2011). "Targeting pathways of C-tail-anchored proteins." Biochimica et Biophysica Acta (BBA)-Biomembranes **1808**(3): 937-946.

Borgese, N., et al. (2001). "Targeting of a tail-anchored protein to endoplasmic reticulum and mitochondrial outer membrane by independent but competing pathways." Molecular biology of the cell **12**(8): 2482-2496.

Bozkurt, G., et al. (2009). "Structural insights into tail-anchored protein binding and membrane insertion by Get3." Proceedings of the National Academy of Sciences **106**(50): 21131-21136.

Bram, R. J. and G. R. Crabtree (1994). "Calcium signalling in T cells stimulated by a cyclophilin B-binding protein." Nature **371**(6495): 355-358.

Bryda, E. C., et al. (2012). "Conditional deletion of calcium-modulating cyclophilin ligand causes deafness in mice." Mammalian genome **23**(3-4): 270-276.

Carvalho, H. J., et al. (2019). "The WRB Subunit of the Get3 Receptor is Required for the Correct Integration of its Partner CAML into the ER." Scientific reports **9**(1): 1-12.

Casson, J., et al. (2017). "Multiple pathways facilitate the biogenesis of mammalian tail-anchored proteins." Journal of Cell Science **130**(22): 3851-3861.

Chang, Y.-W., et al. (2010). "Crystal structure of Get4-Get5 complex and its interactions with Sgt2, Get3, and Ydj1." Journal of Biological Chemistry **285**(13): 9962-9970.

Chartron, J. W., et al. (2010). "Structural characterization of the Get4/Get5 complex and its interaction with Get3." Proceedings of the National Academy of Sciences **107**(27): 12127-12132.

Chitwood, P. J., et al. (2018). "EMC is required to initiate accurate membrane protein topogenesis." Cell **175**(6): 1507-1519. e1516.

Colombo, S. F., et al. (2016). "Tail-anchored protein insertion in mammals function and reciprocal interactions of the two subunits of the TRC40 receptor." Journal of Biological Chemistry **291**(29): 15292-15306.

Connolly, T. and R. Gilmore (1989). "The signal recognition particle receptor mediates the GTP-dependent displacement of SRP from the signal sequence of the nascent polypeptide." Cell **57**(4): 599-610.

Connolly, T., et al. (1991). "Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor." Science: 1171-1173.

Costa, E. A., et al. (2018). "Defining the physiological role of SRP in protein-targeting efficiency and specificity." Science **359**(6376): 689-692.

Cournia, Z., et al. (2015). "Membrane protein structure, function, and dynamics: a perspective from experiments and theory." The Journal of membrane biology **248**(4): 611-640.

Dhanao, P. K., et al. (2010). "Distinct pathways mediate the sorting of tail-anchored proteins to the plastid outer envelope." PloS one **5**(4).

Duncan, O., et al. (2013). "The outer mitochondrial membrane in higher plants." Trends in Plant Science **18**(4): 207-217.

F Colombo, S. and E. Fasana (2011). "Mechanisms of insertion of tail-anchored proteins into the membrane of the endoplasmic reticulum." Current Protein and Peptide Science **12**(8): 736-742.

Fasshauer, D., et al. (1998). "Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs." Proceedings of the National Academy of Sciences **95**(26): 15781-15786.

Fleischer, T. C., et al. (2006). "Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes." Genes & development **20**(10): 1294-1307.

Gilmore, R., et al. (1982). "Protein translocation across the endoplasmic reticulum. I. Detection in the microsomal membrane of a receptor for the signal recognition particle." The Journal of cell biology **95**(2): 463-469.

Gilmore, R., et al. (1982). "Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor." The Journal of cell biology **95**(2): 470-477.

Goder, V. and M. Spiess (2001). "Topogenesis of membrane proteins: determinants and dynamics." FEBS Lett **504**(3): 87-93.

Gristick, H. B., et al. (2014). "Crystal structure of ATP-bound Get3–Get4–Get5 complex reveals regulation of Get3 by Get4." Nature structural & molecular biology **21**(5): 437.

Guna, A., et al. (2018). "The ER membrane protein complex is a transmembrane domain insertase." Science **359**(6374): 470-473.

Halic, M., et al. (2004). "Structure of the signal recognition particle interacting with the elongation-arrested ribosome." Nature **427**(6977): 808-814.

Haßdenteufel, S., et al. (2017). "hSnd2 protein represents an alternative targeting factor to the endoplasmic reticulum in human cells." FEBS Lett **591**(20): 3211-3224.

Hessa, T., et al. (2011). "Protein targeting and degradation are coupled for elimination of mislocalized proteins." Nature **475**(7356): 394-397.

Higy, M., et al. (2004). "Topogenesis of membrane proteins at the endoplasmic reticulum." Biochemistry **43**(40): 12716-12722.

Hwang, Y. T., et al. (2004). "Novel targeting signals mediate the sorting of different isoforms of the tail-anchored membrane protein cytochrome b5 to either endoplasmic reticulum or mitochondria." Plant Cell **16**(11): 3002-3019.

Ichikawa, M., et al. (2014). "Syntaxin of plant proteins SYP123 and SYP132 mediate root hair tip growth in *Arabidopsis thaliana*." Plant and Cell Physiology **55**(4): 790-800.

Inglis, A. J., et al. (2020). "Differential modes of orphan subunit recognition for the WRB/CAML complex." Cell reports **30**(11): 3691-3698. e3695.

Jantti, J., et al. (1994). "Membrane insertion and intracellular transport of yeast syntaxin Sso2p in mammalian cells." Journal of Cell Science **107**(12): 3623-3633.

Jonikas, M. C., et al. (2009). "Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum." Science **323**(5922): 1693-1697.

Kanaji, S., et al. (2000). "Characterization of the signal that directs Tom20 to the mitochondrial outer membrane." The Journal of cell biology **151**(2): 277-288.

Kohl, C., et al. (2011). "Cooperative and independent activities of Sgt2 and Get5 in the targeting of tail-anchored proteins." Biological chemistry **392**(7): 601-608.

Kriechbaumer, V., et al. (2009). "Subcellular distribution of tail-anchored proteins in *Arabidopsis*." Traffic **10**(12): 1753-1764.

Kubota, K., et al. (2012). "Get1 stabilizes an open dimer conformation of get3 ATPase by binding two distinct interfaces." Journal of molecular biology **422**(3): 366-375.

Kutay, U., et al. (1995). "Transport route for synaptobrevin via a novel pathway of insertion into the endoplasmic reticulum membrane." The EMBO journal **14**(2): 217-223.

Kutay, U., et al. (1993). "A class of membrane proteins with a C-terminal anchor." Trends in cell biology **3**(3): 72-75.

Leznicki, P. and S. High (2012). "SGTA antagonizes BAG6-mediated protein triage." Proceedings of the National Academy of Sciences **109**(47): 19214-19219.

Leznicki, P., et al. (2013). "The association of BAG6 with SGTA and tail-anchored proteins." PloS one **8**(3).

Li, L., et al. (2019). "Mitochondrial AAA-ATPase Msp1 detects mislocalized tail-anchored proteins through a dual-recognition mechanism." EMBO reports **20**(4).

Linstedt, A., et al. (1995). "A C-terminally-anchored Golgi protein is inserted into the endoplasmic reticulum and then transported to the Golgi apparatus." Proceedings of the National Academy of Sciences **92**(11): 5102-5105.

Lipka, V., et al. (2007). "SNARE-ware: the role of SNARE-domain proteins in plant biology." Annu. Rev. Cell Dev. Biol. **23**: 147-174.

Lodish, H., et al. (2000). "Molecular cell biology 4th edition." National Center for Biotechnology Information, Bookshelf.

Manu, M., et al. (2018). "Analysis of tail-anchored protein translocation pathway in plants." Biochemistry and biophysics reports **14**: 161-167.

Mariappan, M., et al. (2010). "A ribosome-associating factor chaperones tail-anchored membrane proteins." Nature **466**(7310): 1120-1124.

Mariappan, M., et al. (2011). "The mechanism of membrane-associated steps in tail-anchored protein insertion." Nature **477**(7362): 61.

Mateja, A., et al. (2009). "The structural basis of tail-anchored membrane protein recognition by Get3." Nature **461**(7262): 361.

Meimaridou, E., et al. (2009). "From hatching to dispatching: the multiple cellular roles of the Hsp70 molecular chaperone machinery." Journal of molecular endocrinology **42**(1): 1-9.

Mock, J.-Y., et al. (2015). "Bag6 complex contains a minimal tail-anchor-targeting module and a mock BAG domain." Proceedings of the National Academy of Sciences **112**(1): 106-111.

Moog, D. (2019). "Higher complexity requires higher accuracy: tail-anchored protein targeting to the outer envelope membrane of plant plastids via a specific c-terminal motif." Plant and Cell Physiology **60**(3): 489-491.

Mukhopadhyay, R., et al. (2006). "Targeted disruption of the mouse *Asna1* gene results in embryonic lethality." FEBS Lett **580**(16): 3889-3894.

Nyathi, Y., et al. (2013). "Co-translational targeting and translocation of proteins to the endoplasmic reticulum." Biochimica et Biophysica Acta (BBA)-Molecular Cell Research **1833**(11): 2392-2402.

Okreglak, V. and P. Walter (2014). "The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins." Proceedings of the National Academy of Sciences **111**(22): 8019-8024.

Onishi, M., et al. (2018). "The ER membrane insertase Get1/2 is required for efficient mitophagy in yeast." Biochem Biophys Res Commun **503**(1): 14-20.

Opaliński, Ł., et al. (2014). "Clearing tail-anchored proteins from mitochondria." Proceedings of the National Academy of Sciences **111**(22): 7888-7889.

Park, E. and T. A. Rapoport (2012). "Mechanisms of Sec61/SecY-mediated protein translocation across membranes." Annual review of biophysics **41**: 21-40.

Rabu, C., et al. (2008). "A precursor-specific role for Hsp40/Hsc70 during tail-anchored protein integration at the endoplasmic reticulum." Journal of Biological Chemistry **283**(41): 27504-27513.

Rivera-Monroy, J., et al. (2016). "Mice lacking WRB reveal differential biogenesis requirements of tail-anchored proteins in vivo." Scientific reports **6**: 39464.

Rodrigo-Brenni, M. C., et al. (2014). "Cytosolic quality control of mislocalized proteins requires RNF126 recruitment to Bag6." Molecular cell **55**(2): 227-237.

Rome, M. E., et al. (2014). "Differential gradients of interaction affinities drive efficient targeting and recycling in the GET pathway." Proceedings of the National Academy of Sciences **111**(46): E4929-E4935.

Rome, M. E., et al. (2013). "Precise timing of ATPase activation drives targeting of tail-anchored proteins." Proceedings of the National Academy of Sciences **110**(19): 7666-7671.

Satoh, T., et al. (2015). "dPob/EMC is essential for biosynthesis of rhodopsin and other multi-pass membrane proteins in Drosophila photoreceptors." Elife **4**: e06306.

Schuldiner, M., et al. (2005). "Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile." Cell **123**(3): 507-519.

Schuldiner, M., et al. (2008). "The GET complex mediates insertion of tail-anchored proteins into the ER membrane." Cell **134**(4): 634-645.

Shan, S.-o., et al. (2007). "Conformational changes in the GTPase modules of the signal reception particle and its receptor drive initiation of protein translocation." The Journal of cell biology **178**(4): 611-620.

Shao, S. and R. S. Hegde (2011). "Membrane protein insertion at the endoplasmic reticulum." Annual review of cell and developmental biology **27**: 25-56.

Shurtleff, M. J., et al. (2018). "The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins." Elife **7**: e37018.

Söllner, T., et al. (1993). "SNAP receptors implicated in vesicle targeting and fusion." Nature **362**(6418): 318-324.

Spieß, M. (1995). "Heads or tails—what determines the orientation of proteins in the membrane." FEBS Lett **369**(1): 76-79.

Srivastava, R., et al. (2017). "The GET system inserts the tail-anchored protein, SYP72, into endoplasmic reticulum membranes." Plant physiology **173**(2): 1137-1145.

Stefanovic, S. and R. S. Hegde (2007). "Identification of a targeting factor for posttranslational membrane protein insertion into the ER." Cell **128**(6): 1147-1159.

Stefer, S., et al. (2011). "Structural basis for tail-anchored membrane protein biogenesis by the Get3-receptor complex." Science **333**(6043): 758-762.

Tran, D. D., et al. (2003). "CAML is required for efficient EGF receptor recycling." Developmental cell **5**(2): 245-256.

Uemura, T., et al. (2004). "Systematic analysis of SNARE molecules in Arabidopsis: dissection of the post-Golgi network in plant cells." Cell structure and function **29**(2): 49-65.

Vilardi, F., et al. (2011). "WRB is the receptor for TRC40/Asna1-mediated insertion of tail-anchored proteins into the ER membrane." Journal of Cell Science **124**(8): 1301-1307.

Vilardi, F., et al. (2014). "WRB and CAML are necessary and sufficient to mediate tail-anchored protein targeting to the ER membrane." PloS one **9**(1).

Vogl, C., et al. (2016). "Tryptophan-rich basic protein (WRB) mediates insertion of the tail-anchored protein otoferlin and is required for hair cell exocytosis and hearing." The EMBO journal **35**(23): 2536-2552.

Volkmar, N., et al. (2019). "The ER membrane protein complex promotes biogenesis of sterol-related enzymes maintaining cholesterol homeostasis." Journal of Cell Science **132**(2).

Walter, P. and G. Blobel (1981). "Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes." The Journal of cell biology **91**(2): 557-561.

Wang, F., et al. (2010). "A chaperone cascade sorts proteins for posttranslational membrane insertion into the endoplasmic reticulum." Molecular cell **40**(1): 159-171.

Wang, F., et al. (2014). "The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase." Nature **512**(7515): 441.

Wang, F., et al. (2011). "The mechanism of tail-anchored protein insertion into the ER membrane." Molecular cell **43**(5): 738-750.

Wattenberg, B. and T. Lithgow (2001). "Targeting of C-terminal (tail)-anchored proteins: understanding how cytoplasmic activities are anchored to intracellular membranes." Traffic **2**(1): 66-71.

Wattenberg, B. W., et al. (2007). "An artificial mitochondrial tail signal/anchor sequence confirms a requirement for moderate hydrophobicity for targeting." Bioscience reports **27**(6): 385-401.

Weimbs, T., et al. (1997). "A conserved domain is present in different families of vesicular fusion proteins: a new superfamily." Proceedings of the National Academy of Sciences **94**(7): 3046-3051.

Wereszczynski, J. and J. A. McCammon (2012). "Nucleotide-dependent mechanism of Get3 as elucidated from free energy calculations." Proceedings of the National Academy of Sciences **109**(20): 7759-7764.

Wideman, J. G. (2015). "The ubiquitous and ancient ER membrane protein complex (EMC): tether or not?" F1000Research **4**.

Wohlever, M. L., et al. (2017). "Msp1 is a membrane protein dislocase for tail-anchored proteins." Molecular cell **67**(2): 194-202. e196.

Xing, S., et al. (2017). "Loss of GET pathway orthologs in *Arabidopsis thaliana* causes root hair growth defects and affects SNARE abundance." Proceedings of the National Academy of Sciences **114**(8): E1544-E1553.

Yamamoto, Y. and T. Sakisaka (2012). "Molecular machinery for insertion of tail-anchored membrane proteins into the endoplasmic reticulum membrane in mammalian cells." Molecular cell **48**(3): 387-397.

Yeh, Y.-H., et al. (2014). "Structural and functional characterization of Ybr137wp implicates its involvement in the targeting of tail-anchored proteins to membranes." Molecular and cellular biology **34**(24): 4500-4512.

Zalisko, B. E., et al. (2017). "Tail-anchored protein insertion by a single Get1/2 heterodimer." Cell reports **20**(10): 2287-2293.

Zhuang, X., et al. (2017). "Targeting tail-anchored proteins into plant organelles." Proceedings of the National Academy of Sciences **114**(8): 1762-1764.

## **APPENDICES**

### **I. Xing et al., 2017**

# Loss of GET pathway orthologs in *Arabidopsis thaliana* causes root hair growth defects and affects SNARE abundance

Shuping Xing<sup>a</sup>, Dietmar Gerald Mehlhorn<sup>a</sup>, Niklas Wallmeroth<sup>a</sup>, Lisa Yasmin Asseck<sup>a</sup>, Ritwika Kar<sup>a</sup>, Alessa Voss<sup>a</sup>, Philipp Denninger<sup>b</sup>, Vanessa Aphaia Fiona Schmidt<sup>b</sup>, Markus Schwarzländer<sup>c</sup>, York-Dieter Stierhof<sup>d</sup>, Guido Grossmann<sup>b</sup>, and Christopher Grefen<sup>a,1</sup>

<sup>a</sup>Centre for Plant Molecular Biology, Developmental Genetics, University of Tübingen, 72076 Tuebingen, Germany; <sup>b</sup>Centre for Organismal Studies, CellNetworks Excellence Cluster, University of Heidelberg, 69120 Heidelberg, Germany; <sup>c</sup>Institute of Crop Science and Resource Conservation, University of Bonn, 53113 Bonn, Germany; and <sup>d</sup>Centre for Plant Molecular Biology, Microscopy, University of Tübingen, 72076 Tuebingen, Germany

Edited by Natasha V. Raikhel, Center for Plant Cell Biology, Riverside, CA, and approved December 19, 2016 (received for review November 28, 2016)

**Soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins are key players in cellular trafficking and coordinate vital cellular processes, such as cytokinesis, pathogen defense, and ion transport regulation. With few exceptions, SNAREs are tail-anchored (TA) proteins, bearing a C-terminal hydrophobic domain that is essential for their membrane integration. Recently, the Guided Entry of Tail-anchored proteins (GET) pathway was described in mammalian and yeast cells that serve as a blueprint of TA protein insertion [Schuldiner M, et al. (2008) *Cell* 134(4):634–645; Stefanovic S, Hegde RS (2007) *Cell* 128(6):1147–1159]. This pathway consists of six proteins, with the cytosolic ATPase GET3 chaperoning the newly synthesized TA protein posttranslationally from the ribosome to the endoplasmic reticulum (ER) membrane. Structural and biochemical insights confirmed the potential of pathway components to facilitate membrane insertion, but the physiological significance in multicellular organisms remains to be resolved. Our phylogenetic analysis of 37 GET3 orthologs from 18 different species revealed the presence of two different GET3 clades. We identified and analyzed GET pathway components in *Arabidopsis thaliana* and found reduced root hair elongation in *Atget* lines, possibly as a result of reduced SNARE biogenesis. Overexpression of *AtGET3a* in a receptor knockout (KO) results in severe growth defects, suggesting presence of alternative insertion pathways while highlighting an intricate involvement for the GET pathway in cellular homeostasis of plants.**

GET pathway | TA proteins | SNAREs | ER membrane | root hairs

**P**lants show remarkable acclimation and resilience to a broad spectrum of environmental influences as a consequence of their sedentary lifestyle. On the cellular level, such flexibility requires genetic buffering capacity as well as fine-tuned signaling and response systems. Soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins make a critical contribution toward acclimation (1, 2). Their canonical function facilitates membrane fusion through tight interaction of cognate SNARE partners at vesicle and target membranes (3). This vital process guarantees cellular expansion through addition of membrane material, cell plate formation, and cargo delivery (4, 5). SNARE proteins are also involved in regulating potassium channels and aquaporins (6–8).

Most SNARE proteins are Type II oriented and referred to as tail-anchored (TA) proteins with a cytosolic N terminus and a single C-terminal transmembrane domain (TMD) (9). TA proteins are involved in vital cellular processes in all domains of life, such as chaperoning, ubiquitination, signaling, trafficking, and transcript regulation (10–13). The nascent protein is almost fully translated when the hydrophobic TMD emerges from the ribosome, requiring shielding from the aqueous cytosol to guarantee protein stability, efficient folding, and function (14). One way of facilitating this posttranslational insertion is by proteinaceous components of a Guided Entry of Tail-anchored proteins (GET) pathway that was identified in yeast and mammals (15, 16).

In yeast, recognition of nascent TA proteins is accomplished through a tripartite pretargeting complex at the ribosome consisting of SGT2, GET5, and GET4. This complex binds to the TMD and delivers the TA protein to the cytosolic ATPase GET3 (17, 18). GET3 arranges as zinc-coordinating homodimer and shuttles the client protein to the endoplasmic reticulum (ER) membrane receptors GET1 and GET2, which finalize insertion of the TA protein (15, 19, 20).

This GET pathway is thought to be the main route for TA protein insertion into the ER, but surprisingly, its loss in yeast is only conditionally lethal (15). Conversely, lack of the mammalian GET3 orthologs TRC40 (transmembrane domain recognition complex of 40 kDa) leads to embryo lethality in mice, complicating global physiological analyses (21). Nevertheless, a handful of recent studies have started to analyze individual physiological consequences of the GET pathway *in vivo* using tissue-specific knockout (KO) approaches and observed that its function is required for a diverse range of physiological processes, such as insulin secretion, auditory perception, and photoreceptor function, in animals (22–24). A high degree of evolutionary conservation is often assumed, and it has been recognized that some components of the GET pathway are present in *Arabidopsis thaliana* (25, 26). However, considering the specific physiological roles of the GET pathway observed in yeast and mammals, its significance cannot

## Significance

**Root hairs are unicellular extensions of the rhizodermis, providing anchorage and an increase in surface area for nutrient and water uptake. Their fast, tip-focused growth showcases root hairs as an excellent genetic model to study physiological and developmental processes on the cellular level. We uncovered a root hair phenotype that is dependent on putative *Arabidopsis* orthologs of the Guided Entry of Tail-anchored (TA) proteins (GET) pathway, which facilitates membrane insertion of TA proteins in yeast and mammals. We found that plants have evolved multiple paralogs of specific GET pathway components, albeit in a compartment-specific manner. In addition, we show that differential expression of pathway components causes pleiotropic growth defects, suggesting alternative pathways for TA insertion and additional functions of GET in plants.**

Author contributions: S.X. and C.G. designed research; S.X., D.G.M., N.W., L.Y.A., R.K., A.V., P.D., V.A.F.S., M.S., Y.-D.S., G.G., and C.G. performed research; S.X., D.G.M., M.S., Y.-D.S., G.G., and C.G. analyzed data; and C.G. wrote the paper with input from M.S.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

<sup>1</sup>To whom correspondence should be addressed. Email: christopher.grefen@zmbp.uni-tuebingen.de.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1619525114/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1619525114/-DCSupplemental).

be straightforwardly extrapolated across eukaryotes. A global genetic dissection of the pathway in a multicellular organism, let alone in plants, is currently lacking.

GET3/TRC40 are distant paralogs of the prokaryotic ArsA (arsenite-translocating ATPase), a protein that is part of the arsenic detoxification pathway in bacteria (27). Evidence points toward the GET pathway—albeit at a simpler scale—that exists already in Archaeobacteria (10, 28). Because yeast and mammals are closely related in the supergroup of Opisthokonts (29), limiting any comparative power, we aimed to investigate pathway conservation in other eukaryotes. We also wanted to understand the impact that lack of GET pathway function has on plant development, considering that it started entering the textbooks as a default route for TA protein insertion.

Our results show that loss of GET pathway function in *A. thaliana* impacts on root hair length. This phenotype coincides with reduced protein levels at the plasma membrane of an im-

portant root hair-specific SNARE, conforming to the role of the GET pathway in TA protein insertion. However, similarly to yeast, no global pleiotropic phenotypes were observed, pointing to the existence of functional backup. However, ectopic over-expression of the cytosolic ATPase *AtGET3a* in the putative receptor KO *Atget1* leads to severe growth defects, underscoring pathway conservation while implying an intricate role of the GET pathway in cellular homeostasis of plants.

## Results

**GET3 Paralogs Might Have Evolved as Early as Archaea.** To identify potential orthologs of GET candidates, we used in silico sequence comparison (BLASTp and National Center for Biotechnology Information) of yeast and human GET proteins against the proteome of 16 different species from 13 phyla (Tables 1 and 2). Candidate sequences were assembled in a phylogenetic tree that, surprisingly, reveals that two distinct GET3 clades, which we

**Table 1. Accession numbers of GET3/TRC40/ArsA orthologs of clade a used for the phylogenetic tree in Fig. 1 and their putative GET1/WRB and GET4/TRC40 orthologs identified via BLASTp search**

| Phylum and species                | GET3/TRC40 orthologs         |             | Up-/downstream orthologs |                              |
|-----------------------------------|------------------------------|-------------|--------------------------|------------------------------|
|                                   | Accession no.                | Length (aa) | GET1/WRB                 | GET4/TRC35                   |
| <b>Eubacteria</b>                 |                              |             |                          |                              |
| Proteobacteria                    |                              |             |                          |                              |
| <i>Escherichia coli</i>           | KZO75668                     | 583*        | Not found                | Not found                    |
| Proteoarchaeota                   |                              |             |                          |                              |
| Lokiarchaeota                     |                              |             |                          |                              |
| <i>Lokiarchaeum</i> sp.           | KKK44956                     | 338         | Not found                | Not found                    |
| <b>Opisthokonta</b>               |                              |             |                          |                              |
| Chordata                          |                              |             |                          |                              |
| <i>Homo sapiens</i>               | NP_004308                    | 348         | NP_004618                | NP_057033                    |
| Ascomycota                        |                              |             |                          |                              |
| <i>Saccharomyces cerevisiae</i>   | AAT93183                     | 354         | NP_011495                | NP_014807                    |
| <b>Amoebozoa</b>                  |                              |             |                          |                              |
| Discosea                          |                              |             |                          |                              |
| <i>Acanthamoeba castellanii</i>   | XP_004368068                 | 330         | XP_004353131             | XP_004367722                 |
| Mycetozoa                         |                              |             |                          |                              |
| <i>Dictyostelium purpureum</i>    | XP_003289495                 | 330         | Not found                | XP_003283186                 |
| <b>Archaeplastida</b>             |                              |             |                          |                              |
| Angiospermae                      |                              |             |                          |                              |
| <i>Arabidopsis thaliana</i>       | NP_563640                    | 353         | NP_567498                | NP_201127                    |
| <i>Medicago truncatula</i>        | XP_013444959                 | 358         | XP_003629131             | XP_003591984                 |
| <i>Brachypodium distachyon</i>    | XP_003578462                 | 363         | XP_003564144             | XP_003569076                 |
| <i>Amborella trichopoda</i>       | XP_006857946                 | 353         | XP_006855737             | ERM96291                     |
| Lycopodiophyta                    |                              |             |                          |                              |
| <i>Selaginella moellendorffii</i> | XP_002973461                 | 360         | Not found                | XP_002969945<br>XP_002981415 |
| Marchantiophyta                   |                              |             |                          |                              |
| <i>Marchantia polymorpha</i>      | OAE26618                     | 370         | OAE20217                 | OAE20690                     |
| Bryophyta                         |                              |             |                          |                              |
| <i>Physcomitrella patens</i>      | XP_001758936<br>XP_001774198 | 365<br>365  | XP_001760426             | XP_001760372<br>XP_001758146 |
| Chlorophyta                       |                              |             |                          |                              |
| <i>Chlamydomonas reinhardtii</i>  | XP_001693332                 | 319         | XP_001695038             | XP_001695333                 |
| Rhodophyta                        |                              |             |                          |                              |
| <i>Galdieria sulphuraria</i>      | XP_005708637                 | 706*        | XP_005707118             | XP_005704684                 |
| <b>SAR</b>                        |                              |             |                          |                              |
| Chromerida                        |                              |             |                          |                              |
| <i>Vitrella brassicaformis</i>    | CEM03518                     | 412         | Not found                | CEL97893                     |
| Heterokontophyta                  |                              |             |                          |                              |
| <i>Nannochloropsis gaditana</i>   | EWM27451                     | 370         | EWM21897                 | EWM27335                     |
| Chromalveolata                    |                              |             |                          |                              |
| Cryptophyta                       |                              |             |                          |                              |
| <i>Guillardia theta</i>           | XP_005837457                 | 310         | XP_005829401             | XP_005841994                 |

\*Tandem GET3.

termed GET3a and GET3bc, respectively, exist in Archaeplastida and SAR (supergroup of stramenopiles, alveolates, and Rhizaria) but do not exist in Opisthokonts (yeast and animals) and Amoebozoa. The deep branching of the tree implies that duplication events must have occurred early in the evolution of eukaryotes (Fig. 1A). Interestingly, the recently identified phylum of *Lokiarchaeota*, which is thought to form a monophyletic group with eukaryotes (30), expresses two distinct GET3 orthologs, one of which aligns within the GET3bc clade while lacking some of the important sequence features of eukaryotic GET3 (Fig. S1A). This observation suggests that the last eukaryotic common ancestor had already acquired two copies of GET3.

In Rhodophytes and higher Angiospermae, a third GET3bc paralog branched off. Interestingly, the tandem ATPase motif—likely a consequence of gene duplication in the prokaryotic ArsA and suggested to be a key difference between ArsA and GET3/TRC40 homologs (28)—is not found in either of two *Lokiarchaeota* GET3; conversely, in Rhodophytes and SAR species, GET3 paralogs exist that contain duplications (Tables 1 and 2). Importantly, such repeats are not restricted to the GET3bc clade but also, are found among red algae GET3a orthologs (e.g., XP\_005708637). Comparing sequence conservation of GET3

orthologs reveals that residues important for ATPase function are maintained in all candidates (Fig. S1A and B). However, the sites for GET1 binding and the methionine-rich GET3 motif (31, 32) are only conserved in GET3a candidates of eukaryotes, concurring with the presence of GET1 and GET4 orthologs in most of these species (Table 1).

Strikingly, in silico analysis of the N termini of the identified GET3 orthologs predicts for almost all GET3bc—but not for GET3a candidates—the presence of a transit peptide for mitochondrial or chloroplastic import (Table 2). This observation is also in line with the fact that GET3bc proteins are, on average, larger than their GET3a paralogs (Tables 1 and 2), matching the length range of targeting sequences for the bioenergetic organelles.

#### Distinct Differences in Subcellular Localization of AtGET3 Paralogs.

The three GET3 paralogs of *A. thaliana* were in silico-predicted to localize to the cytosol (*AtGET3a*; At1g01910), chloroplast (*AtGET3b*; At3g10350), and mitochondria (*AtGET3c*; At5g60730), respectively (Tables 1 and 2). To corroborate these predictions, stably transformed, *A. thaliana* *Ubiquitin10* promoter (*P<sub>UBQ10</sub>*)-driven GFP fusions were generated (33). Confocal laser scanning microscopy (CLSM) and transmission electron microscopy (TEM)

**Table 2. Accession numbers of GET3/TRC40/ArsA orthologs of clade bc used for the phylogenetic tree in Fig. 1 and their in silico prediction of an N-terminal signal/transit peptide using three different prediction tools (TargetP 1.1, ChloroP 1.1, and Predotar v1.03)**

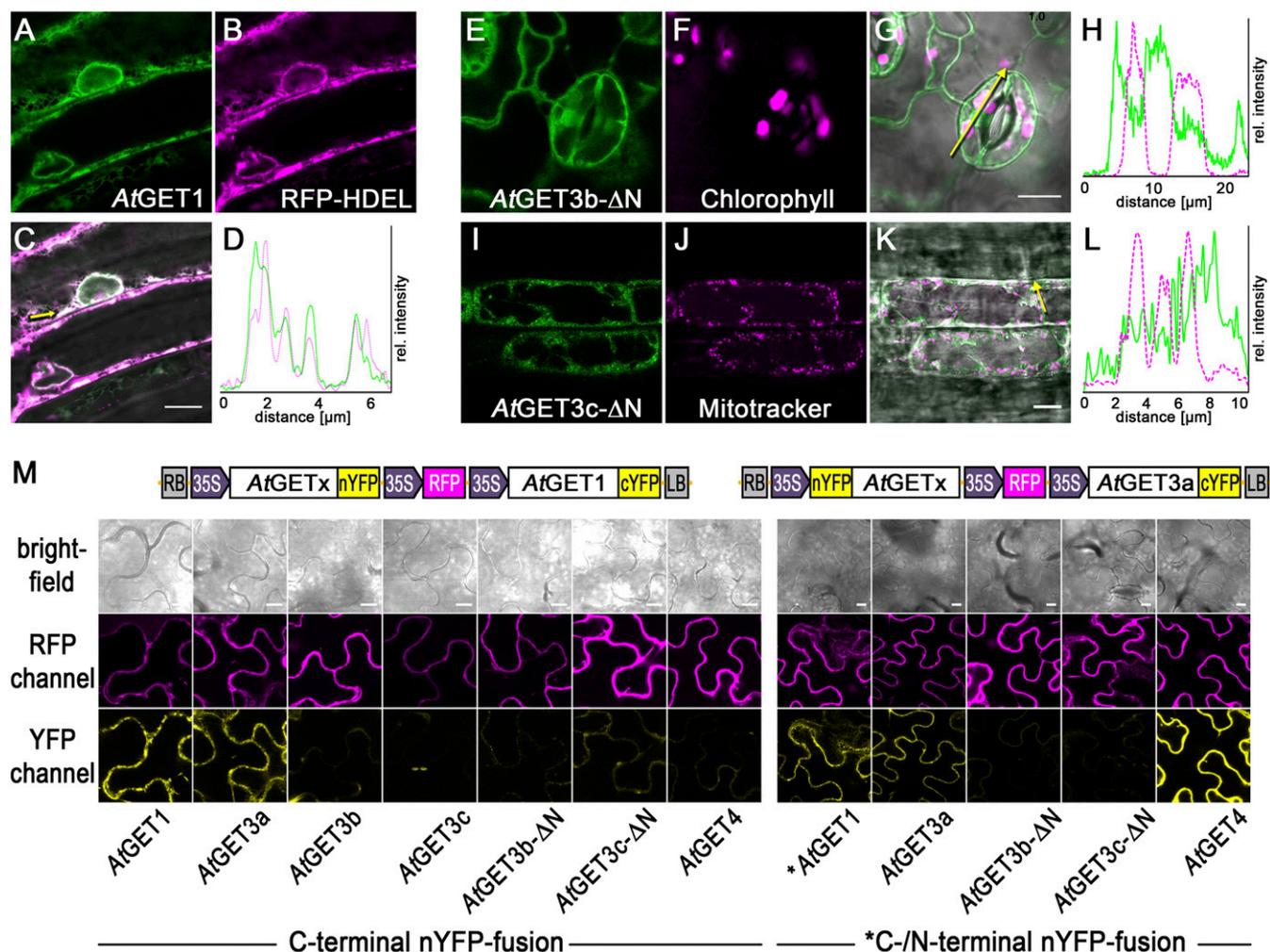
| Phylum and species                | GET3/TRC40 orthologs |                  | Signal/transit peptide prediction |                       |                |
|-----------------------------------|----------------------|------------------|-----------------------------------|-----------------------|----------------|
|                                   | Accession no.        | Length (aa)      | TargetP 1.1                       | ChloroP 1.1           | Predotar v1.03 |
| <b>Eubacteria</b>                 |                      |                  |                                   |                       |                |
| Proteobacteria                    |                      |                  |                                   |                       |                |
| <i>Escherichia coli</i>           | KZO75668             | 583*             |                                   | Non-Eukaryote         |                |
| Proteoarchaeota                   |                      |                  |                                   |                       |                |
| Lokiarchaeota                     |                      |                  |                                   |                       |                |
| <i>Lokiarchaeum</i> sp.           | KKK42590             | 329              |                                   | Non-Eukaryote         |                |
| <b>Archaeplastida</b>             |                      |                  |                                   |                       |                |
| Angiospermae                      |                      |                  |                                   |                       |                |
| <i>A. thaliana</i>                |                      |                  |                                   |                       |                |
|                                   | NP_187646            | 433              | C                                 | C                     | C              |
|                                   | NP_200881            | 391              | M                                 | C                     | M              |
| <i>Medicago truncatula</i>        | XP_003591867         | 406              | C                                 | C                     | Possibly C     |
|                                   | XP_013455984         | 381              | C                                 | C                     | C              |
| <i>Brachypodium distachyon</i>    | XP_003570659         | 403              | M                                 | C                     | M              |
|                                   | XP_010239988         | 371              | M                                 | —                     | M              |
| <i>Amborella trichopoda</i>       | XP_006827440         | 407              | C                                 | C                     | C              |
| Lycopodiophyta                    |                      |                  |                                   |                       |                |
| <i>Selaginella moellendorffii</i> | XP_002974288         | 432              | C                                 | C                     | Possibly M     |
| Marchantiophyta                   |                      |                  |                                   |                       |                |
| <i>Marchantia polymorpha</i>      | OAE21403             | 432              | C                                 | —                     | C              |
| Bryophyta                         |                      |                  |                                   |                       |                |
| <i>Physcomitrella patens</i>      | XP_001781368         | 331              | M                                 | C                     | Possibly M     |
|                                   | XP_001764873         | 359              |                                   | N terminus incomplete |                |
| Chlorophyta                       |                      |                  |                                   |                       |                |
| <i>Chlamydomonas reinhardtii</i>  | XP_001702275         | 513 <sup>†</sup> | M                                 | C                     | C              |
| Rhodophyta                        |                      |                  |                                   |                       |                |
| <i>Galdieria sulphuraria</i>      | XP_005705663         | 481              | —                                 | —                     | Possibly ER    |
|                                   | XP_005703923         | 757*             | M                                 | C                     | Possibly C     |
| <b>SAR</b>                        |                      |                  |                                   |                       |                |
| Heterokontophyta                  |                      |                  |                                   |                       |                |
| <i>Nannochloropsis gaditana</i>   | EWM30283             | 817*             | M                                 | —                     | Possibly C     |
| Chromerida                        |                      |                  |                                   |                       |                |
| <i>Vitrella brassicaformis</i>    | CEM11669             | 809*             | M                                 | —                     | Possibly ER    |
| <b>Chromalveolata</b>             |                      |                  |                                   |                       |                |
| Cryptophyta                       |                      |                  |                                   |                       |                |
| <i>Guillardia theta</i>           | XP_005822752         | 418              | S                                 | C                     | ER             |

C, chloroplast; M, mitochondrion; S, signal peptide.

\*Tandem GET3.

<sup>†</sup>Second P-loop motif at C terminus of protein.





**Fig. 2.** Interaction analysis among *A. thaliana* GET pathway orthologs. (A–D) At4g16444, the putative AtGET1, C-terminally tagged with GFP in stably transformed *A. thaliana* coexpressing the ER marker RFP-HDEL. (D) Line histograms along yellow arrows in C confirm colocalization. (E–L) CLSM analysis of N-terminally truncated AtGET3b and AtGET3c candidates. Counterimaging using autofluorescence of (F) chlorophyll or (J) MitoTracker Orange allows (H and L) line histograms in (G and K) merged images along yellow arrows that corroborate cytosolic retention. (M) Exemplary confocal images of rBiFC analysis of (Left) AtGET1 and (Right) AtGET3a with GET pathway orthologs and truncated constructs. Boxed cartoons show construct design above exemplary images of transiently transformed *Nicotiana benthamiana* leaves. A statistical analysis of the data is in Fig. S3. (Scale bars: 10  $\mu$ m.)

might prevent organellar import and cause their cytosolic retention, an interaction with AtGET1 cannot be observed.

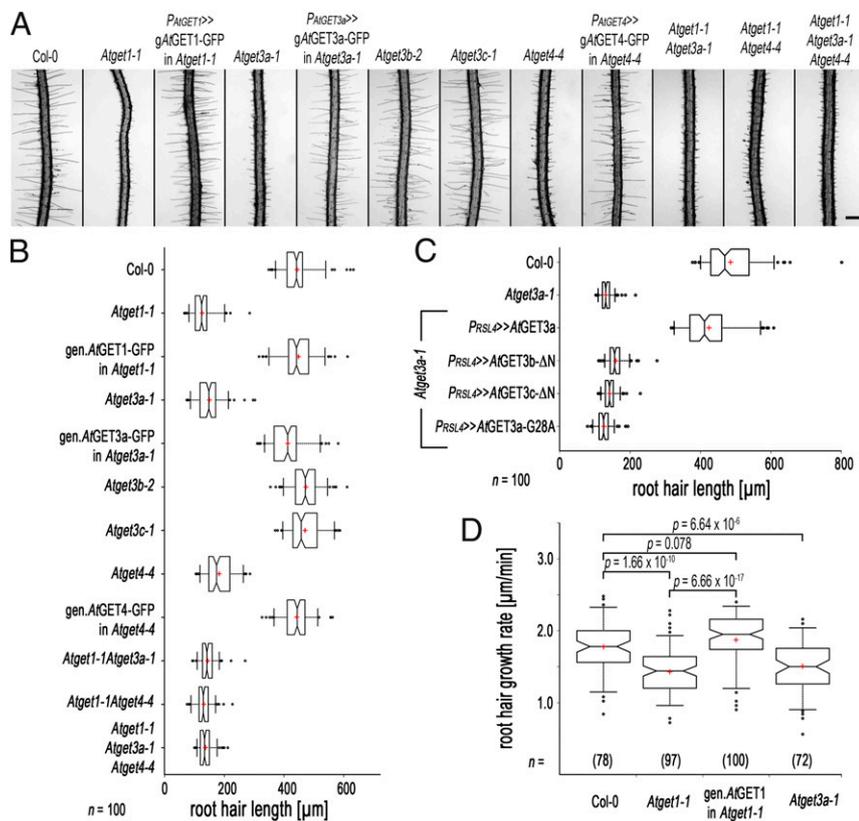
To understand whether the physical separation of AtGET3b/c prevents interaction with AtGET1, we truncated the first 68 aa of AtGET3b and 50 aa of AtGET3c, which lead to their cytosolic localization (Fig. 2 E–L). We applied ratiometric bimolecular fluorescence complementation (rBiFC) (37) to assess whether such artificial mislocalization renders AtGET3b/c susceptible to interaction with AtGET1. Clearly, AtGET1 homodimerizes and interacts with the cytosolic AtGET3a but does not homodimerize or interact with the plastidic AtGET3 paralogs or their transit peptide deletion versions (Fig. 2M and Fig. S3 A and B), confirming that a change in localization does not alter binding behavior. This absence of interaction seems consistent with the lack of a GET1-binding motif (32, 38) in the sequences of AtGET3b/c, further indicating that these likely lack functional redundancy with AtGET3a.

To test this hypothesis before phenotypic complementation, we assessed heterodimerization with AtGET3a. Here, we also included the putative upstream binding partner of AtGET3a, AtGET4 (At5g63220), which we identified through in silico analysis. The expression pattern of AtGET4 resembles that of AtGET3a (Fig. S3E), and the protein localizes to the cytosol (see Fig. S7B). rBiFC analysis substantiates that AtGET3a interacts

with AtGET1, itself, and AtGET4 but fails to heterodimerize with AtGET3b/c. Both proteins were expressed in their truncated, cytosolic form; hence, the lack of interaction cannot be attributed to compartmentalization (Fig. 2M and Fig. S3 A and B). Because dimerization of ScGET3 is a prerequisite for function (31), this result also negates functional redundancy between GET3 paralogs.

**Functional Analyses of *A. thaliana* GET Orthologs.** Loss of function of TRC40, the GET3 ortholog in mammals, causes embryonic lethality befitting of the vital function of TA protein insertion (21). How would loss of GET pathway orthologs impact on survival, growth, and development in plants?

Unexpectedly, multiple different alleles of T-DNA (transfer DNA) insertion lines of each of the five AtGET orthologs identified (Fig. S4 A and B) did not reveal any obvious growth defects. Seeds germinated, and seedlings developed indistinguishable from wild-type (WT) plants. However, a more detailed phenotypic inspection revealed that seedlings of *Atget1*, *Atget3a*, and *Atget4* lines had significantly shorter root hairs compared with Columbia-0 (Col-0) WT plants, whereas *Atget3b* and *Atget3c* did not (Fig. 3 A and B and Fig. S4C). Expressing genomic versions of the GET genes restores near WT-like root hair growth. By contrast, a point



**Fig. 3.** Loss of function of some *A. thaliana* GET orthologs causes root hair growth defects. (A) Exemplary images of root elongation zones of 10-d-old T-DNA insertion lines of *A. thaliana* GET orthologs and genomic complementation. *Atget1-1*, *Atget3a-1*, and *Atget4-4* but not *Atget3b-2* and *Atget3c-1* lines show reduced growth of root hairs compared with WT Col-0 and can be complemented by their respective genomic constructs. Double or triple KO phenocopy single T-DNA insertion lines. Transcript analysis and additional alleles can be found in Fig. S4. (B) Boxplot depicting length of the 10 longest root hairs of 10 individual roots ( $n = 100$ ). Center lines of boxes represent median with outer limits at 25th and 75th percentiles. Notches indicate 95% confidence intervals; Altman whiskers extend to 5th and 95th percentiles, outliers are depicted as black dots, and red crosses mark sample means. (Scale bars: 500 μm.) (C) Boxplot as before, showing root hair length of Col-0 and *Atget3a-1* and complementation thereof using a root hair-specific promoter (RSL4; At1g27740) and N-terminally 3xHA-tagged coding sequences of *AtGET3a*, *AtGET3b-ΔN*, *AtGET3c-ΔN*, and *AtGET3a-G28A*. (D) Boxplot as before, showing root hair growth rates of exemplary T-DNA insertion lines and complemented *Atget1-1* line in micrometers per minute.

mutant of the P loop of the ATPase motif (*AtGET3a-G28A*) expressed under a root hair-specific promoter (*RSL4*) (39) prevents rescue in *Atget3a*, suggesting that ATPase activity of *AtGET3a* is essential for normal root hair growth (Fig. 3C). To substantiate our analysis of the *AtGET3b/c* paralogs, we expressed the transit peptide deletion variants in the *Atget3a* background. The mislocalized *AtGET3b/c* constructs failed to rescue the growth defects, suggesting evolution of alternative functions in the bioenergetic organelles (Fig. 3C).

Multiple crosses between individual T-DNA insertion lines of *AtGET1*, *AtGET3a*, and *AtGET4* did not yield an enhanced phenotype (i.e., further reduction of root hair length compared with their corresponding parental single-KO lines) (Fig. 3A and B), indicating interdependent functionality of all three proteins within a joint pathway. A more detailed kinetic analysis on roots grown in RootChips (40) revealed that the shorter overall root hair length in *Atget1* and *Atget3a* correlates with slowed down growth speed (Fig. 3D).

Root hairs together with pollen tubes are the fastest growing cells in plants and rely on efficient delivery of membrane material to the tip (41). Although we had not observed aberrant segregation ratios of T-DNA insertion lines, which could indicate compromised fertility, we analyzed pollen tube growth in vivo and in vitro but found growth speed as well as final length unaffected in the GET pathway mutants (Fig. S4D and E).

The genetic evidence for function of *AtGET1* and *AtGET3a* in a joint pathway allowing effective root hair growth in *A. thaliana* prompted us to assess their functional conservation. In yeast, *ScGET1* and *ScGET3* are not essential; however, their absence leads to lethality under a range of different abiotic stress conditions (15). We, therefore, tested *A. thaliana* GET orthologs in BY4741 WT and corresponding KO strains for their ability to rescue yeast survival under restrictive conditions. *AtGET1* (Fig. S5A) and to a much lesser extent, *AtGET3a* (Fig. S5B) hardly rescue growth in corresponding KOs, and all other *AtGET3* orthologs—full length or truncated—failed to rescue at all. This result provides strong evidence that the functions of *AtGET1* and

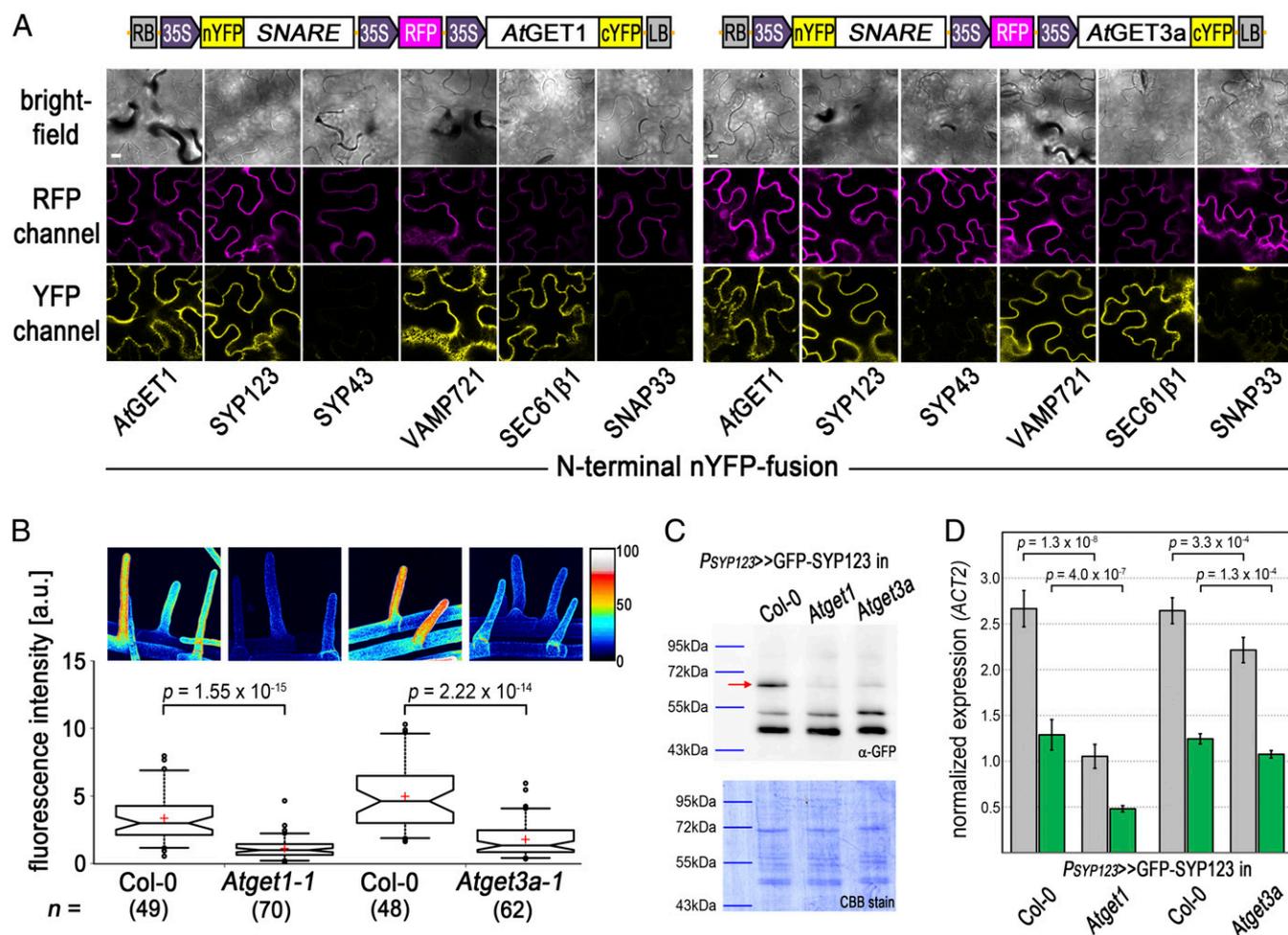
*AtGET3a* may have diverged from yeast, more strongly so for *AtGET3a*.

#### Loss of the GET Pathway Leads to Reduced Protein Levels of SYP123 in Root Hairs.

We compared the predicted “TA-proteome” of *A. thaliana* (13) with the list of interaction partners of *AtGET3a-GFP* from CoIP-MS analysis (Dataset S1). Only 23 TA proteins were detected that coprecipitated with *AtGET3a-GFP* but not GFP alone (Fig. S6B). However, in SUS and rBiFC analysis, *AtGET3a* interacts with a number of candidate TA proteins that we did not find in our CoIP-MS. Among others, the SNARE syntaxin of plants 123 (SYP123) as well as its R-SNARE partner VAMP721 and the TA protein SEC61β, subunit of the SEC61 translocon, interact with both *AtGET1* and *AtGET3a* (Fig. 4A and Fig. S6A and C). The SNARE SYP43 as well as the non-TA SNARE protein SNAP33 failed to interact. SYP123 is a plasma membrane-localized Qa-SNARE that specifically expresses in root hair cells, and its loss results in short root hairs (42). We crossed GFP-SYP123 under its own promoter (42) with our *Atget1-1* and *Atget3a-1* lines to analyze for misinsertion, mislocalization, or cytosolic retention.

CLSM analysis of root hairs expressing SYP123 in WT and mutant backgrounds showed normal distribution of SYP123 in bulge formation and developed root hairs (Fig. S7A). No cytosolic aggregates or increased fluorescence foci were visible in the cytoplasm, which was reminiscent of findings in yeast *get* pathway KOs (15, 43). However, we repeatedly observed differences in GFP signal under identical conditions and settings. GFP fluorescence intensity of root hairs is consistently stronger in the WT than in *Atget1* and *Atget3a* lines (Fig. 4B), suggestive of lower SYP123 protein levels in the plasma membrane of *Atget* lines.

To substantiate this finding, we performed membrane fractionation of protein extracts from roughly 250 roots per line (Fig. 4C). Immunoblot analysis revealed that GFP-SYP123 levels in the membrane fraction of *Atget1* and *Atget3a* lines were strikingly lower than in WT background, suggesting that loss of GET pathway functionality reduces SYP123 abundance



**Fig. 4.** The root hair-specific Qa-SNARE SYP123 shows reduced protein levels in *Atget* lines. (A) rBiFC analysis of (Left) AtGET1 and (Right) AtGET3a with candidate SNARE/TA proteins. Boxed cartoons show construct design above representative images of epidermal cells from transiently transformed *N. benthamiana* leaves. The statistical analysis of the data is presented in Fig. S6C. (Scale bars: 10  $\mu$ m.) (B and C) Analysis of root hairs expressing *P*<sub>SYP123</sub>>>GFP-SYP123 in *Atget1-1*, *Atget3a-1*, or corresponding Col-0 WT. (B) Boxplot of root hair fluorescence intensities of average-intensity z projections (number in parentheses below the x axis). Boxplot as in Fig. 3; *P* values confirm a significant difference in fluorescence intensity between GFP-SYP123 expression in WT (stronger) vs. T-DNA insertion lines (weaker). Heat maps of exemplary z projections are in Upper. (C) Anti-GFP immunoblots of membrane fractions from the marker lines detect a strong GFP-SYP123 band at 62.8 kDa, which is significantly and visibly weaker in *Atget3a* and *Atget1* lines than in WT Col-0. Bands below are likely the result of unspecific cross-reaction of antibody and plant extract. Coomassie brilliant blue staining (CBB stain) of blot confirms equal loading of protein. (D) qRT-PCR analysis of SYP123 transcript levels was performed using either SYP123- (gray) or GFP-specific (green) primers to resolve differences in mRNA levels on Col-0, *Atget1*, or *Atget3a* background. Expression levels were normalized to the Actin2 control. Error bars: SD (*n* = 6).

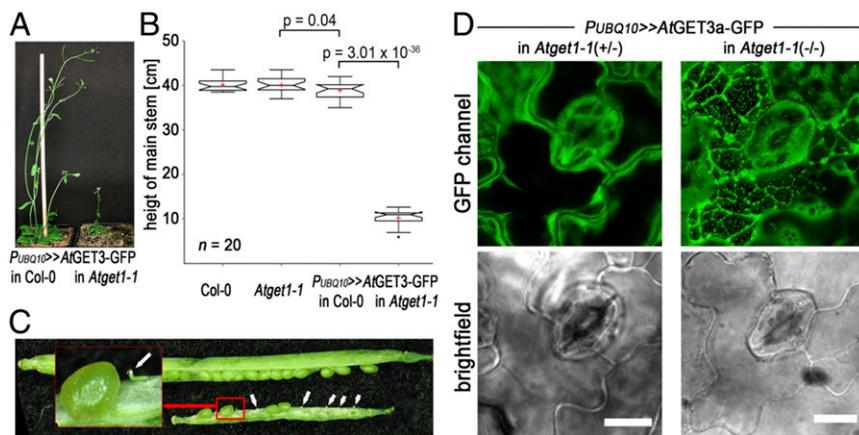
in the membrane. Quantitative RT-PCR (qRT-PCR) analyses further indicated that SYP123 transcript levels are also reduced in both mutants compared with the WT, with a milder transcript reduction in the *Atget3a* than in the *Atget1* background (Fig. 4D). Notably, the differences between endogenous and transgenic levels of transcript remain equal in all lines at roughly 50%, which confirms native expression of the marker construct (44) and suggests regulation of SYP123 in *get* lines also at transcript level.

#### Overexpression of AtGET3a in Atget1 Reveals Severe Growth Defects.

The general viability of *Atget* mutants and the fact that at least part of SYP123 finds its way to the plasma membrane in root hairs of mutants question the role of the GET pathway as the sole route for TA protein insertion in *A. thaliana*. To further understand the physiological importance of the pathway in planta, we crossed the overexpressing AtGET3a-GFP with the *Atget1-1* line. The rationale was to synthetically increase the activity of an upstream player, while limiting downstream capacity of the pathway to enhance phenotypes associated with dysfunction of the pathway.

Such overexpression of the cytosolic AtGET3a in its receptor KO leads to dwarfed plants. Main inflorescence, root, silique, and seed development are severely compromised compared with the parental lines (Fig. 5A–C and Fig. S7C–F). In addition to the obvious aboveground phenotype, the growth of root hairs is impaired more strongly compared with the individual loss of function *Atget1-1* lines (Fig. S7F). Such stronger phenotype might be a consequence of short-circuiting alternative insertion pathways, further depleting vital TA proteins from reaching their site of action.

CLSM analysis of the subcellular expression of AtGET3-GFP in the leaf epidermis of homozygous *Atget1* lines reveals cells with increased GFP fluorescence in foci among cells that resemble the normal cytoplasmic distribution of AtGET3a-GFP (Fig. 5D, Right and Movie S1). Conversely, no cells with GFP foci are present in leaf samples of heterozygous *Atget1*<sup>(+/-)</sup> lines expressing the same construct, and an even cytoplasmic distribution of AtGET3a-GFP is observable instead (Fig. 5D, Left and Movie S2). Foci may be a result of clustering of uninserted TA proteins with multimers of AtGET3a, similar to effects observed in yeast *Agel1* KOs (43, 45). We have also analyzed expression of



**Fig. 5.** Ectopic overexpression of *AtGET3a* in *Atget1* causes severe growth defects. (A) Exemplary images of 6-wk-old *A. thaliana* plants expressing *AtGET3a-GFP* in either Col-0 WT or *Atget1* showing significant differences in growth. (B) Boxplot summarizing the height of the main inflorescences of 20 individual 6-wk-old *A. thaliana* lines as labeled below the x axis. Boxplot as in Fig. 3 but with Tukey whiskers that extend to 1.5× interquartile range. (C) Siliques of mutant plants [*AtGET3a-GFP* in *Atget1* (silique below)] show a high number of aborted embryos in contrast to single *Atget1* lines (silique above). The statistical analysis can be found in Fig. S7C. (D) Maximum projection z stacks of 20 images at 1.1-μm optical slices at 63× magnification showing subcellular localization of *AtGET3a-GFP* in (Left) heterozygous or (Right) homozygous *Atget1-1* lines. Bright-field images below are taken from the 10th image in each stack. The full z stacks are shown in Movies S1 and S2. (Scale bars: 10 μm.)

*AtGET4-mCherry* in an *Atget1-1* background but did not detect similar aggregate-like structures (Fig. S7B).

## Discussion

Numerous biochemical and structural insights from yeast and in vitro systems have convincingly established the ability of the GET pathway to facilitate membrane insertion of TA proteins (reviewed in ref. 46). However, because TRC40 KO mice are embryonic lethal, physiological consequences of GET loss of function in an in vivo context remain insufficiently understood, and those that are available are typically specific to mammalian features. Such findings are in contrast to the high degree of conservation that GET homologs show across the eukaryotic domain, a situation where the model plant *A. thaliana* provides a highly suitable system for additional study.

Phylogenetic analysis of GET pathway components reveals an alternative GET3 clade, which must have evolved before the last eukaryotic ancestor. This hypothesis becomes apparent from the deeply branching phylogenetic tree (Fig. 14) but also, by the presence of a second distinct GET3 homolog in the recently discovered *Lokiarchaeum* sp., which forms a monophyletic group with eukaryotes (30). One of the *LsGET3* copies aligns within the GET3bc clade, with sequences that seem to only exist in Archaeplastida and SAR, whereas Opisthokonts and Amoebozoa may have lost this paralog. GET3bc branched off once more in some red algae and higher plants to evolve another plastidic GET3 paralog. It is unlikely that this third paralog is the result of endosymbiosis, because its sequence homology is too closely related to the other organellar candidate.

Neither root hair nor general growth in *A. thaliana* seem affected by lack of *AtGET3b/c*, and their biological function will require dedicated study in the future. Their localization in the plastid stroma and the mitochondrial matrix; failure to interact with *AtGET1*, *AtGET3a*, or *AtGET4*; absence of obvious downstream candidates to facilitate membrane insertion; lack of conserved sequence motifs for TA binding (Fig. S1); and failure to complement the *AtGET3a*-related growth defects (Fig. 3C) deem it unlikely that *AtGET3b/c* function is related to TA protein insertion.

A previous structural analysis of an archaeal (*Methanocaldococcus jannaschii*) GET3 ortholog inferred some key features that would distinguish GET3 from its prokaryotic ArsA ancestor sequence (28), namely the tandem repeat (exclusive to ArsA) and a conserved CxxC motif (specific for GET3). By contrast, our phylogenetic analysis uncovered the tandem repeat in candidate sequences of both eukaryotic GET3 clades, disproving it as a decisive feature solely of ArsA. Such sequence repeats may explain the presence of a third closely related GET3 paralog in higher plants and red algae as a consequence of an earlier tandem duplication, but this hypothesis requires in-depth analysis of more sequences from different species.

The CxxC motif, which is found in both Metazoa and Fungi GET3 orthologs, also exists in the Amoebozoan and *Lokiarchaeota*

GET3 orthologs and seemingly plays a role in zinc binding/coordination (19). However, this motif is absent in the Archaeplastida and SAR GET3a orthologs, where other invariant cysteines—CVC—some 40 aa upstream of the presumed CxxC motif are present. In contrast to the CxxC motif, the CVC motif can be found in all eukaryotic GET3a orthologs that we analyzed. Nevertheless, the CxxC motif is required for *ScGET3* to act as a general chaperone under oxidative stress conditions, binding unfolded proteins and preventing their aggregation (43, 45). Hence, it is conceivable that GET3bc paralogs—that feature CxxC (Fig. S1B)—have evolved as organellar chaperones with putative thiol-disulfide oxidoreductase function and lost (or never had) the TA insertion capability, whereas GET3a orthologs maintained (or acquired) both functions. Notably, the chaperone function of *ScGET3* is ATP-independent, whereas TA-insertase activity depends on ATP (43). A version of *AtGET3a*, where the ATPase motif is mutated (G28A), fails to rescue the root hair growth phenotype (Fig. 3C), suggesting that it is caused by the TA insertion function of *AtGET3a*, which is dependent on ATPase function (15).

Generally, T-DNA insertion in *AtGET1*, *AtGET3a*, or *AtGET4* leads to a reduction in root hair growth. Complementation with tagged or genomic constructs of the corresponding genes rescues normal growth connecting phenotype with genotype. Interestingly, multiple crosses between loss of function lines of three key players of an *A. thaliana* GET pathway do not lead to a more severe phenotype (i.e., even shorter root hairs than the single T-DNA insertion lines as measured, e.g., in plants overexpressing *AtGET3a-GFP* in *Atget1*) (Fig. S7F). This observation indicates that the three genes act in a linear pathway in *A. thaliana*, which is in agreement with findings in other species (15, 16). Nevertheless, it seems difficult to reconcile our findings with a putative GET pathway as the sole and global route responsible for insertion of TA proteins in plants similar to its proposed role in yeast or mammals (46). Of the estimated 500 TA proteins in *A. thaliana* (13), many are vital for development and survival of the plant. Especially SNARE proteins, which facilitate vesicle fusion to drive processes, such as cytokinesis, pathogen defense, and ion homeostasis (4, 7, 47), require correct and efficient membrane insertion. Inability of the plant to insert TA proteins should yield severe growth defects at least similar to if not stronger than—for example—the *knolle* phenotype caused by an *syp111* loss of function allele (coding for the Qa-SNARE KNOLLE). *Knolle* plants fail to grow beyond early seedling stage because of incomplete cell plate formation (48).

Absence of the root hair-specific Qa-SNARE SYP123 was shown to cause defects in root hair growth (42) as a result of reduced vesicle trafficking. Although lack of *AtGET* pathway components in planta did not lead to complete absence or mislocalization of SYP123 within the plasma membrane of root hairs, a significant reduction of protein levels was observed in vivo. Although this result was also confirmed biochemically, levels of SYP123 mRNA in *Atget1* as well as *Atget3a* lines are also

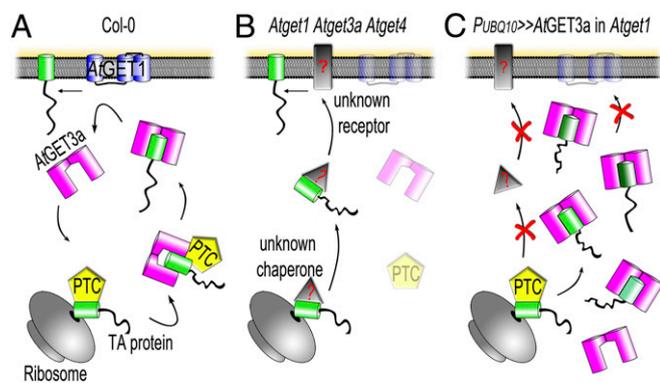
reduced (Fig. 4D), albeit not as strongly as the reduction of protein detected in the membrane fraction of mutants (Fig. 4C). Taken together, our findings indicate feedback control, where loss of *AtGET* function and the resulting failure of SYP123 protein insertion activate inhibition at the transcript level to decrease steady-state levels of both mRNA and protein. Functional cross-talk between the GET pathway and its impact on transcript regulation had been shown previously in other eukaryotes (23, 49).

The fact that lack of GET function can phenotypically only be detected in root hairs might be associated with these requiring fast and efficient trafficking of cargo and membrane material to the tip (42). Hence, slight imbalances in protein biogenesis owing to the absence of one major insertion pathway might strain alternative but unknown insertion systems, at which point lack of the GET pathway becomes rate-limiting. This effect is not reoccurring in the other fast-growing plant cells—pollen tubes—not only suggesting presence of an alternative pathway but also, questioning the monopoly of TA protein insertion of the GET pathway. Nevertheless, our SYP123 case study supports a role of the GET pathway in planta for regulating SNARE abundance. Interaction of *AtGET1* and *AtGET3a* with a wide range of different TA proteins was also shown, but we identified two TA proteins that failed to interact (SYP43 and At5g40510). Also, CoIP-MS analysis of *AtGET3a*-GFP detected only about 23 TA proteins, less than 5% of all TA proteins predicted to be present in *A. thaliana* (13) (Fig. S6B). Although the latter might be attributed to weak or transient binding of the TMD with *AtGET3a* or premature dissolution of binding through experimental conditions, it nevertheless raises questions as to the GET pathway being exclusively engaged in TA protein insertion into the ER. Among the many proteins that were detected in CoIP-MS analysis with *AtGET3a*-GFP, a lot of non-TA proteins but proteins related to trafficking or proteostasis were detected (Dataset S1). If some of these interactions can be confirmed in future studies, functional analyses might uncover alternative roles for *AtGET3a*.

Our findings are summarized in a working model of a presumed GET pathway in plants (Fig. 6). While under normal growth conditions, the GET pathway acts as main route for TA protein insertion into the ER membrane (Fig. 6A), and loss of

either component or a combination thereof brings alternative pathways into play (Fig. 6B). The existence of alternative insertion mechanisms is indicated by not only the relatively mild phenotype but also, the limited number of TA proteins that we found to interact with *AtGET3a*, raising the question of how TA proteins that do not interact with GET pathway components get inserted into membranes. In yeast, it has been shown that some TA proteins can insert unassisted and that chaperoning in the cytosol is facilitated by heatshock proteins (50); however, any alternative receptor remains elusive. Presence of an alternative insertion pathway in *A. thaliana* is also supported by the overexpression of the cytosolic *AtGET3a* in its receptor KO, which has severe phenotypic consequences (Figs. 5 and 6C). This observation corroborates a hierarchical connection of *AtGET3a* and *AtGET1*, because presence of the latter can rescue the growth defects. It further suggests the existence of an alternative pathway for TA insertion with weaker affinity toward pre-targeting factors, such as *AtGET4*, at the ribosome, because the aberrant amounts of *AtGET3a* seem to deplete the alternative pathway. Lastly, the *AtGET3a* foci that can occur in cells of mutant plants (but never in the WT background) (Fig. 5D) and that are similar to aggregates observed in stressed yeast cells (43) suggest additional functions of *AtGET3a* that nonetheless depend on *AtGET1*. The aggregate-like structures were not found in all cells of mutant plants, suggesting a dosage-dependent effect (i.e., if levels of *AtGET3a*-GFP exceed a certain threshold, clustering occurs). Clusters may consist of multimers of *AtGET3a*, complexes of *AtGET3a* bound to TA proteins, or *AtGET3a*/TA proteins bound to the elusive *AtGET2* receptor. In yeast, *ScGET2* is the first contact point at receptor level for the *ScGET3*-TA protein complex before the TA protein is delivered to *ScGET1* (20); hence, lack of *AtGET1* could keep a putative *AtGET3a*/TA protein aggregate stably in the vicinity of the ER.

Future work on this mutant in particular will help to resolve functions of GET components in *A. thaliana*. A current debate about potential cross-talk between GET components in TA protein insertion and protein quality control in yeast and animal cells (51) may be further underpinned by our findings in plants, which provide the fundament to broad comparative investigations in the near future.



**Fig. 6.** Model hypothesizing the subcellular mechanism of *A. thaliana* GET orthologs. (A) In WT Col-0, a pretargeting complex (PTC) likely comprising *A. thaliana* SGT2 and GET5 (both of which revealed many potential orthologs through in silico analyses) as well as the in silico-identified *AtGET4*, which interacts with *AtGET3a* in vivo, might receive nascent TA proteins from the ribosome and deliver these to the homodimer of *AtGET3a*, in turn shuttling the client TA protein to the ER receptor *AtGET1* (an *AtGET2* could not be identified through extensive BLASTp analysis and was left out of the figure). (B) The hypothetical situation in a single *Atget1*, *Atget3a*, or *Atget4* or crosses thereof. In the absence of a functional GET pathway, most TA proteins are delivered by an unknown alternative pathway (depicted as a gray triangle or rectangle with red question marks). (C) Overexpression of *AtGET3a* in absence of a docking station to unload client TA proteins might lead to cytosolic aggregates and block of TA insertion. The affinity between the PTC and *AtGET3a* might be a decisive factor here, because the unknown alternative pathway does not seem to compensate for the aberrant presence of *AtGET3a*.

## Materials and Methods

**Plant Growth Conditions.** Seeds were grown on 1/2 Murashige and Skoog medium including 1% sugar and 0.9% plant agar, pH 5.7. Plants were cultivated in a 16-h light/8-h dark cycle at 18 °C or 23 °C in the growth chamber (SI Materials and Methods).

**Construct Design.** Most constructs were designed by Gateway Recombination Reaction; vectors used for localization analyses can be found in ref. 33. A full list of oligonucleotides and constructs can be found in Tables S1 and S2 (SI Materials and Methods).

**Interaction Analyses.** We performed rBiFC in transiently transformed tobacco according to the work in ref. 37 (SI Materials and Methods).

**Microscopy.** CLSM microscopy was performed using a Leica SP8 at the following laser settings: GFP at 488-nm excitation (ex) and 490- to 520-nm emission (em); YFP at 514-nm ex and 520- to 560-nm em; and RFP/Mitotracker at 561-nm ex and 565- to 620-nm em. Chlorophyll autofluorescence was measured using the 488-nm laser line and em at 600–630 nm. TEM analysis and more details are in SI Materials and Methods.

**T-DNA Lines.** The following T-DNA lines were characterized (Fig. S4 A and B): Sai1\_1210\_E07 (*Atget1-1*), GK\_246D06 (*Atget1-2*), SALK\_033189 (*Atget3a-1*), SALK\_100424 (*Atget3a-2*), SALK\_012980 (*Atget3a-3*), SALK\_017702 (*Atget3b-2*), SALK\_091152 (*Atget3c-1*), SALK\_069782 (*Atget4-1*), and SALK\_121195 (*Atget4-4*). This work suggests new names for *Arabidopsis thaliana* genes previously termed “unknown”: *AtGET1* (At4g16444), *AtGET3a* (At1g01910), *AtGET3b* (At3g10350), *AtGET3c* (At5g60730), and *AtGET4* (At5g63220).

More details and other methods are in SI Materials and Methods.

**Note Added in Proof.** During revision of this article, an analysis of conditional *wrb* KO mice demonstrated that the GET pathway is required for only a subset—but not all—TA proteins *in vivo* (67). Also, an alternative ER insertion pathway was described in yeast (68) and another study reported an ER-stress and early flowering phenotype of the *Atget1-1* and *Atget3a-1* lines (69).

**ACKNOWLEDGMENTS.** We thank Masa H. Sato for sharing the GFP-SYP123 marker line and Gabriel Schaaf for vector pDRF1-GW and yeast strains. MS analysis was done at the Proteomics Centre, University of Tübingen, and we

thank Mirita Franz-Wachtel for her help in interpreting the data. We also thank Gerd Jürgens and Marja Timmermans for critical discussions and comments on the manuscript, Antje Feller and Jaquelynn Mateluna for help with the quantitative PCR analysis, and Eva Schwörzer and Laure Grefen for technical support. This work was supported by an Emmy Noether Fellowship SCHW 1719/1-1 from the Deutsche Forschungsgemeinschaft (DFG; to M.S.), CellNetworks Research Group funds (to G.G.), seed funding through the Collaborative Research Council 1101 (SFB1101; to C.G.), and an Emmy Noether Fellowship GR 4251/1-1 from the DFG (to C.G.).

- Lipka V, Kwon C, Panstruga R (2007) SNARE-ware: The role of SNARE-domain proteins in plant biology. *Annu Rev Cell Dev Biol* 23:147–174.
- Grefen C, Blatt MR (2008) SNAREs—molecular governors in signalling and development. *Curr Opin Plant Biol* 11(6):600–609.
- Rizo J, Südhof TC (2012) The membrane fusion enigma: SNAREs, Sec1/Munc18 proteins, and their accomplices—guiltily as charged? *Annu Rev Cell Dev Biol* 28:279–308.
- Jürgens G, et al. (2015) Plant cytokinesis: A tale of membrane traffic and fusion. *Biochem Soc Trans* 43(1):73–78.
- Surpin M, et al. (2003) The VTI family of SNARE proteins is necessary for plant viability and mediates different protein transport pathways. *Plant Cell* 15(12):2885–2899.
- Hachez C, et al. (2014) Arabidopsis SNAREs SYP61 and SYP121 coordinate the trafficking of plasma membrane aquaporin PIP2;7 to modulate the cell membrane water permeability. *Plant Cell* 26(7):3132–3147.
- Grefen C, et al. (2015) A vesicle-trafficking protein commandeers Kv channel voltage sensors for voltage-dependent secretion. *Nat Plants* 1:15108.
- Zhang B, et al. (2015) The Arabidopsis R-SNARE VAMP721 interacts with KAT1 and KC1 K<sup>+</sup> channels to moderate K<sup>+</sup> current at the plasma membrane. *Plant Cell* 27(6):1697–1717.
- Borgese N, Colombo S, Pedrazzini E (2003) The tale of tail-anchored proteins: Coming from the cytosol and looking for a membrane. *J Cell Biol* 161(6):1013–1019.
- Borgese N, Fasana E (2011) Targeting pathways of C-tail-anchored proteins. *Biochim Biophys Acta* 1808(3):937–946.
- Beilharz T, Egan B, Silver PA, Hofmann K, Lithgow T (2003) Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in *Saccharomyces cerevisiae*. *J Biol Chem* 278(10):8219–8223.
- Kalbfleisch T, Cambon A, Wattenberg BW (2007) A bioinformatics approach to identifying tail-anchored proteins in the human genome. *Traffic* 8(12):1687–1694.
- Kriechbaumer V, et al. (2009) Subcellular distribution of tail-anchored proteins in Arabidopsis. *Traffic* 10(12):1753–1764.
- Hegde RS, Keenan RJ (2011) Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nat Rev Mol Cell Biol* 12(12):787–798.
- Schuldiner M, et al. (2008) The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* 134(4):634–645.
- Stefanovic S, Hegde RS (2007) Identification of a targeting factor for posttranslational membrane protein insertion into the ER. *Cell* 128(6):1147–1159.
- Chang YW, et al. (2010) Crystal structure of Get4-Get5 complex and its interactions with Sgt2, Get3, and Ydj1. *J Biol Chem* 285(13):9962–9970.
- Simpson PJ, Schwappach B, Dohlman HG, Isaacson RL (2010) Structures of Get3, Get4, and Get5 provide new models for TA membrane protein targeting. *Structure* 18(8):897–902.
- Metz J, Wächter A, Schmidt B, Bujnicki JM, Schwappach B (2006) The yeast Arr4p ATPase binds the chloride transporter Gef1p when copper is available in the cytosol. *J Biol Chem* 281(1):410–417.
- Wang F, Chan C, Weir NR, Denic V (2014) The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. *Nature* 512(7515):441–444.
- Mukhopadhyay R, Ho YS, Swiatek PJ, Rosen BP, Bhattacharjee H (2006) Targeted disruption of the mouse Asn1 gene results in embryonic lethality. *FEBS Lett* 580(16):3889–3894.
- Daniele LL, Emran R, Lobo GP, Gaivrin RJ, Perkins BD (2016) Mutation of *wrb*, a component of the guided entry of tail-anchored protein pathway, disrupts photoreceptor synapse structure and function. *Invest Ophthalmol Vis Sci* 57(7):2942–2954.
- Norlin S, Parekh VS, Naredi P, Edlund H (2016) Asn1/TRC40 controls  $\beta$ -cell function and endoplasmic reticulum homeostasis by ensuring retrograde transport. *Diabetes* 65(1):110–119.
- Vogl C, et al. (2016) Tryptophan-rich basic protein (WRB) mediates insertion of the tail-anchored protein otoferlin and is required for hair cell exocytosis and hearing. *EMBO J* 35(23):2536–2552.
- Abell BM, Mullen RT (2011) Tail-anchored membrane proteins: Exploring the complex diversity of tail-anchored-protein targeting in plant cells. *Plant Cell Rep* 30(2):137–151.
- Duncan O, van der Merwe MJ, Daley DO, Whelan J (2013) The outer mitochondrial membrane in higher plants. *Trends Plant Sci* 18(4):207–217.
- Zhou T, Radaev S, Rosen BP, Gatti DL (2000) Structure of the Arsa ATPase: The catalytic subunit of a heavy metal resistance pump. *EMBO J* 19(17):4838–4845.
- Suloway CJ, Rome ME, Clemons WM, Jr (2012) Tail-anchor targeting by a Get3 tetramer: The structure of an archaeal homologue. *EMBO J* 31(3):707–719.
- Adl SM, et al. (2012) The revised classification of eukaryotes. *J Eukaryot Microbiol* 59(5):429–493.
- Spang A, et al. (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521(7551):173–179.
- Mateja A, et al. (2009) The structural basis of tail-anchored membrane protein recognition by Get3. *Nature* 461(7262):361–366.
- Stefer S, et al. (2011) Structural basis for tail-anchored membrane protein biogenesis by the Get3-receptor complex. *Science* 333(6043):758–762.
- Grefen C, et al. (2010) A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. *Plant J* 64(2):355–365.
- Garg SG, Gould SB (2016) The role of charge in protein targeting evolution. *Trends Cell Biol* 26(12):894–905.
- Vilardi F, Lorenz H, Dobberstein B (2011) WRB is the receptor for TRC40/Asn1-mediated insertion of tail-anchored proteins into the ER membrane. *J Cell Sci* 124(Pt 8):1301–1307.
- Grefen C, Obrdlir P, Harter K (2009) The determination of protein-protein interactions by the mating-based split-ubiquitin system (mbSUS). *Methods Mol Biol* 479:217–233.
- Grefen C, Blatt MR (2012) A 2in1 cloning system enables ratiometric bimolecular fluorescence complementation (rBiFC). *Biotechniques* 53(5):311–314.
- Mariappan M, et al. (2011) The mechanism of membrane-associated steps in tail-anchored protein insertion. *Nature* 477(7362):61–66.
- Datta S, Prescott H, Dolan L (2015) Intensity of a pulse of RSL4 transcription factor synthesis determines Arabidopsis root hair cell size. *Nat Plants* 1:15138.
- Grossmann G, et al. (2011) The RootChip: An integrated microfluidic chip for plant science. *Plant Cell* 23(12):4234–4240.
- Grierson C, Nielsen E, Ketelaarc T, Schiefelbein J (2014) Root hairs. *Arabidopsis Book* 12:e0172.
- Ichikawa M, et al. (2014) Syntaxin of plant proteins SYP123 and SYP132 mediate root hair tip growth in Arabidopsis thaliana. *Plant Cell Physiol* 55(4):790–800.
- Voth W, et al. (2014) The protein targeting factor Get3 functions as ATP-independent chaperone under oxidative stress conditions. *Mol Cell* 56(1):116–127.
- Enami K, et al. (2009) Differential expression control and polarized distribution of plasma membrane-resident SYP1 SNAREs in Arabidopsis thaliana. *Plant Cell Physiol* 50(2):280–289.
- Powis K, et al. (2013) Get3 is a holdase chaperone and moves to deposition sites for aggregated proteins when membrane targeting is blocked. *J Cell Sci* 126(Pt 2):473–483.
- Denic V, Dötsch V, Sinning I (2013) Endoplasmic reticulum targeting and insertion of tail-anchored membrane proteins by the GET pathway. *Cold Spring Harb Perspect Biol* 5(8):a013334.
- Kwon C, et al. (2008) Co-option of a default secretory pathway for plant immune responses. *Nature* 451(7180):835–840.
- Lukowitz W, Mayer U, Jürgens G (1996) Cytokinesis in the Arabidopsis embryo involves the syntaxin-related KNOLLE gene product. *Cell* 84(1):61–71.
- Jonikas MC, et al. (2009) Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* 323(5922):1693–1697.
- Johnson N, Powis K, High S (2013) Post-translational translocation into the endoplasmic reticulum. *Biochim Biophys Acta* 1833(11):2403–2409.
- Casson J, McKenna M, High S (2016) On the road to nowhere: Cross-talk between post-translational protein targeting and cytosolic quality control. *Biochem Soc Trans* 44(3):796–801.
- Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33(7):1870–1874.
- Karnik A, Karnik R, Grefen C (2013) SDM-Assist software to design site-directed mutagenesis primers introducing “silent” restriction sites. *BMC Bioinformatics* 14:105.
- Xing S, Wallmeroth N, Berendzen KW, Grefen C (2016) Techniques for the analysis of protein-protein interactions *in vivo*. *Plant Physiol* 171(2):727–758.
- Grefen C (2014) The split-ubiquitin system for the analysis of three-component interactions. *Methods Mol Biol* 1062:659–678.
- Hecker A, et al. (2015) Binary 2in1 vectors improve in planta (co)localization and dynamic protein interaction studies. *Plant Physiol* 168(3):776–787.
- Spitzer M, Wildenhain J, Rappsilber J, Tyers M (2014) BoxPlotR: A web tool for generation of box plots. *Nat Methods* 11(2):121–122.
- Grossmann G, et al. (2012) Time-lapse fluorescence imaging of Arabidopsis root growth with rapid manipulation of the root environment using the RootChip. *J Vis Exp* 65:pii:4290.
- Schindelin J, et al. (2012) Fiji: An open-source platform for biological-image analysis. *Nat Methods* 9(7):676–682.
- Jiang L, et al. (2005) VANGUARD1 encodes a pectin methyltransferase that enhances pollen tube growth in the Arabidopsis style and transmitting tract. *Plant Cell* 17(2):584–596.
- Boavida LC, McCormick S (2007) Temperature as a determinant factor for increased and reproducible *in vitro* pollen germination in Arabidopsis thaliana. *Plant J* 52(3):570–582.
- Roppolo D, et al. (2011) A novel protein family mediates Casparian strip formation in the endodermis. *Nature* 473(7347):380–383.
- Tokuyasu KT (1989) Use of poly(vinylpyrrolidone) and poly(vinyl alcohol) for cryoultramicrotomy. *Histochem J* 21(3):163–171.
- Park M, Touhri S, Müller I, Mayer U, Jürgens G (2012) Sec1/Munc18 protein stabilizes fusion-competent syntaxin for membrane fusion in Arabidopsis cytokinesis. *Dev Cell* 22(5):989–1000.
- Husband AY, Benkovic AH, Nogueira FT, Lodha M, Timmermans MC (2015) The ASYMMETRIC LEAVES complex employs multiple modes of regulation to affect adaxial-abaxial patterning and leaf complexity. *Plant Cell* 27(12):3321–3335.
- Loqué D, Lalonde S, Looger LL, von Wirén N, Frommer WB (2007) A cytosolic trans-activation domain essential for ammonium uptake. *Nature* 446(7132):195–198.
- Rivera-Monroy J, et al. (2016) Mice lacking WRB reveal differential biogenesis requirements of tail-anchored proteins *in vivo*. *Sci Rep* 6:39464.
- Aviram N, et al. (2016) The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. *Nature* 540(7631):134–138.
- Srivastava R, Zalisko BE, Keenan RJ, Howell SH (December 6, 2016) The GET system inserts the tail-anchored SYP72 protein into endoplasmic reticulum membranes. *Plant Physiol*, 10.1104/pp.16.00928.

# Supporting Information

Xing et al. 10.1073/pnas.1619525114

## SI Materials and Methods

**In Silico and Phylogenetic Analysis.** GET pathway orthologs were identified through BLASTp search (National Center for Biotechnology Information) against proteomes of candidate species and using default settings. Multiple sequence alignments were computed using the MUSCLE algorithm with default settings of MEGA7 (52). Evolutionary history was inferred by using the maximum likelihood method based on the Whelan and Goldman + frequency mode, applying 1,000 bootstraps to validate branching. The tree with the highest log likelihood ( $-12,793.272$ ) is shown. Percentages of trees above 70 in which the associated taxa clustered together are shown next to the branches. Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model and then, selecting the topology with superior log-likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites [five categories (+G; parameter = 1.377)]. The rate variation model allowed for some sites to be evolutionarily invariable ([+I]; 4.147% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 37 amino acid sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 279 positions in the final dataset.

**Construct Generation and Plant Transformation.** All  $P_{UBQ10}$  promoter-driven constructs were generated using Gateway technology as described previously (33). Full-length coding sequences of each gene were PCR-amplified; inserted into pDONR207, pDONR221-P1P4, or pDONR221-P3P2 via BP (ThermoFisher) reaction; and confirmed by sequencing (37). A point mutation of  $AtGET3a$  (G28A) was introduced through site-directed mutagenesis as described by ref. 53. Generation of  $P_{AtGET3a} \gg AtGET3a$ -GFP-3xHA,  $P_{AtGET1} \gg AtGET1$ -GFP-3xHA, and  $P_{AtGET4} \gg AtGET4$ -GFP-3xHA was done by conventional cloning from genomic DNA. The genomic fragment from start to stop codon was amplified and inserted into the binary vector  $P_{UBQ10} \gg GFP$ -3xHA 5' of GFP. The 3' UTR of the respective gene was amplified as well and inserted 3' of the 3xHA tag. After verification through sequencing, the promoter region of the gene was amplified and inserted to replace the  $UBQ10$  promoter. These constructs were first transformed into *Agrobacterium tumefaciens* GV3101 and then, dipped with WT (Col-0) plants. Oligonucleotides are listed in Table S1, and all constructs used are in Table S2.

**Interaction Assays.** The mating-based SUS was applied for the detection of protein–protein interactions in yeast (36). Application of methionine decreases Cub/bait-fusion expression. The lower affinity of C-terminal NubA compared with N-terminal NubG fusions was compensated for through the use of low-methionine levels (54). All interaction assays were performed as described previously (55).

The rBiFC (37) was applied to test in planta protein–protein interaction as described previously (56). All boxplots were generated using BoxPlotR (57).

**Plant Growth Conditions.** All mutant (Fig. S4 *A* and *B*) and transgenic lines are in Columbia (Col-0) background and were obtained from the Nottingham *Arabidopsis* Stock Centre ([arabidopsis.info/](http://arabidopsis.info/)). Seeds were imbibed on wet paper and stratified for 2–4 d in the

dark at 4 °C before sowing on soil or surface-sterilized with chlorine gas and plated on 1/2-strength solid Murashige and Skoog medium including 1% sugar and 0.9% plant agar, pH 5.7. Plants were cultivated in a 16-h light/8-h dark cycle at 18 °C or 23 °C in the growth chamber.

**Analysis of Root Hair Growth Kinetics.** Root hair growth kinetics and in part, SYP123 localization were determined on roots grown in RootChips, polydimethylsiloxane-based microfluidic perfusion devices for *Arabidopsis thaliana* root imaging (40). Plant cultivation on RootChips was performed as described elsewhere (58).

Image analysis of root hair growth rate was performed on bright-field time stacks in FIJI (59) as follows; time stacks of  $n$  time points were duplicated and truncated by three time points at the beginning and the end. The absolute difference between the two stacks was calculated using the FIJI image calculator tool. The resulting stack now highlighted the tip of every growing root hair as particle-like signal. The velocity of this tip representation was subsequently analyzed using the FIJI TrackMate plugin.

**Root Hair, Pollen Tube Growth, and CLSM Analysis.** The roots from 10-d-old seedlings grown on 1/2 Murashige and Skoog medium plates were imaged under a light microscope (ZEISS; Axiophot) using 2.5 $\times$  objective. Root hair length was measured using ImageJ. The 10 longest root hairs from 10 individual roots were examined per WT, T-DNA insertion, or complemented line.

Pollination experiments and aniline blue staining for pollen tube growth in pistils were performed as previously described (60). In vitro pollen germination was performed as reported previously (61). Pollen tubes were imaged 7 h after pollen germination on solid medium, and pollen tube length was quantified using ImageJ.

For subcellular localization of the  $AtGETx$ -GFP fusions and GFP-SYP123 in root hairs, roots of 7-d-old seedlings grown on plates or leaves from 2-wk-old plants grown in soil were observed. CLSM images were taken using a Leica SP8 CLSM. To exclude quantitative effects of the genetic background in our GFP fluorescence intensity analysis (Fig. 4*B*), we analyzed descendants of individual heterozygous lines. Macroscopic detection of the root hair phenotype allowed identification of homozygous *get* mutants, which were analyzed for mean fluorescence in root hairs as well as a similar number of randomly picked segregated lines. From at least 15 analyzed roots per line, the 5 with the strongest GFP signals were chosen for fluorescence intensity analysis. Laser settings used are given. GFP signals were measured at 488-nm ex and 490- to 520-nm em, YFP signals were measured at 514-nm ex and 520- to 560-nm em, and RFP/Mitotracker signals were measured at 561-nm ex and 565- to 620-nm em. Chlorophyll autofluorescence was measured using the 488-nm laser line and em at 600–630 nm.

**Immuno-TEM.** Immunogold labeling of ultrathin thawed cryosections was performed as described previously (62). Cotyledons were fixed with 4% (vol/vol) formaldehyde followed by 8% (vol/vol) formaldehyde for 30 and 120 min, respectively. Fixed cotyledons were infiltrated with a mixture of 20% (wt/vol) polyvinylpyrrolidone and 1.8 M sucrose (63) and frozen in liquid nitrogen. Ultrathin cryosections (80–100 nm) were cut with a cryoultramicrotome (UC7/FC7; Leica) at  $-110$  °C. Thawed cryosections mounted on TEM grids were blocked with 0.2% milk powder/0.2% BSA in PBS and incubated with rabbit anti-GFP serum (1:200) for 60 min followed by several washing

steps using blocking buffer. Thereafter, sections were incubated with goat anti-rabbit coupled to ultrasmall gold (Fig. 1 *D* and *G*) (1:50; Aurion) or coupled to 6-nm gold (Fig. 1 *J*) (1:30; Dianova) for 60 min. Gold particles were silver-enhanced using R-Gent (Aurion) for 45 and 35 min. Labeled cryosections were stained with 1% aqueous uranyl acetate and embedded in methyl cellulose containing 0.45% uranyl acetate. Sections were viewed in a JEM-1400plus TEM (Jeol) at 120 kV accelerating voltage, and micrographs were recorded with a TemCam-F416 CMOS Camera (Tietz).

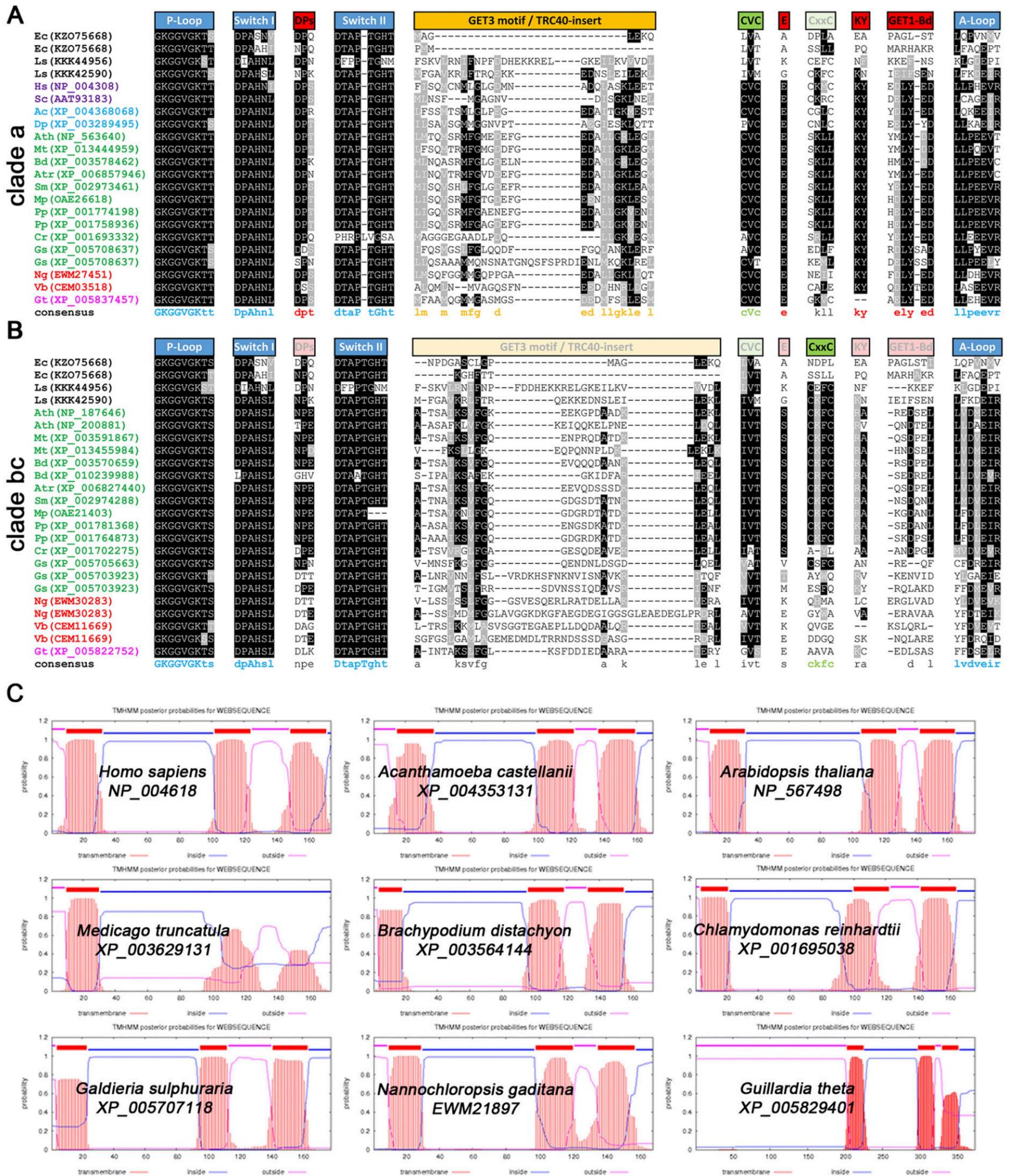
**CoIP-MS Analysis.** Protein extracts of *P<sub>UBQ10</sub>* >> AtGET3a-GFP and as control, *P<sub>UBQ10</sub>* >> GFP seedlings grown under continuous light were harvested after 5 d. Three grams plant tissue was taken for the immunoprecipitation according to the work in ref. 64 with slight modifications. Only the second washing buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Triton X-100) was used, but it was used four times; 60- $\mu$ L GFP-Trap Beads (ChromoTek) were added to each sample. The final precipitate in 2 $\times$  Laemmli buffer was analyzed by MS at the University of Tübingen Proteome Center. Two individual biological replicates were performed, and candidates that interacted with GFP only were omitted from the final list of interaction partners (Dataset S1).

**Membrane Fractionation.** Root samples (0.2–1 g) of 3-wk-old seedlings grown on 1/2 Murashige and Skoog plates (+1% sucrose + 25  $\mu$ g/ml Hygromycin) were harvested and ground on ice. Samples were treated in a ratio of 1:2 with extraction buffer [1 M Tris-HCl, pH 7.5, 1 M MgCl<sub>2</sub>, 1 M DTT, 1/2 tablet protease inhibitor (cOmplete, EDTA-free; ROCHE), 0.5 M sucrose] and homogenized. Separation of membrane and cytosol was achieved

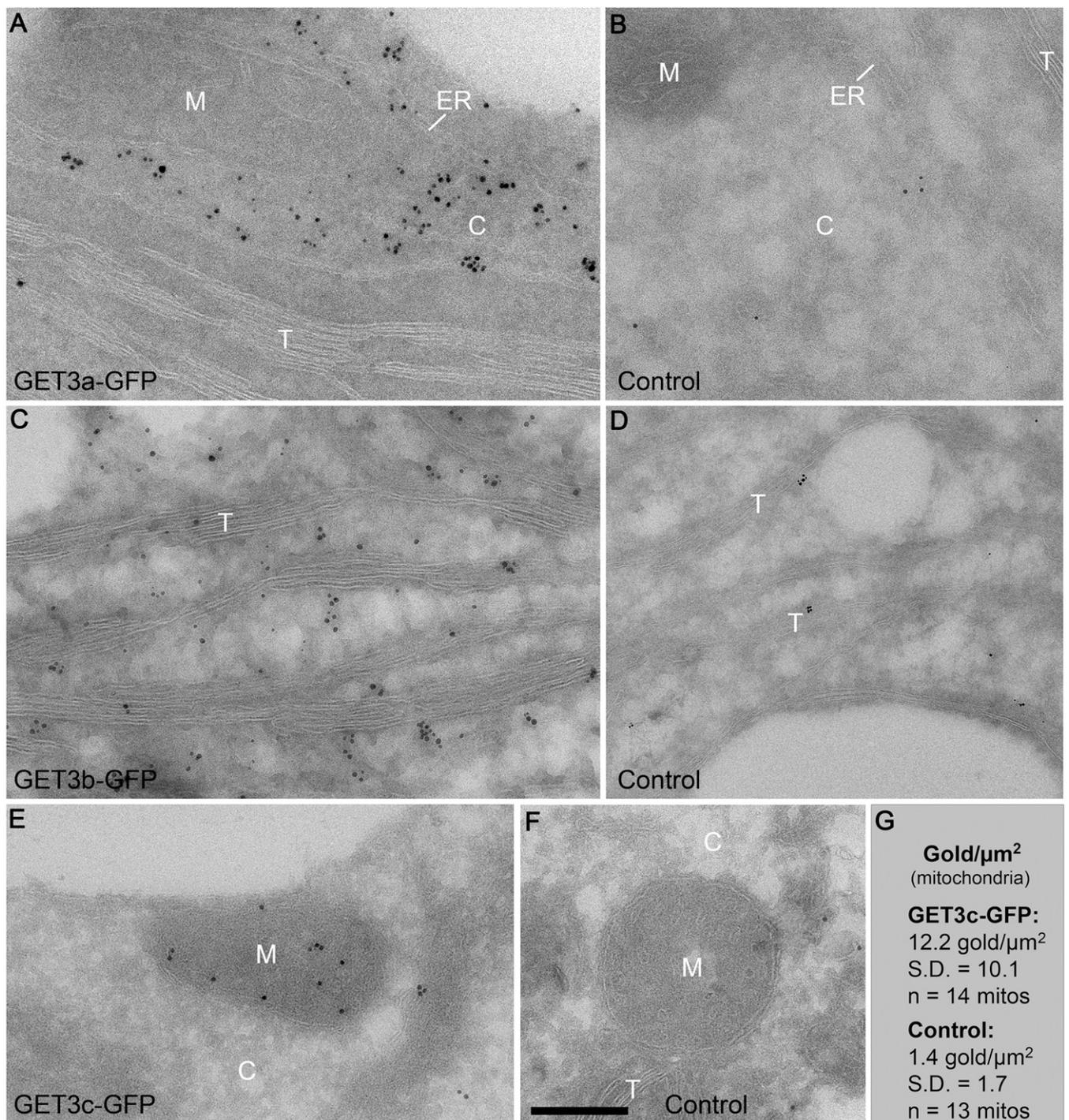
through sequential centrifugation: 10 min at 10,000  $\times$  *g* and 4 °C to purify samples from cell debris followed by 1 h at 100,000  $\times$  *g* and 4 °C. Membrane pellets were resuspended in 50  $\mu$ L fresh extraction buffer and sonicated for 5 s at 60% power, and protein concentration was measured using Bradford reagent prior immunoblotting. Protein samples were adjusted to equal concentration using Laemmli buffer [+3.5% (vol/vol)  $\beta$ -Mercaptoethanol] and boiled for 20 min at 65 °C.

**qRT-PCR Analysis.** Total RNAs were isolated from 100 mg 5-d-old seedlings grown on 1/2 Murashige and Skoog medium by using the Isolate II RNA Plant Kit (Bioline). For cDNA synthesis, ProtoScript II–First Strand cDNA Synthesis Kit (NEB; 1  $\mu$ g RNA) was used. qRT-PCR was performed using oligonucleotides (Table S1) specific to SYP123, GFP, and ACT2 as internal control. iQ SYBR Green Supermix (Bio-Rad) was used and performed on the CFX96 Real-Time PCR System (Bio-Rad). Relative quantification values were calculated using the  $2^{-\Delta C_t}$  method, with the  $\Delta C_t$  of ACT2 as normalization control (65).

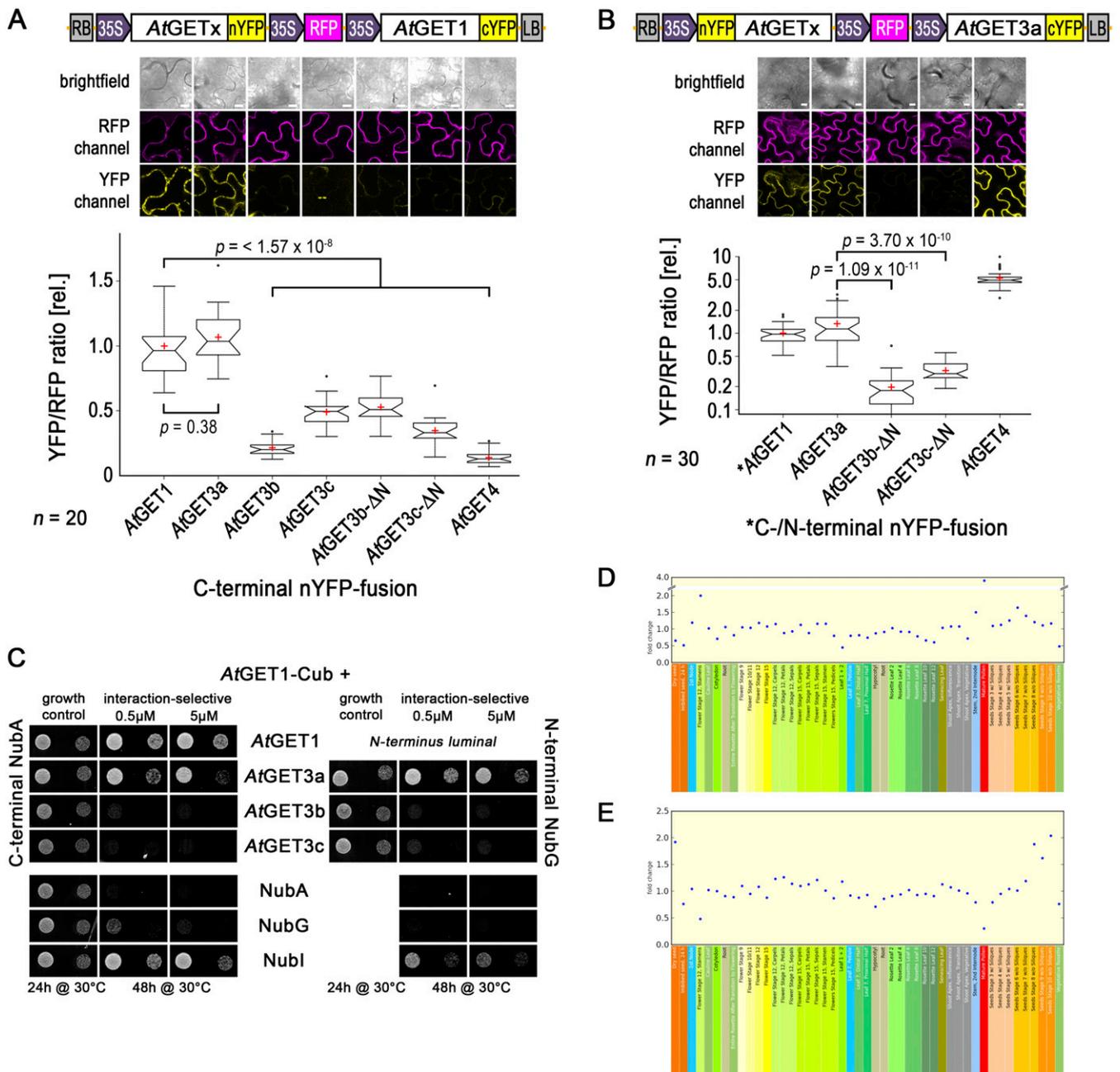
**Yeast Complementation Analysis.** *A. thaliana* genes for the yeast complementation analysis were expressed from a 2 $\mu$  origin plasmid (pYOX1-Dest) under the strong constitutive yeast PMA1 promoter, which was based on the Gateway-compatible pDRf1-GW vector (66). *Get1p* and *get3p* KO and corresponding BY4741 WT strains were originally created by the *Saccharomyces* Genome Deletion Project, Stanford. Yeast was grown and transformed as described for the SUS analysis but using Uracil as selection marker. Yeast was dropped in 10 times OD dilutions on selection media and grown under different temperatures for 3 d.



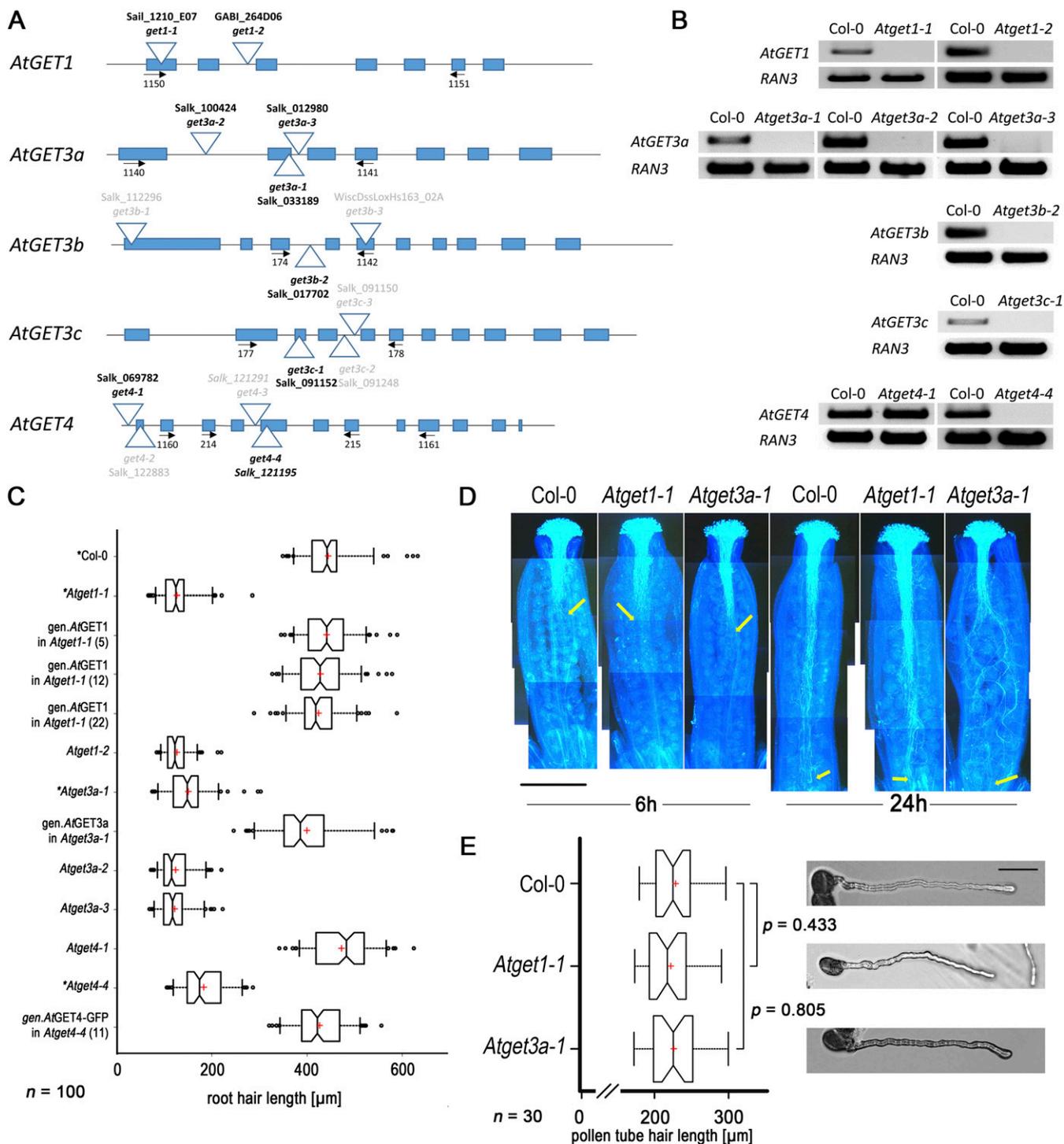
**Fig. S1.** Sequence and structural evaluation of GET orthologs. Excerpts of multiple sequence alignments of (A) clade a and (B) clade bc GET3 orthologs showing conserved motifs. ATPase motifs are in blue (P loop and Switches I and II), and GET1 binding motifs are in red (conserved only in clade a). Cysteine residues (CVC and CxxC motifs) important for metal binding/dimerization are in light green. Absence or partial conservation of motifs is depicted through opaqueness of boxes above the sequences. Tandem sequences were split and treated as two individual GET3 orthologs for accessions: KZO75668, XP\_005708637, XP\_005703923, EWM30283, and CEM11669. Ac, *Acanthamoeba castellanii*; Ath, *A. thaliana*; Atr, *Amborella trichopoda*; Bd, *Brachypodium distachyon*; Cr, *Chlamydomonas reinhardtii*; Dp, *Dictyostelium purpureum*; Ec, *Escherichia coli*; Gs, *Galdieria sulphuraria*; Gt, *Guillardia theta*; Hs, *Homo sapiens*; Ls, *Lokiarchaeum* sp.; Mp, *Marchantia polymorpha*; Mt, *Medicago truncatula*; Ng, *Nannochloropsis gaditana*; Pp, *Physcomitrella patens*; Sc, *Saccharomyces cerevisiae*; Sm, *Selaginella moellendorffii*; Vb, *Vitrella brassicaformis*. (C) Exemplary TMD prediction of membrane domains of ScGET1, HsWRB, and putative orthologs in different eukaryotic species ([www.cbs.dtu.dk/services/TMHMM/](http://www.cbs.dtu.dk/services/TMHMM/)).



**Fig. S2.** Expanded view of localization analysis of AtGET orthologs (original TEM images shown in Fig. 1 *D*, *G*, and *J*). High-resolution images and controls of TEM analysis shown in parts in Fig. 1 *D*, *G*, and *J*. TEM immunogold labeling of GFP in (*A*) AtGET3a-GFP (cytoplasm), (*C*) AtGET3b-GFP (chloroplasts), and (*E*) AtGET3c-GFP (mitochondria) expressing seedlings using ultrathin thawed cryosections of cotyledons. Control experiments using seedlings missing the corresponding fusion protein are shown in *B*, *D*, and *F*. *G* shows a statistical analysis of the relatively weak but specific mitochondrial gold labeling in AtGET3c-GFP seedlings. C, cytoplasm; M, mitochondrion; T, thylakoid. (Scale bar: *A*–*F*, 300 nm.)

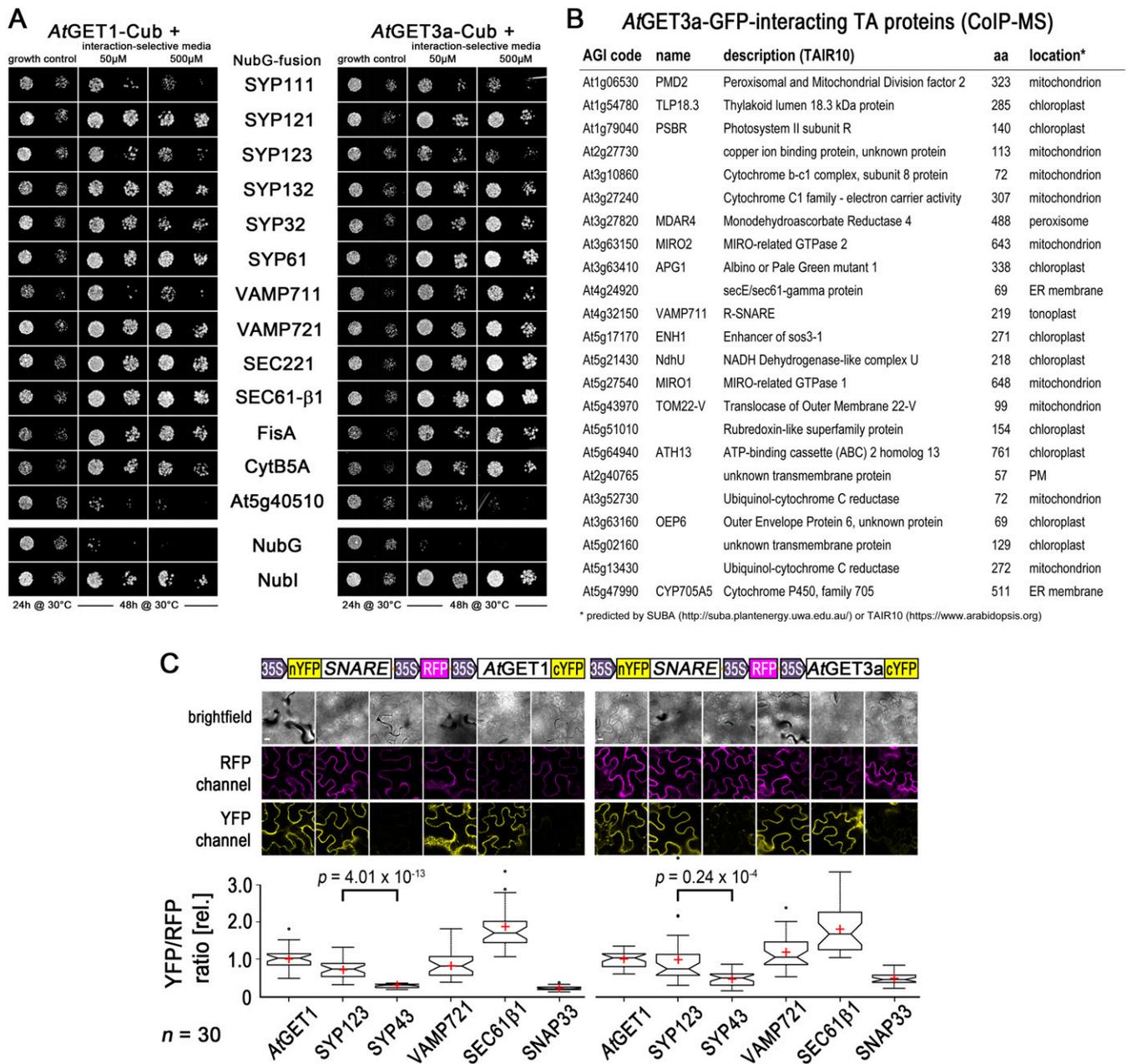


**Fig. S3.** Interaction analysis of AtGET pathway orthologs. (A and B) Complete rBiFC analysis of (A) AtGET1 and (B) AtGET3a with GET pathway orthologs and truncated constructs. Boxed cartoons show construct design above representative images of epidermal cells from transiently transformed *Nicotiana benthamiana* leaves. Larger versions of confocal images are presented in Fig. 2M. YFP/RFP mean fluorescence intensities from 20 different leaf sections were calculated and ratioed against the average YFP/RFP ratio of AtGET1 homodimerization or AtGET3a–AtGET1 interaction. Center lines of boxes represent medians, with outer limits at 25th and 75th percentiles. Notches indicate 95% confidence intervals; Tukey whiskers extend to 1.5× interquartile range, outliers are depicted as black dots, and red crosses mark sample means. (Scale bars: 10 μm.) (C) Split Ubiquitin interaction analysis in yeast. (Left) C-terminally NubA- or (Right) N-terminally NubG-tagged AtGET3 orthologs were coexpressed with AtGET1-Cub in yeast. Untagged NubA, NubG, or Nubi were used as negative (NubG or NubA) or positive (Nubi) controls, respectively. Growth on interaction-selective media was detected for yeast coexpressing AtGET1-Cub and AtGET1-NubA as well as AtGET3a-Nub fusion. The plastidic AtGET3 paralogs do not interact with AtGET1 in yeast in either tag orientation complementing the rBiFC analysis. (D and E) eFP browser screenshots showing fold changes in expression ratios of (D) AtGET1 with AtGET3a and (E) AtGET3a with AtGET4 over different developmental stages from publicly available microarray data ([bar.utoronto.ca/efp\\_arabidopsis/cgi-bin/efpWeb.cgi](http://bar.utoronto.ca/efp_arabidopsis/cgi-bin/efpWeb.cgi)).



**Fig. S4.** Functional analysis of AtGET orthologs in planta and yeast. (A) Cartoon depicting the sequence-verified position of each T-DNA analyzed in this work (in black type font). (B) DNA gels of semi-qRT-PCR corroborate lack of transcript in all mutant lines except *Atget4-1* in line with this being a T-DNA insertion in the 5' UTR. RAN3 (At5g55190) transcript was used as control. (C) Expanded root hair growth analysis showing additional alleles and complementation thereof. Note that the 5' UTR-inserted *Atget4-1* line that still transcribes *AtGET4* shows WT-like root hair growth. \*Values that are also in Fig. 3. (D) Aniline blue staining of pollen tubes (the WT and *Atget* mutants) grown for 6 or 24 h, respectively, after pollination of Col-0 pistils. Yellow arrows point to exemplary pollen tubes termini that have reached ovules. Pictures are composites of individual images along the pistil, and exposure was enhanced to visualize the bright blue pollen tubes against the darker blue background. (E) Growth of pollen tubes was measured in vitro from 30 individual pollen grains 7 h postgermination (representative images in *Right*). Center lines of boxes represent medians, with outer limits at 25th and 75th percentiles. Notches indicate 95% confidence intervals; Tukey whiskers extend to 1.5x interquartile range, and red crosses mark sample means. (Scale bar: 50  $\mu$ m.)





**Fig. 56.** Expanded information on TA-protein interactions. (A) SUS interaction analyses of candidate SNARE/TA proteins with AtGET1 and AtGET3a as Cub/bait fusion. Growth on interaction-selective media (-Ade and -His) was monitored after 2 d, and control plates were monitored after 24 h. OD<sub>600</sub> 1.0 and 0.1 dilutions were dropped, with NubG serving as negative control and Nubl (WT version) serving as positive control, respectively. (B) TA proteins that were identified via CoIP-MS of AtGET3a-GFP-expressing plants that were not detected in GFP-only expressing plants. (C) Complete rBiFC analysis of (Left) AtGET1 and (Right) AtGET3a with candidate SNARE/TA proteins. Boxed cartoons show construct design above exemplary images of transiently transformed *N. benthamiana* leaves. Larger versions of these confocal images are in Fig. 4A. YFP/RFP mean fluorescence intensities from 30 different leaf sections were calculated and ratioed against the average YFP/RFP ratio of AtGET1 homodimerization or AtGET3a-AtGET1 interaction. Center lines of boxes represent medians, with outer limits at 25th and 75th percentiles. Notches indicate 95% confidence intervals; Tukey whiskers extend to 1.5 $\times$  interquartile range, outliers are depicted as black dots, and red crosses mark sample means. (Scale bars: 10  $\mu$ m.)



**Table S1. Oligonucleotides used for cloning and RT-PCR**

| No.  | 5'-3' Sequence   | Purpose                       |
|------|--|-------------------------------|
| 83   | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGCGGCGGATTGCGCGAGG       | pDONR207-AtGET3a              |
| 85   | GGGGACCCTTTGTACAAAGAAAGCTGGGTGGCCACTCTTGACCCGTTTCGAGTTC    | pDONR207-AtGET3a              |
| 104  | GGGGACAAGTTTGTACAAAAAAGCAGGCTGCATGGCGACTCTGTCTTCCATCTG     | pDONR207-AtGET3b              |
| 106  | GGGGACCCTTTGTACAAAGAAAGCTGGGTGTTTCCAAATGATATCGCCCAAGAAG    | pDONR207-AtGET3b              |
| 107  | GGGGACAAGTTTGTACAAAAAAGCAGGCTGCATGGCGGCTTTACTTCTCCTCAATC   | pDONR207-AtGET3c              |
| 109  | GGGGACCCTTTGTACAAAGAAAGCTGGGTGTTTCCAAATGAGATCACCCATGAAC    | pDONR207-AtGET3c              |
| 89   | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGAAGGAGAGAAGCTTATAGAAG   | pDONR207-AtGET1               |
| 91   | GGGGACCCTTTGTACAAAGAAAGCTGGGTGGAATCCACGAACCTACACAC         | pDONR207-AtGET1               |
| 86   | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGTCGAGAGAGAGGATCAAACGTG   | pDONR207-AtGET4               |
| 88   | GGGGACCCTTTGTACAAAGAAAGCTGGGTGGCCATCATCTTGAAGATGTCTCC      | pDONR207-AtGET4               |
| 261  | TCCGAGGTAAAGCAGGTGTTGGGAAG                                 | Introducing G28A in AtGET3a   |
| 262  | TCTTCCCAACACCTGCTTTACCTCCG                                 | Introducing G28A in AtGET3a   |
| 631  | GCGGATTTAAATAGATAAGGCTCTGTTCTTCCC                          | 3' End fragment of AtGET3a    |
| 632  | TGCAGATTATAACGCTTGTACAGATACCCCTCAAC                        | 3' End fragment of AtGET3a    |
| 633  | TGACTGGAGCTCTTAATTAAGGCCTATGGCGGCGGATTGCGCGGAGCGAC         | Genomic fragment of AtGET3a   |
| 634  | GCACTAGTGCCACTCTTGACCCGTTTCGAGTTC                          | A genomic fragment of AtGET3a |
| 554  | TAGTCGTTAATTAATCAGAGGAGAGAGCTAAGTGAAGGG                    | AtGET3a promoter              |
| 553  | TTAGCCCGGGTCTAATTCCTTGTCTGCTCTCTCTTC                       | AtGET3a promoter              |
| 625  | GCGGATTTAAATATCGCATCCCTGAAAAGAGTGAAG                       | 3' End fragment of AtGET1     |
| 626  | TGCAGATTATAATAAGTACACGCGCTTTTAGAATC                        | 3' End fragment of AtGET1     |
| 627  | TGACTGGAGCTCAGGCCTATGGAAGGAGAGAAGCTTATAGAAG                | Genomic fragment of AtGET1    |
| 628  | GCACTAGTGAATCCACGAACCTACACACATATTTG                        | Genomic fragment of AtGET1    |
| 629  | TGACTGGAGCTCGGCGCGCCTTAATTAAGTTGGCCAAAGTAGAAAATGGTTG       | AtGET1 promoter               |
| 630  | GAAGGCCTTAACCCCTTTTGTGATTACTGATTC                          | AtGET1 promoter               |
| 635  | GCGGATTTAAATGGAAGGAGTTTGAAGAGTGAGTTC                       | 3' End fragment of AtGET4     |
| 636  | TGCAGATTATAAGCTCTGTAATACTTCTTGTGTTCCG                      | 3' End fragment of AtGET4     |
| 657  | TGACTGGAGCTCAGGCCTATGTCGAGAGAGAGGATCAAACGTG                | Genomic fragment of AtGET4    |
| 658  | TGCAGCTAGCGCCATCATCTACACAGTTCATGG                          | Genomic fragment of AtGET4    |
| 659  | TGACTGGAGCTCGGCGCGCCTTAATTAACCTCTAATCTCTCCCTAGCTAG         | AtGET4 promoter               |
| 660  | GAAGGCCTGGATCTCAAGGATTTGTTGTTTTTC                          | AtGET4 promoter               |
| 442  | GTAGGCCTATTGTAAATTAACGATCTCATATTG                          | RSL4 promoter                 |
| 443  | TCACTAGTCGCTCTAACTGATCAACTCTTGCC                           | RSL4 promoter                 |
| 761  | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGCTAGCCCAACGGAGACGATTTTC | AtGET3b ΔN68                  |
| 762  | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGCTACTCTTGCTGAAGGAGCTTC  | AtGET3c ΔN50                  |
| 1140 | ATGGCGGCGGATTTGCGGAGCGC                                    | RT-PCR for AtGET3a            |
| 1141 | TCACATCTTTCAAGCCCTCAAGTC                                   | RT-PCR for AtGET3a            |
| 174  | ATAAACCCCTGAGAAGGCTAGGGAAGAG                               | RT-PCR for AtGET3b            |
| 1142 | TCAAGATTTTACCAATGGATGCATC                                  | RT-PCR for AtGET3b            |
| 177  | TGAGATCATTAGCTACTCTTGTCTGAAG                               | RT-PCR for AtGET3c            |
| 178  | TGGGAGCAGTATCAAAAACATACGAG                                 | RT-PCR for AtGET3c            |
| 214  | TCACCGCTCAAAGATTCTCTGAAGC                                  | RT-PCR for AtGET4             |
| 215  | TCTCGGGTCTCAGCTCTAACAAAATG                                 | RT-PCR for AtGET4             |
| 1160 | AGGCAATTACTATGGAGCTTTGC                                    | RT-PCR for AtGET4             |
| 1161 | TCTCATCCATCATAAAGTTTGCATC                                  | RT-PCR for AtGET4             |
| 1150 | GTTAATGGAAGGAGAGAAGCTTATAG                                 | RT-PCR for AtGET1             |
| 1408 | TACATGGCCTGTCTATGTGACCTCC                                  | RT-PCR for AtGET1             |
| 1408 | ATTGGTTTCCCTCTTTTCCCTCGCTCCG                               | RT-PCR for AtGET2             |
| 1410 | AGTGCATCCATTATCTTCTTCCACC                                  | RT-PCR for AtGET2             |
| 1546 | ATGAACGATCTTATCTCAAGCTCATTC                                | RT-PCR for SYP123             |
| 547  | TCAAGGTGCAAGTAGAGTGTAAAG                                   | RT-PCR for SYP123             |
|      | AGAACACCCCATCGGCGAC  | RT-PCR for GFP                |
|      | TGATCGCGCTTCTCGTTGGGGTC                                    | RT-PCR for GFP                |
|      | GCCATCCAAGCTGTTCTCTC                                       | RT-PCR for ACT2               |
|      | CAGTAAGGTACGTCCAGCA  | RT-PCR for ACT2               |

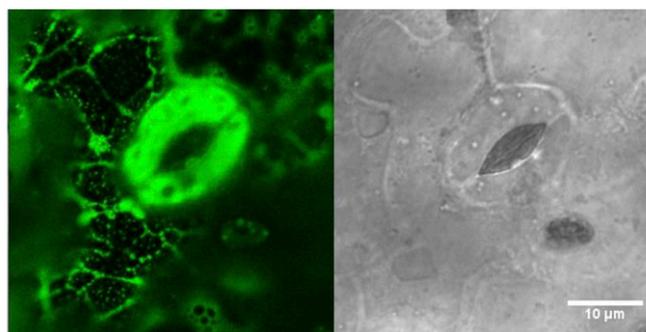
**Table S2. Entry and destination constructs used**

| Int. no. | Name                       | Vector        | Insert      | Purpose         |
|----------|----------------------------|---------------|-------------|-----------------|
| e002     | pDONR207-Syp111-ST         | pDONR207      | At1g08560   | Entry clone     |
| e004     | pDONR207-Syp121-ST         | pDONR207      | At3g11820   | Entry clone     |
| e080     | pDONR207-VAMP711-ST        | pDONR207      | At4g32150   | Entry clone     |
| e081     | pDONR207-VAMP721-ST        | pDONR207      | At1g04750   | Entry clone     |
| e190     | pDONR221-L3L2-VAMP721-ST   | pDONR221-P3P2 | At1g04750   | Entry clone     |
| e192     | pDONR221-L3L2-SNAP33-ST    | pDONR221-P3P2 | At5g61210   | Entry clone     |
| E006     | pDONR207-SYP61-ST          | pDONR207      | At1g28490   | Entry clone     |
| E008     | pDONR207-AtGET3a-ST        | pDONR207      | At1g01910   | Entry clone     |
| E009     | pDONR207-AtGET3a-wo        | pDONR207      | At1g01910   | Entry clone     |
| E011     | pDONR207-AtGET4-wo         | pDONR207      | At5g63220   | Entry clone     |
| E012     | pDONR207-AtGET1-ST         | pDONR207      | At4g16444   | Entry clone     |
| E013     | pDONR207-AtGET1-wo         | pDONR207      | At4g16444   | Entry clone     |
| E014     | pDONR207-SEC221-ST         | pDONR207      | At1g11890   | Entry clone     |
| E101     | pDONR207-AtGET3b-ST        | pDONR207      | At3g10350   | Entry clone     |
| E102     | pDONR207-AtGET3b-wo        | pDONR207      | At3g10350   | Entry clone     |
| E103     | pDONR207-AtGET3c-ST        | pDONR207      | At5g60730   | Entry clone     |
| E104     | pDONR207-AtGET3c-wo        | pDONR207      | At5g60730   | Entry clone     |
| E105     | pDONR207-SYP32-ST          | pDONR207      | At3g24350   | Entry clone     |
| E107     | pDONR221-L3L2-AtSEC221-ST  | pDONR221-P3P2 | At1g11890   | Entry clone     |
| E108     | pDONR221-L1L4-GET3a-wo     | pDONR221-P1P4 | At1g01910   | Entry clone     |
| E109     | pDONR221-L1L4-GET4-wo      | pDONR221-P1P4 | At5g63220   | Entry clone     |
| E120     | pDONR221-L3L2-AtSYP43-ST   | pDONR221-P3P2 | At3g05710   | Entry clone     |
| E124     | pDONR207-ScGET3p-ST        | pDONR207      | YDL100C     | Entry clone     |
| E126     | pDONR207-SYP123-ST         | pDONR207      | At4g03330   | Entry clone     |
| E128     | pDONR207-SYP132-ST         | pDONR207      | At5g08080   | Entry clone     |
| E143     | pDONR207-At5g40510-ST      | pDONR207      | At5g40510   | Entry clone     |
| E154     | pDONR221-L3L2-SYP123-ST    | pDONR221-P3P2 | At4g03330   | Entry clone     |
| E157     | pDONR221-L3L2-AtGET4-ST    | pDONR221-P3P2 | At5g63220   | Entry clone     |
| E195     | pDONR221-L3L2-AtGET1-wo    | pDONR221-P3P2 | At4g16444   | Entry clone     |
| E196     | pDONR221-L1L4-AtGET1-wo    | pDONR221-P1P4 | At4g16444   | Entry clone     |
| E198     | pDONR221-L1L4-AtGET3b-wo   | pDONR221-P1P4 | At3g10350   | Entry clone     |
| E199     | pDONR221-L1L4-AtGET3c-wo   | pDONR221-P1P4 | At5g60730   | Entry clone     |
| E221     | pDONR207-AtGET3bΔN-ST      | pDONR207      | At3g10350   | Entry clone     |
| E222     | pDONR207-AtGET3bΔN-wo      | pDONR207      | At3g10350   | Entry clone     |
| E223     | pDONR207-AtGET3cΔN-ST      | pDONR207      | At5g60730   | Entry clone     |
| E224     | pDONR207-AtGET3cΔN-wo      | pDONR207      | At5g60730   | Entry clone     |
| E243     | pDONR221-L3L2-SEC61β-ST    | pDONR221-P3P2 | At2g45070   | Entry clone     |
| E252     | pDONR221-L1L4-AtGET3bΔN-wo | pDONR221-P1P4 | At3g10350   | Entry clone     |
| E254     | pDONR221-L1L4-AtGET3cΔN-wo | pDONR221-P1P4 | At5g60730   | Entry clone     |
| E265     | pDONR221-L3L2-AtGET3a-wo   | pDONR221-P3P2 | At1g01910   | Entry clone     |
| E289     | pDONR207-FisA-ST           | pDONR207      | At3g57090   | Entry clone     |
| E374     | pDONR207-CYTb5A-ST         | pDONR207      | At5g53560   | Entry clone     |
| D0116    | pDRf1-AtGET1               | pDRf1-GW      | E012        | Complementation |
| D0584    | pZU-LC-ScGET1p             | pZU-LC        | Genomic DNA | Complementation |
| D0512    | pYOX1-AtGET3a              | pYOX1-Dest    | E008        | Complementation |
| D0513    | pYOX1-AtGET3b              | pYOX1-Dest    | E101        | Complementation |
| D0514    | pYOX1-AtGET3bΔN            | pYOX1-Dest    | E221        | Complementation |
| D0515    | pYOX1-AtGET3cΔN            | pYOX1-Dest    | E223        | Complementation |
| D0516    | pYOX1-ScGET3               | pYOX1-Dest    | E124        | Complementation |
| D0520    | pYOX1-AtGET3c              | pYOX1-Dest    | E103        | Complementation |
| D0296    | pMetYC-AtGET1              | pMetYC-Dest   | E013        | SUS             |
| D0076    | pMetOYC-AtGET3a            | pMetOYC-Dest  | E009        | SUS             |
| D0078    | pNX35-AtGET3a              | pNX35-Dest    | E008        | SUS             |
| D0086    | pNX35-AtGET3b              | pNX35-Dest    | E101        | SUS             |
| D0088    | pNX35-AtGET3c              | pNX35-Dest    | E103        | SUS             |
| D0298    | pXNubA22-AtGET3a           | pXNubA22-Dest | E009        | SUS             |
| D0299    | pXNubA22-AtGET3b           | pXNubA22-Dest | E102        | SUS             |
| D0300    | pXNubA22-AtGET3c           | pXNubA22-Dest | E104        | SUS             |
| D0297    | pXNubA22-AtGET1            | pXNubA22-Dest | E013        | SUS             |
| D0081    | pNX35-SEC221               | pNX35-Dest    | E014        | SUS             |
| D0098    | pNX35-SYP32                | pNX35-Dest    | E105        | SUS             |
| d517     | pNX35-SYP121               | pNX35-Dest    | e004        | SUS             |
| d537     | pNX35-VAMP711              | pNX35-Dest    | e080        | SUS             |

**Table S2. Cont.**

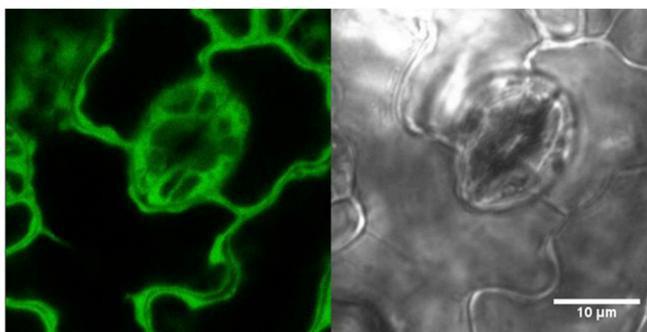
| Int. no. | Name                                   | Vector            | Insert      | Purpose      |
|----------|--|-------------------|-------------|--------------|
| d538     | pNX35-VAMP721                          | pNX35-Dest        | e081        | SUS          |
| D0754    | pNX35-Syp111                           | pNX35-Dest        | e002        | SUS          |
| D0756    | pNX35-Syp61                            | pNX35-Dest        | E006        | SUS          |
| D0786    | pNX35-SYP123                           | pNX35-Dest        | E126        | SUS          |
| D0787    | pNX35-SYP132                           | pNX35-Dest        | E128        | SUS          |
| D0788    | pNX35-At5g40510                        | pNX35-Dest        | E143        | SUS          |
| D0789    | pNX35-SEC61-β1                         | pNX35-Dest        | E228        | SUS          |
| D0790    | pNX35-CYTb5A                           | pNX35-Dest        | E374        | SUS          |
| D0791    | pNX35-FisA                             | pNX35-Dest        | E289        | SUS          |
| D0273    | pBiFct-nYFP-AtGET4-AtGET3a-cYFP        | pBiFct-2in1-NC    | E108 + E157 | rBiFC        |
| D0356    | pBiFct-AtGET1-nYFP-AtGET1-cYFP         | pBiFct-2in1-CC    | E196 + E195 | rBiFC        |
| D0355    | pBiFct-AtGET1-nYFP-AtGET4-cYFP         | pBiFct-2in1-CC    | E109 + E195 | rBiFC        |
| D0354    | pBiFct-AtGET1-nYFP-AtGET3a-cYFP        | pBiFct-2in1-CC    | E108 + E195 | rBiFC        |
| D0545    | pBiFct-AtGET1-nYFP-AtGET3bΔN-cYFP      | pBiFct-2in1-CC    | E252 + E195 | rBiFC        |
| D0546    | pBiFct-AtGET1-nYFP-AtGET3cΔN-cYFP      | pBiFct-2in1-CC    | E254 + E195 | rBiFC        |
| D0361    | pBiFct-AtGET1-nYFP-AtGET3b-cYFP        | pBiFct-2in1-CC    | E198 + E195 | rBiFC        |
| D0362    | pBiFct-AtGET1-nYFP-AtGET3c-cYFP        | pBiFct-2in1-CC    | E199 + E195 | rBiFC        |
| D0965    | pBiFct-AtGET3a-nYFP-AtGET3a-cYFP       | pBiFct-2in1-NC    | E108 + E265 | rBiFC        |
| D0966    | pBiFct-AtGET3a-nYFP-AtGET3bΔN-cYFP     | pBiFct-2in1-NC    | E252 + E265 | rBiFC        |
| D0973    | pBiFct-AtGET3a-nYFP-AtGET3cΔN-cYFP     | pBiFct-2in1-NC    | E254 + E265 | rBiFC        |
| D0123    | pBiFct-nYFP-SYP43-AtGET3a-cYFP         | pBiFct-2in1-NC    | E108 + E120 | rBiFC        |
| D0395    | pBiFct-NC-nYFP-SYP43-ST-AtGET1-cYFP    | pBiFct-2in1-NC    | E196 + E107 | rBiFC        |
| D0980    | pBiFct-NC-nYFP-SYP123-AtGET1-cYFP      | pBiFct-2in1-NC    | E196 + E154 | rBiFC        |
| D0267    | pBiFct-nYFP-SYP123-AtGET3a-cYFP        | pBiFct-2in1-NC    | E108 + E154 | rBiFC        |
| D0371    | pBiFct-NC-nYFP-VAMP721-AtGET1-cYFP     | pBiFct-2in1-NC    | E196 + e190 | rBiFC        |
| D0588    | pBiFct-NC-nYFP-SNAP33-ST-AtGET3a-cYFP  | pBiFct-2in1-NC    | E108 + e192 | rBiFC        |
| D0589    | pBiFct-NC-nYFP-Vamp721-ST-AtGET3a-cYFP | pBiFct-2in1-NC    | E108 + e190 | rBiFC        |
| D0418    | pBiFct-NC-nYFP-Ssβ1-ST-AtGET1-cYFP     | pBiFct-2in1-NC    | E196 + E243 | rBiFC        |
| D0590    | pBiFct-NC-nYFP-Ssβ1-ST-AtGET3a-cYFP    | pBiFct-2in1-NC    | E108 + E243 | rBiFC        |
| D0090    | pUBQ10::AtGET3a-GFP                    | pUBQ10-GW-GFP     | E009        | Localization |
| D0091    | pUBQ10::AtGET3b-GFP                    | pUBQ10-GW-GFP     | E102        | Localization |
| D0092    | pUBQ10::AtGET3c-GFP                    | pUBQ10-GW-GFP     | E104        | Localization |
| D0160    | pUBQ10::AtGET1-GFP                     | pUBQ10-GW-GFP     | E013        | Localization |
| D0504    | pUBQ10::AtGET4-mCherry                 | pUBQ10-GW-mCherry | E011        | Localization |
| D0399    | pUBQ10::AtGET3bΔN-GFP                  | pUBQ10-GW-GFP     | E222        | Localization |
| D0405    | pUBQ10::AtGET3cΔN-GFP                  | pUBQ10-GW-GFP     | E224        | Localization |

ST, native stop codon; wo, without stop codon.



**Movie S1.** CLSM z stack of AtGET3a-GFP in homozygous *Atget1*<sup>(-/-)</sup>.

[Movie S1](#)



Movie S2. CLSM z stack of AtGET3a-GFP in heterozygous *Atget1*<sup>(-/+)</sup>.

[Movie S2](#)

Dataset S1. CoIP-MS raw data of  $P_{UBQ10} \gg$  GET3-GFP interaction partners in WT Col-0 plants from two individual biological replicates (R1 and R2)

[Dataset S1](#)

**II. Asseck et al., submitted**

1 **ER membrane receptors of the GET pathway are conserved throughout**  
2 **eukaryotes**

3

4 Lisa Yasmin Asseck<sup>1,2</sup>, Dietmar Gerald Mehlhorn<sup>1,2</sup>, Jhon Rivera Monroy<sup>3,4</sup>,  
5 Martiniano Maria Ricardi<sup>2</sup>, Holger Breuninger<sup>1</sup>, Niklas Wallmeroth<sup>1</sup>, Kenneth Wayne  
6 Berendzen<sup>1</sup>, Minou Nouwrosian<sup>2</sup>, Shuping Xing<sup>1</sup>, Blanche Schwappach<sup>3</sup>, Martin  
7 Bayer<sup>5</sup>, Christopher Grefen<sup>1,2\*</sup>

8

9 <sup>1</sup>Developmental Genetics, Centre for Plant Molecular Biology, University of  
10 Tuebingen, 72076 Tuebingen, Germany

11 <sup>2</sup>Department of Molecular & Cellular Botany, Ruhr-University Bochum, 44780  
12 Bochum, Germany

13 <sup>3</sup>Department of Molecular Biology, University Medical Center Göttingen,  
14 Humboldtallee 23, 37073 Göttingen, Germany

15 <sup>4</sup>High Complexity Instrumental Laboratory, Universidad de La Salle, 110231 Bogotá,  
16 Colombia

17 <sup>5</sup>Department of Cell Biology, Max Planck Institute for Developmental Biology, 72076  
18 Tuebingen, Germany

19

20 **\*Corresponding author:** Christopher Grefen

21 Correspondence: [christopher.grefen@rub.de](mailto:christopher.grefen@rub.de)

22 @MCBotany

23

24 **Running title:** Analysis of the ER membrane receptors within the GET pathway in  
25 plants

26

27 **Synopsis:** GET pathway in plants

28

29 **Main Question:** Identifying the Archaeplastida GET pathway co-receptor

## 30 **Summary**

31 Type II tail-anchored (TA) membrane proteins are involved in diverse cellular  
32 processes such as protein translocation, vesicle trafficking and apoptosis [1]. They are  
33 characterized by a single C-terminal transmembrane domain (TMD) that mediates  
34 post-translational targeting and insertion into the endoplasmic reticulum (ER) via the  
35 Guided Entry of Tail-anchored Proteins (GET) pathway. The GET system was  
36 originally described in mammals [2] and yeast [3] but was recently shown to be partially  
37 conserved in other eukaryotes such as higher plants [4, 5]. In short, a newly  
38 synthesized TA protein is shielded from the cytosol by a pre-targeting complex and an  
39 ATPase (GET3 (in yeast) / TRC40 (in mammals)) which delivers the protein to the ER  
40 where membrane receptors (GET1 & GET2 / WRB & CAML) facilitate insertion.

41 In the model plant *Arabidopsis thaliana*, most components of the pathway were  
42 identified through *in silico* sequence comparison, however, a functional homolog of the  
43 co-receptor GET2/CAML remained elusive. We performed immunoprecipitation-mass  
44 spectrometry (IP-MS) analysis to detect *in vivo* interactors of AtGET1 and identified a  
45 membrane protein of unknown function that contains structural characteristics of both,  
46 yeast GET2, and mammalian CAML, which we termed GET1-interacting Protein  
47 (G1IP). The protein localises to the ER membrane and coexpresses with AtGET1.  
48 Additional interaction data revealed an intricate relationship with AtGET1 and  
49 AtGET3a. Loss of G1IP in *Arabidopsis* leads to reduced root hair growth phenocopying  
50 previously described GET pathway mutants [5]. Taken together G1IP is most likely the  
51 missing co-receptor of the *Arabidopsis* GET pathway and its protein sequence is an  
52 important puzzle piece in understanding cross-kingdom evolution of the GET pathway.

53 **Results & Discussion**

54 **An unknown transmembrane protein interacts with AtGET1 and AtGET3a in**  
 55 **planta**

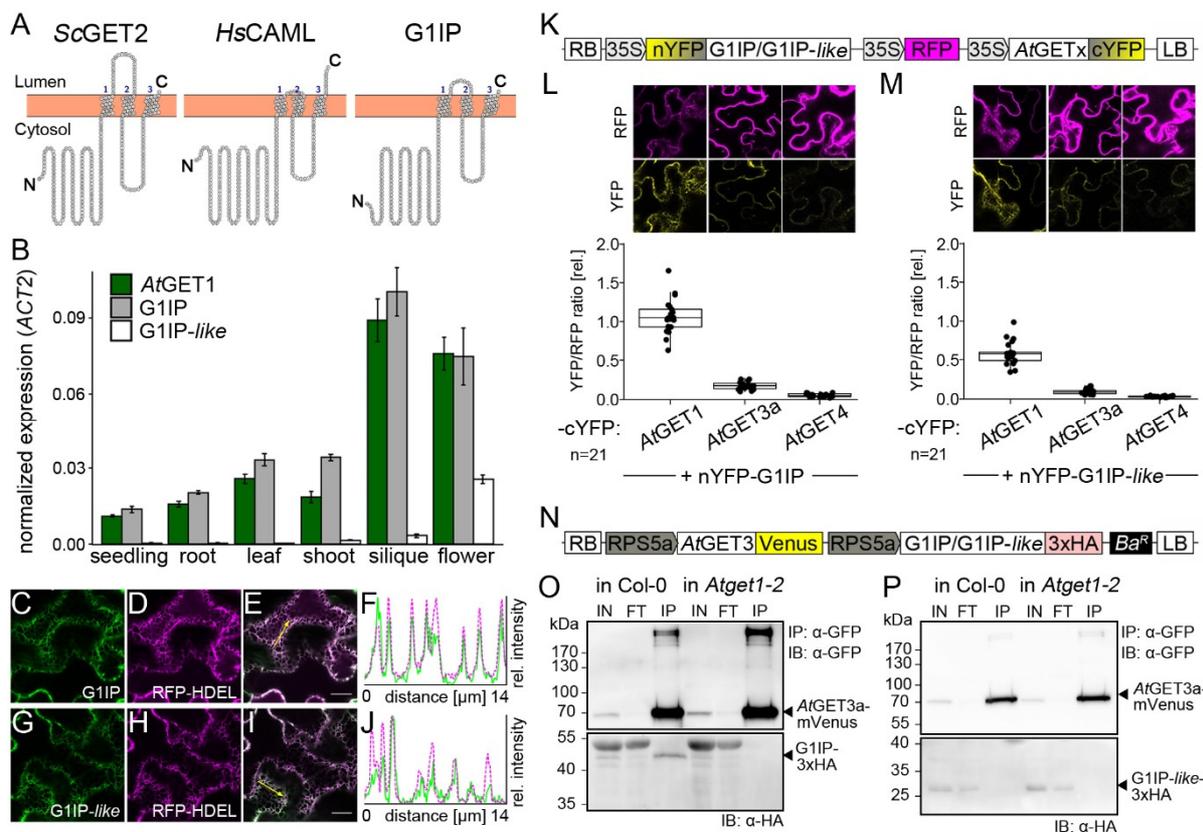
56 Both GET receptor forming protein pairs, Get1 and Get2 in yeast [3, 6] as well as Wrb  
 57 and CAML [7]) in mammalian cells, were shown to co-purify. Hence, we chose affinity  
 58 purification as a promising strategy to identify the elusive co-receptor of AtGET1 and,  
 59 we performed immunoprecipitation of AtGET1-GFP stably expressed in *Arabidopsis*  
 60 *thaliana* wildtype (Col-0) followed by mass spectrometry. Two biological replicates  
 61 were executed and candidates that came up in both experiments and predicted to  
 62 contain TMDs were considered as high-confidence targets (Table 1).

63 **Table 1:** AGI codes and identifier of candidates that were identified in both replicates of AtGET1-GFP  
 64 IP-MS analyses and predicted to contain TMDs

| AGI       | gene name | description                                     | Prediction tool |                |         | also detected via AtGET3a-GFP (Xing et al. 2017) |
|-----------|-----------|---|-----------------|----------------|---------|--|
|           |           |   | Locali-sation*  | number of TMDs |         |  |
|           |           |   | SUBA            | TMHMM          | TMpred  |  |
| AT4G32680 | G1IP      | unknown transmembrane protein                   | Nuc             | 3              | 4 or 3  | yes  |
| AT1G52343 | G1IP-like | unknown transmembrane protein                   | Cyt/Mito        | 2              | 3       | no   |
| AT5G13490 | AAC2      | ADP/ATP carrier 2                               | Mito            | 3              | 5 or 4  | yes  |
| AT5G13430 |           | Ubiquinol-cytochrome C reductase FeS subunit    | Mito            | 0              | 2       | yes  |
| AT1G50200 | ALATS     | Alanyl-tRNA synthetase                          | Mito            | 0              | 1       | yes  |
| AT4G01100 | ADNT1     | adenine nucleotide transporter 1                | Mito            | 0              | 4       | yes  |
| AT5G41670 |           | 6-phosphogluconate dehydrogenase family protein | Mito/Chp        | 0              | 2       | yes  |
| AT2G38040 | CAC3      | carboxyltransferase alpha subunit               | Chp             | 0              | 3       | yes  |
| AT1G64190 |           | 6-phosphogluconate dehydrogenase family protein | Chp             | 0              | 2       | yes  |
| AT1G29900 | CARB      | carbamoil phosphate synthetase B                | Chp             | 0              | 2 or 1  | yes  |
| AT5G30510 | RPS1      | ribosomal protein S1                            | Chp             | 0              | 1       | no   |
| AT5G53480 |           | ARM repeat superfamily protein                  | Cyt/Nuc/Chp     | 0              | 3       | no   |
| AT2G20580 | RPN1A     | 26S proteasome regulatory subunit S2 1A         | Cyt/Nuc         | 0              | 5 or 4  | yes  |
| AT4G24820 |           | 26S proteasome regulatory subunit Rpn7          | Cyt/Nuc         | 0              | 1       | yes  |
| AT2G30490 | C4H       | cinnamate-4-hydroxylase                         | ER              | 0              | 2       | yes  |
| AT5G47990 | CYP705A5  | Cytochrome P450 705A5                           | ER              | 0              | 4       | yes  |
| AT1G07810 | ECA1      | ER-type Ca2 -ATPase 1                           | ER              | 8              | 9       | yes  |
| AT3G51460 | RHD4      | Phosphoinositide phosphatase family protein     | ER              | 2              | 3       | yes  |
| AT1G70770 |           | Protein of unknown function DUF2359             | ER              | 0              | 2 or 1  | yes  |
| AT4G21150 | HAP6      | ribophorin II (RPN2) family protein             | ER              | 4              | 4       | yes  |
| AT1G29310 |           | SecY protein transport family protein           | Golgi           | 10             | 10 or 9 | yes  |
| AT4G25820 | XTH14     | xyloglucan endotransglucosylase/hydrolase 14    | CW              | 1              | 1       | yes  |

65 \* Nuc = Nucleus, Cyt = Cytosol, Mito = Mitochondria, Chp = Chloroplast, ER = Endoplasmic Reticulum, CW = cell wall

66 Since both, Get2 and CAML, contain a C-terminal membrane-anchoring domain with  
 67 three transmembrane helices we focussed on candidates with such structure. We  
 68 identified an unknown membrane protein G1IP (*AtGET1*-Interacting Protein,  
 69 At4g32680) which appeared to match these preferences (Figure 1A). Interestingly,  
 70 G1IP was also detected in our previously published IP-MS results using *AtGET3a*-GFP  
 71 [5] substantiating that this protein may indeed be part of the *Arabidopsis* GET pathway.  
 72 In addition, a close homolog of G1IP exists in *Arabidopsis* (At1g52343) that we termed  
 73 G1IP-like. This protein was identified in both IP-MS analyses of *AtGET1*, but not when  
 74 using *AtGET3a*-GFP as target (Table 1, [5]).



75  
 76 **Figure 1:** G1IP coexpresses with *AtGET1*, localizes to the ER and interacts with *AtGET1* and *AtGET3a*.  
 77 (A) Transmembrane topology prediction of ScGET2, HsCAML and AtG1IP using Protter. (B) Relative  
 78 transcript levels of *AtGET1*, G1IP and G1IP-like in different organs of *A. thaliana* Col-0 plants measured  
 79 by qPCR analysis. ACT2 was used as reference gene. Error bars: SD; (n = 3). (C-J) CLSM analysis of  
 80 the subcellular localization of (C-F) p35S::GFP-G1IP and (G-J) p35S::GFP-G1IP-like in leaves of stably  
 81 transformed *A. thaliana* lines coexpressing the ER marker RFP-HDEL. Line histograms (F, J) along  
 82 yellow arrows in (E, I) confirm colocalization. Scale bars, 10 μm. (K) Schematic of the 2in1 rBiFC  
 83 constructs used in (L, M). (L, M) rBiFC analysis of (L) G1IP and (M) G1IP-like with *Arabidopsis* GET  
 84 pathway components. Exemplary CLSM images of transiently transfected *N. benthamiana* leaves are  
 85 depicted. Mean fluorescence of 21 areas were measured in YFP and RFP channels, ratioed and plotted  
 86 to show YFP complementation. Centre lines of boxes represent median with outer limits at 25th and  
 87 75th percentile. Tukey whiskers extend to 1.5x IQR, all values are depicted as black dots. (N) Schematic  
 88 of the 2in1 Co-IP constructs used in (O, P). Co-IP of *AtGET3a* with (O) G1IP or (P) G1IP-like in Col-0  
 89 and *Atget1-2* mutant background. Protein extracts of *Arabidopsis* seedlings overexpressing *AtGET3a*-  
 90 mVenus and G1IP-3xHA or G1IP-like-3xHA were immunoprecipitated with anti-GFP beads. Protein-  
 91 protein interaction was detected by immunoblotting (IB) using anti-GFP and anti-HA antibody,  
 92 respectively. IN, input; FT, flow-through; IP, immunoprecipitate.

93 Multiple sequence alignment using MegaX shows only low overall similarity between  
94 G1IP and yeast Get2 or mammalian CAML, respectively (Figure S1A). However,  
95 structural comparison revealed that the predicted membrane topology of G1IP  
96 suggests type II orientation with a long cytosolic N-terminus, three transmembrane  
97 helices and a luminal C-terminal region (TMHMM, TMpred and Protter ver.1.0 [8])  
98 closely resembling the structure of yeast Get2 and mammalian CAML (Figure 1A).  
99 Moreover, Phyre2 and HHpred analyses of the sequence maps part of the N-terminus  
100 of G1IP (aa 6-27) with the crystal structure of cytosolic ScGet2 bound to ScGet3  
101 (structures 3ZS9\_D and 3SJD\_E, respectively).

102 The predicted orientation of G1IP was experimentally verified using ratiometric  
103 Bimolecular Fluorescence Complementation (rBiFC, [9]) with the co-receptor AtGET1  
104 (Figure S1B). The putative structure of G1IP-*like* is similar to that of G1IP with a  
105 relatively large N-terminal intracellular region and three transmembrane helices in the  
106 C-terminal domain (predicted via TMHMM and TMpred).

#### 107 **G1IP and AtGET1 share the same expression profile and subcellular localisation**

108 To determine a functional relation between G1IP and AtGET1 we assessed the  
109 expression patterns by quantitative PCR (qPCR). Consistent with expression data of  
110 publicly available microarray and proteomics data [10], qPCR analysis revealed  
111 constitutive coexpression of G1IP and AtGET1 at similar levels across all tissues and  
112 developmental stages supporting the notion of a shared molecular pathway (Figure  
113 1B). In contrast, G1IP-*like* exhibits flower-specific gene expression in both, qPCR and  
114 *in silico* analysis (eFP Browser), indicating functional divergence of the two homologs.  
115 Such expression pattern contradicts a putative housekeeping function that the AtGET1  
116 co-receptor needs to fulfil within the GET pathway. G1IP-*like* may have acquired novel,  
117 flower-specific functions.

118 The AtGET1 receptor was previously described as an ER-localised protein [5].  
119 However, *in silico* prediction suggests a nuclear localisation for G1IP  
120 (<http://suba.live/factsheet.html?id=AT4G32680.1>) which would contradict a potential  
121 ER import function of a GET pathway co-receptor. In order to investigate the  
122 subcellular localisation of G1IP in *A. thaliana* we created stable transgenic plants that  
123 coexpress N-terminally GFP-tagged G1IP with the ER marker secRFP-HDEL. Using  
124 confocal laser scanning microscopy (CLSM) we were able to confirm a subcellular ER

125 localisation for G1IP (Figure 1C-F) as was demonstrated previously for *AtGET1* [5].  
126 Similar to its homolog, *G1IP-like* also localises to the ER membrane (Figure 1G-J).

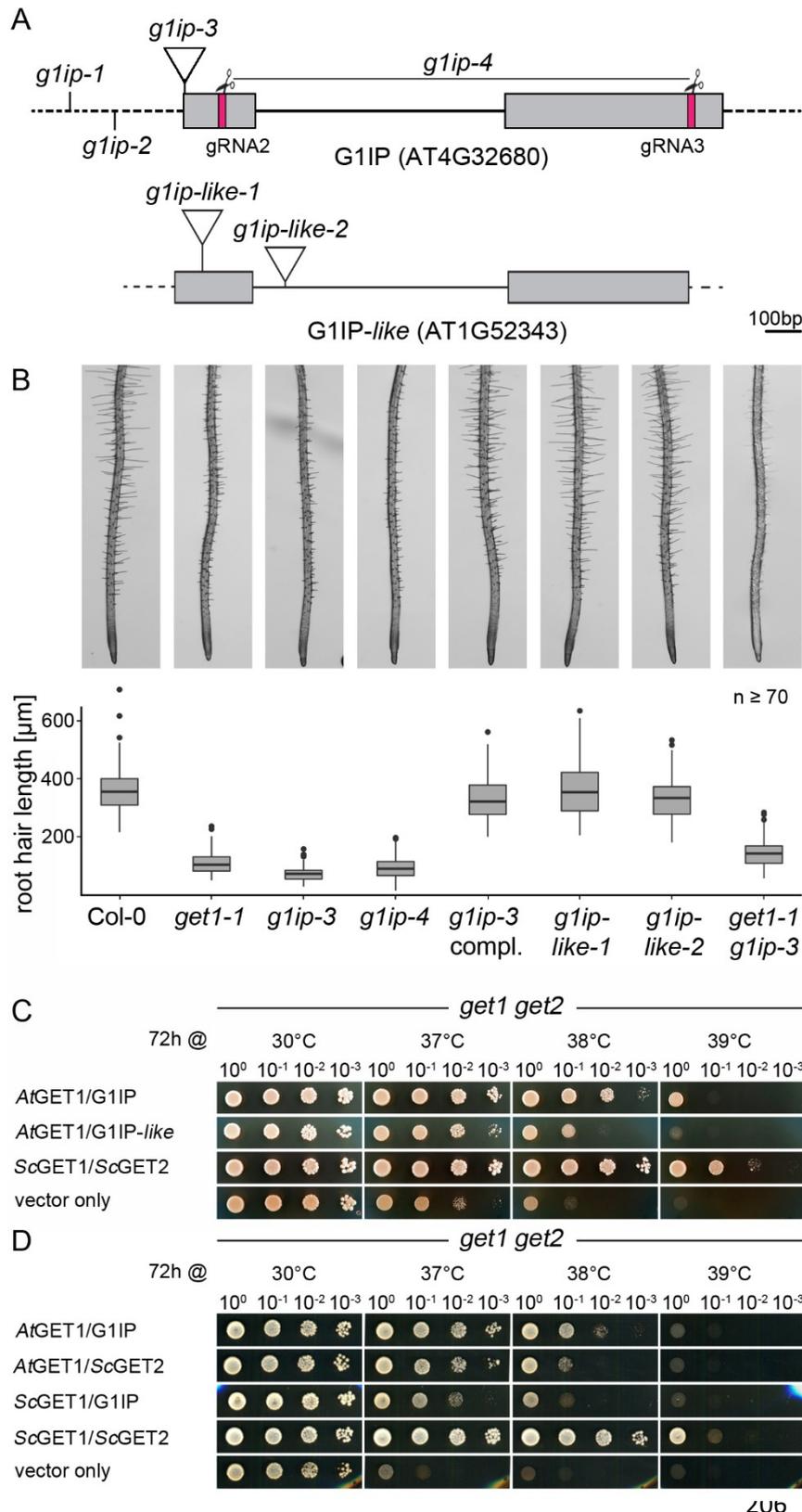
### 127 **G1IP binds *AtGET3a* only in the presence of *AtGET1***

128 To corroborate and expand the analyses of physical interaction of G1IP and *G1IP-like*  
129 with *Arabidopsis* GET pathway components we performed rBiFC [9, 11, 12]) and co-  
130 immunoprecipitation (Co-IP) analyses. Complementation of YFP signal, cue for  
131 physical interaction, was only detected in samples where *AtGET1* was coexpressed  
132 with G1IP or, *G1IP-like* (Figure 1K-M). Residual YFP signal in samples with *AtGET3a*  
133 was comparable to the biological negative control of *AtGET4*, a protein that is found  
134 further upstream of the pathway and unable to interact on its own with the receptors in  
135 yeast and mammals [13, 14]. As we had detected G1IP as binding partner of *AtGET3a*  
136 in our previously published IP-MS analyses [5], lack of an interaction in rBiFC was  
137 somewhat surprising.

138 We therefore generated a new set of Gateway-compatible 2in1 Co-IP vectors allowing  
139 for high constitutive gene coexpression in *Arabidopsis* (Figure 1N). Interestingly,  
140 interaction was only detected in wildtype (Figure 1O) but not in an *Atget1-2* mutant  
141 background (Figure 1P) suggesting that the interaction of *AtGET3a* and G1IP is highly  
142 sensitive to the presence or absence of *AtGET1* (Figure 4C). Recently, it had been  
143 demonstrated that the human *Get1* orthologue *Wrb* is required for protein stability and  
144 correct insertion of CAML, the *Get2* receptor in metazoa [15], however, we did not  
145 observe instability of ectopically expressed G1IP in *Atget1-2* mutants (Figure 1P).

### 146 **G1IP phenocopies GET pathway mutants**

147 We have previously shown that loss of some GET pathway components in *A. thaliana*  
148 leads to reduced root hair elongation under standard growth conditions [5]. To  
149 investigate whether G1IP belongs to the same pathway we analysed the root hair  
150 growth of putative *loss-of-function* lines (Figure 2A). The T-DNA insertion line *g1ip-3*  
151 showed significantly shorter root hairs at seedling level compared to wildtype Col-0  
152 and similar to the *A. thaliana* GET pathway mutant *get1-1* [5] (Figure 2B). Expression  
153 of a genomic version of the *G1IP* gene under the constitutively active VAMP721  
154 promoter (*'g1ip-3 compl.'*) restores wildtype like root hair growth.



**Figure 2:**

G1IP phenocopies GET pathway mutants in Arabidopsis and partially complements a yeast GET receptor mutant in combination with AtGET1. (A) Schematic illustration of the G1IP gene structure. The T-DNA in *g1ip-3* is inserted 5bp downstream of the ATG with an additional insertion of AGTT. In *g1ip-1* and *g1ip-2*, the T-DNA insertion is within the 5' UTR (dotted line), 333bp and 201bp upstream of the ATG, respectively. The *g1ip-4* line lacks the part between the CRISPR target sites indicated in red and symbolized by the scissors above. (B) Representative images of roots of 10-day-old mutant seedlings or complemented lines. Box plots show quantification of root hair length of the 10 longest root hairs from at least 7 seedlings per genotype. Centre lines of boxes represent median with outer limits at 25th and 75th percentile. Tukey whiskers extend to 1.5x IQR, outliers are depicted as black dots. (C, D) Yeast complementation analyses of the yeast  $\Delta get1 get2$  double-deletion strain with different combinations of *A. thaliana* and *S. cerevisiae* proteins. Growth was monitored after 3 days in different temperatures. Genomic fragments of yeast GET1 and GET2 were used as positive control, and empty vectors were used as negative control.

207 Since the T-DNA insertion in *g1ip-3* is located close to the ATG and in order to confirm  
 208 that the observed phenotype is a result of the insertion mutation in *G1IP*, we  
 209 additionally performed CRISPR/Cas9 genome editing to generate a *g1ip* complete

210 deletion mutant (*g1ip-4*). Root hair growth in this line was reduced, phenocopying the  
211 T-DNA line *g1ip-3* and thereby confirming that loss of G1IP leads to the reduced root  
212 hair growth. Simultaneous homozygous knockout of *AtGET1* and G1IP does not  
213 exacerbate the short root hair phenotype indicating that both genes may be part of the  
214 same pathway (Figure 2B).

215 In contrast, *g1ip-like* T-DNA insertion lines exhibit wildtype-like root hair growth without  
216 any significant growth defects at later stages.

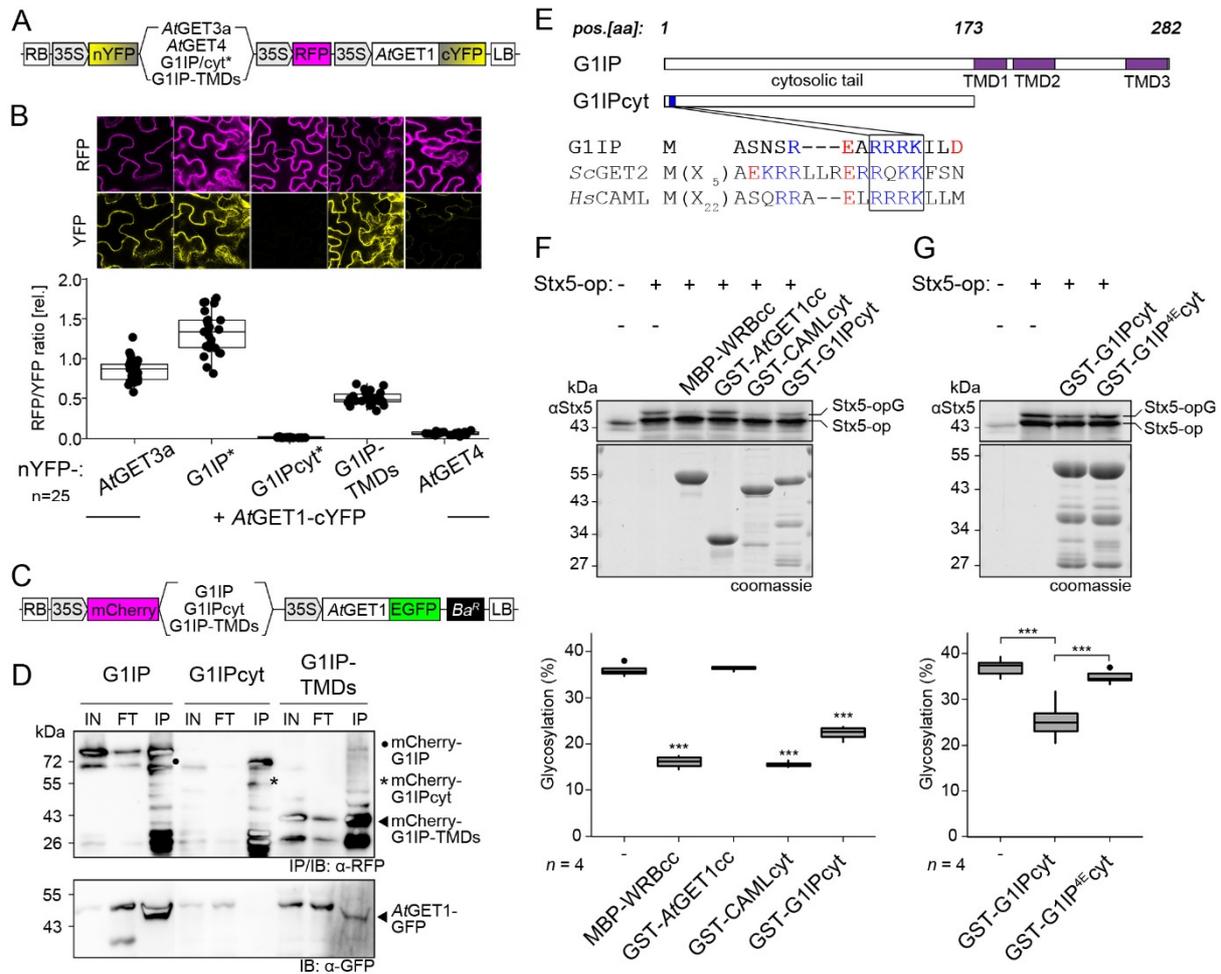
### 217 **G1IP in concert with *AtGET1* can complement yeast GET receptor mutants**

218 It had been demonstrated that loss of GET pathway components in yeast results in a  
219 lack of (heat) stress tolerance [16]. We therefore tested whether G1IP or G1IP-*like* are  
220 able to complement yeast growth under increasing temperatures (Figure 2C,D).  
221 Simultaneous expression of *AtGET1* and G1IP is able to weakly recover the viability of  
222 the  $\Delta get1get2$  strain [17] indicating, at least in part, functional conservation between  
223 the *Arabidopsis* and yeast genes (Figure 2C). However, coexpression of the  
224 *Arabidopsis* homolog G1IP-*like* together with *AtGET1* in  $\Delta get1get2$  is not able to  
225 rescue the lethality of higher temperatures, comparable to the vector only control. Lack  
226 of a noticeable phenotype in *g1ip-like* lines along with the different expression profile  
227 and lack of rescue of  $\Delta get1get2$  yeast strongly suggests that G1IP-*like* has acquired a  
228 novel function independent of the GET pathway.

229 In an additional approach we tested the importance of a hetero- or homologous partner  
230 receptor for yeast rescue (Figure 2D). Mixing the corresponding receptors of the  
231 different species did not rescue as efficiently as the homologous combinations of  
232 *AtGET1*/G1IP or *ScGET1*/*ScGET2*. It seemed however, that the combination of  
233 *ScGET1* with G1IP performed even weaker than the opposite combination with  
234 *AtGET1* and *ScGET2* mirroring an earlier observation with the mammalian GET2  
235 ortholog CAML [16]. This result implies that the eukaryotic Get2/CAML in general may  
236 have undergone more structural changes during evolution making it more specialised  
237 as opposed to the more conserved GET1/WRB.

### 238 **G1IP interacts with the *AtGET1* receptor via its TMDs**

239 Mammalian WRB and CAML were previously shown to associate via interactions  
240 between their TMDs thereby forming a functional receptor complex [7]. We therefore



241  
 242 **Figure 3:** The TMD region of G1IP mediates interaction with *AtGET1* and its cytosolic N-terminus can  
 243 interfere with the mammalian insertion system. (A) Schematic of the 2in1 rBiFC constructs used in (B).  
 244 (B) rBiFC analysis using full-length and truncated versions of G1IP to test for interaction with *AtGET1*.  
 245 Exemplary CLSM images of transiently transfected *N. benthamiana* leaves are depicted. Mean  
 246 fluorescence of at least 25 areas were measured in YFP and RFP channels, ratioed and plotted to show  
 247 YFP complementation. Centre lines of boxes represent median with outer limits at 25th and 75th  
 248 percentile. Tukey whiskers extend to 1.5x IQR, all values are depicted as black dots. (C) Schematic  
 249 of the 2in1 FRET constructs used for co-IP in (D). (D) Co-IP of full-length and truncated G1IP with *AtGET1*,  
 250 transiently expressed in *N. benthamiana* leaves. Protein extracts were immunoprecipitated with anti-  
 251 RFP beads and protein-protein interaction was detected by immunoblotting (IB) using anti-RFP and anti-  
 252 GFP antibody, respectively. IN, input; FT, flow-through; IP, immunoprecipitate. (E) Schematic  
 253 representation of full-length and truncated G1IP, and sequence logos highlighting a conserved cluster  
 254 of positively charged amino acids. (F, G) Insertion assays into microsomal membranes. Stx5-op was  
 255 translated *in vitro* in rabbit reticulocyte lysate and incubated with recombinant cytosolic fragments and  
 256 pancreatic rough microsomes. Protein extracts were immunoblotted with anti-Stx5 antibody and ER  
 257 insertion was monitored via band shift reporting glycosylation. Box plots show quantification of the  
 258 immunoblots from 4 independent experiments. Centre lines of boxes represent median with outer limits  
 259 at 25th and 75th percentile. Tukey whiskers extend to 1.5x IQR, outliers are depicted as black dots. \*\*\*  
 260 p < 0.001; Student's t-test.

261 examined the importance of the transmembrane region of G1IP on binding to *AtGET1*  
 262 using rBiFC and CoIP. We separated the cytosolic tail (1-173aa) of G1IP from its TMD  
 263 region (174-282aa) and tested both domains individually for *AtGET1* interaction  
 264 (Figure 3A,B). Interaction of full length G1IP with *AtGET1* in rBiFC resulted in strong

265 YFP complementation with a YFP to RFP ratio above the positive control *AtGET1* with  
266 *AtGET3a*. While the ratio was lower using the truncated construct G1IP-TMDs it  
267 nonetheless gave a strong signal of YFP complementation. The cytosolic part of G1IP,  
268 however, showed almost complete absence of signal comparable to the biological  
269 negative control of *AtGET1* and *AtGET4*.

270 The rBiFC result was corroborated via co-IP by leveraging a 2in1 FRET construct  
271 transiently transformed in *N. benthamiana* (Figure 3C). Fusion proteins of *AtGET1*-  
272 EGFP coexpressed with either mCherry-G1IP, mCherry-G1IPcyt or mCherry-G1IP-  
273 TMDs were purified from tobacco leaf extracts via the RFP-trap antibody. After  
274 complex elution, immunoblot against GFP revealed the presence of *AtGET1*-GFP in  
275 eluates of G1IP and G1IP-TMDs, but not of G1IPcyt (Figure 3D). Our results indicate  
276 that G1IP acts as binding partner of *AtGET1* via its TMDs.

### 277 **Interference of the cytosolic G1IP N-terminus in TA protein insertion**

278 Despite the low level of sequence similarity between G1IP and yeast Get2 or  
279 mammalian CAML, multiple protein sequence alignment showed that a cluster of  
280 positively charged amino acids near the N-terminus is conserved among the genomes  
281 of vertebrates, plants, and fungal lineages (Figure 3E, Figure S2). This motif is  
282 proposed to be crucial for binding of ScGet3 [18] and its mammalian homolog TRC40,  
283 respectively [7] and has recently been shown to segregate with the membrane-  
284 anchoring domain of Get2/CAML-like proteins in a position-specific iterative (PSI)-  
285 BLAST analysis [19]. To determine the functional effect of this cluster in G1IP, we  
286 performed site-directed substitution mutagenesis to reverse the charge of four amino  
287 acid residues (R9E, R10E, R11E, K12E = G1IP<sup>4E</sup>; Figure 3E).

288 We then *in vitro* expressed/translated the human Syntaxin5 (Stx5) fused to an opsin-  
289 tag (Stx5-op) in TNT reticulocyte lysate and added recombinant cytosolic fragments of  
290 MBP-WRBcc, GST-CAMLcyt, GST-*AtGET1*cc, GST-G1IPcyt and GST-G1IP<sup>4E</sup>cyt  
291 together with pancreatic rough microsomes (RM) to the reaction mix ('cc' refers to the  
292 cytosolic coiled-coil domain in WRB or *AtGET1*, 'cyt' refers to the cytosolic N-terminus  
293 of CAML or G1IP, respectively). The ratio of glycosylated and non-glycosylated Stx5-  
294 op was detected via band shift in immunoblot analyses and revealed that the native  
295 cytosolic domain of G1IP but not the reverse-charged mutant version (G1IP<sup>4E</sup>cyt)  
296 prevents insertion of the *in vitro* translated TA-protein Stx5 into ER-derived

297 microsomes (Figure 3 F, G). The interference of the native G1IP N-terminus with the  
298 mammalian machinery for TA protein insertion suggests a conserved role for this  
299 domain in binding of TRC40/GET3. The coiled-coil motif of *AtGET1*, however, does  
300 not inhibit membrane insertion, indicating that the binding sites or functional residues  
301 may have diverged from those of its orthologue in mammals. These functional  
302 differences are also evident from the yeast complementation assays (Figure 2C,D) and  
303 underpin the importance of the positively charged motif common to yeast *ScGet2*,  
304 mammalian CAML and *Arabidopsis* G1IP.

### 305 **The GET receptor complex shows low evolutionary conservation**

306 While interaction data and the root hair phenotype seem to confirm that *AtGET1* and  
307 G1IP act in the same pathway, sequence conservation of the two receptors compared  
308 to opisthokont candidates is poor (Figure S1A). Similarly, sequence conservation  
309 between fungal *Get2* and mammalian CAML is equally poor which led the authors who  
310 identified the connection to postulate “*mammalian cells have no genes homologous to*  
311 *Get2*” [7].

312 Our finding of G1IP, however, gave us an amino acid sequence with which we were  
313 able to identify numerous archaeplastidic homologs to compare with both fungal *GET2*  
314 and metazoan CAML sequences (Figure 4, Figure S3). The structural similarities of  
315 the cross-kingdom proteins are striking regarding the number of TMDs (three), the  
316 topology of the proteins (cytosolic N-, luminal C-terminus) and most importantly the  
317 positively charged N-terminus (at least 4 Arginine or Lysine residues in a row, see  
318 motifs in Figure 4). A recently published, independent analysis using PSI-BLAST  
319 showed that the N-terminal *Get3* interaction motif and the C-terminal membrane  
320 anchoring domain co-evolve and allow the identification of candidate *GET2* homologs  
321 from distantly related groups including plants [19].

322 Our phylogenetic analysis of (putative) *GET2* homologs from different eukaryotic  
323 groups clearly separates homologs from high-level groups (animals, fungi, plants)  
324 (Figure 4). Somewhat surprisingly, the Brassicales *GET2* homologs are clustered  
325 separately at the base of the eudicots. The *G1IP-like* proteins, which are only found in  
326 the Rosids, cluster as a separate branch. The most striking difference within the N-  
327 terminal *Get3* interaction motif is a conserved Alanine residue in G1IP and *GET2*  
328 orthologs (Figure S4). *G1IP-like* instead, features an additional Glutamic acid residue,



340 although other explanations involving gene duplications and losses cannot be  
341 excluded.

342 Taken together the structural similarities of G1IP with either fungal GET2 or metazoan  
343 CAML, the network of physical interactions with other components of the *Arabidopsis*  
344 GET pathway, complementation of yeast knockouts, as well as the phenocopying of  
345 the *loss-of-function Arabidopsis* mutants strongly suggests that we have indeed  
346 identified the functional ortholog of GET2 in *Arabidopsis*. This discovery is consistent  
347 with a recent, independent bioinformatic analysis [19] presenting candidate  
348 Get2/CAML homologs based on PSI BLAST and allows to recognize GET2/CAML  
349 orthologs in other higher plant species or even basal Archaeplastida (Table S1). In  
350 addition, we have identified a Rosid-lineage specific homolog G1IP-*like* that seems  
351 non-functional in the context of a plant GET pathway. This identification of the missing  
352 GET receptor in plants paves the way for future research into pathway function and  
353 conservation in the eukaryotic domain of life. The absence of a more severe growth  
354 defect in GET pathway mutants of *Arabidopsis* remains puzzling and suggests the  
355 presence of additional membrane targeting pathways and/or alternative functions of  
356 GET in plants.

357

### 358 **Acknowledgments**

359 This project was supported through a Carl Zeiss PhD fellowship to D.G.M. and funds  
360 of the German Research Foundation (DFG) to M.N. (NO 407/7-1), B.S. (SFB1190-  
361 P04), M.B. (SFB1101-B01) and C.G. (Emmy Noether fellowship GR 4251/1-1, and GR  
362 4251/1-2 and SFB1101-A06). We thank Simon Klesen for cloning and providing the  
363 pEC-CAS9 vector and Eva Schwörzer and Laure Grefen for excellent technical  
364 assistance. The IP-MS was performed by the Proteomics facility of the University of  
365 Tübingen and we are grateful to Mirita Franz for helpful advice during data analyses.

366

### 367 **Author Contributions**

368 Conceptualization, L.Y.A., S.X., B.S. and C.G.; Investigation, L.Y.A., D.G.M, J.R.M,  
369 M.M.R., N.W., S.X., M.B. and C.G.; Funding Acquisition, C.G.; Resources, H.B.,  
370 K.W.B., M.N., B.S. and M.B.; Writing, L.Y.A. and C.G. with input from all authors.

371

### 372 **Declaration of Interests**

373 The authors declare no competing interests.

## 374 **STAR Methods**

### 375 **Construct generation and plant transformation**

376 Most constructs were designed using Gateway technology or the Gateway-  
377 compatible cloning system - 2in1 -, respectively [9, 12, 21]. For generation of the  
378 reverse-charged mutation of G1IP, three arginine and one lysine residue at position  
379 9-12 were exchanged with glutamic acid residues by site-directed mutagenesis as  
380 described by [22].

381 *P<sub>VAMP721</sub>* >>GFP-myc-gG1IP was generated by classical cloning. The genomic  
382 fragment of G1IP from start codon to 261bp downstream of the stop codon was PCR-  
383 amplified and inserted into the binary vector *P<sub>VAMP721</sub>*>>GFP-myc 3' of myc.

384 Constructs were transformed into *Agrobacterium tumefaciens* GV3101 and used to  
385 transform Col-0 or respective mutant plants or infiltrated into *Nicotiana benthamiana*  
386 leaves [12]. For the CRISPR construct, annealed oligos (FW: 5'-ATTG + protospacer;  
387 REV: 5'-AAAC + rev-com protospacer) were sequentially ligated into pEn-2xChimera  
388 [23] via BbsI and Esp3I, respectively, followed by Gateway cloning into pEC-CAS9.  
389 Target sites (3'-AAGAAGTAGAATCGGAAGG-5'; 5'-GATGATGGTGAAGAAGATAA-  
390 3') were selected using CRISPR-P 2.0 [24]. Constructs were transformed into Col-0  
391 through floral dipping and T1 plants were selected by red fluorescence.

392

### 393 **Cloning of pEC-CAS9**

394 A modified version of pDe-CAS9 [25] containing pOLE-OLE-tagRFP was digested  
395 using EcoRI. The EC promoter [26] and Cas9-attR1 fragment [25] were PCR-  
396 amplified separately with overlapping ends and combined with the vector backbone  
397 by In-Fusion cloning. The resulting vector pEC-CAS9 was verified by restriction  
398 digest and sequencing.

399

### 400 **Plant material and growth conditions**

401 All mutant and transgenic lines used in this work are in Columbia (Col-0) background.  
402 T-DNA insertion lines were obtained from the Nottingham *Arabidopsis* Stock Centre  
403 (NASC, <http://arabidopsis.info/>) and insertion sites were verified by sequencing:  
404 *Atget1-1* (SAIL\_1210\_E07) [5], *Atget1-2* (GK\_264D06), *g1ip-1* (SALK\_100089), *g1ip-*  
405 *2* (SALK\_119358), *g1ip-3* (SALK\_034959), *g1ip-like-1* (SAIL\_760\_H02), *g1ip-like-2*  
406 (SALK\_045533).

407 The CRISPR based mutant line was generated with a dual sgRNA approach and  
408 screened by using a visual selection marker (FAST-Red). Expression of Cas9 was  
409 driven by the egg cell-specific promoter EC1. Large-fragment deletion mutants were  
410 identified by PCR-based genotyping and verified by sequencing. Primer sets used for  
411 genotyping are listed in Table S2.

412 Plants were grown at 22°C under long day conditions (16h light/8h dark) in soil or ½  
413 strength Murashige and Skoog (MS) agar plates (1%, pH 5.7). Seeds were surface-  
414 sterilized with chlorine gas and stratified at 4°C for 2-3 days in darkness to equalize  
415 germination.

416

### 417 **Ratiometric Bimolecular Fluorescence Complementation (rBiFC)**

418 Coding sequences were cloned into binary 2in1 rBiFC vectors [9] and transformed  
419 into *N. benthamiana* through syringe-mediated infiltration as described in [12].

420 Fluorescence intensities were measured 3 days post-infiltration using a Leica SP8  
421 confocal laser scanning microscope (YFP at 514nm excitation (ex) and 520 to 560nm  
422 emission (em); RFP at 561nm ex and 565 to 620nm em). YFP/RFP ratios were  
423 calculated from at least 21 different leaf regions and plotted using BoxPlotR  
424 (<http://shiny.chemgrid.org/boxplotr/>).

425

### 426 **Subcellular localisation analysis**

427 Coding sequences were cloned into the Gateway vector pH7WGF2 [27] and co-  
428 transformed with an ER membrane marker (CD3-959 or CD3-960) into Col-0 through  
429 floral dipping. T1 plants were selected on hygromycin and leaves were imaged using  
430 a Leica SP8 confocal laser scanning microscope (GFP at 488nm ex and 490 to  
431 520nm em; RFP at 561nm ex and 565 to 620nm em).

432

### 433 **Root hair imaging and measurements**

434 Roots from 10-d-old seedlings grown on ½ MS agar plates were imaged with a  
435 ZEISS Axio Zoom.V16 light microscope, and the length of the 10 longest root hairs  
436 from at least 7 seedlings per genotype were measured using ImageJ ( $n \geq 70$ ).

437

438 **qPCR analysis**

439 Total RNA was isolated from various plant tissues (100mg) using the GeneMATRIX  
440 Universal RNA Purification Kit (roboklon). 1µg of each sample was converted into  
441 complementary DNA (cDNA) by using the Protoscript II-First Strand cDNA Synthesis  
442 Kit (NEB). cDNA was diluted 1:5 and quantified on the CFX96 Real-Time PCR  
443 System (Bio-Rad) using GoTag® qPCR Master Mix (promega) with SYBR green.  
444 Transcript levels were calculated by the  $2^{-\Delta Ct}$  method and normalized to ACT2  
445 expression. Primer sets used for qPCR are listed in Table S2.

446

447 **Yeast complementation assay**

448 *S. cerevisiae* genes with part of the 5' and 3' flanking regions (~0.5kb) were cloned  
449 into low copy ARS/CEN vectors. *A. thaliana* genes (full-length CDS) were  
450 constitutively expressed from 2µ origin plasmids using the yeast PMA1 promoter.  
451 The *get/get2* double-deletion mutant (MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*  
452 *ygl020c::Kan<sup>R</sup> yer083c::Nat<sup>R</sup>*, [3]) was co-transformed as described in [28] and  
453 dropped in 10-fold serial dilutions on vector selective media (complete supplement  
454 media (CSM) L-, U-) and grown at different temperatures for 3 days.

455

456 **Creation of 2in1 Co-IP vectors (mVenus/3xHA)**

457 The new set of Gateway-compatible 2in1 Co-IP vectors (pCoIP-2in1-NN, -NC, -CN, -  
458 CC) was generated by classical cloning. RPS5a driven N- and C-terminally 3xHA-  
459 tagged R3R2 expression cassettes were generated by replacing the 35S promoter in  
460 pUC57-Tec-N-HA and pUC35S-R3R2-3xHA (NarI/HpaI), respectively, with the  
461 RPS5a promoter (1684bp) which was PCR amplified and flanked by NarI-StuI/NaeI  
462 (blunt end, like HpaI) restriction sites (pUC-RPS-HA-lacZ and pUC-RPS-lacZ-HA).  
463 The resulting expression cassettes were excised via StuI and inserted into pBBb  
464 [21] via EcoICRI (blunt end, like StuI) to yield the intermediate vectors pCoIP-intA and  
465 pCoIP-intB, respectively. Another pUC helper vector (pUC-RPS5a::R1R4) was  
466 created by introducing the RPS5a promoter via NarI/NaeI into pUC57-Tec-N-myc.  
467 mVenus was PCR amplified (NaeI/SpeI) and inserted via NaeI/HpaI 5' of the R1R4  
468 expression cassette (pUC-RPS5-Ven-R1R4). To introduce mVenus at the C-  
469 terminus, PCR-amplified mVenus-TGA (NaeI/PsiI) was inserted into pUC-  
470 RPS5a::R1R4 via PsiI (pUC-RPS5-R1R4-Ven). For the final 2in1 vector assembly,

471 the intermediate vectors pCoIP-intA and pCoIP-intB were linearized via AfeI and the  
472 R3R2 and R1R4 expression cassettes were inserted (StuI/FspI). All vectors were  
473 verified by restriction digest and sequencing.

474

#### 475 **Co-IP analysis – stable gene expression in *Arabidopsis***

476 3 grams of *Arabidopsis* seedlings were harvested after 10 days under continuous  
477 light. Cells were lysed by mortar grinding in liquid nitrogen and thawed in lysis buffer  
478 (50mM Tris pH 7.5, 150mM NaCl, 1% Triton X-100; 1.43ml per gram) supplemented  
479 with protease inhibitor cocktail (cOmplete EDTA-free®, Roche). Cell debris was  
480 removed by centrifugation and filtration through two layers of Miracloth. 2.5 ml  
481 supernatant were mixed with 2ml lysis buffer and incubated with anti-GFP beads  
482 (25µl, GFP-trap, Chromotek) for 2 hours at 4°C under mild rotation. Beads were  
483 collected by centrifugation, transferred on spin columns and washed six times using  
484 washing buffer (50mM Tris pH7.5, 150mM NaCl; 0.5% Triton X-100) supplemented  
485 with protease inhibitor cocktail. (Co-) immunoprecipitated proteins were eluted with  
486 2x Laemmli buffer (+3% β-mercaptoethanol) at 80°C for 5 min, separated by SDS-  
487 PAGE and detected by Western blotting (anti-HA-peroxidase from rat IgG1, Roche,  
488 1:1000; anti-GFP from mouse IgG1κ, Roche, 1:1000; anti-mouse IgG (Fc specific)  
489 produced in goat, Sigma, 1:10000).

490

#### 491 **Co-IP analysis – transient gene expression in *N. benthamiana***

492 FRET 2in1 destination vectors containing monomeric enhanced green fluorescent  
493 protein (mEGFP) and mCherry (pFRETgc-2in1) were used to transiently express  
494 recombinant proteins in *N. benthamiana* for Co-IP analysis [12, 21]. Leaf material  
495 (150-600mg) was harvested 3 days post-infiltration and homogenized after freezing  
496 in liquid nitrogen. Lysis buffer (25mM Tris pH 8.0, 150mM NaCl, 1% NP-40, 0.5% Na-  
497 deoxycholate(DOC)) supplemented with protease inhibitor cocktail and 2mM DTT  
498 was added and incubated for 1 hour at 4°C with mild rotation. After centrifugation, the  
499 supernatant was mixed with 20-25µl RFP-beads (RFP-trap, Chromotek) and  
500 incubated for 1 hour at 4°C with mild rotation. Beads were collected by centrifugation,  
501 transferred on spin columns and rinsed twice with lysis buffer followed by six washes  
502 with wash buffer (25mM Tris pH 8.0, 150mM NaCl). (Co-) immunoprecipitated  
503 proteins were eluted with 2x Laemmli buffer (+3.5% β-mercaptoethanol) and heated

504 at 65°C for 15 min (membrane proteins) or 95°C for 5 min (soluble proteins). Proteins  
505 were separated by SDS-PAGE and detected by Western blotting (anti-RFP from  
506 mouse, Chromotek, 1:2500; anti-GFP from mouse IgG1κ, Roche, 1:1000; anti-mouse  
507 IgG (Fc specific) produced in goat, Sigma, 1:10000).

508

### 509 **Protein purification**

510 *E. coli* BL21 DE3 cells were transformed with GST-tagged versions of the cytosolic  
511 portions of *AtGET1* and *G1IP*. Expression was induced with 200 μM IPTG in 1l of  
512 2YT-3% glycerol cultures at 0.5 OD600. Cell pellet was collected after 3 h at 30°C  
513 and lysed by sonification in ice-cold purification buffer (20 mM HEPES, 2% glycerol,  
514 150 mM potassium acetate, 5 mM magnesium acetate, 1mM EDTA, 1mM DTT, 1mM  
515 PMSF, pH 7.4). The lysate was cleared at 100.000 xg for 30 min and incubated with  
516 Glutathione Sepharose resin (GE Healthcare). After 1 hour of binding, the resin was  
517 washed sequentially with purification buffer, purification buffer containing 5 mM ATP  
518 and purification buffer for 10 min. GST-tagged protein was eluted with purification  
519 buffer containing 20 mM glutathione.

520 Expression of the N terminal domain of CAML (GST-CAMLcyt) and the WRB coiled-  
521 coil domain (MBP-WRBcc) was carried out as previously described [7, 29].

522

### 523 **Stx5op in vitro transcription/translation and insertion assay into microsomes**

524 Reactions were performed in the TnT Quick Coupled Transcription/Translation  
525 System (Promega) as previously described [30, 31] with some modifications. Stx5op  
526 synthesis was induced with 100 ng of pGem3z-Stx5op in 4.5 μL of TNT reticulocyte  
527 lysate for 90 min at 30°C. Where indicated, equimolar amounts (5 μM) of  
528 recombinant cytosolic fragments (MBP-WRBcc, GST-CAMLcyt, GST-*AtGET1*, GST-  
529 *G1IP*cyt and GST-*G1IP*<sup>4E</sup>cyt) and pancreatic rough microsomes (RM) were added to  
530 the reaction mix after Stx5 translation was completed. After 90 min of incubation at  
531 30°C with the RM, the reaction was stopped with SDS loading buffer and analysed by  
532 western blot with rabbit anti-Stx5 antibody (Synaptic Systems Cat. No. 110053).

533

### 534 **Multiple alignments and construction of phylogenetic trees**

535 Multiple alignments were generated with Muscle in MEGA6.06 [32, 33]. Phylogenetic  
536 analyses were performed with MrBayes 3.2.7a with 500,000 generations [34].

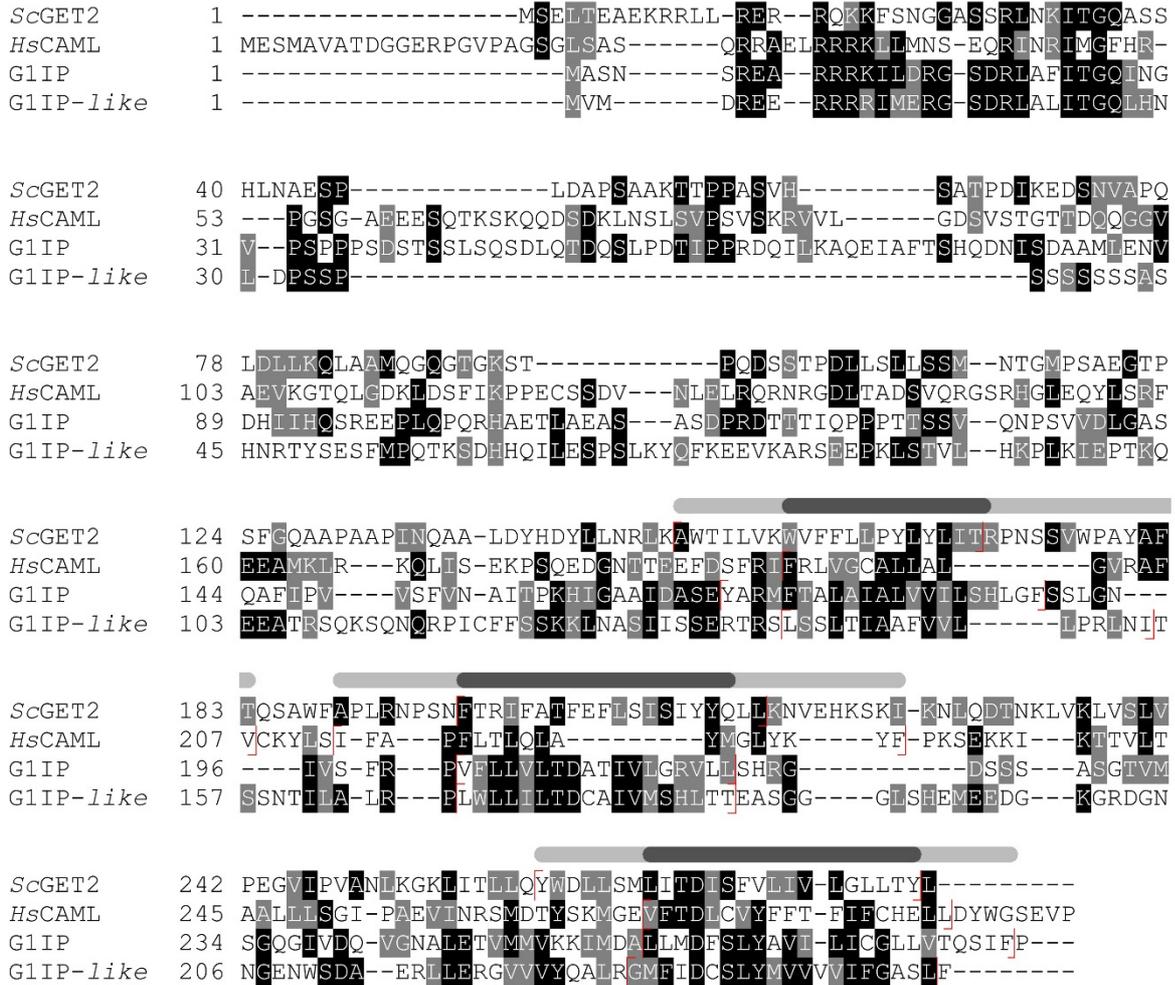
537 **References**

- 538 1. Borgese, N., Colombo, S., and Pedrazzini, E. (2003). The tale of tail-anchored proteins:  
539 coming from the cytosol and looking for a membrane. *The Journal of cell biology* *161*, 1013-  
540 1019.
- 541 2. Stefanovic, S., and Hegde, R.S. (2007). Identification of a targeting factor for posttranslational  
542 membrane protein insertion into the ER. *Cell* *128*, 1147-1159.
- 543 3. Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., Schwappach, B.,  
544 and Weissman, J.S. (2008). The GET complex mediates insertion of tail-anchored proteins  
545 into the ER membrane. *Cell* *134*, 634-645.
- 546 4. Srivastava, R., Zalisko, B.E., Keenan, R.J., and Howell, S.H. (2017). The GET System Inserts the  
547 Tail-Anchored Protein, SYP72, into Endoplasmic Reticulum Membranes. *Plant physiology* *173*,  
548 1137-1145.
- 549 5. Xing, S., Mehlhorn, D.G., Wallmeroth, N., Asseck, L.Y., Kar, R., Voss, A., Denninger, P.,  
550 Schmidt, V.A., Schwarzlander, M., Stierhof, Y.D., et al. (2017). Loss of GET pathway orthologs  
551 in *Arabidopsis thaliana* causes root hair growth defects and affects SNARE abundance.  
552 *Proceedings of the National Academy of Sciences of the United States of America* *114*,  
553 E1544-E1553.
- 554 6. Auld, K.L., Hitchcock, A.L., Doherty, H.K., Fietze, S., Huang, L.S., and Silver, P.A. (2006). The  
555 conserved ATPase Get3/Arr4 modulates the activity of membrane-associated proteins in  
556 *Saccharomyces cerevisiae*. *Genetics* *174*, 215-227.
- 557 7. Yamamoto, Y., and Sakisaka, T. (2012). Molecular machinery for insertion of tail-anchored  
558 membrane proteins into the endoplasmic reticulum membrane in mammalian cells.  
559 *Molecular cell* *48*, 387-397.
- 560 8. Omasits, U., Ahrens, C.H., Muller, S., and Wollscheid, B. (2014). Protter: interactive protein  
561 feature visualization and integration with experimental proteomic data. *Bioinformatics* *30*,  
562 884-886.
- 563 9. Grefen, C., and Blatt, M.R. (2012). A 2in1 cloning system enables ratiometric bimolecular  
564 fluorescence complementation (rBiFC). *BioTechniques* *53*, 311-314.
- 565 10. Mergner, J., Frejno, M., List, M., Papacek, M., Chen, X., Chaudhary, A., Samaras, P., Richter,  
566 S., Shikata, H., Messerer, M., et al. (2020). Mass-spectrometry-based draft of the *Arabidopsis*  
567 proteome. *Nature* *579*, 409-414.
- 568 11. Xing, S., Wallmeroth, N., Berendzen, K.W., and Grefen, C. (2016). Techniques for the Analysis  
569 of Protein-Protein Interactions in Vivo. *Plant physiology* *171*, 727-758.
- 570 12. Mehlhorn, D.G., Wallmeroth, N., Berendzen, K.W., and Grefen, C. (2018). 2in1 Vectors  
571 Improve In Planta BiFC and FRET Analyses. *Methods in molecular biology* *1691*, 139-158.
- 572 13. Chang, Y.W., Chuang, Y.C., Ho, Y.C., Cheng, M.Y., Sun, Y.J., Hsiao, C.D., and Wang, C. (2010).  
573 Crystal structure of Get4-Get5 complex and its interactions with Sgt2, Get3, and Ydj1. *The*  
574 *Journal of biological chemistry* *285*, 9962-9970.
- 575 14. Chang, Y.W., Lin, T.W., Li, Y.C., Huang, Y.S., Sun, Y.J., and Hsiao, C.D. (2012). Interaction  
576 surface and topology of Get3-Get4-Get5 protein complex, involved in targeting tail-anchored  
577 proteins to endoplasmic reticulum. *The Journal of biological chemistry* *287*, 4783-4789.
- 578 15. Carvalho, H.J.F., Del Bondio, A., Maltecca, F., Colombo, S.F., and Borgese, N. (2019). The WRB  
579 Subunit of the Get3 Receptor is Required for the Correct Integration of its Partner CAML into  
580 the ER. *Scientific reports* *9*, 11887.
- 581 16. Vilaridi, F., Stephan, M., Clancy, A., Janshoff, A., and Schwappach, B. (2014). WRB and CAML  
582 are necessary and sufficient to mediate tail-anchored protein targeting to the ER membrane.  
583 *PloS one* *9*, e85033.
- 584 17. Powis, K., Schrul, B., Tienson, H., Gostimskaya, I., Breker, M., High, S., Schuldiner, M., Jakob,  
585 U., and Schwappach, B. (2013). Get3 is a holdase chaperone and moves to deposition sites

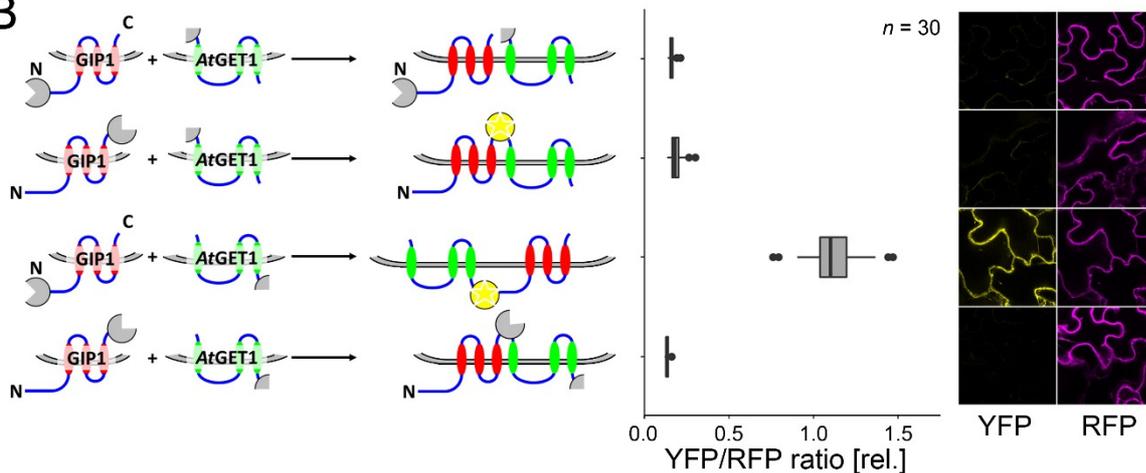
- 586 for aggregated proteins when membrane targeting is blocked. *Journal of cell science* *126*,  
587 473-483.
- 588 18. Stefer, S., Reitz, S., Wang, F., Wild, K., Pang, Y.Y., Schwarz, D., Bomke, J., Hein, C., Lohr, F.,  
589 Bernhard, F., et al. (2011). Structural basis for tail-anchored membrane protein biogenesis by  
590 the Get3-receptor complex. *Science* *333*, 758-762.
- 591 19. Borgese, N. (2020). Searching for Remote Homologs of Caml among Eukaryotes. *Traffic*.
- 592 20. Barker, M.S., Vogel, H., and Schranz, M.E. (2009). Paleopolyploidy in the Brassicales: analyses  
593 of the Cleome transcriptome elucidate the history of genome duplications in Arabidopsis and  
594 other Brassicales. *Genome Biol Evol* *1*, 391-399.
- 595 21. Hecker, A., Wallmeroth, N., Peter, S., Blatt, M.R., Harter, K., and Grefen, C. (2015). Binary  
596 2in1 vectors improve in planta (co-) localisation and dynamic protein interaction studies.  
597 *Plant physiology* *168*, 776-787.
- 598 22. Karnik, A., Karnik, R., and Grefen, C. (2013). SDM-Assist software to design site-directed  
599 mutagenesis primers introducing "silent" restriction sites. *BMC bioinformatics* *14*, 105.
- 600 23. Durr, J., Papareddy, R., Nakajima, K., and Gutierrez-Marcos, J. (2018). Highly efficient  
601 heritable targeted deletions of gene clusters and non-coding regulatory regions in  
602 Arabidopsis using CRISPR/Cas9. *Scientific reports* *8*, 4443.
- 603 24. Liu, H., Ding, Y., Zhou, Y., Jin, W., Xie, K., and Chen, L.L. (2017). CRISPR-P 2.0: An Improved  
604 CRISPR-Cas9 Tool for Genome Editing in Plants. *Molecular plant* *10*, 530-532.
- 605 25. Fauser, F., Schiml, S., and Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases  
606 can be used efficiently for genome engineering in Arabidopsis thaliana. *The Plant journal : for*  
607 *cell and molecular biology* *79*, 348-359.
- 608 26. Wang, Z.P., Xing, H.L., Dong, L., Zhang, H.Y., Han, C.Y., Wang, X.C., and Chen, Q.J. (2015). Egg  
609 cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for  
610 multiple target genes in Arabidopsis in a single generation. *Genome Biol* *16*, 144.
- 611 27. Karimi, M., Depicker, A., and Hilson, P. (2007). Recombinational cloning with plant gateway  
612 vectors. *Plant physiology* *145*, 1144-1154.
- 613 28. Asseck, L.Y., Wallmeroth, N., and Grefen, C. (2018). ER Membrane Protein Interactions Using  
614 the Split-Ubiquitin System (SUS). *Methods in molecular biology* *1691*, 191-203.
- 615 29. Vilaridi, F., Lorenz, H., and Dobberstein, B. (2011). WRB is the receptor for TRC40/Asna1-  
616 mediated insertion of tail-anchored proteins into the ER membrane. *Journal of cell science*  
617 *124*, 1301-1307.
- 618 30. Pfaff, J., Rivera Monroy, J., Jamieson, C., Rajanala, K., Vilaridi, F., Schwappach, B., and  
619 Kehlenbach, R.H. (2016). Emery-Dreifuss muscular dystrophy mutations impair TRC40-  
620 mediated targeting of emerin to the inner nuclear membrane. *Journal of cell science* *129*,  
621 502-516.
- 622 31. Rivera-Monroy, J., Musiol, L., Unthan-Fechner, K., Farkas, A., Clancy, A., Coy-Vergara, J.,  
623 Weill, U., Gockel, S., Lin, S.Y., Corey, D.P., et al. (2016). Mice lacking WRB reveal differential  
624 biogenesis requirements of tail-anchored proteins in vivo. *Scientific reports* *6*, 39464.
- 625 32. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular  
626 Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution* *30*, 2725-2729.
- 627 33. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high  
628 throughput. *Nucleic acids research* *32*, 1792-1797.
- 629 34. Ronquist, F., and Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under  
630 mixed models. *Bioinformatics* *19*, 1572-1574.

631

A



B



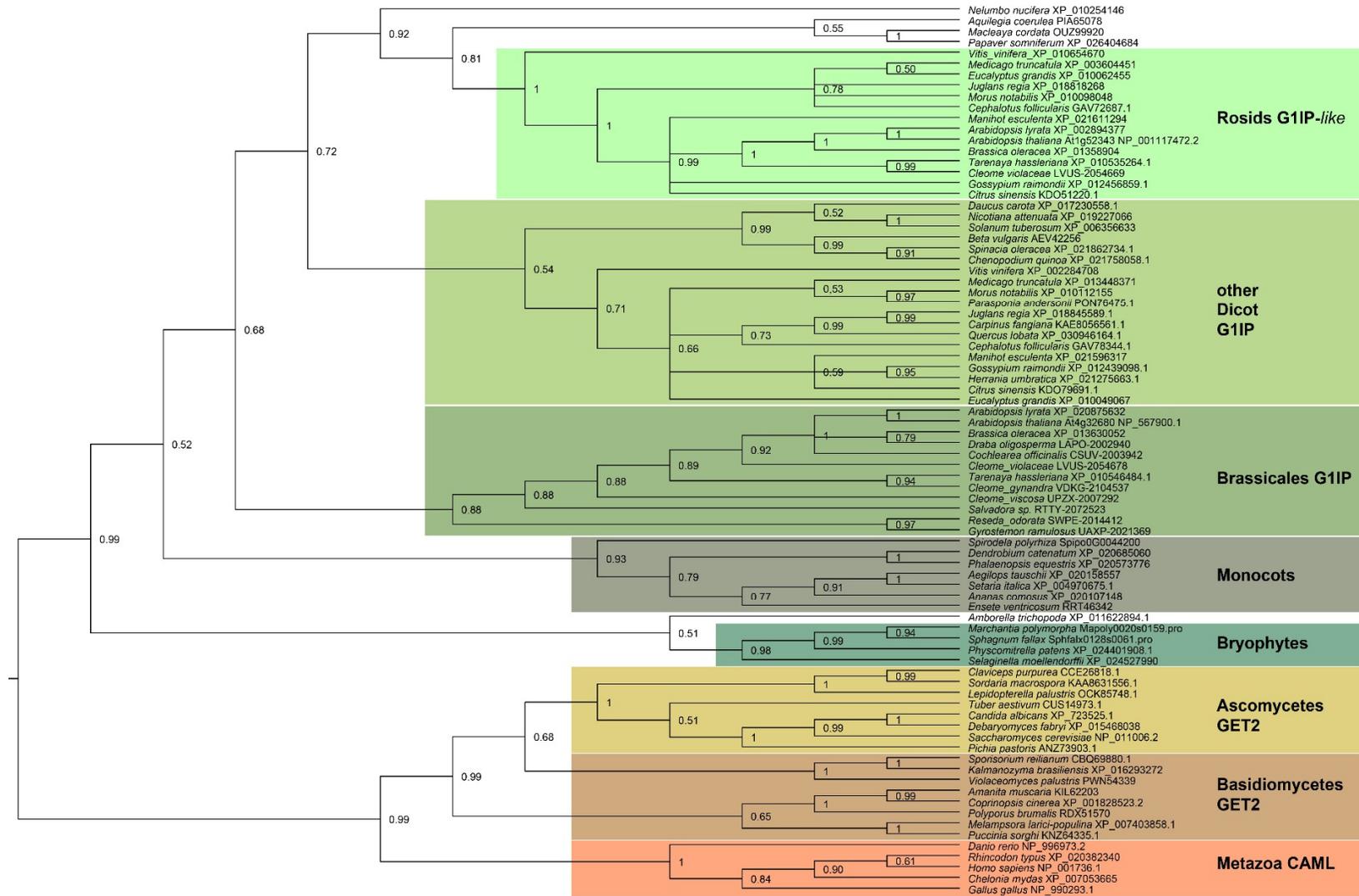
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12

**Supplemental Figure S1:** (A) Multiple sequence alignment of *HsCam1*, *ScGET2*, *Arabidopsis* *G1IP* and *G1IP-like*. The grey bars above the aligned sequences represent the predicted transmembrane domain helices (Protter). Red horizontal lines within each sequence mark beginning and end of the predicted helices. (B) Topological analysis of the putative receptor pair *AtGET1* and *G1IP*. Four different orientations for N-terminal or C-terminal tagging of YFP halves were fused to either *AtGET1* or *G1IP* and analysed via rBiFC for complementation of signal to verify presence of both termini in either cytosolic or luminal side of the ER membrane. In principle, two combinations should show fluorescence, however, only N-terminally tagged *G1IP* and C-terminally tagged *AtGET1* yielded significant YFP complementation. Failure of fluorescence complementation in the reciprocal interaction pair may be due to a different pH or redox state in the ER lumen or masking of the *AtGET1* N-terminus which may lead to incomplete or aberrant membrane insertion.

|                                |                             |                                     |                             |
|--------------------------------|-----------------------------|-------------------------------------|-----------------------------|
| G1IP-like                      | Rosids                      | <i>Vitis vinifera</i>               | SREA--RRKKIMERE-SDRLALITGR  |
|                                |                             | <i>Cephalotus follicularis</i>      | SRDE--RRRRIVERG-TDRLALITGR  |
|                                |                             | <i>Medicago truncatula</i>          | NTEE--RRRRIQERG-SDRMALITGR  |
|                                |                             | <i>Juglans regia</i>                | GSEE--RRRRILERG-SDRMALITGQ  |
|                                |                             | <i>Manihot esculenta</i>            | VRQE--RRRKIVERG-GDRLALITGQ  |
|                                |                             | <i>Morus notabilis</i>              | NREE--RRRRIVERG-SDRMALITGR  |
|                                |                             | <i>Arabidopsis lyrata</i>           | DREE--RRRRIMERG-SDRLALITGQ  |
|                                |                             | <i>Arabidopsis thaliana</i>         | DREE--RRRRIMERG-SDRLALITGQ  |
|                                |                             | <i>Brassica oleracea</i>            | DREE--RRRRIMERG-SDRLALITGQ  |
|                                |                             | <i>Tarenaya hassleriana</i>         | DREE--RRRRIMERG-SDRLAMITGQ  |
|                                |                             | <i>Gossypium raimondii</i>          | EKFE--RKRRAEKG-LNRMSQIRSA   |
|                                |                             | <i>Eucalyptus grandis</i>           | SREE--RRRKIVDGG-SDRLALITGH  |
|                                |                             | <i>Citrus sinensis</i>              | SREE--RRKRILDRG-SDRLALISGR  |
|                                |                             | G1IP                                | Dicots                      |
| <i>Solanum tuberosum</i>       | GRDT--RRRRIVERG-NDRLALITGR  |                                     |                             |
| <i>Beta vulgaris</i>           | GREA--RRRRILKGG-QDRLALITGR  |                                     |                             |
| <i>Cephalotus follicularis</i> | SREG--RRRKIVERG-SDRLALITGR  |                                     |                             |
| <i>Vitis vinifera</i>          | SREA--RRRRILERG-SDRLALITGR  |                                     |                             |
| <i>Medicago truncatula</i>     | SREAQRRRRRILQQG-SDRLAFITKGH |                                     |                             |
| <i>Quercus lobata</i>          | VREA--RRRKIMERG-SDRLALITGQ  |                                     |                             |
| <i>Carpinus fangiana</i>       | VREA--RRRKIVERG-ADRLALIAGR  |                                     |                             |
| <i>Juglans regia</i>           | VREA--RRRKILERG-SDRLALIAGR  |                                     |                             |
| <i>Manihot esculenta</i>       | AREY--RRKKILDRG-ADRLAFIAGR  |                                     |                             |
| <i>Parasponia andersonii</i>   | SREA--RRRKILERG-SDRLALITGQ  |                                     |                             |
| <i>Morus notabilis</i>         | SREA--RRKKIMERS-SDRLALITSR  |                                     |                             |
| <i>Arabidopsis lyrata</i>      | SREA--RRRKILERG-SDRLAFITGQ  |                                     |                             |
| <i>Arabidopsis thaliana</i>    | SREA--RRRKILDRG-SDRLAFITGQ  |                                     |                             |
| <i>Brassica oleracea</i>       | SREA--RRRKILERG-SDRLAFITGQ  |                                     |                             |
| <i>Tarenaya hassleriana</i>    | SREA--RRRKIIERG-SDRLAFITGQ  |                                     |                             |
| <i>Herrania umbratica</i>      | TRES--RRRKIVERG-SDRLAFIKGR  |                                     |                             |
| <i>Gossypium raimondii</i>     | ARES--RRRKILDRG-SDRLAYITGQ  |                                     |                             |
| <i>Eucalyptus grandis</i>      | NRES--RRRRIVDRG-SDRLALITGR  |                                     |                             |
| <i>Citrus sinensis</i>         | SRES--RRRKILERG-SDRLAFVTGR  |                                     |                             |
|                                | Monocots                    | <i>Dendrobium catenatum</i>         | ARES--RRKRLAERK-TDRLAFITGQ  |
|                                |                             | <i>Aegilops tauschii</i>            | PRANT-RRRRLVERG-SDRLAFITGQ  |
|                                |                             | <i>Ananas comosus</i>               | ARDA--RRRRILERG-SNRLAFITGE  |
|                                | Gymnosperms et al.          | <i>Amborella trichopoda</i>         | EREA--RRKRIMERG-SDRLAFITGQ  |
|                                |                             | <i>Ginkgo biloba</i>                | AREA--RRKKIMERG-ADRLAFITGE  |
|                                |                             | <i>Picea engelmannii</i>            | AREA--RRRKILERG-TDRLALITGE  |
|                                |                             | <i>Metasequoia glyptostroboides</i> | AREV--RRRKILERG-ADRLAFITSGD |
|                                | Bryophytes                  | <i>Marchantia polymorpha</i>        | AREA--RRRKILARG-ADRLAYITGE  |
|                                |                             | <i>Physcomitrella patens</i>        | LREA--RRRKILERG-NDRLAFITGQ  |
|                                |                             | <i>Selaginella moellendorffii</i>   | AREA--RRQRILRRG-ADRLAFITGE  |
|                                |                             | <i>Sphagnum fallax</i>              | AREA--RRRKILERG-SDRLAFITGE  |
| GET2                           | Fungi                       | <i>Saccharomyces cerevisiae</i>     | LRER--RQKKFSNGGASSRLNKITGQ  |
|                                |                             | <i>Pichia pastoris</i>              | LRER--RKAKLSQGGGLDRLKKITGE  |
|                                |                             | <i>Sordaria macrospora</i>          | RKER--REAKIRANA-GNRLNRTIGL  |
|                                |                             | <i>Tuber aestivum</i>               | RRER--RQAKVKEGG-SSRLNRTIGT  |
|                                |                             | <i>Coprinopsis cinerea</i>          | RAEA--RRRAILSRG-SDRLAKLITTS |
|                                |                             | <i>Sporisorium reilianum</i>        | KREA--RKARILSKG-SDRLARITNT  |
|                                |                             | <i>Puccinia sorghi</i>              | RAEA--RKQKILSKG-NDRLAKITGA  |
| CAML                           | Metazoa                     | <i>Danio rerio</i>                  | QRRAEIRRRKLLNNS-EDRMNRIVGF  |
|                                |                             | <i>Rhincodon typus</i>              | RRRAELRRRKLLNNS-EERLNRIMGF  |
|                                |                             | <i>Chelonia mydas</i>               | QRRAELRRRKLLMNS-EERINRIMGF  |
|                                |                             | <i>Gallus gallus</i>                | QRRAELRRRKLLMNS-EERINRIMGF  |
|                                |                             | <i>Homo sapiens</i>                 | QRRAELRRRKLLMNS-EQRINRIMGF  |

13  
14  
15  
16

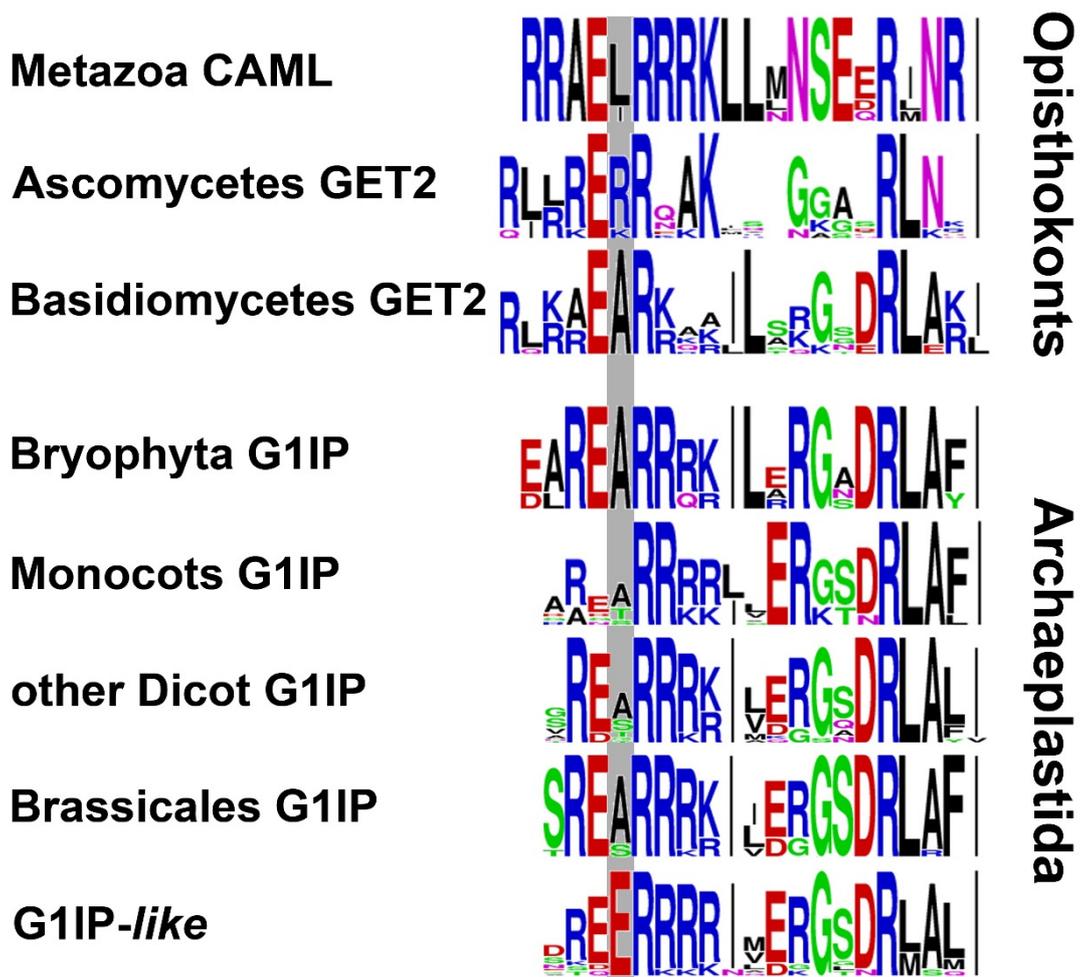
**Supplemental Figure S2:** Multiple sequence alignment of the conserved N-terminal motif of Archaeplastida G1IP and homologous fungal GET2 and metazoan CAML proteins



17

18 **Supplemental Figure S3:** Phylogenetic tree of GET2 and G1IP-like homologous proteins. A multiple alignment was generated with Muscle and the phylogenetic tree  
 19 generated with MrBayes. Bayesian probabilities are given at the branches.

20 The tree was rooted with the branch leading to the animal/fungal (ascomycota and basidiomycota) lineages. Accession numbers or locus tag numbers or sequence  
 21 numbers from the 1000 plant transcriptomes initiative [1] are given after the species names.



22  
 23 **Supplemental Figure S4:** N-terminal sequence motifs of GET2, CAML, G1IP and G1IP-like homologs  
 24 assembled from the alignment in Figure S2 and visualised using Weblogo  
 25 (<https://weblogo.berkeley.edu/logo.cgi>). The grey vertical bar underneath the motifs highlights the conserved  
 26 Glutamic acid residue in G1IP-like sequences opposed to all other eukaryotic homologs.  
 27  
 28  
 29

**Table S1.** Accession numbers of sequences used for multiple alignments and phylogenetic analyses of GET2, CAML, G1IP and G1IP-like proteins.

| Group                 | Species                        | Accession number or locus tag number |
|-----------------------|--------------------------------|--------------------------------------|
| Rosida<br>G1IP-like   | <i>Arabidopsis lyrata</i>      | XP_002894377                         |
|                       | <i>Arabidopsis thaliana</i>    | NP_001117472.2                       |
|                       | <i>Brassica oleracea</i>       | XP_01358904                          |
|                       | <i>Cephalotus follicularis</i> | GAV72687.1                           |
|                       | <i>Citrus sinensis</i>         | KDO51220.1                           |
|                       | <i>Cleome violaceae</i>        | LVUS-2054669 <sup>1</sup>            |
|                       | <i>Eucalyptus grandis</i>      | XP_010062455                         |
|                       | <i>Gossypium raimondii</i>     | XP_012456859.1                       |
|                       | <i>Juglans regia</i>           | XP_018818268                         |
|                       | <i>Manihot esculenta</i>       | XP_021611294                         |
|                       | <i>Medicago truncatula</i>     | XP_003604451                         |
|                       | <i>Morus notabilis</i>         | XP_010098048                         |
|                       | <i>Tarenaya hassleriana</i>    | XP_010535264.1                       |
|                       | <i>Vitis vinifera</i>          | XP_010654670                         |
| Brassicales<br>G1IP   | <i>Arabidopsis lyrata</i>      | XP_020875632                         |
|                       | <i>Arabidopsis thaliana</i>    | NP_567900.1                          |
|                       | <i>Brassica oleracea</i>       | XP_013630052                         |
|                       | <i>Cleome gynandra</i>         | VDKG-2104537 <sup>1</sup>            |
|                       | <i>Cleome violaceae</i>        | LVUS-2054678 <sup>1</sup>            |
|                       | <i>Cleome viscosa</i>          | UPZX-2007292 <sup>1</sup>            |
|                       | <i>Cochlearea officinalis</i>  | CSUV-2003942 <sup>1</sup>            |
|                       | <i>Draba oligosperma</i>       | LAPO-2002940 <sup>1</sup>            |
|                       | <i>Gyrostemon ramulosus</i>    | UAXP-2021369 <sup>1</sup>            |
|                       | <i>Reseda odorata</i>          | SWPE-2014412 <sup>1</sup>            |
|                       | <i>Salvadora sp.</i>           | RTTY-2072523 <sup>1</sup>            |
|                       | <i>Tarenaya hassleriana</i>    | XP_010546484.1                       |
| other Dicot<br>G1IP   | <i>Beta vulgaris</i>           | AEV42256                             |
|                       | <i>Carpinus fangiana</i>       | KAE8056561.1                         |
|                       | <i>Cephalotus follicularis</i> | GAV78344.1                           |
|                       | <i>Chenopodium quinoa</i>      | XP_021758058.1                       |
|                       | <i>Citrus sinensis</i>         | KDO79691.1                           |
|                       | <i>Daucus carota</i>           | XP_017230558.1                       |
|                       | <i>Eucalyptus grandis</i>      | XP_010049067                         |
|                       | <i>Gossypium raimondii</i>     | XP_012439098.1                       |
|                       | <i>Herrania umbratica</i>      | XP_021275663.1                       |
|                       | <i>Juglans regia</i>           | XP_018845589.1                       |
|                       | <i>Manihot esculenta</i>       | XP_021596317                         |
|                       | <i>Medicago truncatula</i>     | XP_013448371                         |
|                       | <i>Morus notabilis</i>         | XP_010112155                         |
|                       | <i>Nicotiana attenuata</i>     | XP_019227066                         |
|                       | <i>Parasponia andersonii</i>   | PON76475.1                           |
|                       | <i>Quercus lobata</i>          | XP_030946164.1                       |
|                       | <i>Spinacia oleracea</i>       | XP_021862734.1                       |
|                       | <i>Solanum tuberosum</i>       | XP_006356633                         |
| <i>Vitis vinifera</i> | XP_002284708                   |                                      |
|                       | <i>Aquilegia coerulea</i>      | PIA65078                             |
|                       | <i>Macleaya cordata</i>        | OUZ99920                             |
|                       | <i>Nelumbo nucifera</i>        | XP_010254146                         |
|                       | <i>Papaver somniferum</i>      | XP_026404684                         |
| Monocots              | <i>Aegilops tauschii</i>       | XP_020158557                         |
|                       | <i>Ananas comosus</i>          | XP_020107148                         |
|                       | <i>Dendrobium catenatum</i>    | XP_020685060                         |
|                       | <i>Ensete ventricosum</i>      | RRT46342                             |
|                       | <i>Phalaenopsis equestris</i>  | XP_020573776                         |
|                       | <i>Setaria italica</i>         | XP_004970675.1                       |
|                       | <i>Spirodela polyrhiza</i>     | Spipo0G0044200                       |
|                       | <i>Ginkgo biloba</i>           | SGTW-2038521 <sup>1</sup>            |

|                     |                                     |                           |
|---------------------|-------------------------------------|---------------------------|
| Gymnosperms<br>etc. | <i>Metasequoia glyptostroboides</i> | NRXL-2062375 <sup>1</sup> |
|                     | <i>Picea engelmannii</i>            | AWQB-2010070 <sup>1</sup> |
|                     | <i>Amborella trichopoda</i>         | XP_011622894.1            |
| Bryophyta           | <i>Marchantia polymorpha</i>        | Mapoly0020s0159           |
|                     | <i>Physcomitrella patens</i>        | XP_024401908.1            |
|                     | <i>Sphagnum fallax</i>              | Sphfalx0128s0061          |
|                     | <i>Selaginella moellendorffii</i>   | XP_024527990              |
| Ascomycetes         | <i>Candida albicans</i>             | XP_723525.1               |
|                     | <i>Claviceps purpurea</i>           | CCE26818.1                |
|                     | <i>Debaryomyces fabryi</i>          | XP_015468038              |
|                     | <i>Lepidopterella palustris</i>     | OCK85748.1                |
|                     | <i>Pichia pastoris</i>              | ANZ73903.1                |
|                     | <i>Saccharomyces cerevisiae</i>     | NP_011006.2               |
|                     | <i>Sordaria macrospora</i>          | KAA8631556.1              |
|                     | <i>Tuber aestivum</i>               | CUS14973.1                |
| Basidiomycetes      | <i>Amanita muscaria</i>             | KIL62203                  |
|                     | <i>Coprinopsis cinerea</i>          | XP_001828523.2            |
|                     | <i>Kalmanozyma brasiliensis</i>     | XP_016293272              |
|                     | <i>Melampsora larici-populina</i>   | XP_007403858.1            |
|                     | <i>Polyporus brumalis</i>           | RDX51570                  |
|                     | <i>Puccinia sorghi</i>              | KNZ64335.1                |
|                     | <i>Sporisorium reilianum</i>        | CBQ69880.1                |
|                     | <i>Violaceomyces palustris</i>      | PWN54339                  |
| Metazoa             | <i>Chelonia mydas</i>               | XP_007053665              |
|                     | <i>Danio rerio</i>                  | NP_996973.2               |
|                     | <i>Gallus gallus</i>                | NP_990293.1               |
|                     | <i>Homo sapiens</i>                 | NP_001736.1               |
|                     | <i>Rhincodon typus</i>              | XP_020382340              |

32 <sup>1</sup>Sequence from the 1000 plant transcriptomes initiative [1]

33

34 **Table S2: List of primers used in this study**

| #    | 5'-3' Sequence   | Purpose                           |
|------|--|-----------------------------------|
| 439  | ATGGAAGGAGAGAAGCTTATAGAAG                                      | qRT-PCR for <i>AtGET1</i>         |
| 134  | AGCCTCTCTCAAAGCTGCTTAATTTT                                     | qRT-PCR for <i>AtGET1</i>         |
| 1408 | ATTGGTTTCCTCTTTTCCTCGCTCCG                                     | qRT-PCR for G1IP                  |
| 1799 | GCCGTTGATCTGACCAGTGATA   | qRT-PCR for G1IP                  |
| 1781 | ATGGTGATGGATAGAGAAGAAAGG                                       | qRT-PCR for G1IP-like             |
| 1953 | GAGAAGCCGATGATGAGGAAGA   | qRT-PCR for G1IP-like             |
| 1672 | GCCATCCAAGCTGTTCTCTC   | qRT-PCR for ACT2                  |
| 1673 | CAGTAAGGTCACGTCCAGCA   | qRT-PCR for ACT2                  |
| 2316 | ATTGAAGAAGTAGAATCGGAAGG  | CRISPR of G1IP (gRNA2)            |
| 2317 | AAACCCTCCGATTCTACTTCTT   | CRISPR of G1IP (gRNA2)            |
| 2318 | ATTGATGATGGTGAAGAAGATAA  | CRISPR of G1IP (gRNA3)            |
| 2319 | AACTTATCTTCTTACCATCAT  | CRISPR of G1IP (gRNA3)            |
| 1249 | TACTGGGCCCATGGCGTCCAACAGCAGAGAAGCC                             | Genomic fragment of G1IP          |
| 1250 | GGACTAGTAATCTCAAACAAGAAAAAATACAC                               | Genomic fragment of G1IP          |
| 133  | TGAAGGCTTCAAATTTCTGTGAATCC                                     | Genotyping of <i>get1-1</i>       |
| 134  | AGCCTCTCTCAAAGCTGCTTAATTTT                                     | Genotyping of <i>get1-1</i>       |
| 1093 | TTGCAGCGATTGCATCTCCCTCTC                                       | Genotyping of <i>g1ip-3</i>       |
| 1094 | CGATTTCTTGAGCTTTAAGAATCTG                                      | Genotyping of <i>g1ip-3</i>       |
| 2350 | ACACTTGAATTGGCCCGTTAAGAAG                                      | Genotyping of <i>g1ip-4</i>       |
| 2351 | GCAAAACACAAATCTACCGAGCACA                                      | Genotyping of <i>g1ip-4</i>       |
| 1434 | TTCTTCCCTGCTTTGATGGATG   | Genotyping of <i>g1ip-4</i>       |
| 1496 | GAACATAGGGAAGAATTCATCTTTC                                      | Genotyping of <i>g1ip-like1/2</i> |
| 1497 | TGAAGAACAGTCGAGAGTTTTGGTTC                                     | Genotyping of <i>g1ip-like1/2</i> |
| 1888 | CAGAGAAGCCGAGGAGGAGATTCTAGATAGAGGATCTG                         | SDM on G1IP                       |
| 1889 | CTATCTAGAATCTCCTCCTCCTCGGCTTCTCTGCTG                           | SDM on G1IP                       |
| 1056 | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGCGTCGAA<br>CAGCAGAG         | pDONR207-G1IP                     |
| 1134 | GGGGACCACTTTGTACAAGAAAGCTGGGTCTAAGGAAAGAT<br>GCTTTGGGTGAC      | pDONR207-G1IP                     |
| 1594 | GGGGACCACTTTGTACAAGAAAGCTGGGTCTCGAGTTACATC<br>CGTGCGTATTCCGAAG | pDONR207-G1IP <sub>cyt</sub>      |
| 2043 | GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGTTACACAGC<br>TCTTGCGATTG     | pDONR207-G1IP-TMDs                |
| 1582 | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGTGATGGA<br>TAGAGAAGAAAGGA   | pDONR207-G1IP-like                |
| 1583 | GGGGACCACTTTGTACAAGAAAGCTGGGTTTAAAAAAGAGAG<br>GCTCCAAAAATAACA  | pDONR207-G1IP-like                |

36 **References**

- 37 1. One Thousand Plant Transcriptomes, I. (2019). One thousand plant transcriptomes and the  
38 phylogenomics of green plants. *Nature* 574, 679-685.

39

### **III. Asseck and Grefen, 2018**



## Detecting Interactions of Membrane Proteins: The Split-Ubiquitin System

Lisa Yasmin Asseck and Christopher Grefen

### Abstract

The *in vivo* analysis of protein–protein interactions (PPIs) is a critical factor for gaining insights into cellular mechanisms and their biological functions. To that end, a constantly growing number of genetic tools has been established, some of which are using baker’s yeast (*Saccharomyces cerevisiae*) as a model organism. Here, we provide a detailed protocol for the yeast mating-based split-ubiquitin system (mbSUS) to study binary interactions among or with full-length membrane proteins in their native subcellular environment. The system is based on the reassembly of two autonomously non-functional ubiquitin moieties attached to proteins of interest (POIs) into a native-like molecule followed by the release of a transcription factor. Upon its nuclear import, the activation of reporter gene expression gives a visual output via growth on interaction-selective media. Additionally, we apply a modification of the classical split-ubiquitin technique called CytoSUS that detects interactions of non-membrane/soluble proteins in their full-length form via translational fusion of an ER membrane anchor.

**Key words** Protein–protein interaction, Yeast, Split-ubiquitin, mbSUS, CytoSUS, PCA, Membrane proteins, Gateway

---

## 1 Introduction

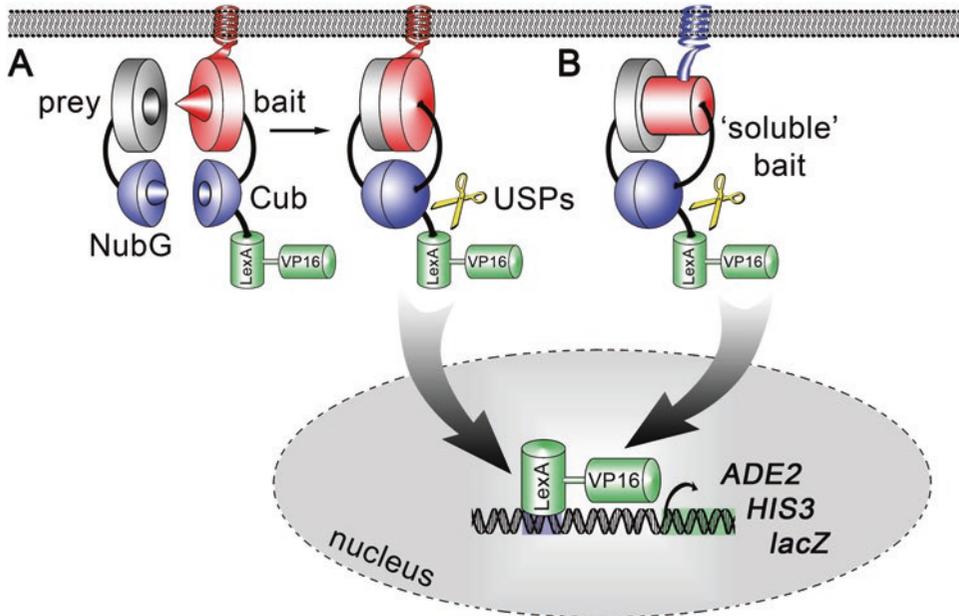
Protein–protein interactions (PPIs) are crucial to various aspects of cellular functions such as signaling, transport, metabolism, and catabolism. Nowadays a multitude of tools is available to characterize complex protein networks for a better understanding of cellular mechanisms [1].

The first and still one of the most prominent *in vivo* technique for detecting PPIs is the yeast two-hybrid system (Y2H) invented in 1989 [2] and eponymous for this book. This method relies on the reconstitution of the yeast Gal4p transcription factor that is separated into two autonomously functional protein fragments: a DNA-binding and activation domain. Upon interaction of two proteins of interest (POIs), which are fused to these domains, a chimeric transcription factor upstream of the

reporter genes is created. The readout of the activated reporters is monitored by either growth on depleted medium (medium without adenine and/or histidine, respectively) or colorimetric assays (*lacZ*). The inherent functionality which both domains maintain despite their truncation is prerequisite for a “two-hybrid” technique in contrast to later developed Protein-Fragment Complementation Assays (PCAs) [1, 3]. However, the domains are only functional in the nucleus requiring the interaction to be monitored there, which is one of the biggest drawbacks of the Y2H system as this necessitates truncation and mislocalization of integral or membrane-associated proteins, factors that might create artifactual results.

An alternative in vivo method to identify potential interactions among or with full-length membrane proteins in their native cellular context is the split-ubiquitin system (SUS) [4]. This method is not a two-hybrid approach but a PCA using two non-functional domains as probes. Here, ubiquitin is split into two fragments, a N-terminal Nub (amino acids 1–34) and a C-terminal Cub (amino acids 35–76) which is linked to the artificial transcription factor PLV (ProteinA-LexA-VP16) [5]. The Cub moiety is fused to the cytosolic terminus of a membrane-attached or -integrated protein (“bait”) and the Nub moiety is conjugated to putative binding partners (“preys”) that can either be membrane-associated or soluble (*see* **Notes 1–3**). Spontaneous reassembly of the two ubiquitin moieties is inhibited by a single point mutation of Ile-13 in the N-terminal fragment to either Gly (NubG) or Ala (NubA). When brought into close proximity via interacting proteins fused to Cub and Nub, respectively, the reconstituted ubiquitin molecule is recognized by ubiquitin-specific proteases (USPs) subsequently leading to the release of the LexA-VP16 transcript activator into the cytosol. The transcription factor is then translocated into the cell nucleus to induce transcriptional activation of reporter genes allowing auxotrophy selection (*ADE2*, *HIS3*) and quantification of the relative interaction strength (*lacZ*) (Fig. 1). The SUS has also been used in a mating-based approach (mbSUS, [6]). This facilitates not only investigating the interaction between two known proteins [7–10] but is particularly useful for high-throughput screening of protein binding partners [11] (*see* **Note 4**).

The CytoSUS is an adaption of the classical SUS to determine the interaction with soluble baits [12]. Here, an OST4p (Oligosaccharyltransferase 4) transmembrane domain is attached to the N-terminus of the Cub fusion to artificially anchor the protein to the ER membrane, thus preventing diffusion into the nucleus and activation of reporter genes due to its PLV fusion (Fig. 1, *see* **Note 5**). We had previously modified the SUS bait vector pMetYC-Dest [11] and inserted the coding sequence of the OST4p membrane anchor between the methionine repressible promoter *MET25* and the Gateway cassette [13].



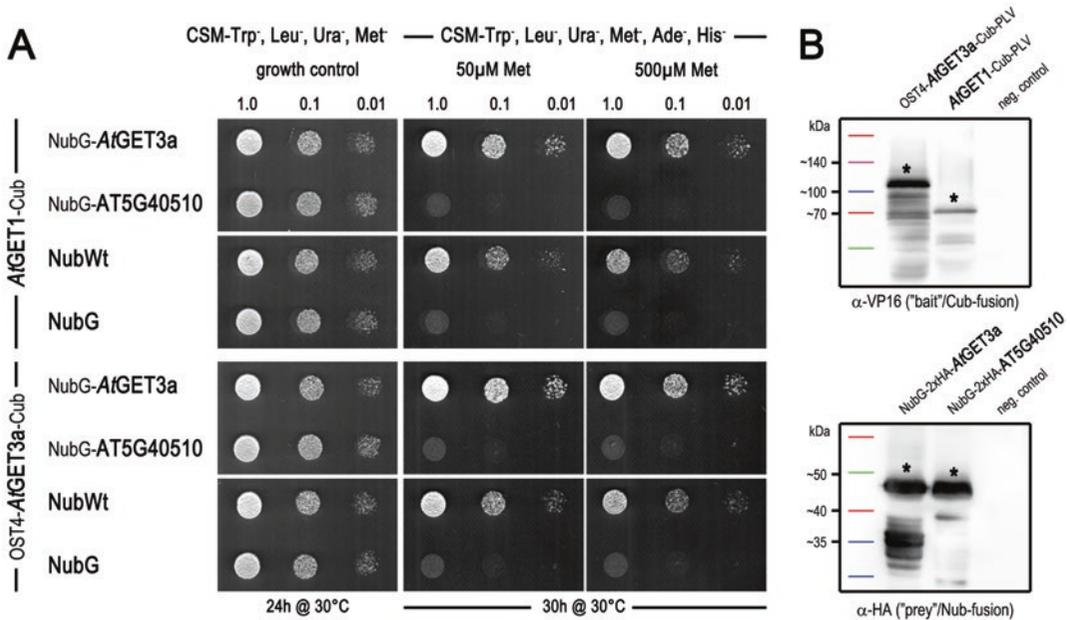
**Fig. 1** Schematic illustration of the classical SUS and CytoSUS. The ubiquitin moieties NubG and Cub (blue half-spheres) are fused to two POIs, whereby the bait protein needs to be attached or integrated into a membrane either through an intrinsic transmembrane domain (red helix; **a**) or an artificial N-terminal membrane anchor domain (the transmembrane domain of Oligosaccharyltransferase 4, OST4p, blue helix; **b**). Interaction of bait (red) and prey (grey) enables reconstitution of functional ubiquitin leading to the release of the LexA-VP16 transcription factor via cleavage by ubiquitin-specific proteases (USPs) and initiation of reporter gene transcription (*ADE2*, *HIS3*, *lacZ*) upon nuclear import. The prototrophic markers *ADE2* and *HIS3* allow qualitative evaluation of PPIs via growth on selective medium whereas the *lacZ* gene enables semiquantitative readout via blue/white coloring of colonies. (Figure modified from [1])

In this chapter, we detail the application of both mbSUS and CytoSUS using the ER receptor *AtGET1* and the cytosolic ATPase *AtGET3a*, respectively, as examples. Both proteins belong to the recently identified 'Guided-Entry of Tail-anchored proteins (GET) pathway' in *Arabidopsis thaliana*, which mediates insertion of tail-anchored (TA) proteins into the ER membrane [10]. We demonstrate that the SUS approach can be used with both membrane (*AtGET1*) and soluble proteins (*AtGET3a*) as bait (Fig. 2).

## 2 Materials

### 2.1 Vectors and Strains

A list of Gateway-compatible (exception: pNubWt-Xgate) mbSUS and CytoSUS vectors is given in Table 1. Maps and sequences of these vectors can be downloaded from <http://www.zmbp.uni-tuebingen.de/dev-genetics/grefen/resources/yeast-vectors.html> or [https://www.addgene.org/Christopher\\_Grefen/](https://www.addgene.org/Christopher_Grefen/). Table 2 shows genotypes of yeast strains used in this book chapter. Plasmids are available through Addgene, yeast strains via ABRC ([www.arabidopsis.org](http://www.arabidopsis.org), stock-# CD3-808 and CD3-809).



**Fig. 2** Mating-based SUS and CytoSUS analysis of *AtGET* pathway orthologs. **(a)** Growth assay of diploid yeast expressing the indicated fusion proteins. Yeast was dropped in serial dilutions (of OD<sub>600</sub> from 1.0 to 0.01) on vector-selective (CSM-Leu<sup>-</sup>, Trp<sup>-</sup>, Ura<sup>-</sup>) and interaction-selective (CSM-Leu<sup>-</sup>, Trp<sup>-</sup>, Ura<sup>-</sup>, Met<sup>-</sup>, Ade<sup>-</sup>, His<sup>-</sup>) media with different methionine concentrations. NubWt (=Nubl) was used as positive control, NubG as negative control (see **Notes 5** and **9**). **(b)** Western blot analysis of haploid yeast shown in **(a)** using antibodies against the VP16 domain within PLV and the HA-tag, respectively. OST4-*AtGET3a*-Cub-PLV (~99 kDa), *AtGET1*-Cub-PLV (~74 kDa), NubG-2xHA-*AtGET3a* (~47 kDa), NubG-2xHA-*AT5G40510* (~44 kDa)

## 2.2 Growth and Transformation of Yeast

1. YPD media: 2% peptone, 2% glucose, 1% yeast extract; adjust pH to 6–6.3 with KOH before adding 2% oxid agar.
2. Sterile deionized water (ddH<sub>2</sub>O).
3. Sterile PCR strips/lids and PCR cyler.
4. 1 M lithium acetate (LiAc): dissolve LiAc in ddH<sub>2</sub>O. Adjust the pH to 7.5 with acetic acid, sterilize by filtration.
5. 50% polyethylene glycol 3350 (PEG 3350): dissolve PEG 3350 in ddH<sub>2</sub>O to a final concentration of 50% (w/v), sterilize by filtration. Avoid water loss through autoclaving or during storage as this significantly decreases the transformation efficiency.
6. Single-stranded carrier DNA (ssDNA): dissolve 10 mg/ml ssDNA in ddH<sub>2</sub>O, sonicate, and/or boil for 10 min following cooling on ice before use.
7. CSM-Ade<sup>-</sup>, His<sup>-</sup>, Leu<sup>-</sup>, Met<sup>-</sup>, Trp<sup>-</sup>, Ura<sup>-</sup> as dropout.
8. Chemicals for auxotrophy selection, each dissolved in 100 ml water and sterilized by filtration; store in darkness at 4 °C:
  - ADE: 0.4 g of adenine sulfate (add 5 ml per liter media).
  - HIS: 0.4 g of L-histidine-HCl (add 5 ml per liter media).

**Table 1**  
**Destination vectors used for mbSUS and CytoSUS**

| Plasmid name  | Promoter | Origin         |         | Selection      |       | Function   | References |
|---------------|----------|----------------|---------|----------------|-------|--|------------|
|               |          | <i>E. coli</i> | Yeast   | <i>E. coli</i> | Yeast |  |            |
| pMetYC-Dest   | MET25    | pUC            | ARS/CEN | Amp, Cm        | LEU2  | Met-repressible fusion protein with C-terminal Cub-PLV                             | [11]       |
| pMetOYC-Dest  | MET25    | pUC            | ARS/CEN | Amp, Cm        | LEU2  | Met-repressible fusion protein with N-terminal OST4p anchor and C-terminal Cub-PLV | [13]       |
| pNX35-Dest    | ADH1     | pUC            | 2 $\mu$ | Amp, Cm        | TRP1  | Constitutive expression with N-terminal NubG-2xHA                                  | [9]        |
| pXNubA22-Dest | ADH1     | pUC            | 2 $\mu$ | Amp, Cm        | TRP1  | Constitutive expression with C-terminal NubA-3xHA                                  | [14]       |
| pNubWt-Xgate  | ADH1     | pUC            | 2 $\mu$ | Amp, Cm        | TRP1  | Positive control vector, NubWt peptide; not a Gateway vector                       | [6]        |

**Table 2**  
**Yeast strain genotypes used for mbSUS and CytoSUS**

| Name        | Organism             | Genotype  | Function  | References |
|-------------|----------------------|---|---|------------|
| THY.<br>AP4 | <i>S. cerevisiae</i> | MAT $\alpha$ ; <i>ade2</i> <sup>-</sup> , <i>his3</i> <sup>-</sup> , <i>leu2</i> <sup>-</sup> , <i>trp1</i> <sup>-</sup> , <i>ura3</i> <sup>-</sup> ; <i>lexA::ADE2</i> , <i>lexA::HIS3</i> , <i>lexA::lacZ</i> | Reporter yeast strain, used for transformation of Cub-clones                      | [15]       |
| THY.<br>AP5 | <i>S. cerevisiae</i> | MAT $\alpha$ ; <i>ade2</i> <sup>-</sup> , <i>his3</i> <sup>-</sup> , <i>leu2</i> <sup>-</sup> , <i>trp1</i> <sup>-</sup>  | Used for transformation of Nub-clones; mate with THY. AP4 for binary interactions | [15]       |

LEU: 2.0 g of L-leucine (add 5 ml per liter media).

TRP: 1.0 g of L-tryptophan (add 5 ml per liter media).

URA: 0.4 g of uracil (add 5 ml per liter media).

MET: 1.5 g of L-methionine (equals a 100 mM stock; add appropriate amount to obtain 0.5, 5, 50, and 500  $\mu$ M final concentrations).

9. Selection media: 0.17% YNB (without amino acids), 0.5% ammonium sulfate, 2% glucose, 0.056% CSM-dropout mix; adjust pH to 6–6.3 with KOH before adding 2% oxid agar; add appropriate auxotrophy selection chemicals before or after autoclaving, e.g., ADE, HIS, TRP, and URA for transformation of THY.AP4 in the mbsUS assay.

### **2.3 Western Blot Analysis**

1. Lyse and load (LL-) buffer: 50 mM Tris (pH 6.8—HCl), 2% SDS, 7 M urea, 30% glycerol, 0.1 M DTT, 0.04% bromophenol blue; store at  $-20^{\circ}\text{C}$ .
2. Acid-washed glass beads (diam.  $\sim 0.25\text{--}0.5$  mm).
3. SDS-PAGE resolving gel (10%): 3.4 ml  $\text{H}_2\text{O}$ , 4.0 ml acrylamide mix (30%), 4.5 ml bottom buffer (1 M Tris—HCl pH 8.8, 0.27% SDS; sterilize by filtration), 0.1 ml  $(\text{NH}_4)_2\text{S}_2\text{O}_8$  (10%), 0.008 ml TEMED.
4. SDS-PAGE stacking gel (4.5%): 1.4 ml  $\text{H}_2\text{O}$ , 0.6 ml acrylamide mix (30%), 2.0 ml upper buffer (0.25 M Tris—HCl pH 6.8, 0.2% SDS; sterilize by filtration), 0.02 ml  $(\text{NH}_4)_2\text{S}_2\text{O}_8$  (10%), 0.004 ml TEMED.
5. 10 $\times$  SDS running buffer: 250 mM Tris, 1.9 M glycine, 1.5% SDS.
6. 100% methanol.
7. PVDF membrane.
8. Transfer buffer: 25 mM Tris, 190 mM glycine, 20% EtOH.
9. 10 $\times$  TBS: 500 mM Tris, 1.5 M NaCl; adjust pH 7.5 (HCl).
10. Washing buffer, 1 $\times$  TBS—Tween: 100 ml 10 $\times$  TBS, 900 ml  $\text{H}_2\text{O}$ , 0.1% Tween 20.
11. Blocking buffer: 1 $\times$  TBS—Tween, 5% milk powder.
12. Antibodies
  - (a) Primaries:  $\alpha$ -VP16 (rabbit),  $\alpha$ -HA-HRP (dilute 1:1000 in 1 $\times$  TBS—Tween, add 0.1%  $\text{NaN}_3$ ).
  - (b) Secondaries: goat anti-rabbit IgG-HRP (dilute 1:25000 in 1 $\times$  TBS—Tween, add 0.1%  $\text{NaN}_3$ ).
13. Chemiluminescent substrate for detecting horseradish peroxidase (HRP)-conjugated antibodies.

---

## **3 Methods**

### **3.1 Yeast Transformation**

1. Streak THY.AP4 and THY.AP5 yeast strains out on YPD plates and incubate for 2 days at  $30^{\circ}\text{C}$ .
2. Prepare precultures by separately inoculating 5 ml YPD liquid media with a single colony of each strain and grow overnight at  $30^{\circ}\text{C}$  while shaking.

3. Transfer 2 ml of the precultured yeast to 100 ml of fresh YPD each and grow for 3–5 h at 30 °C while shaking until an OD<sub>600</sub> of 0.5–0.8 is reached.
4. Centrifuge the cells (5 min at 2000 × *g*) using sterile 50 ml tubes and discard the supernatant.
5. Wash with 20 ml of sterile ddH<sub>2</sub>O and pellet the cells by centrifugation (5 min at 2000 × *g*). Discard the supernatant.
6. Resuspend the cells in 1 ml of 0.1 M LiAc and transfer to a 2 ml tube. Spin down (2 min at 1000 × *g*) and remove the supernatant.
7. Resuspend the cell pellet in an appropriate amount of 0.1 M LiAc (20 μl per transformation) and incubate at room temperature for 30 min.
8. Meanwhile prepare sterile PCR strips with 10 μl ssDNA (boiled and cooled on ice) and 5 μl of plasmid DNA for each transformation (*see* **Notes 6** and **7**).
9. Make a master mix by combining 70 μl of 50% PEG (viscous—pipette slowly!), 10 μl 1 M LiAc, and 20 μl of competent yeast cells (**step 7**) for each transformation. Calculate for one extra transformation reaction and mix well until the solution is homogenous.
10. Add 100 μl of the master mix to each PCR tube and mix carefully with the prepared DNA mixture.
11. Incubate for 20 min at 30 °C using a PCR cycler. Mix the reactions by gently pipetting up and down several times with a multichannel pipette or by briefly vortexing the tubes (3–5 s).
12. Incubate for an additional 10 min at 30 °C.
13. Heat-shock the cells at 43 °C for 15 min.
14. Spin down briefly. Carefully remove the supernatant using a pipette.
15. *Optional*: Wash the pellet with 100 μl of sterile ddH<sub>2</sub>O.
16. Resuspend the cells in 100 μl of sterile ddH<sub>2</sub>O.
17. Plate 100 μl of the transformation mixture on appropriate selective minimal media (THY.AP4: CSM-Leu<sup>-</sup>, THY.AP5: CSM-Trp<sup>-</sup>, Ura<sup>-</sup>) using sterile glass beads (diam. ~2.85–3.45 mm).
18. Seal the plates and incubate for 2–4 days at 30 °C.

### 3.2 Mating

1. Pool several colonies of the transformed yeast and grow shaking overnight in 5 ml of appropriate selective medium at 30 °C.
2. Harvest 2 ml each by centrifugation for western blot analysis (*see* Subheading 3.4). Discard the supernatant and store the pellet at –20 °C (*see* **step 1** of Subheading 3.4).

3. Harvest another 2 ml each by centrifugation (5 min at  $1000 \times g$ ) and remove the supernatant. Gently resuspend the cell pellet in 200  $\mu$ l YPD. Scale up the volume for a higher number of crossings (20  $\mu$ l per mating).
4. Mix 20  $\mu$ l each, bait (THY.AP4) and prey (THY.AP5) of any desired combination in sterile PCR strips. Be careful not to cross-contaminate neighboring samples.
5. Carefully drop 5  $\mu$ l of each mating onto an YPD plate.
6. Incubate (right side up) for approximately 6–8 h at 30 °C.
7. Transfer mated yeast on CSM-Leu<sup>-</sup>, Trp<sup>-</sup>, Ura<sup>-</sup> plates using sterile pipette tips or a replicator stamp. Be careful not to transfer YPD medium with the cells as this will allow growth of non-mated/haploid parental cells.
8. Incubate overnight at 30 °C (*see Note 8*).

### 3.3 Detection Assay

1. Use overnight culture from **step 8** Subheading 3.2 to inoculate 2 ml selective media (CSM-Leu<sup>-</sup>, Trp<sup>-</sup>, Ura<sup>-</sup>) and grow shaking overnight at 30 °C. Be careful not to take too much cell material as this will lead to unspecific background growth associated with a high number of dead cells (*see Note 9*).
2. Pipette 100  $\mu$ l in a 1.5 ml tube and 100  $\mu$ l into a cuvette containing 900  $\mu$ l H<sub>2</sub>O.
3. Determine 1:10 diluted OD<sub>600</sub>. Note the values.
4. Harvest cells in the tubes by centrifugation (2 min at  $2000 \times g$ ). Remove the supernatant by pipetting. Be careful not to aspirate the cell pellet.
5. Add the appropriate volume of sterile ddH<sub>2</sub>O to reach a final OD<sub>600</sub> of 1.0 (e.g., 1:10 dilution has an OD<sub>600</sub> value of 0.450; resuspend yeast pellet in 450  $\mu$ l sterile ddH<sub>2</sub>O).
6. Make tenfold serial dilutions (1:10, 1:100): Prepare 2.0 ml tubes with 900  $\mu$ l of sterile ddH<sub>2</sub>O. Add 100  $\mu$ l of the appropriate yeast (OD<sub>600</sub> = 1.0) and mix well by shaking by hand. For the 1:100 dilution transfer 100  $\mu$ l of the 1:10 dilution into another tube containing 900  $\mu$ l ddH<sub>2</sub>O.
7. Drop 7  $\mu$ l of each dilution on selective plates containing increasing methionine concentrations as well as on vector-selective media (CSM-Leu<sup>-</sup>, Trp<sup>-</sup>, Ura<sup>-</sup>) as growth control. Allow the drops to dry until the liquid is completely evaporated (*see Notes 3, 5 and 9*).
8. Seal the plates and incubate for 1–3 days at 30 °C (depending on the expression and interaction strength of bait and prey fusions). Remove the vector-selective growth control after 24 h of incubation to see the gradient of the yeast ODs.
9. Document results by scanning or photography using a black background.

### 3.4 Western Blot (See Note 10)

1. Harvest 2 ml of overnight yeast culture by centrifugation or use deep-frozen aliquots (*see step 2* of Subheading 3.2).
2. Add ~50  $\mu$ l glass beads (diam. ~0.25–0.5 mm).
3. Resuspend cells in 100  $\mu$ l LL-buffer by vortexing for approx. 2 min.
4. Incubate shaking for 10 min at 65 °C.
5. Centrifuge at 16,500  $\times g$  for 10 min.
6. Transfer the supernatant to a fresh tube.
7. Load 10  $\mu$ l on a SDS-PAGE gel (appropriate gel percentage depends on protein sizes). *Optional*: Store at –80 °C.
8. Run gel with appropriate conditions (e.g., ~1 h at 130 V, constant voltage).
9. Run western blot (e.g., transfer on PVDF membrane via wet blot overnight at 30 V, constant voltage).
10. Block membrane in blocking buffer on a shaker for 1 h at room temperature.
11. Wash three times, 10 min each with 1 $\times$  TBST.
12. Transfer membrane into primary antibody solution. Incubate for at least 1 h at room temperature.
13. Wash three times, 10 min each with 1 $\times$  TBST.
14. Detection of membranes incubated in primary antibodies directly conjugated to HRP.
15. Transfer membrane into secondary antibody solution. Incubate for at least 1 h at room temperature.
16. Wash three times, 10 min each with 1 $\times$  TBST.
17. Detection of membranes incubated in HRP-conjugated secondary antibodies.

---

## 4 Notes

1. Please note that both ubiquitin moieties have to be located in the cytosol as the ubiquitin-specific proteases are cytosolic and the released transcription factor needs to be able to migrate into the cell nucleus to activate the reporter genes. Additionally, the bait protein has to be membrane-attached (either through an intrinsic transmembrane domain or an artificial N-terminal OST4p membrane anchor) to prevent leakage into the nucleus [5, 12].
2. The orientation of the Nub moiety is an additional factor to consider, as masking of leader sequences or signal peptides can lead to artificial mislocalization and/or protein aggregation, misfolding, and degradation [1, 9]. Prior to PPI studies,

sequence analysis of the POIs using *in silico* tools such as PSORT and/or TargetP can be used to determine tag orientation. C-terminal Nub fusions show reduced reassembly with Cub fusions probably due to steric effects of the ubiquitin split sites, which is why we recommend using C-terminal NubA instead of NubG. Affinity of NubA to Cub is higher than of NubG compensating for the weaker performance of a C-terminal Nub [16]. However, we repeatedly noted that NubG fusions (including N-terminal double HA epitope tag) are more reliably detected via immunoblot.

3. The bait proteins are cloned into low-copy ARS/CEN vectors containing a *MET25* promoter (pMetYC-DEST, pMetOYC-DEST) [13, 14]. The *MET25* promoter is tightly repressed in the presence of methionine allowing for stringent control of protein expression and high selectivity. Balancing the expression of recombinant proteins is of great importance because artificial, high concentrations may lead to unspecific interactions with Nub fusions or can cause mutant phenotypes or even lethality in yeast. We usually test interactions on different methionine concentrations of up to 500  $\mu\text{M}$ . However, it is important to note that commercially available, complete supplement mixture (CSM) usually contains 134  $\mu\text{M}$  methionine (20 mg/l) which would already significantly reduce gene expression and thus can lead to the suppression of weak or transient interactions.
4. The SUS can also be used for the detection of multimeric interactions using the so-called SUB (SUS bridge assay) [9, 17] or for screening approaches [18].
5. We recommend testing bait proteins for potential toxicity and self-activation prior to the actual PPI analysis via mating with soluble NubG and NubWt (=NubI) peptides, respectively. This includes both growth assay and immunoblot analysis of the POIs.
6. Cloning problems in *Escherichia coli* due to gene toxicity/instability—especially of eukaryotic membrane proteins—or mutations caused by the insertion of transposon elements can be circumvented by performing recombination-based *in vivo* cloning in yeast or use of a specialized *E. coli* strain which reduces copy number [7].
7. The (co-)transformation efficiency can be enhanced by increasing the amount of plasmid DNA or by upscaling of the transformation mixture. We recommend to use at least 1  $\mu\text{g}$  of plasmid for a single transformation reaction to obtain an appropriate number of colonies.
8. The color phenotype of diploid cells on nonselective media can be used as a preliminary tool to estimate the outcome of the

PPI analysis. The effect is based on the *ADE2* reporter gene that encodes an enzyme in the adenine biosynthetic pathway. Nonactivation of *ADE2* due to the lack of interaction between bait and prey peptides leads to accumulation of a red-colored intermediate, whereas positive PPIs result in white colonies.

9. To correctly analyze and interpret the PPI data it is essential to include appropriate controls in each experiment. The selected proteins should be expressed in the same cells and localize in the same compartment under native conditions. The ideal negative control is a closely related protein or a mutated version of the POI using site-directed mutagenesis [19]. Oligonucleotides for this can easily be designed using our SDM-assist software (<http://www.psrg.org.uk/sdm-assist.html>).
10. Biochemical verification of protein expression by immunoblot analysis is highly recommended especially in case of negative results. Always consider that expressing recombinant proteins in heterologous expression systems might result in low translation efficiency due to several aspects such as suboptimal codon usage, incorrect post-translational modifications, altered co-translational folding or protein instability. In some cases, adjustment of the codon bias to the heterologous expression host might positively influence translational efficiency [1].

## References

1. Xing S, Wallmeroth N, Berendzen KW, Grefen C (2016) Techniques for the analysis of protein-protein interactions in vivo. *Plant Physiol* 171(2):727–758
2. Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. *Nature* 340(6230):245–246
3. Stynen B, Tournu H, Tavernier J, Van Dijck P (2012) Diversity in genetic in vivo methods for protein-protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system. *Microbiol Mol Biol Rev* 76(2):331–382
4. Johnsson N, Varshavsky A (1994) Split ubiquitin as a sensor of protein interactions in vivo. *Proc Natl Acad Sci U S A* 91(22):10340–10344
5. Stagljar I, Korostensky C, Johnsson N, te Heesen S (1998) A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proc Natl Acad Sci U S A* 95(9):5187–5192
6. Obrdlik P et al (2004) K<sup>+</sup> channel interactions detected by a genetic system optimized for systematic studies of membrane protein interactions. *Proc Natl Acad Sci U S A* 101(33):12242–12247
7. Grefen C et al (2008) Subcellular localization and in vivo interactions of the Arabidopsis thaliana ethylene receptor family members. *Mol Plant* 1(2):308–320
8. Grefen C et al (2010) A novel motif essential for SNARE interaction with the K<sup>(+)</sup> channel KC1 and channel gating in Arabidopsis. *Plant Cell* 22(9):3076–3092
9. Grefen C, Blatt MR (2012) Do calcineurin B-like proteins interact independently of the serine threonine kinase CIPK23 with the K<sup>+</sup> channel AKT1? Lessons learned from a menage a trois. *Plant Physiol* 159(3):915–919
10. Xing S et al (2017) Loss of GET pathway orthologs in Arabidopsis thaliana causes root hair growth defects and affects SNARE abundance. *Proc Natl Acad Sci U S A* 114(8):E1544–E1553
11. Grefen C, Lalonde S, Obrdlik P (2007) Split-ubiquitin system for identifying protein-protein interactions in membrane and full-length proteins. *Curr Protoc Neurosci*. Chapter 5:Unit 5 27
12. Möckli N et al (2007) Yeast split-ubiquitin-based cytosolic screening system to detect interactions between transcriptionally active proteins. *BioTechniques* 42(6):725–730

13. Karnik R et al (2015) Binding of SEC11 indicates its role in SNARE recycling after vesicle fusion and identifies two pathways for vesicular traffic to the plasma membrane. *Plant Cell* 27(3):675–694
14. Grefen C, Obrdlik P, Harter K (2009) The determination of protein-protein interactions by the mating-based split-ubiquitin system (mbSUS). *Methods Mol Biol* 479:217–233
15. Ludewig U et al (2003) Homo- and hetero-oligomerization of ammonium transporter-1 NH<sub>4</sub> uniporters. *J Biol Chem* 278(46):45603–45610
16. Raquet X, Eckert JH, Muller S, Johnsson N (2001) Detection of altered protein conformations in living cells. *J Mol Biol* 305(4):927–938
17. Grefen C (2014) The split-ubiquitin system for the analysis of three-component interactions. *Methods Mol Biol* 1062:659–678
18. Jones AM et al (2014) Border control--a membrane-linked interactome of Arabidopsis. *Science* 344(6185):711–716
19. Karnik A, Karnik R, Grefen C (2013) SDM-assist software to design site-directed mutagenesis primers introducing “silent” restriction sites. *BMC Bioinformatics* 14:105

# CURRICULUM VITAE

## Personal information

Name: Lisa Yasmin Asseck  
Date of birth: 10.02.1990  
Place of birth: Göppingen

## Education

2015 – now Ph.D. in Biology  
University of Tübingen  
Centre for Plant Molecular Biology (ZMBP)  
Department of Developmental Genetics

2013 – 2015 M.Sc. Biology  
University of Tübingen  
Master thesis in molecular plant biology:  
"Identifizierung von Interaktionspartnern des Brassinosteroid-  
Rezeptorkomplexes"

2009 – 2013 B.Sc. Molecular Medicine  
University Ulm  
Bachelor thesis in transfusion medicine:  
"S100A4 induziert die Proliferation regulatorischer T-Zellen"

2006 – 2009 General higher education entrance qualification (Abitur)  
Justus-von-Liebig Schule Göppingen

## Publications and manuscripts

**Asseck, L.Y.**, Mehlhorn, D.G., Rivera Monroy, J., Ricardi, M.M., Breuninger, H., Wallmeroth, N., Berendzen, K.W., Nouwrosian, M., Xing, S., Schwappach, B., Bayer, M., Grefen, C. "ER membrane receptors of the GET pathway are conserved throughout eukaryotes." Submitted.

Denninger, P., Reichelt, A., Schmidt, V.A.F., Mehlhorn, D.G., **Asseck, L.Y.**, Stanley, C.E., Keinath, N.F., Evers, J.F., Grefen, C., Grossmann, G. (2019). "Distinct RopGEFs successively drive polarization and outgrowth of root hairs." *Current Biology* 29(11): 1854-1865. e1855.

Neu, A., Eilbert, E., **Asseck, L.Y.**, Slane, D., Henschen, A., Wang, K., Bürgel, P., Hildebrandt, M., Musielak, T.J., Kolb, M., Lukowitz, W., Grefen, C., Bayer, M. (2019). "Constitutive signaling activity of a receptor-associated protein links fertilization with embryonic patterning in *Arabidopsis thaliana*." *Proceedings of the National Academy of Sciences* 116(12): 5795-5804.

**Asseck, L. Y.** and Grefen, C. (2018). Detecting interactions of membrane proteins: the split-ubiquitin system. *Two-Hybrid Systems*, Springer: 49-60.

**Asseck, L. Y.**, Wallmeroth, N., Grefen, C. (2018). ER membrane protein interactions using the split-ubiquitin system (SUS). *The Plant Endoplasmic Reticulum*, Springer: 191-203.

Xing, S., Mehlhorn, D., Wallmeroth, N., **Asseck, L.**, Kar, R., Voss, A., Denninger, P., Schmidt, V.A.F., Schwarzländer, M., Stierhof, Y.D., Grossmann, G., Grefen, C. (2017). "Loss of GET pathway orthologs in *Arabidopsis thaliana* causes root hair growth defects and affects SNARE abundance." *Proceedings of the National Academy of Sciences* 114(8): E1544-E1553.

Lotfi, R., Wiegmann, D.S., **Asseck, L.**, Erle, A., Yildiz, T., Jahrsdörfer, B., Schrezenmeier, H. (2012). "Necrosis-associated factors (DAMPs) including S100A4 used to pulse dendritic cells (DCs) induce regulatory T cells." *J Cell Sci Ther* 3(134): 2.

### **Poster presentations and workshops**

**Asseck, L.**, Mehlhorn, D.G., Xing, S., Wallmeroth, N., Grefen, C. "Analysis of GET pathway orthologues in *Arabidopsis*: TA protein receptors at the ER membrane." 30<sup>th</sup> Conference on Molecular Biology of Plants, Dabringhausen 2017.

**Asseck, L.**, Mehlhorn, D.G., Wallmeroth, N., Xing, S., Grefen, C. "IP-MS identifies a membrane receptor of the *Arabidopsis* GET pathway." SEB GARNet symposium - From proteome to phenotype: role of post-translational modifications, Edinburgh 2017.

GARNet workshop on plant proteomics, Edinburgh 2017.

## **ACKNOWLEDGEMENTS**

First, I would like to thank my supervisor, Prof. Dr. Christopher Grefen, for giving me the opportunity to work in his group and for his support and guidance throughout my PhD.

Another thank you goes to Prof. Dr. Gerd Jürgens for providing a second opinion on this thesis.

Thanks too to Prof. Dr. Klaus Harter and Dr. Martin Bayer for being part of my thesis advisory committee and for helpful discussions and suggestions.

Many thanks to my former and present colleagues, especially Prof. Dr. Shuping Xing, Dietmar Mehlhorn, Niklas Wallmeroth and Dr. Martiniano Ricardi who worked with me on this project. Special thanks are also due to Eva Schwörzer and Laure Grefen for their excellent technical support and invaluable assistance. Thanks too to Dr. Holger Breuninger and Simon Klesen for help with CRISPR/Cas9.

Moreover, especial thanks to Dr. Sabine Müller and her group, who welcomed me warmly to their lab over the past few months.

Last, I would like to thank my family and Simon for their infinite support, love and encouragement.