At1g78620

—

Phytyl Phosphate Kinase in Arabidopsis

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Christian Matthias Plohmann
aus Solingen

November 2015

aus dem Institut für Biochemie der Pflanzen
Heinrich-Heine-Universität Düsseldorf

Gedruckt mit Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. A. P. M. Weber
Korreferent: Apl. Prof. Dr. P. Jahns

Tag der mündlichen Prüfung: 18.01.2016

EIDESSTATTLICHE ERKLÄRUNG

Ich versichere an Eides Statt, dass diese Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Die Dissertation habe ich in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht.

Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Solingen, 29.11.2015

Christian Plohmann

Contents

ABBREVIATIONS

Abbreviation	Meaning
2-OG	2-oxoglutarate
2-PG	2-phosphoglycolate
3-PGA	3-phosphogylcerate
°C	degree Celsius
µl	microliter
µm	micrometer
amiRNA	artificial micro RNA
Amp^R	ampicillin resistance
bp	base pairs
C_2-cycle	photorespiration
CAT	catalase
cDNA	complementary DNA/copy DNA
cm	centimeter
Col-0	Columbia 0
CTP	cytidine triphosphate
cTP	chloroplast targeting transit peptide
d	day
DCT	dicarboxylate transporter
DiT	dicarboxylate translocator
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E	Einstein
EDTA	ethylenediaminetetraacetic acid
FW	fresh-weight
GC/MS	gas chromatography-mass spectrometry
GDC	glycine decarboxylase complex
gDNA	genomic DNA
$Gent^R$	gentamycin resistance
GGPS	geranylgeranyl pyrophosphate synthase 1
GGR	geranylgeranyl reductase
GGT	glutamate:glyoxylate aminotransferase
GLYK	glycerate 3-kinase
GOGAT	ferredoxin-dependent glutamate synthase
GOX	glycolate oxidase
GS	glutamine synthetase
h	hour
HC	high CO2
HPR1	hydroxypyruvate reductase
IF	infiltration
Kan^R	kanamycin resistance
LHC	light harvesting complex
LSM	laser scanning microscope

MES	2-(N-morpholino)ethanesulfonic acid
miRNA	microRNA
mm	millimeter
mM	millimolar
MR	mutual rank
mRNA	messenger RNA
MS	Murashige & Skoog
NADPH	reduced form of nicotinamide adenine dinucleotide phosphate
NC	normal CO2 / ambient CO2 conditions
nm	nanometer
npq	nonphotochemical quenching
nt	nucleotides
OD	optical density
ONC	overnight colony
p	p-value
PCR	polymerase chain reaction
PGP	phosphoglycolate phosphatase
PLGG1	plastidic glycolate glycerate transporter 1
pm	picomol
ppm	parts per million
PSII	photosystem II
RBL	rhomboid-like protein
RNA	ribonucleic acid
ROS	reactive oxygen species
RP-HPLC	reverse phase high pressure liquid chromatography
RT-PCR	reverse transcriptase polymerase chain reaction
Rubisco	ribulose-1,5-bisphosphate carboxylase/oxygenase
RuBP	ribulose-1,5- bisphosphate
SGT	serine:glyoxylateaminotransferase
$Spec^R$	spectinomycin resistance
Sul^R	sulfadiazine resistance
TAE	tris-acetate-EDTA
T-DNA	transfer DNA
TGF	transforming growth factor
TMD	transmembrane domain
UBQ	ubiquitin
UTR	untranslated region
UV	ultraviolet
V	volume
V+A+Z	violaxanthin/antheraxanthin/zeaxanthin
vs.	versus
VTE2-1	homogentisate phytyltransferase 2, vitamin E pathway 2
VTE5	phytol kinase, vitamin E pathtway 5
w/v	weight per volume
w/w	weight per weight
WGCNA	weighted correlation network analysis

wt	wild-type
YFP	yellow fluorescent protein

PREFACE

This PhD thesis presented here, consists of three individual manuscripts. The first manuscript “Co-expression analysis as a tool for the discovery of transport proteins in photorespiration” (Bordych et al., 2013) was written as a review, which shows in parts initial results of ongoing work. It was published in 2013 in the Plant Biology special issue “Photorespiration”. It describes how At1g78620 was chosen as candidate gene for a putative photorespiratory transport protein.

In the second part a publication is presented which deals with the characterization of the candidate gene At1g78620, which originated from the work, which was performed for the first manuscript. Noticeably, VTE6 (At1g78620) was not mentioned as a prime target in the first publication to keep ongoing work hidden from other research groups. The characterization was realized together with the group of Prof. P. Dörmann (University of Bonn, Germany) and published as “Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis” (vom Dorp et al., 2015) in Plant Cell.

The last part of this thesis is written as a draft for future publication. It summarizes unpublished work which was performed for the characterization of VTE6.

All materials and methods which did not fit into the manuscripts are added as “Supplementary Material and Methods” to this thesis.

The manuscripts are preceded by a section which describes the “Aim of the thesis” ,abstracts in German and English for the whole thesis in general, and an introduction into this thesis.

REFERENCES

Bordych, C., Eisenhut, M., Pick, T.R., Kuelahoglu, C., Weber, A.P. (2013) Co-expression analysis as tool for the discovery of transport proteins in photorespiration. Plant Biol (Stuttg), **15**(4), 686-693.

vom Dorp, K., Holzl, G., Plohmann, C., Eisenhut, M., Abraham, M., Weber, A.P., Hanson, A.D., Dormann, P. (2015) Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis. The Plant cell.

AIM OF THE THESIS

The aim of this thesis was originally to discover and characterize missing genes of photorespiration while placing the emphasis on transport processes between occupied organelles (plastid, peroxisome, and mitochondrion). While the enzymatic steps have been identified on the genomic and characterized on the biochemical level for a long time, several transport processes could not be narrowed down to a certain gene locus at the time the work for this thesis was initiated.

Given the fact that photorespiration is often considered as a wasteful process (beside functions in photoprotection) there is a chance to modulate photorespiration in the future. Modulation can ultimately increase plant yield to deal with the increasing worldwide population and stagnation in agricultural area gain. Complete understanding of the plant photorespiratory cycle is a prerequisite for potential improvements in agricultural yield, through redirection of photorespiratory intermediates, adjustments to connected metabolic pathways, and potential integration of foreign metabolic pathways (Maurino and Peterhansel, 2010).

The majority of enzymatic processes during photorespiration were identified by Somerville and co-workers by an elegant forward genetic approach (Somerville, 2001). However, forward genetics seemed to be insufficient to identify all genes which code for transport processes during photorespiration. Consequently, work as described here is based on reverse genetics utilizing co-expression analysis.

Experiments in the starting phase aimed at the generation of a list of candidate genes which are likely a) co-regulated with known genes of photorespiration and b) predicted proteins which facilitate shuttling through membranes. This initial work was based on available *in silico* data in forms of publicly available expression data for *Arabidopsis thaliana* and computational data and experiments to predict localization and membrane topology of candidate genes' expressed proteins.

Later, the focus was applied to the characterization of one candidate gene (At1g78620) which was shown to be co-expressed with known photorespiratory genes. It should be clarified if At1g78620 and the corresponding protein is indeed occupied in the photorespiratory cycle.

REFERENCES

Maurino, V.G. and Peterhansel, C. (2010) Photorespiration: current status and approaches for metabolic engineering. Curr Opin Plant Biol, **13**(3), 249-256.

Somerville, C.R. (2001) An early Arabidopsis demonstration. Resolving a few issues concerning photorespiration. Plant physiology, **125**(1), 20-24.

SUMMARY

The photorespiratory cycle has been deciphered on the enzymatic level in detail for a long time. However, the components facilitating the shuttling of intermediates between compartments have not been identified completely by the start of this work. The first part of this work has been summarized in "Co-expression analysis as a tool for the discovery of transport proteins in photorespiration" (Bordych et al., 2013). It shows how publicly available expression data of *Arabidopsis thaliana* genes is utilized to generate hypotheses about the corresponding proteins' functions. A workflow was generated in which the initial part was compromised of co-expression analyses on the atted-II platform (Obayashi et al., 2009) to generate a list of candidate genes for photorespiratory transport proteins. Subsequent parts of this publication could show how to reduce the list of candidate genes to a number of genes which can be tested in *in vivo* experiments. It was shown how certain filters and tools can be applied to increase the likelihood of picking a photorespiratory transporter gene out of the big list of genes which are co-regulated with photorespiratory genes.

The publication concludes with the comparison to another standard method (WGCNA analysis) for investigation of potential functions of genes of interest (Langfelder and Horvath, 2008). It is shown that this method supports the results obtained from co-expression analyses and filtering. This publication led ultimately to a list of candidate genes for photorespiratory transport which should be tested in further experiments for their potential role in the plant C_2-cycle.

Screening for photorespiratory mutants with T-DNA lines based on this list, revealed no direct evidence for any of the tested genes being occupied in photorespiration. Still, data from the literature and public databases (Myouga et al., 2010) were encouraging enough to work further on the candidate gene At1g78620. At1g78620 T-DNA insertion lines stuck out, due to absence of homozygous insertions of the T-DNA, which points towards an essential function of the corresponding protein in *Arabidopsis thaliana*.

Analyses of a generated amiRNA mutant (Schwab et al., 2006) for At1g78620 revealed that At1g78620 is not involved in photorespiratory transport. However, results from the analyses strongly suggested a role for At1g78620 in tocopherol metabolism. Together with results from the group of Prof. Dörmann (University of

Bonn) it was finally possible to identify At1g78620 as the missing phytyl phosphate kinase (VTE6) in *Arabidopsis* (vom Dorp et al., 2015). This kinase works orchestrated together with phytol kinase VTE5 to generate phytyl diphosphate based on phytol (Ischebeck et al., 2006) for tocopherol synthesis.

Ongoing work aims at an in-depth characterization of the generated amiRNA line for At1g78620 (*amiR-vte6*) to get further insights into the important function of VTE6.

REFERENCES

Bordych, C., Eisenhut, M., Pick, T.R., Kuelahoglu, C., Weber, A.P. (2013) Co-expression analysis as tool for the discovery of transport proteins in photorespiration. Plant Biol (Stuttg), **15**(4), 686-693.

Ischebeck, T., Zbierzak, A.M., Kanwischer, M., Dormann, P. (2006) A salvage pathway for phytol metabolism in Arabidopsis. The Journal of biological chemistry, **281**(5), 2470-2477.

Langfelder, P. and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics, **9**, 559.

Myouga, F., Akiyama, K., Motohashi, R., Kuromori, T., Ito, T., Iizumi, H., Ryusui, R., Sakurai, T., Shinozaki, K. (2010) The Chloroplast Function Database: a large-scale collection of Arabidopsis Ds/Spm- or T-DNA-tagged homozygous lines for nuclear-encoded chloroplast proteins, and their systematic phenotype analysis. The Plant journal : for cell and molecular biology, **61**(3), 529-542.

Obayashi, T., Hayashi, S., Saeki, M., Ohta, H., Kinoshita, K. (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic acids research, **37**(Database issue), D987-991.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N., Weigel, D. (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. The Plant cell, **18**(5), 1121-1133.

vom Dorp, K., Holzl, G., Plohmann, C., Eisenhut, M., Abraham, M., Weber, A.P., Hanson, A.D., Dormann, P. (2015) Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis. The Plant cell.

ZUSAMMENFASSUNG

Enzyme welche an der pflanzlichen Photorespiration beteiligt sind, sind sowohl genetisch als biochemisch charakterisiert. Das Wissen über die Transportprozesse zwischen den Organellen der Photorespiration war bei Beginn dieser Dissertation jedoch stark begrenzt. Einzelne Transportprozesse konnten auf genetischer Ebene noch nicht identifiziert werden. Der erste Teil dieser Arbeit wurde in der Publikation "*Co-expression analysis as a tool for the discovery of transport proteins in photorespiration*" (Bordych et al., 2013) zusammenfassend beschrieben. Es wird gezeigt wie Expressionsdaten von *Arabidopsis thaliana* nutzen, um Hypothesen zu generieren, die dabei helfen die Funktion von nicht näher charakterisierten Genen zu beschreiben. Ein Arbeitsablauf wurde erstellt in dessen erster Phase mithilfe der atted-II (Obayashi et al., 2009) Plattform Kandidaten-Gene identifiziert wurden die ähnliche Expressionsmuster wie bekannte Gene der Photorespiration zeigen. Spätere Arbeitsabschnitte dieser Veröffentlichung zeigen wie die Liste an Kandidaten-Genen reduziert wird, so dass eine Anzahl an Kandidaten-Genen vorliegt, welche in *in vivo* Experimenten untersucht werden kann. Durch das Festlegen von Kriterien die das Gen oder das zugehörige Protein erfüllen musste, um für den Transport von photorespiratorischen Zwischenprodukten in Frage zu kommen, konnten gezielt geeignete Kandidaten-Gene isoliert werden.

Im letzten Teil der ersten Veröffentlichung werden die generierten Daten anhand einer anderen bioinformatischen Herangehensweise validiert und können durch WGCNA Analyse (Langfelder and Horvath, 2008) bestätigt werden.

Im Zuge dieses ersten Teils ist eine Liste mit Kandidaten-Genen entstanden, welche hinsichtlich ihrer Rolle in der Photorespiration untersucht werden sollten.

Arabidopsis T-DNA Linien der Kandidaten-Gene wurden untersucht, um zu prüfen ob ein möglicher Gen-Knockout des Kandidaten-Genes zu einer Beeinträchtigung der Pflanze hinsichtlich der Photorespiration führt. Keine der untersuchten T-DNA Linien zeigte dabei einen typischen Phänotyp auf, der mit Photorespiration in Verbindung gebracht werden konnte. Publizierte Ergebnisse und Ergebnisse in Datenbanken (Myouga et al., 2010) veranlassten jedoch dazu, weiter an der Charakterisierung von At1g78620 zu arbeiten. At1g78620 T-DNA Linien verstärkten diesen Eindruck, da über mehrere Generationen keine Mutanten identifiziert werden konnten, in denen die

T-DNA Insertion homozygot vorlag. Diese Tatsache unterstreicht die essentielle Rolle von At1g78620 für *Arabidopsis*.

Versuche mit einer neu erstellten amiRNA Linie (Schwab et al., 2006) für At1g78620 zeigten, dass At1g78620 jedoch im engen Sinne nicht mit Photorespiration in Verbindung gebracht werden kann. Die Ergebnisse weiterer Versuche und die Zusammenarbeit mit der Arbeitsgruppe von Prof. Dörmann (Universität Bonn) konnten dann zeigen, dass At1g78620 die Phytylphosphatkinase (VTE6) ist, welche die Umwandlung von Phytylphosphat zu Phytyldiphosphat katalysiert und von essentieller Bedeutung während der Synthese von Tocopherol ist (vom Dorp et al., 2015). In Interaktion zwischen VTE6 und VTE5 wird die Umwandlung von Phytol zu Phytyldiphosphat katalysiert (Ischebeck et al., 2006).

Die momentanen Arbeiten zielen darauf ab, die erstellte Mutante für At1g78620 (*amiR-vte6*) ausführlich zu charakterisieren, um die wichtige Rolle von VTE6 in *Arabidopsis* weiter zu entziffern.

REFERENZEN

Bordych, C., Eisenhut, M., Pick, T.R., Kuelahoglu, C., Weber, A.P. (2013) Co-expression analysis as tool for the discovery of transport proteins in photorespiration. Plant Biol (Stuttg), **15**(4), 686-693.

Ischebeck, T., Zbierzak, A.M., Kanwischer, M., Dormann, P. (2006) A salvage pathway for phytol metabolism in Arabidopsis. The Journal of biological chemistry, **281**(5), 2470-2477.

Langfelder, P. and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics, **9**, 559.

Myouga, F., Akiyama, K., Motohashi, R., Kuromori, T., Ito, T., Iizumi, H., Ryusui, R., Sakurai, T., Shinozaki, K. (2010) The Chloroplast Function Database: a large-scale collection of Arabidopsis Ds/Spm- or T-DNA-tagged homozygous lines for nuclear-encoded chloroplast proteins, and their systematic phenotype analysis. The Plant journal : for cell and molecular biology, **61**(3), 529-542.

Obayashi, T., Hayashi, S., Saeki, M., Ohta, H., Kinoshita, K. (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic acids research, **37**(Database issue), D987-991.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N., Weigel, D. (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. The Plant cell, **18**(5), 1121-1133.

vom Dorp, K., Holzl, G., Plohmann, C., Eisenhut, M., Abraham, M., Weber, A.P., Hanson, A.D., Dormann, P. (2015) Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis. The Plant cell.

INTRODUCTION

Plant photorespiration

The plant enzyme Rubisco (Ribulose-1,5-bisphosphate carboxylase/oxygenase) catalyzes the first step of the dark reaction during photosynthesis. This reaction is the carboxylation of ribulose-1,5-bisphosphate (RuBP) and yields two molecules of 3-phosphogylcerate (3-PGA), which is further metabolized in the Calvin cycle (Calvin and Benson, 1948). Beside carboxylation reaction, Rubisco is also capable of catalyzing an oxygenation reaction yielding 2-phosphoglycolate (2-PG), which was shown by radioactive labeling of RuBP and O_2 (Andrews et al., 1973; Lorimer et al., 1973). This phenomena serves as an explanation for an observation made in 1920, which is thought to be the first description of the oxygenase reaction of Rubisco. Otto Warburg observed that O_2 accumulation can decrease CO_2 fixation in *Chlorella* (Warburg, 1928). Later, John P. Decker showed the evolution of CO_2 of leaves upon stimulation with O_2 in the light (Decker, 1955).The first theory of photorespiration was made when Ogren and Bowes linked inhibition of photosynthesis by O_2 to O_2 stimulation of glycolate synthesis, and CO_2 evolution (Ogren and Bowes, 1971).

Accumulating 2-PG is toxic for the photosynthetic active cell as it reduces photosynthetic efficiency through inhibition of triose phosphate isomerases (Anderson, 1971) and thereby RuBP regeneration. In addition it was shown that 2-PG inhibits phosphofructokinase, a central enzyme of starch breakdown in plants (Kelly and Latzko, 1976). Plants recycle 3-phosphoglycerate (3-PGA) from 2-PG, which can be fed back into the Calvin cycle, during the photorespiratory cycle. This metabolic pathway consists of several catalytic reactions and is located to chloroplast, mitochondrion, peroxisome, and the cytosol.

Oxygenation reaction of Rubisco in the chloroplast yields one molecule 3-PGA and one molecule of 2-PG. Two molecules of 2-PG are converted to glycolate through dephosphorylation by 2-phosphoglycolate phosphatase (PGP) in the chloroplast. Glycolate is shuttled into the peroxisome by a transport protein. The plastidal glycolate glycerate transporter (PLGG1) shuttles glycolate out of the chloroplast and into the peroxisome (Pick et al., 2013). Once transported to the peroxisome, glycolate is converted to glyoxylate and H_2O_2 by the enzyme glycolate oxidase (GOX). H_2O_2 is

detoxified in the peroxisome by catalase (CAT), while glyoxylate is further converted. After transamination by glutamate:glyoxylate aminotransferase (GGT), the resulting glycine is shuttled out of the peroxisome into the mitochondrion. Two molecules of glycine are interconverted to serine by the combined action of serine hydroxymethyltransferase and the glycine decarboxylase complex (GDC), releasing ammonia (NH_4^+) and CO_2 (Bauwe et al., 2010). This release of previously fixed carbon and nitrogen is the main reason why photorespiration is at first sight considered a wasteful process (Peterhansel and Maurino, 2011).

The C_3 compound serine is shuttled from the mitochondrion back to the peroxisome, where it forms hydroxypyruvate catalytically driven by serine:glyoxylateaminotransferase (SGT). During this interconversion an amine group is released and used to generate glycine from glyoxylate by GGT. Reduction of hydroxypyruvate by the peroxisomal hydroxypyruvate reductase (HPR1) yields glycerate, which is translocated back into the chloroplast in exchange for glycerate by PLGG1. The final conversion from glycerate to 3-PGA in the chloroplast is catalyzed by the enzyme glycerate 3-kinase (GLYK). The resulting 3-PGA is reintroduced into the Calvin cycle, ultimately generating RuBP for primary carbon fixation through Rubisco. The complete photorespiratory cycle, its components and its origin is reviewed in great detail by Bauwe and colleagues (Bauwe et al., 2010).

NH_4^+, released in the mitochondrion during the conversion of glycine to serine, is shuttled into the chloroplast. In the chloroplast it is refixed under ATP consumption by an enzyme complex built up of glutamate synthetase (GS2) and the ferredoxin-dependent glutamate synthase (GOGAT). GS2 combines glutamate and NH_4^+ to yield glutamine. Glutamine is directly used by GOGAT to generate glutamate from 2-oxoglutarate (2OG). 2OG is utilized for glycine synthesis in the peroxisome and NH_4^+ refixation (Bauwe et al., 2010). Glutamate and 2OG needed for these reactions are transported over the plastidal membrane in exchange for malate by the dicarboxylate translocator system consisting of dicarboxylate translocator 1 (DiT1) and dicarboxylate translocator 2 (DiT2) (Weber et al., 1995).

Photorespiratory enzymes and transport proteins

The enzymes of the photorespiratory cycle were mostly identified by mutant analyses. In a very elegant, yet simple screen Chris and Shauna Somerville shifted mutagenized populations of *Arabidopsis thaliana* from high CO_2 conditions (1% CO_2, HC), which suppressed photorespiration, to ambient air conditions (0.03%, normal CO_2 conditions; NC). Mutants which grew under HC but suffered under NC, were considered to be prime targets for further investigations (Somerville, 2001). Feeding experiments with radioactive labeled CO_2 resulted in the identification of five different key enzymes of the photorespiratory cycle between 1979 and 1982. The very first enzyme which was identified and characterized was PGP (Somerville and Ogren, 1979). Within three years SGT (Somerville and Ogren, 1980b), GS (Somerville and Ogren, 1980a), SHMT (Somerville and Ogren, 1981), and GDC (Somerville and Ogren, 1982) were characterized. Still, it took more than 20 years until GLYK, the last missing photorespiratory enzyme, was identified and characterized (Boldt et al., 2005). Today, the core photorespiratory cycle is well described. However, several processes needed for transport of photorespiratory intermediates are still not characterized on the genetic and the molecular level (Eisenhut et al., 2013).

A major step towards closing this gap of knowledge was made by the identification of PLGG1 on the genetic level (Pick et al., 2013) and its subsequent characterization. The plastidic transport of glycerate and glyoxylate was already described biochemically earlier (Howitz and McCarty, 1985; Howitz and McCarty, 1991), without linking it to a certain gene or protein.

Photorespiration has beneficial effects- Photoprotection

The release of fixed carbon (among NH_3, and the consumption of ATP and NADH) through photorespiration, results in a net decrease of photosynthetic efficiency of about 25%, because one fourth of reactions performed by Rubisco are oxygenation reactions (Sharkey, 1988). Photorespiration is therefore often considered as a wasteful process (Peterhansel and Maurino, 2011).

Even before the core (enzymatic) photorespiratory cycle was completely deciphered on the genetic level with the characterization of GLYK (Boldt et al., 2005), it was

shown that plants can actually benefit from increased rates of photorespiration. Overexpression of GS2 in tobacco increases the capacity for photorespiration of the plant and renders plants more tolerant to light stress in comparison to the wild type control. Tobacco lines, in which GS2 activity is repressed, suffer more from light stress than the respective wild type control (Kozaki and Takeba, 1996). In fact, a photoprotective function by consuming an excess of reducing power generated during photosynthesis under high light conditions, was already suggested several years earlier (Tolbert, 1971). Investigating photorespiratory mutants for GOGAT, SHMT1, DiT1, DiT2, and GLYK showed that these mutants suffer from accelerated photoinhibition of photosystem II (PSII) by suppression of the repair of photodamaged PSII (Takahashi et al., 2007). Considering its photoprotective function during photosynthesis, photorespiration is orchestrated together with several other mechanisms enabling plants to grow under high light or drought conditions.

Photoprotection in plants

During photosynthesis reactive oxygen species are produced at all light intensities. In case the amount of absorbed light energy exceeds the capacity for its utilization, the potential for photooxidative damage is exacerbated (Niyogi, 1999). Several mechanisms to deal with this damage potential are described for plants.

First, plants are capable of modulating the antennae size of chlorophyll in PSII and photosystem I (PSI) (Melis, 1991) to balance light absorption.

Second, a very well characterized process to cope with excessive amounts of absorbed light is the elimination of photons by thermal dissipation. This effect is measured as nonphotochemical quenching (npq) (Horton et al., 1994; Niyogi, 1999). Thermal light dissipation is tightly interconnected with the xanthophyll cycle (Eskling et al., 1997), during which violaxantin is converted into zeaxanthin. Zeaxanthin is needed to induce a conformational change in the light harvesting complex (LHC) which is necessary for thermal dissipation (Niyogi, 1999). Dissecting mutants defective in xanthophyll metabolism showed its essential function during photoprotective energy dissipation (Niyogi et al., 1998).

Third, carotenoids (including xanthophylls) have a vital purpose during photoprotection as membrane-bound antioxidants through quenching of triplet chlorophyll and singlet O_2 (Kuhlbrandt et al., 1994). Singlet oxygen in the PSII reaction center can be efficiently quenched by β-carotene, as shown by Telfer and co-workers (Telfer et al., 1994). Despite their important function in photoprotection, single types of carotenoids are not essential for plant photoprotection but are functional redundant as shown by elimination of combinations of xanthophylls in double mutants (Niyogi et al., 1997).

Fourth, tocopherols prevent lipid peroxidation in the membrane through scavenging of superoxide, singlet oxygen, and hydroxide ion (Foyer et al., 1994). α-tocopherol, for example, is not bound to proteins like xantophylls but can move freely in the lipid matrix and control membrane stability (Fryer, 1992; Niyogi, 1999).

This brief summary about photoprotective mechanisms (photorespiration, npq) and compounds (carotenoids, tocopherols) in plants shows that a close network exists, which helps in the protection against high light intensities, to maintain photosynthesis.

Understanding Photorespiration

Photorespiration is not restricted to being a wasteful process as it is tightly integrated into a network of photoprotective mechanisms and compounds to serve as a crucial part of photosynthesis itself. Still, photorespiration is not deciphered completely on the genomic level as some transport processes are yet not narrowed down to single genes (Eisenhut et al., 2013). Forward genetic approaches were sufficient to identify the enzymatic part of photorespiration (Somerville, 2001), but failed largely to identify genes coding for photorespiratory transport proteins.

As mentioned above, a major step towards the full characterization of transport processes in the photorespiratory cycle was made by the characterization of PLGG1 (Pick et al., 2013). In this work reverse genetic approaches were utilized to find missing links in photorespiration. Co-expression analysis was utilized to find candidate genes for further characterization by reverse genetic techniques. Complete understanding of photorespiration and connected transport processes is of major importance for the future. Two reasons are presented here briefly.

First, deciphering the photorespiratory cycle on the molecular and biochemical level offers the opportunity to search for valuable targets to modulate the cycle, ultimately increasing photosynthetic efficiency and thereby plant yield. Initial experimental steps to generate plants with increased photorespiratory efficiency are already described (Peterhansel and Maurino, 2011). The increasing amount of food needed, combined with limited availability of agricultural space on earth, highlights the necessity to generate better performing agricultural plants. Redesign of the photorespiratory cycle in model organisms like *Arabidopsis thaliana* is a very first step towards high-performance agricultural plants.

The second reason is closely connected with the first one. Understanding photorespiration and its function beside detoxification of 2-PG can help to further increase positive aspects of photorespiration, such as photoprotection. Plants with increased capacity for photoprotection are capable of acclimatization to harsher environments concerning light intensity, eventually increasing agricultural space on earth, by yet not agriculturally conquered areas. Complete understanding of the photorespiratory cycle is crucial to evaluate, if photorespiration is a valuable target to modulate for better photoprotective function.

A reverse genetic approach to shed light on photorespiratory transport

The thesis presented here aims at the characterization of additional genes involved in photorespiratory transport or closely connected pathways through reverse genetic approaches. The term “reverse genetics” describes the attempt to unravel the function of a certain gene. The question arises how suitable candidate genes can be found to investigate their function in a reverse genetic approach.

Nowadays, a very valuable first step to find potential target genes is based on genomics and transcriptomics. The vast amount of genomic and transcriptomic data available (Brazma and Vilo, 2000) paired with the ongoing gain in calculation power renders such an approach very promising. Visualizing, ordering, and integrating these data into online databases was the first and an important step to render scientist able to use this data. Eisen and colleagues described how hierarchical clustering can be used to let the raw data speak: converting ‘data into information and then information into knowledge’ (Eisen et al., 1998).

Genes showing similar expression patterns under various different conditions, are thought to be likely operating in the same metabolic pathway or in closely related ones, based on the "guilt-by-association" paradigm (Usadel et al., 2009). Thus, identifying genes, showing to be highly co-expressed with genes of a known function in a pathway under investigation might very likely yield potential candidate genes for missing links in this pathway.

In this context, photorespiratory transport genes are thought to be co-expressed with the core genes of photorespiration coding for the enzymatic part of the photorespiratory cycle. Photorespiratory genes coding for the peroxisomal enzymes of photorespiration were already shown to be co-expressed with each other (Reumann and Weber, 2006). In the first part of this thesis it is described how co-expression analysis can be utilized as a candidate discovery tool for photorespiratory transporter genes. The corresponding workflow has been described in detail in the first manuscript, which constitutes one part of this thesis (Bordych et al., 2013).

Co-expression analysis – Important considerations

Utilizing co-expression analysis to generate hypotheses about the function of certain genes, respectively the corresponding proteins, bears risks which need to be considered, when working with co-expression data.

Genes, which show orchestrated expression patterns are not necessarily occupied in the same metabolic pathway, as metabolic pathways in *Arabidopsis* are intertwined to a large extend. Thus, candidate genes which are chosen for further investigations might also be assigned to reactions or pathways in the plant, which show only minor connectivity to plant photorespiration. Photorespiratory genes very likely are also co-expressed with photosynthetic genes and genes which are coding for important factors in light protection, as photorespiratory reactions are tightly integrated into plant metabolism.

Preventing tunnel vision, is of major importance to consider alternative protein functions too, which cannot be linked directly to the core photorespiratory cycle. This consideration does not only apply to results from co-expression analysis alone. Co-

expression analysis was only one major component of the workflow which was described (Bordych et al., 2013).

Since co-expression as utilized in this work returns a large amount of putative candidate genes, the list of candidate genes has to be pruned to yield a number of suitable targets, which can be actually tested in wet laboratory experiments. One major step in reducing the list co-expressed genes down to a manageable list is filtering. Co-expressed genes need manual validation by usage of publicly available information about the corresponding gene or (predicted) protein. An important prerequisite for a (photorespiratory) transport protein is its tight integration into the membrane of occupied organelles to facilitate shuttling of compounds through the membrane. Integration of transport proteins is achieved through a network of (hydrophobic) transmembrane domains distributed over the whole length of the protein. Public available algorithms, integrated on internet platforms such as ConPredII and SOSUI (Arai et al., 2004; Hirokawa et al., 1998) can be used easily to predict such transmembrane domains. Still, these prediction algorithms can only predict and visualize membrane topologies of an amino acid sequence which was used as query. While further increasing the likelihood of focusing on true transporters, one has to take into account that the presence of transmembrane domains *per se* does not automatically render the candidate protein a true transport protein.

Phrasing this differently, one has to consider that not all proteins, which are tightly integrated into plant lipid membranes through several transmembrane domains, are necessarily transport proteins. The following paragraph is going to summarize knowledge about catalytically active enzymes which are tightly inserted into the plant membrane via several transmembrane domain.

Enzymes can be tightly inserted into plant membranes like transport proteins

Transport proteins are tightly inserted into the lipid membrane to facilitate shuttling of compounds over the membrane. Besides the vast majority of transport proteins, enzymes can also be found to be integrated into the membrane by several transmembrane domains. So far this requirement for multiple transmembrane domains for enzymes is rather rare, but has been published for example for rhomboid proteases (Lemberg et al., 2005; Wang et al., 2006). Rhomboid proteases are deeply integrated

into membranes and their respective active site resides within the membrane. Consequently, rhomboid proteases cleave other proteins within their transmembrane domain or in closely connected regions (Freeman, 2008). Pioneer work about rhomboid proteases was performed in *Drosophila melanogaster* by characterization of Rhomboid-1 the first characterized intramembrane serine protease. Rhomboid-1 was shown to cleave the membrane-anchored TGFα-like growth factor Spitz within its transmembrane anchor (Urban et al., 2001).

Interestingly, rhomboid proteases are also described for plants. *Arabidopsis* was shown to harbor Rhomboid-like protein 2 (RBL2) which shows the same capability to cleave the membrane-anchored TGFα-like growth factor Spitz and other *Drosophila* ligands (Kanaoka et al., 2005). With RBL-10 Thompson and co-workers described a rhomboid protease in *Arabidopsis*, which is integrated into the chlorplast membrane. It affects flower development through cleavage of an unknown ligand (Thompson et al., 2012).

Besides rhomboid proteases, an enzymatic component of the vitamin E biosynthesis pathway is located to the membrane of the chloroplast. Here, vitamin E describes the group of compounds which show α-tocopherol activitity, including the complete series of tocochromanols: α-, β-, γ-, and δ-tocopherol, aswell as α-, β-, γ-, and δ-tocotrienol (Colombo, 2010; DellaPenna and Pogson, 2006; Mayer et al., 1967). They are synthesized exlusivly by photosynthetic organisms and fulfill multifarious functions as antioxidants in both, plants (Falk and Munne-Bosch, 2010; Maeda and DellaPenna, 2007; Munne-Bosch and Falk, 2004) and animals (Reiter et al., 2007; Tucker and Townsend, 2005).

The enzyme VTE5 (vitamin E pathway 5, phytol kinase) is an integral membrane protein in *Arabidopsis*. Valentin an co-workers present evidence from an in-depth bioinformatic analysis for VTE5 being an integral membrane protein of the chloroplast. They showed that VTE5 forms six independent transmembrane helices. Further, homology to the PFAM domain of cytidylyltransferases was found. Members of the cytidylyltransferase family are integral proteins of membranes itself (Valentin et al., 2006).

In addition, it was shown in the same year that phytol kinase (VTE5) activity is located to the chloroplast envelope membrane together with phytyl phosphate kinase activity

by an uncharacterized protein. VTE5 is part of the enzymatic machinery which converts phytol originating from chlorophyll breakdown into phytyl diphosphate, which is in turn used for vitamin E synthesis (Ischebeck et al., 2006).

At1g78620 - Candidate gene for photorespiratory transport

At1g78260 is the gene of interest in this work, which originated from co-expression analysis together with additional candidate genes. At1g78620 fulfils all criteria, which were set up during the initial part of this work, while searching for photorespiratory transport candidate genes. It is co-expressed with characterized enzymes of the core photorespiratory cycle in *Arabidopsis*. These co-expression results for At1g78620 have not been made publicly available in the first manuscript of this work, which is shown in the following section (Bordych et al., 2013). At1g78620 was the prime target for further investigation. It was intended to keep ongoing work hidden from the eyes of other scientists.

Public available data for At1g78620, accumulated during the initial phase of the project described in this thesis, shows that the predicted peptide for AT1g78620 consists of 333 amino acids according to the Plant Proteom Database (Sun et al., 2009). Additionally, six transmembrane domains (TMDs) were predicted by the TMpred software (Hofmann and Stoffel, 1993). Such a distribution of transmembrane domains serves as a prerequisite for the protein to be inserted into membranes, to act as a transport protein. According to TargetP (Emanuelsson et al., 2007) a chloroplast transit peptide was anticipated for the n-terminal part of the predicted protein, which argues towards localization of the protein, coded by At1g78620, to one of the organelles which is important in photorespiration. Moreover, the Plant Proteom Database predicted a cytidylyltransferase family domain according to PFAM (Finn et al., 2014). All information given above is schematically depicted in Figure 2.

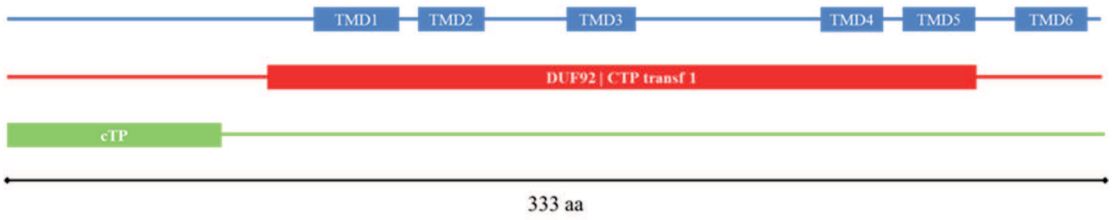


Figure 1 **Predicted At1g78620 protein.** Transmembrane Domain (TMD) patterning (blue), predicted domain of unknown function (DUF)/Cytidylyltransferase family domain (red), and the predicted chloroplast targeting peptide (green) are depicted along with the overall peptide length (black).

The work as presented here, shows how At1g78620 was chosen as a prime target for photorespiratory transport based on co-expression analysis. This work is summarized in the first publication which follows in the next section. Later parts of this thesis (manuscript 2 and manuscript 3) describe the functional characterization of At1g78620, answering the question if At1g78620 is indeed occupied in the photorespiratory cycle as a transport protein.

REFERENCES

Anderson, L.E. (1971) Chloroplast and cytoplasmic enzymes. II. Pea leaf triose phosphate isomerases. Biochimica et biophysica acta, **235**(1), 237-244.

Andrews, T.J., Lorimer, G.H., Tolbert, N.E. (1973) Ribulose diphosphate oxygenase. I. Synthesis of phosphoglycolate by fraction-1 protein of leaves. Biochemistry, **12**(1), 11-18.

Arai, M., Mitsuke, H., Ikeda, M., Xia, J.X., Kikuchi, T., Satake, M., Shimizu, T. (2004) ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. Nucleic acids research, **32**(Web Server issue), W390-393.

Bauwe, H., Hagemann, M., Fernie, A.R. (2010) Photorespiration: players, partners and origin. Trends in plant science, **15**(6), 330-336.

Boldt, R., Edner, C., Kolukisaoglu, U., Hagemann, M., Weckwerth, W., Wienkoop, S., Morgenthal, K., Bauwe, H. (2005) D-GLYCERATE 3-KINASE, the last unknown enzyme in the photorespiratory cycle in Arabidopsis, belongs to a novel kinase family. The Plant cell, **17**(8), 2413-2420.

Bordych, C., Eisenhut, M., Pick, T.R., Kuelahoglu, C., Weber, A.P. (2013) Co-expression analysis as tool for the discovery of transport proteins in photorespiration. Plant Biol (Stuttg), **15**(4), 686-693.

Brazma, A. and Vilo, J. (2000) Gene expression data analysis. FEBS letters, **480**(1), 17-24.

Calvin, M. and Benson, A.A. (1948) The Path of Carbon in Photosynthesis. Science, **107**(2784), 476-480.

Colombo, M.L. (2010) An update on vitamin E, tocopherol and tocotrienol-perspectives. Molecules, **15**(4), 2103-2113.

Decker, J.P. (1955) A Rapid, Postillumination Deceleration of Respiration in Green Leaves. Plant physiology, **30**(1), 82-84.

DellaPenna, D. and Pogson, B.J. (2006) Vitamin synthesis in plants: tocopherols and carotenoids. Annu Rev Plant Biol, **57**, 711-738.

Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America, **95**(25), 14863-14868.

Eisenhut, M., Pick, T.R., Bordych, C., Weber, A.P. (2013) Towards closing the remaining gaps in photorespiration--the essential but unexplored role of transport proteins. Plant Biol (Stuttg), **15**(4), 676-685.

Emanuelsson, O., Brunak, S., von Heijne, G., Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc, **2**(4), 953-971.

Eskling, M., Arvidsson, P.O., Åkerlund, H.E. (1997) The xanthophyll cycle, its regulation and components. Physiol Plant, **100**(4), 806-816.

Falk, J. and Munne-Bosch, S. (2010) Tocochromanol functions in plants: antioxidation and beyond. J Exp Bot, **61**(6), 1549-1566.

Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L., Tate, J., Punta, M. (2014) Pfam: the protein families database. Nucleic acids research, **42**(Database issue), D222-230.

Foyer, C., Descourvieres, P., Kunert, K. (1994) Protection against oxygen radicals: an important defence mechanism studied in transgenic plants. Plant, cell & environment, **17**(5), 507-523.

Freeman, M. (2008) Rhomboid proteases and their biological functions. Annu Rev Genet, **42**, 191-210.

Fryer, M.J. (1992) The Antioxidant Effects of Thylakoid Vitamin-E(Alpha-Tocopherol). Plant Cell and Environment, **15**(4), 381-392.

Hirokawa, T., Boon-Chieng, S., Mitaku, S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics, **14**(4), 378-379.

Hofmann, K. and Stoffel, W. (1993) TMbase-A database of membrane spanning protein segments. Biol. Chem. Hoppe-Seyler, **374**, 166.

Horton, P., Ruban, A.V., Walters, R.G. (1994) Regulation of Light Harvesting in Green Plants (Indication by Nonphotochemical Quenching of Chlorophyll Fluorescence). Plant physiology, **106**(2), 415-420.

Howitz, K.T. and McCarty, R.E. (1985) Kinetic characteristics of the chloroplast envelope glycolate transporter. Biochemistry, **24**(11), 2645-2652.

Howitz, K.T. and McCarty, R.E. (1991) Solubilization, partial purification, and reconstitution of the glycolate/glycerate transporter from chloroplast inner envelope membranes. Plant physiology, **96**(4), 1060-1069.

Ischebeck, T., Zbierzak, A.M., Kanwischer, M., Dormann, P. (2006) A salvage pathway for phytol metabolism in Arabidopsis. The Journal of biological chemistry, **281**(5), 2470-2477.

Kanaoka, M.M., Urban, S., Freeman, M., Okada, K. (2005) An Arabidopsis Rhomboid homolog is an intramembrane protease in plants. FEBS letters, **579**(25), 5723-5728.

Kelly, G.J. and Latzko, E. (1976) Inhibition of spinach-leaf phosphofructokinase by 2-phosphoglycollate. FEBS letters, **68**(1), 55-58.

Kozaki, A. and Takeba, G. (1996) Photorespiration protects C3 plants from photooxidation.

Kuhlbrandt, W., Wang, D.N., Fujiyoshi, Y. (1994) Atomic model of plant light-harvesting complex by electron crystallography. Nature, **367**(6464), 614-621.

Lorimer, G.H., Andrews, T.J., Tolbert, N.E. (1973) Ribulose diphosphate oxygenase. II. Further proof of reaction products and mechanism of action. Biochemistry, **12**(1), 18-23.

Maeda, H. and DellaPenna, D. (2007) Tocopherol functions in photosynthetic organisms. Curr Opin Plant Biol, **10**(3), 260-265.

Mayer, H., Metzger, J., Isler, O. (1967) [The stereochemistry of natural gamma-tocotrienol (plastochromanol-3), plastochromanol-8 and plastochromenol-8]. Helv Chim Acta, **50**(5), 1376-1393.

Melis, A. (1991) Dynamics of photosynthetic membrane composition and function. Biochimica et Biophysica Acta (BBA)-Bioenergetics, **1058**(2), 87-106.

Munne-Bosch, S. and Falk, J. (2004) New insights into the function of tocopherols in plants. Planta, **218**(3), 323-326.

Niyogi, K.K. (1999) PHOTOPROTECTION REVISITED: Genetic and Molecular Approaches. Annu Rev Plant Physiol Plant Mol Biol, **50**, 333-359.

Niyogi, K.K., Bjorkman, O., Grossman, A.R. (1997) The roles of specific xanthophylls in photoprotection. Proceedings of the National Academy of Sciences of the United States of America, **94**(25), 14162-14167.

Niyogi, K.K., Grossman, A.R., Bjorkman, O. (1998) Arabidopsis mutants define a central role for the xanthophyll cycle in the regulation of photosynthetic energy conversion. The Plant cell, **10**(7), 1121-1134.

Ogren, W.L. and Bowes, G. (1971) Ribulose diphosphate carboxylase regulates soybean photorespiration. Nat New Biol, **230**(13), 159-160.

Peterhansel, C. and Maurino, V.G. (2011) Photorespiration redesigned. Plant physiology, **155**(1), 49-55.

Pick, T.R., Brautigam, A., Schulz, M.A., Obata, T., Fernie, A.R., Weber, A.P. (2013) PLGG1, a plastidic glycolate glycerate transporter, is required for photorespiration and defines a unique class of metabolite transporters. Proceedings of the National Academy of Sciences of the United States of America, **110**(8), 3185-3190.

Reiter, E., Jiang, Q., Christen, S. (2007) Anti-inflammatory properties of alpha- and gamma-tocopherol. Mol Aspects Med, **28**(5-6), 668-691.

Reumann, S. and Weber, A.P. (2006) Plant peroxisomes respire in the light: some gaps of the photorespiratory C2 cycle have become filled--others remain. Biochimica et biophysica acta, **1763**(12), 1496-1510.

Sharkey, T.D. (1988) Estimating the rate of photorespiration in leaves. Physiol Plant, **73**(1), 147-152.

Somerville, C.R. (2001) An early Arabidopsis demonstration. Resolving a few issues concerning photorespiration. Plant physiology, **125**(1), 20-24.

Somerville, C.R. and Ogren, W.L. (1979) Phosphoglycolate Phosphatase-Deficient Mutant of Arabidopsis. Nature, **280**(5725), 833-836.

Somerville, C.R. and Ogren, W.L. (1980a) Inhibition of photosynthesis in Arabidopsis mutants lacking leaf glutamate synthase activity. Nature, **286**(5770), 257-259.

Somerville, C.R. and Ogren, W.L. (1980b) Photorespiration mutants of Arabidopsis thaliana deficient in serine-glyoxylate aminotransferase activity. Proceedings of the National Academy of Sciences of the United States of America, **77**(5), 2684-2687.

Somerville, C.R. and Ogren, W.L. (1981) Photorespiration-deficient Mutants of Arabidopsis thaliana Lacking Mitochondrial Serine Transhydroxymethylase Activity. Plant physiology, **67**(4), 666-671.

Somerville, C.R. and Ogren, W.L. (1982) Mutants of the cruciferous plant Arabidopsis thaliana lacking glycine decarboxylase activity. The Biochemical journal, **202**(2), 373-380.

Sun, Q., Zybailov, B., Majeran, W., Friso, G., Olinares, P.D., van Wijk, K.J. (2009) PPDB, the Plant Proteomics Database at Cornell. Nucleic acids research, **37**(Database issue), D969-974.

Takahashi, S., Bauwe, H., Badger, M. (2007) Impairment of the photorespiratory pathway accelerates photoinhibition of photosystem II by suppression of repair but not acceleration of damage processes in Arabidopsis. Plant physiology, **144**(1), 487-494.

Telfer, A., Dhami, S., Bishop, S.M., Phillips, D., Barber, J. (1994) beta-Carotene quenches singlet oxygen formed by isolated photosystem II reaction centers. Biochemistry, **33**(48), 14469-14474.

Thompson, E.P., Smith, S.G., Glover, B.J. (2012) An Arabidopsis rhomboid protease has roles in the chloroplast and in flower development. J Exp Bot, **63**(10), 3559-3570.

Tolbert, N. (1971) Microbodies-peroxisomes and glyoxysomes. Annual review of plant physiology, **22**(1), 45-74.

Tucker, J.M. and Townsend, D.M. (2005) Alpha-tocopherol: roles in prevention and therapy of human disease. Biomed Pharmacother, **59**(7), 380-387.

Urban, S., Lee, J.R., Freeman, M. (2001) Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases. Cell, **107**(2), 173-182.

Usadel, B., Obayashi, T., Mutwil, M., Giorgi, F.M., Bassel, G.W., Tanimoto, M., Chow, A., Steinhauser, D., Persson, S., Provart, N.J. (2009) Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. Plant, cell & environment, **32**(12), 1633-1651.

Valentin, H.E., Lincoln, K., Moshiri, F., Jensen, P.K., Qi, Q., Venkatesh, T.V., Karunanandaa, B., Baszis, S.R., Norris, S.R., Savidge, B., Gruys, K.J., Last, R.L. (2006) The Arabidopsis vitamin E pathway gene5-1 mutant reveals a critical role for phytol kinase in seed tocopherol biosynthesis. The Plant cell, **18**(1), 212-224.

Warburg, O. (1928) *Über die Geschwindigkeit der photochemischen Kohlensäurezersetzung in lebenden Zellen* Springer.

Weber, A., Menzlaff, E., Arbinger, B., Gutensohn, M., Eckerskorn, C., Flugge, U.I. (1995) The 2-oxoglutarate/malate translocator of chloroplast envelope membranes: molecular cloning of a transporter containing a 12-helix motif and expression of the functional protein in yeast cells. Biochemistry, **34**(8), 2621-2627.

MANUSCRIPT 1

Co-expression analysis as tool for the discovery of transport proteins in photorespiration

Co-expression analysis as tool for the discovery of transport proteins in photorespiration

Christian Bordych, Marion Eisenhut, Thea R. Pick, Canan Kuelahoglu & Andreas P.M. Weber*

Institute of Plant Biochemistry, Center of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-University, Universitätsstraße 1, D-40225 Düsseldorf

Running Head: Candidates for photorespiratory transport

*Corresponding author:

Andreas P.M. Weber

Institute of Plant Biochemistry, Center of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-University, Universitätsstraße 1, D-40225 Düsseldorf

Phone: +49 211 81 12347

Fax: +49 211 81 13706

Keywords: co-expression, photorespiration, transporters, Arabidopsis

plant biology

Plant Biology ISSN 1435-8603

REVIEW ARTICLE

Co-expression analysis as tool for the discovery of transport proteins in photorespiration

C. Bordych, M. Eisenhut, T. R. Pick, C. Kuelahoglu & A. P. M. Weber

Institute of Plant Biochemistry, Center of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-University, Düsseldorf, Germany

Keywords
Arabidopsis; co-expression; photorespiration; transporters.

Correspondence
A. P. M. Weber, Institute of Plant Biochemistry, Center of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-University, Universitätsstraße 1, D-40225 Düsseldorf, Germany.
E-mail: andreas.weber@uni-duesseldorf.de

Editor
H. Rennenberg

Received: 17 December 2012; Accepted: 25 February 2013

doi:10.1111/plb.12027

ABSTRACT

Shedding light on yet uncharacterised components of photorespiration, such as transport processes required for the function of this pathway, is a prerequisite for manipulating photorespiratory fluxes and hence for decreasing photorespiratory energy loss. The ability of forward genetic screens to identify missing links is apparently limited, as indicated by the fact that little progress has been made with this approach during the past decade. The availability of large amounts of gene expression data and the growing power of bioinformatics, paired with availability of computational resources, opens new avenues to discover proteins involved in transport of photorespiratory intermediates. Co-expression analysis is a tool that compares gene expression data under hundreds of different conditions, trying to find groups of genes that show similar expression patterns across many different conditions. Genes encoding proteins that are involved in the same process are expected to be simultaneously expressed in time and space. Thus, co-expression data can aid in the discovery of novel players in a pathway, such as the transport proteins required for facilitating the transfer of intermediates between compartments during photorespiration. We here review the principles of co-expression analysis and show how this tool can be used for identification of candidate genes encoding photorespiratory transporters.

INTRODUCTION

Starting in 1978, several key enzymes of photorespiratory metabolism (reviewed in Timm & Bauwe, this issue) were identified by Chris and Shauna Somerville and Bill Ogren, utilising a forward genetic approach (reviewed in Somerville 2001). Shifting mutagenised *Arabidopsis thaliana* populations from elevated CO_2 concentrations (1%) to atmospheric conditions (0.03%) revealed mutants, which were not able to cope with ambient CO_2 concentrations. By labelling photosynthetic products with $^{14}CO_2$ and subsequent enzymatic tests, five individual players in the photorespiratory C_2 cycle were discovered, namely phosphoglycolate phosphatase (Somerville & Ogren 1979), serine:glyoxylate aminotransferase (Somerville & Ogren 1980b), glutamate synthase (Somerville & Ogren 1980a), serine hydroxymethyltransferase (Somerville & Ogren 1981) and glycine decarboxylase (Somerville & Ogren 1982). Based on this work, a significant part of the photorespiratory cycle was deciphered at the biochemical level. Since the detection of glycerate kinase (Boldt *et al.* 2005), all key enzymes of the core C_2 cycle have now been identified at the genetic level in the model plant *A. thaliana* and have been characterised biochemically.

The C_2 cycle is a highly compartmentalised pathway localised to chloroplasts, peroxisomes, mitochondria and the cytosol. Thus, some intermediates of the cycle need to be transported out of a compartment and re-imported into another one *via* transport proteins for further reactions to occur (for review see Eisenhut *et al.* 2012). A large number of these transport processes in photorespiration between chloroplasts, peroxisomes and mitochondria remain uncharacterised, even more than 30 years after discovery of the first enzyme involved in photorespiration. Forward genetic screens have, to date, contributed little to closing the gaps in knowledge about the transport of photorespiratory intermediates between compartments.

In this article, we review co-expression analysis and explore it as a novel approach to discover these missing players in photorespiration. We demonstrate that co-expression analysis can serve as a candidate discovery tool for reverse genetic approaches to identify as yet unknown proteins involved in transport process during photorespiration.

IDENTIFICATION OF CANDIDATE GENES THROUGH CO-EXPRESSION ANALYSIS

In order to discover unknown transport proteins involved in photorespiratory metabolism, we chose a novel contemporary approach, the combined use of co-expression analysis and reverse genetic studies of deduced candidate genes. The subsequent workflow is shown in Fig. 1 and is explained below.

The principle idea of co-expression analysis

Within the last 15 years it became possible to measure the expression level of every single gene of an organism under various conditions in a high-throughput manner, utilising microarrays, and thereby creating vast amounts of data (Brazma & Vilo 2000). It was a major challenge to visualise, to order and to integrate these data into public available databases, which had to be developed to make efficient use

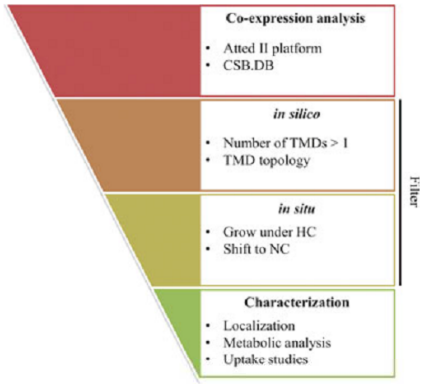


Fig. 1. Discovery of new genes involved in photorespiratory transport can be divided into four major parts; each part severely reduces the number of possible candidate genes. Co-expression analysis is carried out on the Atted-II platform. All hits are filtered afterwards for transmembrane proteins with more than one predicted transmembrane domain (TMD). Distribution of predicted TMDs can be compared with known transport proteins in the transport database (TransportDB; www.membranetransport.org). Finally, T-DNA lines are screened for a photorespiratory phenotype; mutants showing a positive result are subsequently characterised for the gene function.

of these resources. Exploratory multivariate statistical tools, such as hierarchical clustering, became the technique of choice for analysing these large datasets to finally let the raw data speak: converting 'data into information and then information into knowledge' (Eisen *et al.* 1998). Basically, in these approaches a similarity score is assigned to each possible gene pair, comparing their expression states under many different conditions. On the basis of these scores a distance matrix is generated, and in the case of clustering approaches, a dendrogram is produced in which the genes showing the most similar expression patterns are located next to each other (Eisen *et al.* 1998).

Clustered genes within a dendrogram are thought likely to participate in the same process (*e.g.,* a metabolic pathway). As proof of principle, Persson *et al.* (2005) provided evidence for a functional relationship between co-expressed genes involved in cellulose synthesis during secondary cell wall formation. At least three distinct cellulose synthase genes (*CESA4*, *CESA7* and *CESA8*) are required for cellulose synthesis during secondary cell wall formation. Knocking-out each one of these genes caused, in each case, a defect in cell wall formation, indicating that these genes code for subunits that build up a cellulose synthase complex. All three proteins have been shown to be present at the same time in the same tissue in *Arabidopsis*, pointing to orchestrated gene expression of corresponding genes (Taylor *et al.* 2003). Based on this observation Persson *et al.* (2005) hypothesised that additional partners involved in secondary cell wall synthesis are also co-expressed with these known genes. To identify some of these genes, microarray data were screened for further co-expressed genes. The study was based solely on expression data from Affymetrix ATH1 microarrays to avoid distortions caused by data from different microarray platforms. To exclude those chips representing potential outlier data, quality control was performed prior to clustering. Basically, data from each chip were tested to fit a model derived from data of all remaining chips in the set ('Deleted Residuals'). Using this 'quality control', 95 out of 503 available Affymetrix chips were rejected for future studies.

Results from subsequent clustering showed that *CESA* genes, which are involved in primary and secondary cell wall formation, 'exhibit a high degree of co-expression based on correlation coefficients' (Persson *et al.* 2005). This observation is consistent with the finding that all corresponding proteins are expressed simultaneously in a time- and space-dependent manner (Taylor *et al.* 2003), and serves as a prerequisite for further analyses. *CESA4, 7* and *8* were top of the list of co-regulated genes, with the three *CESA* genes 'confirming tight co-regulation of the three CESA subunits' (Persson *et al.* 2005). Four different genes with no functional annotation from the list of the 40 most co-regulated genes of *CESA4, 7* and *8* were investigated for their function in secondary cell wall synthesis. *Arabidopsis* knockout mutants for two of these genes showed a phenotype resembling a defect in secondary cellulose synthesis, indicating a functional relationship between the proteins encoded by the co-expressed genes (Persson *et al.* 2005).

Yonekura-Sakakibara *et al.* (2007) took advantage of co-expression analysis to identify a gene involved in specific glycosylation of flavonol in *Arabidopsis* out of 107 candidate genes from a multigene family. Investigation of glycosylation patterns of flavonoid structures demanded five additional genes be involved in glycosylation in *Arabidopsis*, besides the four already known genes. Correlation between known structural and regulatory genes involved in flavonoid biosynthesis in *Arabidopsis* and all 107 members of the multigene family were analysed. Showing that the already identified and functionally characterised genes of flavonoid biosynthesis cluster together, these authors proved that the functional relationship of these genes is mirrored by their expression state. For five out of the initial 107 candidate genes a significant positive correlation with the bait genes was observed. Two independent knockout mutants for one of these genes were analysed, demonstrating a deficiency in synthesis of a specific flavonol. This phenotype could be rescued by re-introduction of the intact gene (Yonekura-Sakakibara *et al.* 2007). The combination of co-expression analysis, metabolic profiling and reverse genetics was shown to 'be a versatile tool for functional identification of genes that belong to a multigene family and to complete the model of a particular metabolic pathway in *Arabidopsis*' (Yonekura-Sakakibara *et al.* 2007).

These and a number of other studies (*e.g.,* Stuart *et al.* 2003; Lee *et al.* 2004) demonstrate the power of co-expression analysis as a candidate discovery tool, which encouraged us to explore this approach for the identification of yet unknown genes putatively involved in the transport of photorespiratory intermediates between the compartments. This idea gained further support from a study that demonstrated that genes for peroxisomal enzymes involved in photorespiration are co-expressed (Reumann & Weber 2006), which was later confirmed through mutant analyses, demonstrating a role of peroxisomal malate dehydrogenases in photorespiration (Cousins *et al.* 2008).

Application of co-expression analysis tools

During the last few years several tools have been developed to make the results of co-expression analyses publicly available (for a detailed review on popular web tools and databases see Aoki *et al.* 2007). However, for efficiently using such web tools to connect yet unknown genes with functionally annotated ones, one should be aware of all features offered by the web tool. The following paragraphs focus on different features and their consequences for co-expression results.

To measure co-expression between two genes, one has to apply a measure for co-expression for each gene pair. Most online tools use Pearson's correlation coefficient (PCC). PCC values range between −1 (anti-correlated, gene A is up-regulated when gene B is down-regulated in many different experiments/conditions) and +1 (correlated, increased expression of gene A goes together with increased expression of gene B over many samples). PCC has been shown to be prone to outliers, rendering gene pairs correlated instead of non-correlated. To circumvent this effect, several tools utilise Spearman's correlation coefficient (SCC), which does not use the expression values of each gene directly for calculation, but the ranks which are assigned to each gene under each condition beforehand (Usadel *et al.* 2009). In the case of the web tool Atted, mutual rank (MR) is used as a measure of co-expression, which resembles the geometric mean of the correlation ranks (which are based on PCC) between two genes (Obayashi & Kinoshita 2009). These authors state that a functional relationship between genes is not always mirrored by a high PCC, thus true relationships sometimes cannot be resolved. To circumvent this problem, a new correlation measurement was introduced. MR puts emphasis on the fact that the correlation rank for the relationship of gene A to gene B differs from the rank for gene B to gene A, and represents the geometric mean of both rankings (Obayashi & Kinoshita 2009).

Several co-expression web tools (*e.g.,* Atted) accept more than one query gene. Co-expression analysis with groups of bait genes results in a list of genes that are co-expressed with the entire group of queries. This option turned out to be useful if it is not known which query serves as the best, leading to more reliable candidate lists of co-expressed genes (Usadel *et al.* 2009). Further, a multi-query run can confirm results for runs with a single query to provide additional support. Since co-expression analysis using web tools with pre-calculated co-expression scores is a very rapid method, this additional step with groups of query genes can easily be included in the analysis, when offered by the platform.

The most available online tools include the option to preselect conditions (condition-dependent) under which co-expression should be investigated. Thereby, co-expression analysis can be restricted to a defined set of microarrays (thus conditions or experiments) (Obayashi *et al.* 2011). This feature is very useful, because one can limit the search to single tissues or special conditions to further improve the quality of the results. Some popular online tools add the option of individualising conditions by allowing a user-specified set or even upload data that are not publicly available. Nonetheless, a condition-independent approach (using all available microarray data) can still prove very fruitful in finding overall gene relationships or obtaining an initial overview (Usadel *et al.* 2009). Furthermore, a condition-independent approach combined with a condition-dependent approach has been reported to be effective in discovering new interaction partners (Hirai *et al.* 2007).

The number of online tools for co-expression analysis is still increasing, so that only a limited overview of their features can be discussed here. However, running co-expression analysis *via* an online platform requires profound knowledge about all given features in order to select the right platform, subsequently the appropriate features. Usadel *et al.* (2009) have published a clear and detailed article about different online tools dealing with co-expression.

Co-expression results

Once the right parameters for co-expression analysis are set up, one can obtain most co-expressed genes within a few minutes. However, the main work is still to probe the list of genes for promising candidate genes, thus further manual filtering using *a priori* knowledge is required. As an example for such a filtering approach, a common workflow from our lab to identify candidate genes coding for photorespiratory transporters is described here.

Initial co-expression analysis was carried out on the Atted-II platform (atted.jp; Obayashi *et al.* 2007) using 13 genes coding for soluble proteins involved in photorespiratory metabolism (Table 1), resulting in a list of 300 ranked co-expressed genes per query gene (in total 3,900 single hits). As previously shown,

Table 1. Co-expression of photorespiratory genes. Co-expression analysis was carried out on the Atted platform (www.atted.jp) using each AGI of column 1. The resulting TOP300 co-expressed genes were screened for photorespiratory genes from column 1. Column 'TOP300' shows how often each photorespiratory gene is present, while 'TOP50' presents how frequent the gene is among the first 50 ranked genes.

locus	description	TOP300	TOP50
At1 g11860	Glycine decarboxylase T protein (GDT1)	7	3
At1 g23310	Glutamate:Glyoxylate aminotransferase 2 (GGT2)	7	5
At1 g32470	Glycine decarboxylase H protein 3 (GDH3)	9	4
At1 g48030	Glycine decarboxylase L protein (mLPD1)	8	5
At1 g68010	Hydroxypyruvate reductase 1 (HPR1)	7	4
At2 g26080	Glycine decarboxylase P protein 2 (GDP2)	6	2
At2 g35370	Glycine decarboxylase H protein 1 (GDH1)	8	4
At4 g33010	Glycine decarboxylase P protein 1 (GDP1)	10	5
At4 g37930	Serine hydroxymethyltransferase (SHM1)	8	4
At5 g04140	Fd-dependent glutamate synthase (Fd-GOGAT)	12	6
At5 g12860	Dicarboxylate transporter 1 (DiT1)	5	3
At5 g35630	Glutamine synthase 2 (GS2)	11	5
At5 g64290	Dicarboxylate transporter 2.1 (DiT2.1)	4	2

genes for photorespiratory enzymes in peroxisomes are simultaneously expressed in a variety of experiments and form a cluster (Reumann & Weber 2006). This finding was already reflected in the initial results obtained from the Atted-II platform, as many of the genes coding for these enzymes were frequently placed among the TOP300 genes co-expressed for each query, while most of these hits were ranked within the TOP50 of all hits (Table 1).

Filtering

Subsequent filtering of co-expression analysis results strongly depends on the initial question asked. In this example, we were looking for additional players in photorespiration, specifically for transport proteins. In order to restrict the list of genes co-expressed with known genes of photorespiration, we applied different filters to increase the probability of identifying genes coding for transporters. The filtering procedure did not aim at generating a complete list of all photorespiratory transporters already present in the original list generated through co-expression analysis, but rather an essential list of the most promising candidate genes that should be analysed in detail through follow-up experiments, such as reverse genetic analyses.

Criterion 1: Prediction of transmembrane domains (in silico)
First, a list of membrane proteins was generated through discarding all genes for which the encoded protein had been experimentally proven to be soluble. The online platform ARAMEMNON (http://aramemnon.botanik.uni-koeln.de; Schwacke *et al.* 2007) served as the basis for this analysis. ARAMEMNON merges data from up to 17 individual prediction programs, and generates a consensus prediction for transmembrane domains (Schwacke *et al.* 2003). Further support is added from experimental data (if available) *via* linking to corresponding publications. Second, genes encoding proteins with less than two predicted transmembrane domains (TMDs) are discarded. Single TMDs often render proteins membrane-bound, rather than enabling them to transport metabolites or cofactors. It cannot be ruled out that some genuine transport proteins were lost in this step, but the list of putative candidates was considerably shortened, which therefore made this step reasonable. Third, focus centred on all proteins showing typical distributions of TMDs, as frequently found in known transport proteins. While there is no single common pattern known for the distribution of TMDs along the entire peptide chain, the transporter database (www.
membranetransport.org; Ren *et al.* 2004, 2007) provides a good overview of how TMDs are distributed in proteins that have been characterised as transporters. ARAMEMNON visualises data from transmembrane helice prediction programs and generates a consensus prediction (TmConsens), including the number of predicted TMDs and the amino acid sequences that contribute to the estimated transmembrane helices. For additional support, the subcellular proteomic database (SUBA; http://suba.plantenergy.uwd.edu.au) is linked and gives hydropathy plots (Heazlewood *et al.* 2007) for comparison with the obtained TmConsens data.

ARAMEMNON is not only helpful in terms of identification of transmembrane proteins, but also in discovering genes that are members of multi-gene families. We decided to exclude genes that are members of larger families because of possible redundant functions of the gene family members, which would complicate reverse genetic analyses. Clearly, this filter criterion might exclude genes that play a role in the transport of photorespiratory intermediates; however, it helps focus on single-copy genes whose effect on photorespiration is easier to study using reverse genetic screens. That is, multi-gene family members were given lower priority for experimental testing, but were not excluded from the screen.

Criterion 2: Photorespiratory phenotype (in vivo)
The second step of filtering was to check public seed stocks for availability of corresponding T-DNA lines. Today, *Arabidopsis* T-DNA insertion lines are available for most known *Arabidopsis* genes, and can be ordered from the stock centres: NASC (Nottingham Arabidopsis Stock Centre, Nottingham University, UK) and ABRC (Arabidopsis Biological Resource Center, Ohio State University, USA). Whenever possible, at least two individual T-DNA lines were ordered to check if the knockout of a candidate gene results in a photorespiratory phenotype, as for example observed for plants deficient in serine hydroxymethyltransferase activity (Voll *et al.* 2006; Jamai *et al.* 2009). Several photorespiratory phenotypes are reviewed in this Special Issue in the paper of Timm & Bauwe (2012).

T-DNA lines were grown under high CO_2 conditions (HC; 0.3% CO_2) and selected for homozygous plants. For shifting experiments, homozygous seeds were spotted on solid MS medium, together with wild-type plants as a negative control, and grown for approximately 2 weeks under HC conditions. One half of these plant sets was shifted to normal CO_2 conditions (NC; 0.038% CO_2), while the other half remained under HC conditions as a control. Plants were checked continuously for a photorespiratory phenotype, *e.g.*, leaf chlorosis or stunted growth, for a maximum of 10 days. To investigate effects in later developmental stages, plants grown under HC conditions were transferred to soil and further grown at HC and used for shifts to NC at later stages.

In some rare cases, a homozygous knockout line was not available for analysis due to embryo lethality. However, in such cases it is possible to check the candidate by applying the miRNA (Bartel 2004) approach, *e.g.*, through generation of artificial miRNA (amiRNA) lines (Schwab *et al.* 2006). A common approach in our lab is to repress the transcript amount of the corresponding genes with amiRNA under control of either a constitutive promoter or an inducible one. Silencing in the constitutive lines can reveal a phenotype in photorespiratory shifting experiments as long as plants are viable. Controlling the amiRNA expression level using an inducible promoter is particularly helpful in cases where repression from a constitutive promoter fails, *e.g.*, if constitutive down-regulation of the candidate gene prevents establishment of viable seedlings.

SCREENING OF CANDIDATE GENES FOR POSSIBLE FUNCTION IN PHOTORESPIRATION

Arabidopsis T-DNA lines adversely affected in NC conditions, having pale green or yellow leaves, were further characterised. The first step in the detailed characterisation of these selected candidates was localisation of the corresponding protein inside the plant cell. It was necessary to verify that this protein is indeed targeted to the membrane of the predicted organelle for it to function there as transporter.

In a second step, the metabolic response of mutants in the candidate genes to a shift from HC to NC conditions was studied. Levels of metabolites associated with photorespiration were measured using gas chromatography–mass spectrometry (GC/MS) to test for altered levels of specific metabolites between mutant and wild-type plants. Accumulation of a specific metabolite might indicate a deficiency in transport of this metabolite across an organellar membrane. This hypothesis could then be tested through recombinant expression of the candidate protein in *Escherichia coli*, *Saccharomyces cerevisiae* or a cell-free expression system (Nozawa *et al.* 2007; Haferkamp & Linka 2012), followed by reconstitution into proteoliposomes and radiolabelled flux assay (for review see Haferkamp & Linka 2012). In some cases, heterologous expression of proteins and reconstitution in artificial liposome systems fail, so that the transported compound or corresponding transport kinetics remain elusive. In such cases, transport assays with isolated organelles (Werdan & Heldt 1972) has afforded an alternative approach to assess possible impaired transport processes, through comparison of uptake rates of organelles from mutant lines with those of wild-type organelles.

Further, a search for homologue proteins in *E. coli*, *S. cerevisiae* or cyanobacteria could reveal closely related genes that facilitate transport of the same compound. Because the deletion of a single gene in these organisms is relatively easy and straightforward, mutant lines for homologous genes can be generated to investigate the impact of the mutation. Growth studies in different media supplemented with various nitrogen or carbon sources can provide information about the transported substrate.

IDENTIFICATION OF CANDIDATE GENES INVOLVED IN PHOTORESPIRATORY TRANSPORT

We employed the above protocol in the search for photorespiratory transporters and generated ten promising candidates according to our defined criteria (Table 2). We focused on these genes, which (i) are co-regulated with several genes of the photorespiratory pathway; and (ii) show more than one predicted transmembrane domain and transporter-like distribution of transmembrane domains, according to the transporter database (Ren *et al.* 2004, 2007). T-DNA lines and amiRNA lines for these candidates are currently under investigation to select for those with a photorespiratory phenotype (*in vivo* filtering).

Table 2. Genes co-expressed with soluble photorespiratory enzymes. Putative candidates after *in silico* analysis are shown. Co-expression analyses were carried out for all identified genes of photorespiration. Most abundant genes were manually filtered for the existence of transmembrane domains and transporter-like distribution of transmembrane domains. For all noted genes, a protein with more than one transmembrane domain is predicted. Description of the locus is according to the ARAMEMNON platform (http://aramemnon.botanik.uni-koeln.de/).

locus	description
At1 g32080	LrgB-like membrane protein, required for chloroplast development (AtLrgB)
At5 g59250	putative VGT subfamily sugar transporter
At3 g61870	putative membrane protein of unknown function
At1 g44920	putative membrane protein of unknown function
At2 g21960	protein of unknown function
At2 g42770	putative peroxisomal PMP22-type protein of unknown function
At4 g35760	VKOR-type thiol oxidoreductase (AtLTO1/AtVKOR-DsbA)
At1 g22850	putative vesicle transport SNARE-associated protein
At1 g49380	putative cytochrome c biogenesis factor
At2 g35260	putative CAAX-type-II prenyl protease

COMPARISON OF ATTED WITH WEIGHTED GENE CORRELATION NETWORK ANALYSIS

The chances of identifying a true positive hit in the search for photorespiratory transporters will clearly increase when focusing on those genes that encode proteins containing predicted TMDs. Indeed, this strategy has proved to be a powerful tool for the identification of new transport proteins involved in photorespiration (see section 'Proof of principle'). To evaluate robustness of the Atted-based approach in possibly identifying additional candidate genes, we have also employed weighted gene correlation network analysis (WGCNA; Langfelder & Horvath 2008). This approach was based on data from a subset of Affymetrix ATH1 arrays (Schmid *et al.* 2005) with *Arabidopsis* wild-type expression data from 64 different experiments. Expression values were normalised (arithmetic mean of three replicates calculated, then log2-transformed and finally median-centred) and WGCNA performed on the R platform, with a calculated soft threshold of seven and a static tree cut height of 0.8 for module detection. Without exception, the photorespiratory key enzymes (Table 1) were assigned to one distinct module, along with all of the TOP10 candidates listed in Table 2, and selected after co-expression analysis. The probability that ten randomly drawn genes (Table 2) would all be members of the WGCNA module that represents photorespiratory key genes is $P = 8.1 \times 10^{-11}$, further strengthening our selection. The size of this module (2241 genes) exceeds the list of results from Atted before application of the filter (1246 genes, excluding all duplicates). Interestingly, 78% of the genes obtained from Atted are present in the WGCNA module containing the photorespiratory key enzymes (Fig. 2). This indicates that both methods are largely consistent with each other. However, the number of putative candidates from both systems requires additional filters to reduce the list of genes. Since the module generated with WGCNA analysis is rather large, further analysis of this module might add additional can-

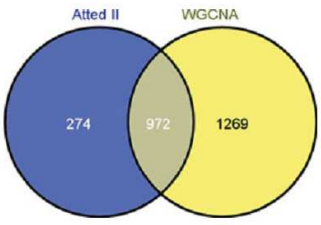


Fig. 2. Comparison between Atted-II and WGCNA results. A total of 972 genes are present in results from both analyses, constituting roughly 80% of all genes obtained from Atted-II.

didates to our list of interest. This work is currently ongoing in our lab.

Both co-expression and WGCNA analysis, performed individually, are not sufficient to yield lists of candidate genes that are short enough to investigate *in vivo*. Nevertheless, they are reasonable first steps to provide a first suggestion of putative candidates. In combination with the right setup of filters, such *in silico* experiments provide a powerful but still simple tool for the initial phase of our studies.

PROOF OF PRINCIPLE

We have demonstrated that co-expression analysis can actually function as a viable tool for the discovery of candidate genes involved in photorespiratory transport. We were able to identify three new genes coding for photorespiratory transporters using the workflow described above. These genes had remained hidden in previous attempts using forward genetic approaches.

Identification of *AtBOU* as photorespiratory transporter

The A BOUT DE SOUFFLE (BOU) protein was successfully identified as a transporter involved in shuttling intermediates in the photorespiratory C_2 cycle (Eisenhut *et al.* 2013). Since *BOU* was already under investigation, it has been excluded from the list of TOP10 candidates (Table 1). Nevertheless, it reached one of the highest ranks before the filter was applied to co-expression results. Out of 13 possible hits, *BOU* was found ten times in the list of co-expressed genes.

The idea of screening *Arabidopsis bou* mutants for a photorespiratory phenotype originated from initial co-expression results showing that *BOU* mRNA is simultaneously expressed with genes for photorespiratory key enzymes. *BOU* gene expression is apparently particularly coordinated with the expression of genes coding for the GDC multi-enzyme complex, which is central in the photorespiratory glycine-to-serine conversion. *BOU* promoter activity is triggered by light and restricted to leaf tissues and cotyledons, supporting a pivotal role in photosynthesis and photorespiration (Eisenhut *et al.* 2013).

Indeed, *bou* knockout plants suffer in ambient air, having chlorotic leaves and growth retardation, but growing much like the wild type when kept under HC. Despite these morphological alterations, further data support the role of BOU in photorespiration. Gas exchange measurements demonstrated that knockout plants have strongly reduced CO_2 fixation rates, and therefore have elevated CO_2 compensation points in comparison to the wild type under HC conditions, which increase upon a shift to NC conditions. Moreover, photorespiratory metabolism is disturbed in the *bou* mutant, as indicated in metabolic analysis. Most remarkably, the glycine level in the mutant is greatly increased in comparison to that of wild-type plants. Together with the finding that mitochondrial glycine degradation is strongly reduced in *bou*, it is anticipated that the function of BOU will be connected with GDC activity. We discriminate between direct transport of glycine into mitochondria by BOU from shuttling of a GDC cofactor (Eisenhut *et al.* 2013). Studies with spinach mitochondria revealed the ability of glycine to diffuse through the mitochondrial membrane if the glycine concentration exceeds 0.5 mM, which is the case in photosynthetic tissues (Yu *et al.* 1983). In the *bou* mutant, the concentration of glycine is even higher, which makes the cofactor hypothesis more likely (Eisenhut *et al.* 2013). To date, the specific substrate transported *via* BOU has not been identified, but ongoing research is focusing on GDC cofactors.

Identification of *DiT1* as photorespiratory transporter

The plastidial 2-oxoglutarate (2-OG)/malate transporter (*DiT1*, *AtDiT1*, *OMT1*) is an interesting candidate, due to its clustering with the already described *DiT2.1* (Taniguchi *et al.* 2002; Renne *et al.* 2003) in a co-expression analysis approach and its sequence homology with *DiT2.1* (*AtpDCT1*). *dit1* mutant plants were shown to suffer under NC conditions, resulting in retarded development, small leaf size, frequently emerging shoots and a decrease in chlorophyll content; however, the mutant still survives under NC conditions, while *DiT2.1*-deficient plants are only viable under HC conditions (Kinoshita *et al.* 2011). DiT1 supplies the chloroplast with 2-OG, which is utilised by FD-GOGAT in the chloroplast and forms a double transporter system together with DiT2.1, further participating in the export of synthesised glutamate and re-fixation of ammonium ions originating from the photorespiratory cycle (Schneidereit *et al.* 2006; Kinoshita *et al.* 2011).

Identification of *AtPLGG1* as photorespiratory transporter

The gene *PLGG1* (*At1 g32080*) was ranked as a most promising candidate according to our co-expression analysis (Table 1). Analysed *plgg1-1* knockout plants develop chlorotic regions along the leaf lamina when grown under ambient air (NC) conditions. This phenotype can be complemented to almost wild type-like when *plgg1-1* knockout plants are kept under HC. We very recently demonstrated that PLGG1 functions as the long-sought plastidic glycerate/glycolate transporter (Pick *et al.* 2013).

Experiments showed that PLGG1 is the translocator mediating glycolate export and glycerate import over the chloroplast membrane, linking chloroplasts to peroxisomes during photorespiration. Metabolite analysis revealed that glycolate and glycerate accumulate in *plgg1-1* plants in response to a shift from HC to NC. Labelling experiments demonstrated that the export of glycolate is blocked in *plgg1-1* plants because, after an $^{18}O_2$ pulse, the level of labelled glycolate increases rapidly compared to the wild type. In contrast to this observation, a slow accumulation of labelled glycerate in *plgg1-1* plants compared to the wild type was measured, pointing to a block in glycerate import. Further experiments linking glycerate import to O_2 evolution at photosystem II revealed reduced O_2 generation in *plgg1-1* chloroplasts compared to wild-type chloroplasts, pointing again to impaired import of glycerate into PLGG1-deficient chloroplasts. These data mirror and complete previous observations demonstrating that glycolate and glycerate are transported by the same protein without facilitating strict counter-exchange (Howitz & McCarty 1986), as needed in photorespiration.

PERSPECTIVE

In this review we describe a four-step protocol that integrates *in silico* and *in vivo* experiments to identify new transport proteins for the photorespiratory cycle. The effectiveness of this

approach is demonstrated through the identification of BOU, DiT1 and PLGG1 as photorespiratory transporters. Thus, the co-expression screen offers a promising method to shed light on yet unknown transport processes between chloroplasts, peroxisomes and mitochondria. Complete understanding of these processes is essential in future attempts to remodel photorespiratory carbon flux in the effort to decrease photorespiratory carbon loss, which can account for up to 20% of net CO_2 assimilation (Cegelski & Schaefer 2006).

The workflow outlined here for the discovery of candidate genes involved in photorespiration can be used, with minor individual adjustments, to develop hypotheses about gene functions not related to photorespiration (Reumann & Weber 2006). To find missing links in other pathways, the initial query genes used for co-expression analysis must be adjusted, together with filtering criteria in the second step, to match the needs of the particular query. Since co-expression analysis through publicly available online resources is now fast and simple, it provides a valuable first step in the search for currently unknown gene functions and for annotating orphan steps in biological pathways and processes.

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (PROMICS Research Group, WE 2231/8-1).

REFERENCES

Aoki K., Ogata Y., Shibata D. (2007) Approaches for extracting practical information from gene co-expression networks in plant biology. *Plant and Cell Physiology*, **48**, 381–390.

Bartel D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.

Boldt R., Edner C., Kolukisaoglu U., Hagemann M., Weckwerth W., Wienkoop S., Morgenthal K., Bauwe H. (2005) D-GLYCERATE 3-KINASE, the last unknown enzyme in the photorespiratory cycle in Arabidopsis, belongs to a novel kinase family. *The Plant Cell*, **17**, 2413–2420.

Brazma A., Vilo J., (2000) Gene expression data analysis. *FEBS Letters*, **480**, 17–24.

Cegelski L., Schaefer J. (2006) NMR determination of photorespiration in intact leaves using *in vivo* $^{13}CO_2$ labeling. *Journal of Magnetic Resonance*, **178**, 1–10.

Cousins A.B., Pracharoenwattana I., Zhou W., Smith S.M., Badger M.R. (2008) Peroxisomal malate dehydrogenase is not essential for photorespiration in Arabidopsis but its absence causes an increase in the stoichiometry of photorespiratory CO_2 release. *Plant Physiology*, **148**, 786–795.

Eisen M.B., Spellman P.T., Brown P.O., Botstein D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences USA*, **95**, 14863–14868.

Eisenhut M., Pick T., Bordych C., Weber A. (2012) Towards closing the remaining gaps in photorespiration – the essential but unexplored role of transport proteins. *Plant Biology*, **15**, 676–685.

Eisenhut M., Planchais S., Cabassa C., Guivarc'h A., Justin A.M., Taconnat L., Renou J.P., Linka M., Gagneul D., Timm S., Bauwe H., Carol P., Weber A.P. (2013) Arabidopsis A BOUT DE SOUFFLE is a putative mitochondrial transporter involved in photorespiratory metabolism and is required for meristem growth at ambient CO_2 levels. *The Plant Journal*, **73**, 836–849.

Haferkamp I., Linka N. (2012) Functional expression and characterisation of membrane transport proteins. *Plant Biology*, **14**, 675–690.

Heazlewood J.L., Verboom R.E., Tonti-Filippini J., Small I., Millar A.H. (2007) SUBA: the Arabidopsis Subcellular Database. *Nucleic Acids Research*, **35** (Database issue), D213–D218.

Hirai M.Y., Sugiyama K., Sawada Y., Tohge T., Obayashi T., Suzuki A., Araki R., Sakurai N., Suzuki H., Aoki K., Goda H., Nishizawa O.I., Shibata D., Saito K. (2007) Omics-based identification of Arabidopsis Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proceedings of the National Academy of Sciences USA*, **104**, 6478–6483.

Howitz K.T., McCarty R.E. (1986) d-glycerate transport by the pea chloroplast glycolate carrier: studies on [1-C]d-glycerate uptake and d-glycerate dependent O_2 evolution. *Plant Physiology*, **80**, 390–395.

Jamai A., Salome P.A., Schilling S.H., Weber A.P., McClung C.R. (2009) Arabidopsis photorespiratory serine hydroxymethyltransferase activity requires the mitochondrial accumulation of ferredoxin-dependent glutamate synthase. *The Plant Cell*, **21**, 595–606.

Kinoshita H., Nagasaki J., Yoshikawa N., Yamamoto A., Takito S., Kawasaki M., Sugiyama T., Miyake H., Weber A.P., Taniguchi M. (2011) The chloroplastic 2-oxoglutarate/malate transporter has dual function as the malate valve and in carbon/nitrogen metabolism. *The Plant Journal*, **65**, 15–26.

Langfelder P., Horvath S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 559.

Lee H.K., Hsu A.K., Sajdak J., Qin J., Pavlidis P. (2004) Coexpression analysis of human genes across many microarray data sets. *Genome Research*, **14**, 1085–1094.

Nozawa A., Nanamiya H., Miyata T., Linka N., Endo Y., Weber A.P., Tozawa Y. (2007) A cell-free translation and proteoliposome reconstitution system for functional analysis of plant solute transporters. *Plant and Cell Physiology*, **48**, 1815–1820.

Obayashi T., Kinoshita K. (2009) Rank of correlation coefficient as a comparable measure for biological significance of gene coexpression. *DNA Research*, **16**, 249–260.

Obayashi T., Kinoshita K., Nakai K., Shibaoka M., Hayashi S., Saeki M., Shibata D., Saito K., Ohta H. (2007) ATTED-II: a database of co-expressed genes and cis elements for identifying co-regulated gene groups in Arabidopsis. *Nucleic Acids Research*, **35** (Database issue), D863–D869.

Obayashi T., Nishida K., Kasahara K., Kinoshita K. (2011) ATTED-II updates: condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. *Plant and Cell Physiology*, **52**, 213–219.

Persson S., Wei H., Milne J., Page G.P., Somerville C.R. (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proceedings of the National Academy of Sciences USA*, **102**, 8633–8638.

Pick T.R., Bräutigam A., Schulz M.A., Obata T., Fernie A.R., Weber A.P.M. (2013) PLGG1, a plastidic glycolate glycerate transporter, is required for photorespiration and defines a unique class of metabolite transporters. *Proceedings of the National Academy of Sciences USA*, doi: 10.1073/pnas.1215142110.

Ren Q., Kang K.H., Paulsen I.T. (2004) TransportDB: a relational database of cellular membrane transport systems. *Nucleic Acids Research*, **32**(Database issue), D284–D288.

Ren Q.H., Chen K.X., Paulsen I.T. (2007) TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic Acids Research*, **35**, D274–D279.

Renne P., Dressen U., Hebbeker U., Hille D., Flugge U.I., Westhoff P., Weber A.P. (2003) The Arabidopsis mutant *dct* is deficient in the plastidic glutamate/malate translocator DiT2. *The Plant Journal*, **35**, 316–331.

Reumann S., Weber A.P. (2006) Plant peroxisomes respire in the light: some gaps of the photorespiratory C2 cycle have become filled – others remain. *Biochimica et Biophysica Acta*, **1763**, 1496–1510.

Schmid M., Davison T.S., Henz S.R., Pape U.J., Demar M., Vingron M., Scholkopf B., Weigel D., Lohmann J.U. (2005) A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics*, **37**, 501–506.

Schneidereit J., Hausler R.E., Fiene G., Kaiser W.M., Weber A.P. (2006) Antisense repression reveals a crucial role of the plastidic 2-oxoglutarate/malate translocator DiT1 at the interface between carbon and nitrogen metabolism. *The Plant Journal*, **45**, 206–224.

Schwab R., Ossowski S., Riester M., Warthmann N., Weigel D. (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. *The Plant Cell*, **18**, 1121–1133.

Schwacke R., Schneider A., van der Graaff E., Fischer K., Catoni E., Desimone M., Frommer W.B., Flugge U.I., Kunze R. (2003) ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. *Plant Physiology*, **131**, 16–26.

Schwacke R., Fischer K., Ketelsen B., Krupinska K., Krause K. (2007) Comparative survey of plastid and mitochondrial targeting properties of transcription factors in Arabidopsis and rice. *Molecular Genetics and Genomics*, **277**, 631–646.

Somerville C.R. (2001) An early Arabidopsis demonstration. Resolving a few issues concerning photorespiration. *Plant Physiology*, **125**, 20–24.

Somerville C.R., Ogren W.L. (1979) A phosphoglycolate phosphatase-deficient mutant of Arabidopsis. *Nature*, **280**, 833–836.

Somerville C.R., Ogren W.L. (1980a) Inhibition of photosynthesis in Arabidopsis mutants lacking leaf glutamate synthase activity. *Nature*, **286**, 257–259.

Somerville C.R., Ogren W.L. (1980b) Photorespiration mutants of *Arabidopsis thaliana* deficient in serine-glyoxylate aminotransferase activity. *Proceedings of the National Academy of Sciences USA*, **77**, 2684–2687.

Somerville C.R., Ogren W.L. (1981) Photorespiration-deficient mutants of *Arabidopsis thaliana* lacking mitochondrial serine transhydroxymethylase activity. *Plant Physiology*, **67**, 666–671.

Somerville C.R., Ogren W.L. (1982) Mutants of the cruciferous plant *Arabidopsis thaliana* lacking glycine decarboxylase activity. *The Biochemical Journal*, **202**, 373–380.

Stuart J.M., Segal E., Koller D., Kim S.K. (2003) A gene-coexpression network for global discovery of conserved genetic modules. *Science*, **302**, 249–255.

Taniguchi M., Taniguchi Y., Kawasaki M., Takeda S., Kato T., Sato S., Tabata S., Miyake H., Sugiyama T. (2002) Identifying and characterizing plastidic 2-oxoglutarate/malate and dicarboxylate transporters in *Arabidopsis thaliana*. *Plant and Cell Physiology*, **43**, 706–717.

Taylor N.G., Howells R.M., Huttly A.K., Vickers K., Turner S.R. (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. *Proceedings of the National Academy of Sciences USA*, **100**, 1450–1455.

Timm S., Bauwe H. (2012) The variety of photorespiratory phenotypes – employing the current status for future research directions on photorespiration. *Plant Biology*, **15**, 737–747.

Usadel B., Obayashi T., Mutwil M., Giorgi F.M., Bassel G.W., Tanimoto M., Chow A., Steinhauser D., Persson S., Provart N.J. (2009) Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. *Plant, Cell & Environment*, **32**, 1633–1651.

Voll L.M., Jamai A., Renne P., Voll H., McClung C.R., Weber A.P. (2006) The photorespiratory Arabidopsis *shm1* mutant is deficient in SHM1. *Plant Physiology*, **140**, 59–66.

Werdan K., Heldt H.W. (1972) Accumulation of bicarbonate in intact chloroplasts following a pH gradient. *Biochimica et Biophysica Acta*, **283**, 430–441.

Yonekura-Sakakibara K., Tohge T., Niida R., Saito K. (2007) Identification of a flavonol 7-O-rhamnosyltransferase gene determining flavonoid pattern in Arabidopsis by transcriptome coexpression analysis and reverse genetics. *Journal of Biological Chemistry*, **282**, 14932–14941.

Yu C., Claybrook D.L., Huang A.H.C. (1983) Transport of glycine, serine, and proline into spinach leaf mitochondria. *Archives of Biochemistry and Biophysics*, **227**, 180–187.

Authors' contribution to manuscript 1

C.B. wrote the manuscript and performed co-expression analysis and filtering

M.E. participated in drafting the manuscript

T.R.P performed experiments for PLGG1, provided data about PLGG1 and participated in drafting the manuscript's paragraph about PLGG1

C.K. and C.B. performed WGCNA analysis

A.P.M.W participated in drafting the manuscript

MANUSCRIPT 2

Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of *Arabidopsis*

Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of *Arabidopsis*

Katharina vom Dorp,[a] Georg Hölzl,[a] Christian Plohmann,[b] Marion Eisenhut,[b] Marion Abraham,[c] Andreas Weber,[b] Andrew D. Hanson,[d] and Peter Dörmann[a, 1]

[a] Institute of Molecular Physiology and Biotechnology of Plants, University of Bonn, Karlrobert-Kreiten-Straße 13, 53115 Bonn, Germany;

[b] Institute of Plant Biochemistry, University of Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany;

[c] Leibniz Institute for Baltic Sea Research, Seestraße 15, 18119 Rostock, Germany;

[d] Horticultural Sciences Department, University of Florida, Gainesville, Florida 32611, USA.

[1]Address correspondence to**:** doermann@uni-bonn.de

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is): Peter Dörmann (doermann@uni-bonn.de).

This article is a *Plant Cell* Advance Online Publication. The date of its first appearance online is the official date of publication. The article has been edited and the authors have corrected proofs, but minor changes could be made before the final version is published. Posting this version online reduces the time to publication by several weeks.

Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis

Katharina vom Dorp,[a] Georg Hölzl,[a] Christian Plohmann,[b] Marion Eisenhut,[b] Marion Abraham,[c] Andreas P. M. Weber,[b] Andrew D. Hanson,[d] and Peter Dörmann[a,1]

[a] Institute of Molecular Physiology and Biotechnology of Plants, University of Bonn, 53115 Bonn, Germany
[b] Institute of Plant Biochemistry, Cluster of Excellence on Plant Science (CEPLAS) Heinrich-Heine-University, 40225 Düsseldorf, Germany
[c] Leibniz Institute for Baltic Sea Research, 18119 Rostock, Germany
[d] Horticultural Sciences Department, University of Florida, Gainesville, Florida 32611

ORCID IDs: 0000-0003-0970-4672 (A.P.M.W.); 0000-0003-2585-9340 (A.D.H.); 0000-0002-5845-9370 (P.D.)

Phytol from chlorophyll degradation can be phosphorylated to phytyl-phosphate and phytyl-diphosphate, the substrate for tocopherol (vitamin E) synthesis. A candidate for the phytyl-phosphate kinase from *Arabidopsis thaliana* (At1g78620) was identified via a phylogeny-based approach. This gene was designated *VITAMIN E DEFICIENT6* (*VTE6*) because the leaves of the Arabidopsis *vte6* mutants are tocopherol deficient. The *vte6* mutant plants are incapable of photoautotrophic growth. Phytol and phytyl-phosphate accumulate, and the phytyl-diphosphate content is strongly decreased in *vte6* leaves. Phytol feeding and enzyme assays with Arabidopsis and recombinant *Escherichia coli* cells demonstrated that VTE6 has phytyl-P kinase activity. Overexpression of *VTE6* resulted in increased phytyl-diphosphate and tocopherol contents in seeds, indicating that *VTE6* encodes phytyl-phosphate kinase. The severe growth retardation of *vte6* mutants was partially rescued by introducing the phytol kinase mutation *vte5*. Double mutant plants (*vte5 vte6*) are tocopherol deficient and contain more chlorophyll, but reduced amounts of phytol and phytyl-phosphate compared with *vte6* mutants, suggesting that phytol or phytyl-phosphate are detrimental to plant growth. Therefore, VTE6 represents the missing phytyl-phosphate kinase, linking phytol release from chlorophyll with tocopherol synthesis. Moreover, tocopherol synthesis in leaves depends on phytol derived from chlorophyll, not on de novo synthesis of phytyl-diphosphate from geranylgeranyl-diphosphate.

INTRODUCTION

Tocochromanols (vitamin E), a group of prenyl quinol lipids accumulating in cyanobacteria and chloroplasts of plants and green algae, are involved in the protection against oxidative stress and in adaptation to low-temperature conditions (Sattler et al., 2004; Havaux et al., 2005; Maeda et al., 2006). Depending on the side chain, three classes of tocochromanols can be distinguished, with tocopherols carrying a phytyl chain, tocotrienols harboring a geranylgeranyl chain, and plastochromanol-8 (PC-8) with a solanesyl chain. The biosynthesis of tocopherols includes the condensation of homogentisate with phytyl-diphosphate (phytyl-PP) catalyzed by homogentisate phytyl transferase (VTE2) (Collakova and DellaPenna, 2001; Savidge et al., 2002). In some plants, in particular monocotyledons, geranylgeranyl-diphosphate (GG-PP) is used for condensation with homogentisate by an alternative enzyme, homogentisate geranylgeranyl transferase, leading to the synthesis of tocotrienols (Cahoon et al., 2003; Yang et al., 2011). Subsequent methylation reactions and closure of the second ring by tocopherol cyclase (VTE1) result in the production of the four forms of tocopherol (α, β, γ, and δ) differing by the numbers and positions of the methyl groups on the chromanol ring (Dörmann, 2007; Hunter and Cahoon, 2007; Maeda and DellaPenna, 2007). PC-8 is produced from plastoquinol-9 via cyclization by VTE1 (Kleinig and Liedvogel, 1978; Mène-Saffrané et al., 2010; Zbierzak et al., 2010).

While the biosynthesis of the tocopherol head group has been studied in detail, less is known about the origin of the phytol moiety for tocopherol synthesis. Phytyl-PP is believed to largely originate from isoprenoid de novo synthesis via reduction of three double bonds in GG-PP. The corresponding enzyme, geranylgeranyl reductase (GGR), from plants has been characterized (Soll et al., 1983; Keller et al., 1998). Later, an alternative pathway for phytyl-PP production via phosphorylation of free phytol to phytyl-monophosphate (phytyl-P) and phytyl-PP was described (Ischebeck et al., 2006; Valentin et al., 2006). Free phytol can be derived from chlorophyll turnover and breakdown, in particular during senescence or chlorotic stress. Chlorophyll degradation starts with the removal of the magnesium cation, yielding pheophytin (Schelbert et al., 2009). Subsequently, phytol is cleaved from pheophytin by pheophytin pheophorbide hydrolase (PPH). Interestingly, GGR from *Arabidopsis thaliana* is capable of reducing the geranylgeranyl moiety in both GG-PP and the geranylgeranylated form of chlorophyll (Keller et al., 1998). Chlorophyll synthase can utilize phytyl-PP or GG-PP for prenylation of chlorophyllide (Soll et al., 1983). As the pool sizes and turnover rates of the isoprenyl-diphosphates in plants are unknown, the exact routes for chlorophyll and tocopherol synthesis are unclear. Increasing attention has been paid to the link of chlorophyll degradation and tocopherol accumulation. Overexpression of PPH in seeds resulted in a modest increase in tocopherol (Zhang et al., 2014).

[1] Address correspondence to doermann@uni-bonn.de.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Peter Dörmann (doermann@uni-bonn.de).
www.plantcell.org/cgi/doi/10.1105/tpc.15.00395

These results suggested that manipulation of expression of genes involved in chlorophyll breakdown reveals minor effects on seed tocopherol levels and that a PPH-independent pathway for chlorophyll dephytylation might exist in seeds.

The first enzyme of phytol phosphorylation, phytol kinase, was isolated from Arabidopsis (*VTE5*, At5g04490) and *Synechocystis* (slr1652) (Valentin et al., 2006). The Arabidopsis *vte5* mutant contains 65 and 85% of tocopherol in leaves and seeds, respectively, compared with the wild type. However, it remained unclear to what extent phytol phosphorylation or the GG-PP-based de novo pathway contributes to overall phytyl-PP synthesis. Furthermore, the identity of the second enzyme of the phosphorylation pathway, phytyl-P kinase, remained unknown.

In prokaryotes, the genes of a common biosynthetic pathway are often localized in close proximity in the genome, e.g., in operons. Therefore, the sequences of enzymes of one pathway in plants can be identified by analyzing the prokaryotic genome structure, provided that the pathway of interest is present in both plants and bacteria. The phytol kinase is predicted to be present in all plants, algae, and bacteria containing chlorophyll and carrying out oxygenic photosynthesis (Ischebeck et al., 2006; Valentin et al., 2006). In agreement with this prediction, genes encoding phytol kinase were identified in Arabidopsis and *Synechocystis*, and related sequences were found in many other plants and photosynthetic bacteria (Valentin et al., 2006). The SEED database, harboring sequences of a large number of bacterial and plant genomes, provides the means to search for genes in subsystems (groups of genes with a related function or participating in a common pathway) (Overbeek et al., 2005; Seaver et al., 2014). Analysis of the subsystem encompassing functions involved in the metabolism of long-chain isoprenoids, including phytol kinase *VTE5*, undecaprenyl pyrophosphate synthetase (UppS), GGR, and chlorophyll synthase, revealed the presence of genes encoding a protein of unknown function (COG1836, pfam01940) that localize close to other isoprenoid genes in many bacterial genomes. Among Archaea, COG1836 sequences cluster with different genes of isoprenoid synthesis. In some cyanobacteria (*Anabaena* and *Nostoc*), orthologs of COG1836 are located in the genome close to phytol kinase, and in certain bacteria (*Pelodyction*, *Symbiobacterium*, and *Thermoplasma*) they are C-terminally fused to an ortholog of phytol kinase. Sequence orthologs of COG1836 are also present in *Synechocystis* (sll0875) and in Arabidopsis (At1g78620) (Seaver et al., 2014).

Here, we describe the characterization of the Arabidopsis At1g78620 gene and the corresponding insertional mutant plants. Biochemical and physiological analyses reveal that At1g78620 encodes phytyl-P kinase and that the phytol phosphorylation pathway represents the most relevant route of phytyl-PP synthesis for tocopherol production in leaves. Furthermore, we show that the phytol phosphorylation pathway is essential for plant growth and development.

RESULTS

Tocopherol Accumulation in Leaves Depends on Phytol Released from Chlorophyll

Tocopherol accumulates in leaves during stress or senescence (Collakova and DellaPenna, 2003). Phytyl-PP required for tocopherol synthesis can be derived from GG-PP from plastidial isoprenoid de novo synthesis or by phosphorylation of free phytol from chlorophyll hydrolysis. To determine the relative contribution of the phytol phosphorylation pathway for phytyl-PP synthesis destined for tocopherol production in leaves, the tocopherol contents in Arabidopsis wild type and the *pheophorbide a oxygenase1* (*pao1*) mutant were determined during senescence. Chlorophyll degradation in the *pao1* mutant is significantly retarded, due to a mutation in pheophorbide a oxygenase (PAO1) (Pruzinská et al., 2003, 2005) (Figure 1A). While total tocochromanol strongly accumulates during dark-induced senescence of detached leaves in the wild type, it remains at a very low amount in *pao1* leaves (Figure 1B), indicating that all tocopherol synthesis during leaf senescence is based on phytol derived from chlorophyll degradation. In *pao1* mutant seeds, the tocopherol content is reduced by 60% (Figure 1C). This result suggests an important role for phytyl-PP derived from chlorophyll degradation in the seeds, in accordance with previous findings (Valentin et al., 2006).

The COG1836 Sequence At1g78620 Shows Similarity with Cytidylyltransferases

At1g78620 from Arabidopsis shows high similarity with COG1836 sequences from the genomes of cyanobacteria (*Anabaena* and *Nostoc*) and other photosynthetic prokaryotes (*Archaeoglobus*) where it clusters in close proximity with genes of isoprenoid metabolism (Seaver et al., 2014). COG1836 amino acid sequences from bacteria and plants can be organized into different groups. The first group includes sequences from Streptophyta (including At1g78620), Chlorophyta, and Cyanobacteria. A second group contains a second Arabidopsis sequence (At5g19930) and COG1836 sequences that form translational fusions with VTE5 proteins from Firmicutes and Chlorobi (Supplemental Figure 1). The two archaeal COG1836 sequences from *Archaeoglobus* and *Thermoplasma* can be placed between the At1g78620- and At5g19930-related COG1836 sequences. The At1g78620 protein sequence is more closely related to *Synechocystis* sll0875 (43.8% identity) than to At5g19930, a gene encoding a protein of unknown function (30.5%). At1g78620 belongs to the DUF92 family of proteins of unknown function harboring several transmembrane α-helices (Seaver et al., 2014). BLAST searches in the GenBank protein database revealed that At1g78620 shows low similarity to cytidylyltransferases, including At3g45040 (24.3%), a putative dolichol kinase (Valentin et al., 2006) (Supplemental Figure 1). At1g78620 is weakly related to phytol kinases from different organisms, including Arabidopsis (At5g04490/VTE5, 17.5% identity). The phytol kinase/VTE5 sequences cluster into a cyanobacterial/plant group containing At5g04490 (VTE5) and a related sequence (At5g58560, VTE5-like) (Fitzpatrick et al., 2011) and the second branch containing the phytol kinase domains that form fusion proteins with COG1836 sequences from Chlorobi and Firmicutes (Supplemental Figure 1) (Valentin et al., 2006).

The At1g78620 Protein Localizes to Chloroplast Envelopes

The At1g78620 sequence harbors an ~70-amino-acid-long N-terminal extension absent from other COG1836 proteins (e.g.,

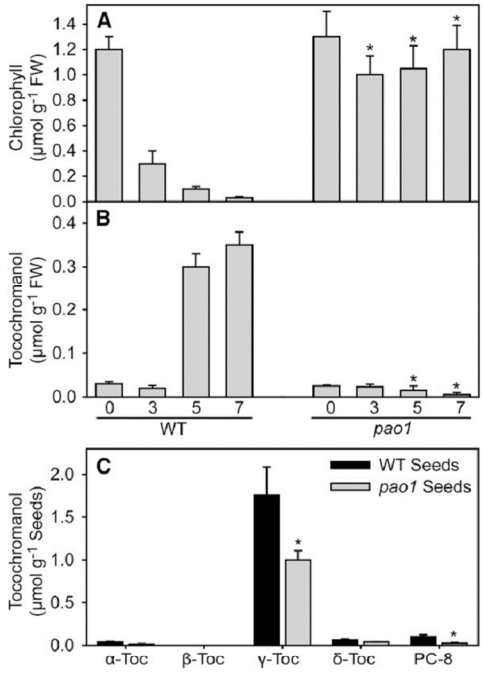


Figure 1. Tocopherol Accumulation in Leaves and Seeds of the Wild Type and *pao1* Mutant.

(A) Chlorophyll contents in leaves of the wild type and *pao1* mutant. Detached leaves were incubated on wet filter paper in the darkness for 3, 5, and 7 d.

(B) Tocochromanol (tocopherol and plastochromanol-8) contents in leaves of the wild type and *pao1*.

(C) Tocochromanol (tocopherol and plastochromanol-8) contents in seeds of the wild type and *pao1*.

Data represent mean and SD of three to four measurements. Values significantly different from the wild type; *P < 0.05; Student's *t* test. Toc, tocopherol.

Synechocystis sll0875) and At5g19930 (Supplemental Figure 1). Sequence analysis with ChloroP1.1 (http://www.cbs.dtu.dk/) (Emanuelsson et al., 1999) revealed the presence of a putative N-terminal transit peptide of 65 amino acids predicted to target the protein to the chloroplast. To test the subcellular localization experimentally, the full-length At1g78620 sequence was C-terminally fused to YFP, and this construct introduced into *Agrobacterium tumefaciens*. After infiltration into *Nicotiana benthamiana* leaves, subcellular localization was observed in protoplasts by confocal laser scanning microscopy. Yellow fluorescence was observed as a ring-like structure surrounding the chlorophyll autofluorescence, indicating that At1g78620 localizes to the chloroplasts envelopes (Supplemental Figure 2). This result is in agreement with the identification of the spinach (*Spinacia oleracea*) ortholog of At1g78620 in isolated chloroplast envelopes by a proteomics approach (Ferro et al., 2002).

Arabidopsis At1g78620 Insertional Mutant Plants Are Incapable of Photoautotrophic Growth

To study the role of At1g78620 in growth and phytol lipid synthesis, two Arabidopsis insertional mutant lines were obtained. The mutant plants (pst15134 and pst00121) carry transposons in the first exon of the At1g78620 gene (Figure 2A). Expression analysis by RT-PCR revealed that the At1g78620 mRNA was not detectable in the two mutant lines, suggesting that they represent null alleles (Figure 2B). The heterozygous plants of the two mutant lines grow similarly to the wild type. Homozygous pst15134 and pst00121 mutant plants are only viable on Murashige and Skoog (MS) medium (Murashige and Skoog, 1962) supplemented with sucrose (Figure 2C). The plants are characterized by a stunted growth and short petioles and show a pale-green leaf color. They have a weakly developed or missing root system. Later, the plants produce a bushy rosette with small leaves but do not bolt and flower and therefore do not produce seeds. The plants die shortly after transfer to soil.

At1g78620 Mutant Plants Are Tocopherol-Deficient

To study the role of At1g78620 in tocopherol metabolism, the amounts of the different forms of tocopherol and of PC-8 were measured by fluorescence HPLC in leaves of homozygous At1g78620 mutant plants grown on MS medium with sucrose for 6 weeks. Arabidopsis wild-type leaves mainly contain α-tocopherol and low amounts of the other tocopherol forms and of PC-8 (Figure 2D). In contrast, tocopherol and PC-8 levels were significantly reduced in the leaves of the two At1g78620 mutant lines, mostly due to the reduction in α-tocopherol by ~98 to 100% (Figures 2D and 4A, respectively). Therefore, the mutation in the gene At1g78620 causes severe tocochromanol deficiency, and the corresponding mutant lines were renamed *vte6-1* (pst15134) and *vte6-2* (pst00121) for *vitamin E deficient6*.

Longevity of Homozygous *vte6* Seeds Is Compromised

Previous studies showed that seed longevity of the tocopherol-deficient mutants *vte1* and *vte2* was strongly compromised (Sattler et al., 2004). This defect was associated with increased nonenzymatic oxidation of storage lipids in the tocopherol-deficient seeds of *vte2*. To study the seed longevity of the *vte6* mutants, ~250 seeds each from heterozygous *vte6-1* and *vte6-2* plants were either freshly harvested or stored for more than 3 months at room temperature. The seeds were germinated on sucrose-supplemented MS medium. The average germination rate was calculated and the genotype of the germinated plants was determined by PCR of genomic DNA. The total germination rate was 96% ± 3%, 76% ± 6%, and 91% ± 3% for freshly harvested seeds from wild-type and heterozygous *vte6-1* and *vte6-2* plants, respectively. It was significantly reduced in segregating *vte6-2* seeds compared with the wild type after storage, when 67% ± 8% (*vte6-2*) and 89% ± 5% (wild type) of seeds germinated. The germination rate in freshly harvested segregating *vte6-1* seeds was significantly lower (76% ± 6%) than that of wild-type seeds, yet was not altered after storage (76% ± 7%).

A segregation of 50% heterozygous plants, 25% wild-type, and 25% homozygous mutant plants would be expected for the seeds

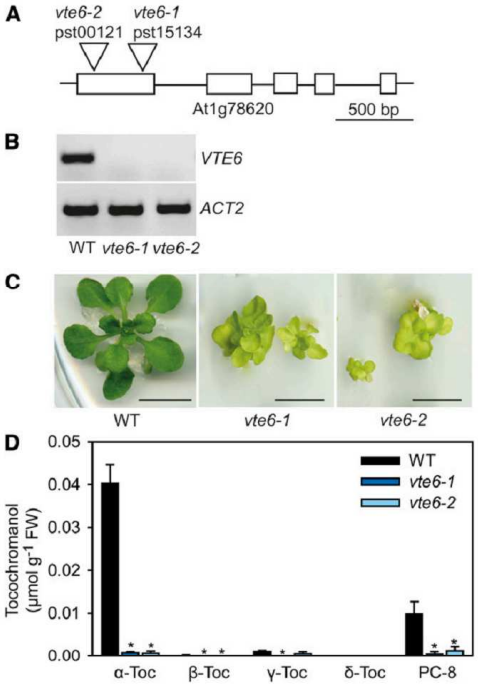


Figure 2. Isolation and Tocochromanol Content of the *vte6* Mutant from Arabidopsis.

(A) Map of the *VTE6* (At1g78620) locus of Arabidopsis with exons (boxes) showing the positions of transposon insertions in the *vte6* mutant alleles. Bar = 500 bp.

(B) Expression analysis of *VTE6* in Arabidopsis *vte6* mutant plants. The RT-PCR products were separated by agarose gel electrophoresis and stained with ethidium bromide. Actin (*ACT2*) was used as control.

(C) Growth of Arabidopsis *vte6-1* and *vte6-2* mutant plants. Plants of the wild type and the two *vte6* mutant alleles were grown on MS medium with sucrose for 6 weeks. The *vte6* mutant plants are homozygous as revealed by PCR of genomic DNA.

(D) Different forms of tocopherol (α-Toc, β-Toc, γ-Toc, and δ-Toc) and PC-8 were measured by fluorescence HPLC in leaves of the wild type, *vte6-1*, and *vte6-2* grown on MS medium with sucrose for 6 weeks. Data show mean and SD of three measurements. Values significantly different from the wild type; *P < 0.05; Student's *t* test.

harvested from heterozygous *vte6-1* or *vte6-2* plants. However, only 11% of homozygous seedlings was retrieved from freshly harvested segregating *vte6-1* or *vte6-2* seeds, indicating that homozygous mutant seeds show significantly reduced viability (Figure 3). When seeds of heterozygous *vte6-1* or *vte6-2* plants were stored for 3 months at room temperature, germination of homozygous seeds was abolished. Only heterozygous or wild-type segregant seeds were able to germinate after storage. Therefore, similar to *vte1* and *vte2*, homozygous *vte6* seeds reveal a strong reduction in seed longevity, indicating that the *VTE6* gene is essential for maintenance of seed viability.

The *vte6* Mutation Affects Chlorophyll Content and Results in the Accumulation of Phytol and Fatty Acid Phytyl Esters

The strong reduction in tocopherol in leaves of the *vte6-1* and *vte6-2* mutant plants (Figures 2D and 4A) was expected to affect the amounts of other phytol-derived lipids, including chlorophyll. In agreement with the yellowish leaf color, the chlorophyll contents in *vte6-1* and *vte6-2* were significantly reduced as compared with the wild type (Figure 4B). The measurement of free phytol revealed a significant increase (~3-fold) in the *vte6-1* and *vte6-2* leaves (Figure 4C). Previous results showed that free phytol can be used for the synthesis of fatty acid phytyl esters in chloroplasts of Arabidopsis. Quantification of fatty acid phytyl esters showed a significant increase in different molecular species, in particular 12:0-phytol, 14:0-phytol, and 16:0-phytol in *vte6-1* and *vte6-2* (Figure 4D).

VTE6 Harbors Phytyl-P Kinase Activity

The in vivo activity of VTE6 was investigated by feeding phytol to seedlings of Arabidopsis wild type and *vte6-1*. Seedlings were incubated in buffer (control) or in buffer containing phytol, then seedlings were harvested and phytyl-P and phytyl-PP quantified by liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Figure 5A). Control plants incubated in buffer contained low amounts of phytyl-P. The amounts of phytyl-PP were high in the wild type, but extremely low in *vte6-1* plants. Phytol feeding resulted in a strong increase of phytyl-P in *vte6-1*, but not in the wild type. In contrast, the phytyl-PP content in the wild type strongly increased (>4-fold) after phytol feeding, but not in *vte6-1*, where phytol feeding had almost no effect on the amount of phytyl-PP. In fact, the phytyl-PP content in *vte6-1* was only 5% compared with the wild type.

Next, total protein was isolated from leaves of Arabidopsis wild-type, *vte6-1*, and *vte6-2* plants. Phytyl-P kinase assays were done employing phytyl-P and CTP as substrates, and phytyl-PP was quantified by LC-MS/MS. The phytyl-P kinase activity with wild-type protein was significantly (3 to 6 times) higher compared with *vte6-1* and *vte6-2* (Figure 5B).

Phytol can be taken up and converted into phytyl-P by *Escherichia coli* cells expressing VTE5 (Valentin et al., 2006); this system was employed to demonstrate the activity of recombinant VTE6. To this end, the two cDNAs for VTE5 and VTE6 were introduced into *E. coli* as single constructs or in combination. After induction, the expression of VTE5 and VTE6 proteins was detected by immunoblot analysis (Supplemental Figure 3). Phytol was added to the *E. coli* cultures, and after 3 h, phytyl-P and phytyl-PP were extracted and quantified by LC-MS/MS. The amount of phytyl-P in VTE6-expressing cells was very low, suggesting that VTE6 might not harbor phytol kinase activity (Figure 5C). Cells expressing both VTE5 and VTE6 produced phytyl-P, albeit less than cells expressing only VTE5. The amounts of phytyl-PP in cells expressing either VTE5 or VTE6 were very low. However, when the two proteins were coexpressed in *E. coli*, large amounts of phytyl-PP accumulated (Figure 5C).

Coupled enzyme assays with protein extracts from *E. coli* expressing either VTE5 or VTE6 or expressing both proteins were performed employing phytol and CTP as substrates. Only minor amounts of phytyl-PP were produced with protein extracts from

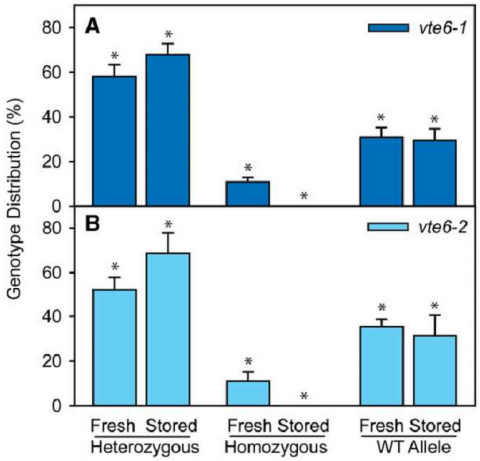


Figure 3. Decreased Germination Capacity of *vte6* Seeds after Storage.

(A) Germination of seeds from a heterozygous *vte6-1* plant.
(B) Germination of seeds from a heterozygous *vte6-2* plant.
The seeds were germinated on MS medium with sucrose. Freshly harvested seeds stored for 3 months at room temperature were used. The genotype of the seedlings was determined by PCR of genomic DNA. The P values (asterisks) calculated from χ^2 tests for the deviations of the genotype distributions of fresh or stored *vte6-1* or *vte6-2* seeds from the expected segregation (50:25:25) were <0.01.

VTE5- or VTE6-expressing cells, indicating the presence of low phytol kinase and phytyl-P kinase activity in *E. coli*. Protein from cells coexpressing VTE5 and VTE6 showed significantly (~3-fold) higher activity (Figure 5D). Taken together, phytol feeding experiments and enzyme assays with Arabidopsis wild-type and *vte6* mutant plants, and with recombinant *E. coli* cells expressing VTE5/VTE6, demonstrate that VTE6 harbors phytyl-P kinase activity.

The Growth Defect of *vte6-1* Is Alleviated in the *vte5-2 vte6-1* Double Mutant

The strong growth retardation of the *vte6-1* and *vte6-2* mutant plants is in contrast with other tocopherol-deficient mutants (*vte1*, *vte2*, *vte4*, and *vte5*) because growth of these mutants under normal conditions is very similar to the wild type (Porfirova et al., 2002; Bergmüller et al., 2003; Maeda et al., 2006; Valentin et al., 2006). As an exception, the *vte3-2* mutant shows reduction of both tocopherol and plastoquinol-9 due to a mutation in the methyltransferase VTE3; therefore, *vte3-2* plants are incapable of photoautotrophic growth (Cheng et al., 2003). The tocopherol-deficient mutants *vte1*, *vte2*, *vte4*, and *vte5* can grow on soil and produce fertile seeds, while *vte6-1* and *vte6-2* cannot grow photoautotrophically and are infertile. Therefore, the strong growth reduction of *vte6* mutant plants cannot be explained by tocopherol deficiency per se, but must be caused by other factors. For example, it is possible that due to the block in phytyl-P phosphorylation, the accumulation of phytyl-P in *vte6* plants might have negative effects on chloroplast development and growth, in contrast to other tocopherol-deficient plants, where such accumulation would not be expected. To investigate this potential scenario, and to suppress the accumulation of phytyl-P in *vte6* plants, the *vte5* mutation (phytol kinase) was introduced into the *vte6-1* background. The Arabidopsis *vte5-1* mutant carrying a premature stop codon in the *VTE5* gene was previously isolated from a chemically mutagenized population (Valentin et al., 2006). An independent mutant allele (pst12490, *vte5-2*) was obtained that carries a transposon insertion in the first exon behind the start codon of the *VTE5* gene (Figure 6A). Analysis of gene expression by RT-PCR revealed the absence of *VTE5* mRNA in *vte5-2* (Figure 6B), indicating that the *vte5-2* mutation represents a null allele. After crossing of homozygous *vte5-2* and heterozygous *vte6-1* plants, double homozygous plants were identified by PCR in the F2 population. When grown on MS medium with sucrose, *vte5-2 vte6-1* plants produce larger and greener leaves in contrast with *vte6-1* plants (Figure 6C). The double mutant plants survive after transfer to soil. Figures 6D and 6E show 6-, 9-, and 12-week-old *vte5-2 vte6-1* plants grown on soil. The morphology of the plants is stunted and bushy with many leaves. Double mutant plants can grow on soil for longer than 4 months. At this stage, wild-type and *vte5-2* plants have died. Therefore, *vte5-2 vte6-1* plants show a stay-green phenotype with strongly delayed senescence and extended lifetime. Old leaves do not turn yellow, but remain green and eventually wither and die (Figure 6E). These results indicate that the strong growth retardation and the incapability to grow on soil of the *vte6-1* single mutant can be partially complemented by introducing the *vte5-2* mutation into the genetic background of *vte6-1*.

Tocopherol, Phytol, and Fatty Acid Phytyl Esters in the *vte5-2 vte6-1* Double Mutant

Wild-type, *vte5-2*, *vte6-1 vte6-2*, and *vte5-2 vte6-1* plants were grown on MS medium with sucrose for 6 weeks, and leaves were harvested for the measurement of tocochromanols, chlorophyll, phytol, and fatty acid phytyl esters. Leaves of *vte5-2* contain ~50% less tocopherol than wild-type leaves under these conditions (Figure 4A). Tocopherol levels in leaves of *vte6-1*, *vte6-2*, and *vte5-2 vte6-1* are below the detection limit. The reduced amount of tocopherol in *vte6-1*, *vte6-2*, and in the double mutant *vte6-1 vte5-2* was accompanied by a significant decrease in PC-8 levels compared with the wild type (Figures 2D and 4A). The chlorophyll content is only slightly decreased in *vte5-2 vte6-1* compared with the wild type and *vte5-2*, but it much is higher than in *vte6-1* and *vte6-2* (Figure 4B). Free phytol accumulates in leaves of *vte6-1* and *vte6-2* to ~3 nmol g^{-1} fresh weight (FW), while the amounts of phytol are unchanged in *vte5-2* and *vte5-2 vte6-1* compared with the wild type (Figure 4C). In addition, the different forms of fatty acid phytyl esters are increased in leaves of *vte5-2*, *vte6-1*, *vte6-2*, and *vte5-2 vte6-1* compared with the wild type, especially 12:0-phytol, 14:0-phytol, and 16:0-phytol (Figure 4D).

Phytyl-P Accumulates in the *vte6-1* Mutant, but Not in the *vte5-2 vte6-1* Double Mutant Plants

The severe growth retardation of *vte6-1* was partially rescued in the *vte5-2 vte6-1* double mutant, suggesting that alterations in phytol, phytyl-P, or phytyl-PP metabolism might be causal for the growth reduction of *vte6-1*. Phytyl-P, phytyl-PP, and GG-PP were

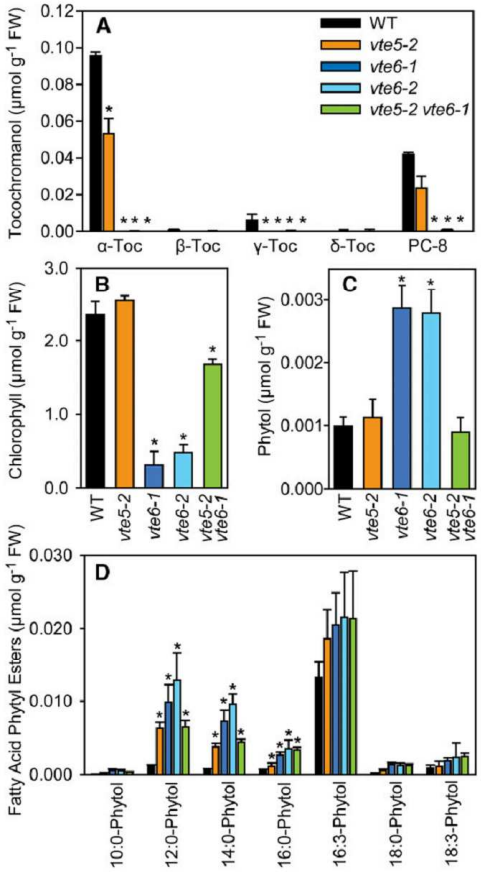


Figure 4. Tocopherol, Chlorophyll, Phytol, and Fatty Acid Phytyl Esters in the Wild Type, *vte5-2*, *vte6-1*, *vte6-2*, and *vte5-2 vte6-1*.

(A) Tocopherol (α-Toc, β-Toc, γ-Toc, and δ-Toc) and PC-8 were measured by fluorescence HPLC.
(B) Chlorophyll was measured photometrically.
(C) The amount of free phytol was quantified by gas chromatography-mass spectrometry.
(D) Fatty acid phytyl esters were measured by direct infusion quadrupole time-of-flight tandem mass spectrometry.
Plants (wild type, *vte5-2*, *vte6-1*, *vte6-2*, and *vte5-2 vte6-1*) were grown on MS medium containing sucrose for 6 weeks. Leaves were harvested and analyzed for their lipid content. Mean and SD of four measurements are shown. Significantly different from the wild type; *P < 0.05. Chlorophyll content in *vte5-2 vte6-1* was also significantly different from *vte6-1* (Figure 4B); Student's *t* test.

quantified by LC-MS/MS in the leaves of plants grown on MS medium with sucrose. Arabidopsis leaves of wild-type plants contain around 15 nmol g^{-1} FW of phytyl-PP. The phytyl-P content in the wild type is extremely low (0.1 nmol g^{-1} FW). The phytyl-PP contents are significantly decreased in *vte5-2*, *vte6-1*, *vte6-2*, and *vte5-2 vte6-1* mutant leaves to 6 to 8 nmol g^{-1} FW (Figure 7). The phytyl-P contents in *vte6-1* and *vte6-2* are significantly increased by ~4- and 10-fold, respectively, compared with the wild type, *vte5-2*, and *vte5-2 vte6-1*. These results show that the increase in the amount of phytyl-P in *vte6-1* is suppressed in the *vte5-2 vte6-1* double mutant, suggesting a negative impact of phytyl-P accumulation on plant growth. The amounts of GG-PP were ~2 nmol g^{-1} FW for all lines, indicating that the content of GG-PP is much lower than that of phytyl-PP and that it is not affected by changes in phytyl-P or phytyl-PP metabolism. Furthermore, these results are consistent with the presumed function of VTE6 as phytyl-P kinase because a block in phytyl-P kinase activity in *vte6-1* and *vte6-2* is expected to cause an accumulation of phytyl-P associated with a decrease in phytyl-PP.

Overexpression of *VTE6* Results in Increased Phytyl-PP and Tocopherol Levels in Seeds

VTE6 was overexpressed under the control of the CaMV 35S promoter in Arabidopsis wild-type plants, and the seeds of these plants were analyzed for isoprenyl-phosphates and tocochromanols. In the VTE6-overexpressing lines VTE6#3 and VTE6#4, phytyl-PP levels were significantly increased (509 and 514 nmol g^{-1} seeds) compared with the empty vector control line (247 nmol g^{-1} seeds) (Figure 8A). Phytyl-P and GG-PP levels were unaffected by overexpression of VTE6 (Figure 8A). Interestingly, the γ-tocopherol content was significantly increased in the VTE6 overexpression lines (1.5 and 1.6 µmol g^{-1} seeds) compared with the empty vector line (1.3 µmol g^{-1} seeds) (Figure 8B). Therefore, overexpression of *VTE6* in seeds results in the accumulation of phytyl-PP, which can be used for prenylation of homogentisate by VTE2, finally causing an increase in tocopherol biosynthesis.

DISCUSSION

Tocopherol (vitamin E) is an important antioxidant in plants and animals and an essential component of the human diet. Work in recent years has established the pathways for tocopherol and tocotrienol synthesis, including the prenylation of homogentisate with a phytyl or geranylgeranyl group, respectively, with subsequent methylation and cyclization reactions (Hunter and Cahoon, 2007; Mène-Saffrané and DellaPenna, 2010). For a long time it was believed that the phytyl moiety for tocopherol synthesis exclusively originates from GG-PP via reduction by GGR. However, in the past years it became clear that tocopherol synthesis also depends on phytyl-PP derived from phytol released during chlorophyll breakdown (Figure 9). Two kinase reactions are required for the conversion of phytol into phytyl-PP (Ischebeck et al., 2006). While the gene for phytol kinase (VTE5) was characterized (Valentin et al., 2006), the identity of the second kinase remained unknown. Here, we present the characterization of a novel protein as candidate for phytyl-P kinase (VTE6). The sequence for *VTE6* was identified based on its colocalization with orthologs of *VTE5* and other genes of isoprenoid lipid metabolism in genomes of prokaryotic photosynthetic organisms, its coexpression pattern with genes of tocopherol metabolism, and its predicted localization in the chloroplast (Seaver et al., 2014). Furthermore, the VTE6 sequence displays weak similarity to cytidylyltransferases and to isoprenoid lipid kinases (Supplemental Figure 1).

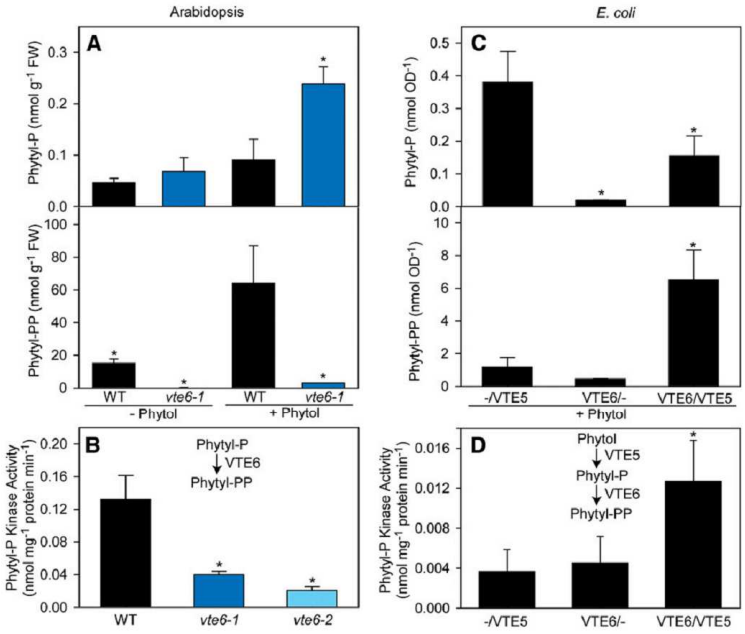


Figure 5. VTE6 Harbors Phytyl-P Kinase Activity.

(A) Phytol feeding experiments with Arabidopsis plants. Wild-type and *vte6-1* plants were incubated in buffer without phytol or with 0.1% (v/v) phytol for 48 h, and isoprenyl-phosphates were isolated.
(B) Phytyl-P kinase assay with protein extracts from Arabidopsis wild-type, *vte6-1*, and *vte6-2* leaves employing phytyl-P and CTP as substrates.
(C) Phytol feeding experiment with recombinant *E. coli* cells. Protein expression in *E. coli* cells harboring plasmids with either VTE5 (-/VTE5), VTE6 (VTE6/-) or both cDNAs (VTE6/VTE5) was induced and the cells incubated in the presence of phytol before extraction of isoprenyl-phosphates.
(D) Coupled phytol kinase/phytyl-P kinase assay with protein extracts from recombinant *E. coli* cells. Total protein was isolated from *E. coli* expressing VTE5 (-/VTE5), VTE6 (VTE6/-), or both proteins (VTE6/VTE5) and used for a coupled phytyl kinase/phytyl-P kinase assay employing phytol and CTP as substrates. Isoprenyl-phosphates (phytyl-P and phytyl-PP) were measured by LC-MS/MS. Data represent mean and SD of three to four measurements. Significantly different from the wild type (**[A]** and **[B]**) or to *E. coli* expressing only VTE5 (**[C]** and **[D]**); *$P < 0.05$; Student's *t* test.

Phytyl-P Kinase Activity

Phytyl-P kinase catalyzes the conversion of phytyl-P into phytyl-PP. Feeding experiments with Arabidopsis wild-type and *vte6-1* seedlings showed that phytol is taken up by the plants and converted into phytyl-P by VTE5. Phytyl-PP levels were strongly increased in wild-type seedlings, while phytyl-PP levels were extremely low in *vte6-1* plants after phytol feeding (Figure 5). On the other hand, phytyl-P, the precursor of phytyl-PP, increased in *vte6-1*, but not in the wild type. Furthermore, enzyme assays with Arabidopsis leaf protein showed that phytyl-P kinase activity in *vte6-1* is ~3- to 4-fold lower compared with the wild type.

Feeding experiments and enzyme assays with recombinant *E. coli* cells confirmed the results obtained with Arabidopsis plants. After phytol feeding, phytyl-P but not phytyl-PP accumulated in VTE5-expressing cells. In VTE5/VTE6 coexpressing cells, the accumulation of phytyl-P was attenuated, while phytyl-PP accumulated, indicating that a certain proportion of phytyl-P was converted into phytyl-PP by VTE6. In cells expressing only VTE6, neither phytyl-P nor phytyl-PP accumulated, indicating that VTE6 does not harbor phytol kinase activity, but is dependent on the presence of phytyl-P, which can only be provided by VTE5 but not by other *E. coli* proteins in this system.

A coupled enzyme assay with total protein extracts from *E. coli* and phytol and CTP as substrates showed that phytyl-PP synthesis depends on the presence of VTE6. Taken together, these experiments conclusively demonstrate that VTE6 converts phytyl-P into phytyl-PP, in agreement with the hypothesis that *VTE6* encodes phytyl-P kinase.

Phytyl-PP for Tocopherol Biosynthesis in Leaves Is Almost Exclusively Provided by the Phytol Phosphorylation Pathway

Two mutant alleles, *vte6-1* and *vte6-2*, were isolated which both showed reduced growth and changes in phytol lipid contents. Furthermore, the *vte6-2* mutant was backcrossed to the wild type (see Methods), and a homozygous F2 plant was obtained that showed the same growth and lipid phenotypes as *vte6-1*. These experiments demonstrate that the strong growth retardation and

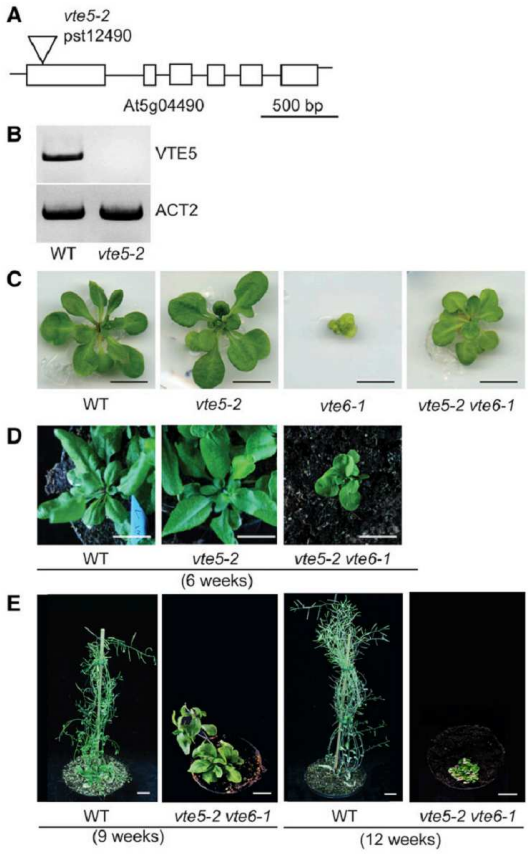


Figure 6. The *vte5-2 vte6-1* Double Mutant.

(A) The *vte5-2* mutant carries a transposon insertion in the first exon of *VTE5* (At5g04490, phytol kinase). Exons are indicated by boxes. Bar = 500 bp.
(B) Expression analysis of the *vte5-2* mutant. The RT-PCR products for *VTE5* and *ACT2* (control) were separated by agarose gel electrophoresis and stained with ethidium bromide.
(C) Growth of wild-type, *vte5-2*, *vte6-1*, and *vte5-2 vte6-1* plants on MS medium with sucrose. Plants are 6 weeks old. Bar = 1 cm.
(D) Growth of wild-type, *vte5-2*, and *vte5-2 vte6-1* plants on soil for 6 weeks. Bar = 2 cm.
(E) Wild-type and *vte5-2 vte6-1* plants grown on soil for 9 or 12 weeks. Bar = 2 cm

the phytol lipid changes are genetically linked with the *vte6* mutations.

Phytyl-phosphates and geranylgeranyl-phosphates are important intermediates for the synthesis of chlorophyll, tocopherol, phylloquinol, and carotenoids in plant leaves, but their measurement was hindered by their low abundance. Employing an LC-MS/MS-based strategy, the amounts of phytyl-P, phytyl-PP, and GG-PP were measured in leaves and seeds of Arabidopsis. Phytyl-P levels in wild-type leaves are extremely low (~0.1 nmol g^{-1} FW), while phytyl-PP and GG-PP amount to 15 or 2 nmol g^{-1} FW, respectively (Figure 7). Phytyl-P accumulates in the *vte6-1* mutant, in accordance with the role of VTE6 as phytyl-P kinase, while the phytyl-PP level is reduced to 50% compared with the wild type. As *vte6-1* represents a null mutation (Figure 2), the residual amount of phytyl-PP presumably is derived from GG-PP reduction by GGR. Interestingly, the reduction of phytyl-PP to 50% of wild-type levels is associated with a severe decrease in tocopherol contents in leaves (by ~98 to 100% compared with the wild type), highlighting the relevance of VTE6 for tocopherol synthesis. Therefore, the residual content of phytyl-PP in *vte6* cannot be used for tocopherol synthesis. Maybe different pools of phytyl-PP exist in the chloroplast, derived from phytol phosphorylation or from GGR, and only the phytyl-PP derived from phytol phosphorylation might be accessible for tocopherol synthesis. The amount of GG-PP was not decreased in leaves of *vte6-1* but was much lower than that of phytyl-PP, indicating that GG-PP reduction to phytyl-PP cannot compensate for the loss of phytyl-P phosphorylation activity to provide phytyl-PP as substrate for tocopherol synthesis. Therefore, GG-PP produced via isoprenoid de novo synthesis plays a negligible role in providing the phytyl-moiety for tocopherol synthesis in leaves.

Phytyl-PP Synthesized by VTE6 Is Used for Tocopherol Biosynthesis in Seeds

Seeds of the Arabidopsis phytol kinase mutant *vte5-1* contain only 20% of wild-type tocopherol levels (Valentin et al., 2006). Therefore, it was considered that the bulk of tocopherol is synthesized from phytyl-PP that originates from the phytol phosphorylation pathway. Intriguingly, overexpression of *VTE6* in Arabidopsis results in a 2-fold increase in phytyl-PP and in increased tocopherol levels in seeds. This finding is in agreement with the hypothesis that *VTE6* encodes phytyl-P kinase and that it provides phytyl-PP for seed tocopherol synthesis. Furthermore, this result suggests that phytyl-PP might be limiting for tocopherol synthesis in seeds.

Longevity in *vte6* Mutant Seeds Is Impaired

Seeds of the *vte2* mutant are completely devoid of tocopherol and exhibit a severe reduction in longevity. Accelerated ageing at high temperature and high humidity causes a strong defect in germination rate of *vte2* seeds, which was associated with non-enzymatic peroxidation of seed lipids, caused by the lack of antioxidant capacity (Sattler et al., 2004). However, apart from a delay in seedling development and rare cotyledon defects, *vte2* plants develop normally, grow to full size, and produce seeds. The *vte5-1* mutant contains 20 and 35% of tocopherol in the seeds and leaves, respectively, compared with the wild type and also exhibits normal plant development and growth (Valentin et al., 2006). Similarly, germination of *vte5-2* seeds was not affected. In contrast, *vte6-1* and *vte6-2* plants can only grow on sucrose-supplemented MS medium, are pale-green and dwarfed, and do not produce seeds. Homozygous *vte6* seeds do not germinate after storage for more than 3 months. These plants are infertile, and homozygous *vte6* mutant seeds cannot be selected from seeds of a heterozygous plant because the seeds cannot be distinguished. Therefore, it was not possible to analyze homozygous *vte6* seeds for lipid contents. Attempts to measure tocopherol, phytol, or isoprenyl-phosphates in

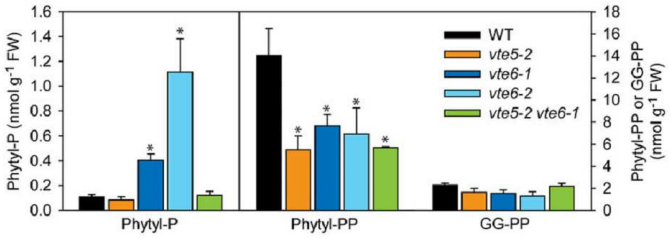


Figure 7. Quantification of Isoprenyl-Phosphates in Wild-Type, *vte5-2*, *vte6-1*, *vte6-2*, and *vte5-2 vte6-1* Plants.

Plants were grown on MS medium with sucrose and leaves were harvested after 6 weeks. Isoprenyl-phosphates were measured by LC-MS/MS. Data show mean and SD of four to five measurements. Significantly different from the wild type; *P < 0.05; Student's *t* test.

single, segregating seeds of a heterozygous *vte6* plant were not successful. However, the fact that longevity is affected in *vte6* seeds, similar to *vte2* seeds, suggests that *vte6* seeds also suffer from a severe reduction in tocopherol content.

The Deficiency in Growth and Development of *vte6* Plants Is Unrelated to the Degree of Membrane Lipid Unsaturation

Tocopherol plays a crucial role during the adaptation of growth under low-temperature conditions (Maeda et al., 2006). Tocopherol-deficient mutants including *vte2* show increased sugar and carbohydrate contents in the leaves, accompanied by callose deposition at plasmodesmata in phloem parenchyma transfer cells, when grown at low temperature. Biochemical and genetic evidence showed that this effect is unrelated to oxidative stress, but it is associated with reduced desaturation of linoleic acid (18:2) to α-linolenic acid (18:3), particularly in phosphatidylcholine (PC). Therefore, when grown under normal or low temperature, PC in *vte2* contains a reduced amount of 18:3 compared with the wild type. The mechanism of how tocopherol deficiency affects lipid desaturation remains unclear. It was possible that the severe growth retardation of the *vte6* mutant was related to the low temperature phenotype of *vte2* plants. Measurements of membrane lipids revealed that the proportion of monogalactosyldiacylglycerol is reduced in *vte6-1* and *vte6-2*, while the relative amount of PC is increased (Supplemental Figure 4). This result indicates a general decrease in the amounts of thylakoid lipids, albeit much less severe as the decrease in chlorophyll. The molecular species of 36:6 (18:3/18:3) and 36:5 (18:3/18:2) are the predominant PC molecules carrying two C18 fatty acids. The *vte6-1* and *vte6-2* plants contain higher amounts of 36:6 PC than the wild type, while the amounts of 36:5 PC remain similar. Therefore, *vte6* plants contain a higher degree of unsaturated C18 fatty acids in PC than the wild type, in contrast to *vte2*, which contains lower amounts of 36:6 compared with the wild type (Maeda et al., 2006). This result indicates that the severe growth retardation of *vte6* is unrelated to the low-temperature phenotype of the *vte2* mutant.

Phytyl Lipids in the *vte5* and *vte6* Mutants

The *vte5-1* mutant of Arabidopsis was described previously (Valentin et al., 2006). We used a second mutant allele (*vte5-2*), which grows like the wild type and shows normal chlorophyll contents. The amount of tocopherol is reduced, as was shown for *vte5-1*, indicating that the insertion in *vte5-2* is causal for the tocopherol deficiency, in analogy to *vte5-1*. The amount of tocopherol is strongly reduced in leaves of *vte6* (<5%), while *vte5* leaves contain considerable more tocopherol (*vte5-1*, 35% of the wild type, Valentin et al., 2006; *vte5-2*, 50% of the wild type, Figures 4 and 6). As *vte5-2* represents a null allele, a second phytol kinase activity might exist in the chloroplasts, which could explain the residual capacity for the conversion of phytol into phytyl-P and the absence of phytol accumulation in *vte5-2*. On the other hand, VTE6 presumably is the only phytyl-P kinase in chloroplasts. Therefore, the complete block of the phytol phosphorylation pathway in *vte6* might lead to the accumulation of phytyl-P and at the same time of phytol because phytol phosphorylation by VTE5 (and additional phytol kinase activities) might be downregulated, possibly by a feedback mechanism, to avoid the uncontrolled accumulation of phytyl-P.

The Severe Impairment of Growth and Development in *vte6* Mutants Is Alleviated in *vte5 vte6* Double Mutants

To test the hypothesis whether the block in sequestration of phytol to phytyl-P and subsequently to phytyl-PP is related to the strong growth retardation of the *vte6-1* mutant, a genetic test was conducted by generating a double mutant of *vte5-2* and *vte6-1*. Interestingly, growth of homozygous *vte5-2 vte6-1* plants is strongly improved compared with *vte6-1* single mutant plants. Similar to *vte6-1*, *vte5-2 vte6-1* double mutant plants show strongly reduced amounts of tocopherol (Figure 4A).

The amount of phytyl-P is increased in *vte6-1* leaves (0.4 nmol g^{-1} FW) but is reduced in *vte5-2 vte6-1* to levels similar to the wild type and *vte5-2* (around 0.1 nmol g^{-1} FW) (Figure 7). The amounts of free phytol follow a similar pattern, because phytol is unaffected in *vte5-2* and *vte5-2 vte6-1*, while it is strongly accumulated in *vte6-1* compared with the wild type (Figure 4C). Therefore, the introduction of the *vte5-2* mutation into the *vte6-1* mutant background strongly improves growth, associated with the suppression of the accumulation of phytol and phytyl-P. This pattern of changes is in accordance with the expected result of combining mutations in phytol kinase and phytyl-P kinase activities in a single plant.

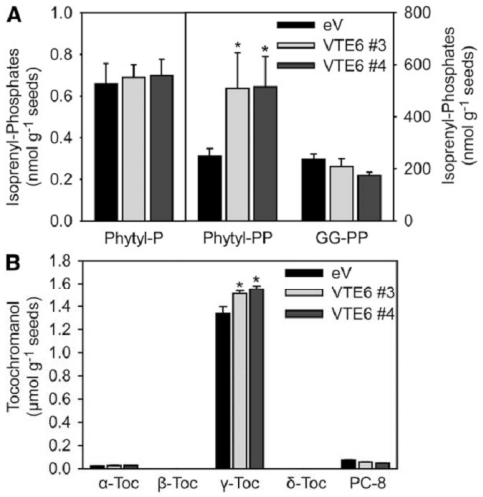


Figure 8. Quantification of Isoprenyl-Phosphates and Tocopherol in *VTE6* Overexpression Lines.

Seeds of Arabidopsis wild-type plants, empty vector (eV) control, and two lines overexpressing *VTE6* (VTE6#3 and VTE6#4) were harvested from plants simultaneously grown on soil. Data show mean and SD of three to four measurements. Significantly different from the wild type; *P < 0.05; Student's *t* test.

(A) Isoprenyl-phosphates were isolated from 20 mg of seeds and quantified by LC-MS/MS.

(B) Tocopherol (α-Toc, β-Toc, γ-Toc, and δ-Toc) and PC-8 were measured by fluorescence HPLC.

The fact that phytyl-PP in *vte6-1* and *vte5-2 vte6-1* is reduced compared with the wild type, but that chlorophyll content in *vte5-2 vte6-1* is strongly increased, indicates that the reduction in phytyl-PP content per se does not become limiting for chlorophyll synthesis in *vte5-2 vte6-1*, as well as in *vte6-1*. It is possible that phytol or phytyl-P levels, which are increased in *vte6-1*, affect chlorophyll synthesis activity, e.g., by competing with phytyl-PP for binding to chlorophyll synthase or by interfering with other important chloroplast pathways.

The relative amounts of galactolipids in *vte6* are reduced to only a minor extent (monogalactosyldiacylglycerol drops from 42 to 33 to 36%, while digalactosyldiacylglycerol and sulfoquinovosyldiacylglycerol are not changed; Supplemental Figure 4), indicating that the overall effect on chloroplast membrane lipids is much less severe than the decrease in chlorophyll content. Furthermore, not all chloroplast lipids are decreased. Some lipids increased in *vte6*, in particular phytol, phytyl-P, and fatty acid phytyl esters. Therefore, the block in *vte6* preferentially affects phytol-containing lipids. On the other hand, the amounts of PC-8 in *vte6-1* and *vte6-2* were reduced. PC-8 synthesis is not directly linked with phytyl-PP metabolism. It is possible that the block in phytol phosphorylation might affect tocochromanol synthesis (head group and/or prenyl side chain production) by feedback inhibition (see above), which would explain the general decrease in the levels of all tocochromanols in *vte6*. In line with this scenario, PC-8 also decreased in the *vte5-2* mutant, which has normal chlorophyll levels (Figure 4), and in the seeds of the *pao1* mutant (Figure 1).

The *vte5-2 vte6-1* plants contain significantly higher chlorophyll contents than *vte6-1* (Figure 4B). Therefore, the strong reduction in *vte6* growth is primarily not caused by tocopherol deficiency or reduction in chlorophyll synthesis capacity, but might rather be associated with the accumulation of free phytol or phytyl-P. It has previously been shown that the accumulation of free alcohols is detrimental to membrane integrity (Löbbecke and Cevc, 1995). Therefore, the accumulation of phytol might destabilize chloroplast membranes affecting chlorophyll and galactolipid synthesis in *vte6-1*. In addition, it is possible that phytyl-P is toxic per se or that the sequestration of phytol into the phytyl-P pool is harmful for

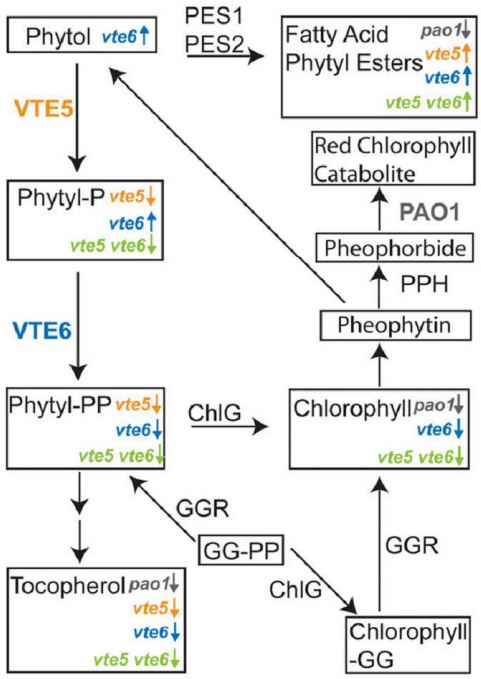


Figure 9. Phytol Metabolism in Arabidopsis.

Phytol is derived from chlorophyll degradation via PPH or by an unknown dephytylase. PAO1 is involved in further pheophorbide catabolism. Phosphorylation by VTE5 and VTE6 results in the production of phytyl-PP, which serves as precursor for tocopherol synthesis. Phytol is also used for the synthesis of fatty acid phytyl esters by phytyl ester synthases (PES1 and PES2) (Lippold et al., 2012). GG-PP or phytyl-PP can be used to synthesize geranylgeranyl-chlorophyll (chlorophyll-GG) or chlorophyll, respectively, by chlorophyll synthase (ChlG). Chlorophyll-GG is reduced to chlorophyll by GGR. The contribution of de novo synthesis via direct reduction of GG-PP to phytyl-PP by GGR is low. Alterations in metabolite or lipid contents in the respective mutants are indicated by arrows (upregulation or downregulation).

plant metabolism. In this regard, it is interesting to note that phytyl-P represents an alcohol phosphate analogous to sphingobase-1-phosphates (e.g., sphingosine-1-phosphate), which are known signaling molecules in plants and animals (Lynch et al., 2009). It is also possible that phytyl-P-derived compounds accumulate in *vte6* with detrimental consequences for plant growth.

Synechocystis harbors a protein sequence (sll0875) highly related to Arabidopsis VTE6 (Supplemental Figure 1). A gene knockout strategy was employed to mutagenize sll0875 in *Synechocystis* by homologous recombination. However, screening of transformed *Synechocystis* cells after repeated rounds of selection on antibiotic medium revealed the absence of fully segregated homozygous mutant cells, indicating that this gene is essential. Therefore, the phosphorylation pathway of phytol is crucial for Arabidopsis and *Synechocystis*, and this essential function is independent of tocopherol deficiency.

METHODS

Plant Materials and Growth Conditions

The *Arabidopsis thaliana* transposon insertion lines *vte6-1* (pst15134, 13-0604-1) and *vte6-2* (pst00121, 54-0912-1) were obtained from the RIKEN Bioresource Center (Tsukuba, Japan). Seeds of heterozygous *vte6-1* and *vte6-2* plants were germinated on MS medium containing 2% (w/v) sucrose (Murashige and Skoog, 1962). Homozygous seedlings were selected by PCR using the primers bn233 and bn234 (genomic *VTE6* locus, 1255 bp); bn232 and bn234 (*vte6-1*, 970 bp); and bn130 and bn234 (*vte6-2*, 1230 bp). The phenotype of the line *vte6-2* (pst00121) was complicated by the fact that heterozygous plants were severely stunted, originating from a second site mutation. After backcrossing this plant with the RIKEN transposon donor line Ds13 (N8521), a *vte6-2* plant devoid of the secondary mutation was obtained.

The *vte5-2* (At5g04490; pst12490, 11-6074-1) transposon mutant allelic to *vte5-1* deficient in phytol kinase (Valentin et al., 2006) was obtained from the RIKEN Bioresource Center. Homozygous *vte5-2* plants were isolated by PCR screening (genomic locus, bn771, bn772, 1352 bp; insertion, bn771, bn232, 1080 bp). Heterozygous *vte6-1* and homozygous *vte5-2* plants were crossed, and plants heterozygous for *vte6-1* and homozygous for *vte5-2* were selected in the F2 progeny, and double homozygous *vte5-2 vte6-1* plants in the F3, by PCR screening.

The *pao1* mutant (At3g44880, SALK_111333) was obtained from Stefan Hörtensteiner, University of Zürich, Switzerland (Pruzinská et al., 2005). Plants were grown at 150 µmol m^{-2} s^{-1} light at 16 h light per day, 20°C, and 55% relative humidity.

RT-PCR

RNA was extracted from leaves, and cDNA synthesis and RT-PCR were performed as described previously (Lippold et al., 2012). Oligonucleotides used for RT-PCR of *VTE6*, *VTE5*, and *ACT2* are described in Supplemental Table 1.

Generation of Overexpression Lines

To generate transgenic Arabidopsis lines overexpressing *VTE6*, the full-length coding sequence of At1g78620/*VTE6* was amplified from *VTE6* cDNA (pda00492, RIKEN) by PCR adding restriction sites using the oligonucleotides bn327 (*Xma*JI) and bn328 (*Sal*I) (Supplemental Table 1). The PCR products were ligated into the cloning vector pGEM-T Easy (Promega). The plasmid was digested with *Xma*JI and *Sal*I (*VTE6*) and the insert ligated into the *Spe*I and *Sal*I sites of the binary vector pLH9000 (DNA Cloning Service), yielding pLH9000-VTE6 for expression in plants under the control of the CaMV 35S promoter. The plasmid was transferred into *Agrobacterium tumefaciens* (GV3101-pMP90) by electroporation. Arabidopsis wild-type Col-0 plants were transformed by floral dipping (Clough and Bent, 1998). Transgenic seeds were selected by fluorescence microscopy based on the expression of the DsRed marker gene. Overexpression of *VTE6* was confirmed by RT-PCR.

Measurement of Chlorophyll, Tocopherol, Phytol, and Fatty Acid Phytyl Esters

Chlorophyll was quantified photometrically (Porra et al., 1989). For the extraction of tocopherol, fatty acid phytyl esters, and free phytol, 50 mg of leaf tissue was frozen in liquid nitrogen and ground to a fine powder (Precellys homogenizer). The ground tissue was resuspended in 500 µL diethylether and phase separation achieved by centrifugation after adding 250 µL of 0.2 M H_3PO_4/1 M KCl. The organic phase was harvested and dried under air flow. Tocopherol was measured by fluorescence HPLC on a diol column (Zbierzak et al., 2010). Free phytol was quantified by gas chromatography-mass spectrometry after derivatization with *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide using oleyl alcohol (Sigma-Aldrich) as internal standard (Ischebeck et al., 2006).

Fatty acid phytyl esters were quantified using 17:0-phytol, synthesized from heptadecanoic acid (17:0) and phytol (Ischebeck et al., 2006). Fatty acid phytyl esters were detected by scanning for the neutral loss of the mass of 278.2974 ($C_{20}H_{38}$, [phytol-$H_2$0]) by direct infusion-tandem mass spectrometry experiments (Agilent 6530 Accurate Mass Q-TOF). The samples were infused via an HPLC-ChipCube nanospray ion source in chloroform/methanol/300 mM ammonium acetate (300:665:35, v/v/v) at 1 µL min^{-1} (Welti et al., 2002). The Q-TOF mass spectrometer was operated in positive ion mode with a fragmentor voltage of 270 V. Membrane glycerolipids were isolated and quantified as described (Gasulla et al., 2013).

Quantification of Isoprenyl-Phosphates by Liquid Chromatography-Mass Spectrometry

Isoprenyl-phosphates were extracted from leaves (50 mg) using isopropanol/50 mM KH_2PO_4/acetic acid (1:1:0.025, v/v/v) (Larson and Graham, 2001). Internal standards (0.1 nmol each of decanyl-phosphate and decanyl-diphosphate) were synthesized and purified (Joo et al., 1973) and added to the lipid extracts. Isoprenyl-phosphates were separated on a Nucleodur C8 column (50 mm × 4.6 mm, 1.8 µm; Macherey and Nagel), using a gradient of acetonitril and 5 mM aqueous ammonium acetate. The concentration of acetonitril was increased from 20 to 100% in 25 min. Mass spectrometry was performed on an Agilent 6530 Accurate Mass Q-TOF instrument with JetStream ESI source. Isoprenyl-phosphates were analyzed in negative ion mode after collision-induced dissociation, by precursor ion scanning for the phosphate group (HPO_3^-, *m/z* 78.9591). The parental ions of the isoprenyl-phosphates $[M-H]^-$ were phytyl-P (*m/z* 375.2664), phytyl-PP (*m/z* 455.2328), and GG-PP (*m/z* 449.1858) as previously described (Valentin et al., 2006).

Expression of VTE6 and VTE5 Proteins in *Escherichia coli*

The coding sequences of *VTE6* and *VTE5* were amplified from Arabidopsis cDNA, and specific restriction sites were added using the primer pairs CB1/CB2 and CB3/CB4 (Supplemental Table 1), respectively. The fragments were subcloned into pJet2.1 (ThermoFisher Scientific) and verified by sequencing. After digestion with *Bam*HI/*Not*I (*VTE6* fragment) or *Nde*I/*Fse*I (*VTE5* fragment), the fragments were ligated into the multiple cloning sites 1 or 2 of the expression vector pETDuet-1 (Novagen, Merck Millipore), which is designed for coexpression of two genes. Thus, *VTE6* was N-terminally fused to a 6x His-tag sequence, while *VTE5* was C-terminally fused to an S-tag. In addition to the pETDuet-1 plasmid containing both the *VTE6* and *VTE5* cDNAs, plasmids carrying either *VTE6* or *VTE5* were also generated. The plasmids were transformed into *E. coli* Rosetta (DE3) cells, and

ampicillin-resistant transformants were selected. Cells were grown in Luria-Bertani medium at 37°C and 200 rpm until an OD_{600} of 0.7 was reached. Protein expression was induced with 1 mM isopropylthio-β-galactoside. The fusion proteins were detected by immunodetection after blotting using horseradish peroxidase-conjugated antibodies specific for His-tag (Miltenyi Biotec) and S-tag (Merck Millipore) for VTE6 and VTE5, respectively.

In Vivo Phytol Feeding

Arabidopsis seedlings (the wild-type, *vte6-1*, and *vte6-2*) were grown on MS medium with 2% (w/v) sucrose for 6 weeks and then transferred to 20 mL of 20 mM MES-KOH, pH 6.5, containing 0.1% (v/v) phytol (Sigma-Aldrich) in an Erlenmeyer flask. The flasks were incubated under continuous light (100 μmol photons m^{-2} s^{-1}) for 48 h while shaking before extraction of isoprenyl-phosphates.

Feeding of *E. coli* cells with synthetic phytol was done as described by Valentin et al. (2006) with some modifications. Briefly, protein expression was induced using 1 mM isopropylthio-β-galactoside and the cells grown overnight at 16°C. Cells were harvested, washed, and transferred to 5 mL of Luria-Bertani medium containing carbenicillin (50 μg/mL), 5 mM phytol, and 0.2% toluene. The cells were incubated for 3 h at 30°C and harvested by centrifugation and the cell pellet used for isoprenyl-phosphate extraction.

In Vitro Phytol Kinase and Phytyl-P Kinase Assays

Proteins were isolated from Arabidopsis leaves and *E. coli* cell pellets as previously described (Ischebeck et al., 2006). For the phytyl-P kinase assay, 10 nmol of phytyl-P (Isoprenoids) in ethanol were transferred to a 1.5-mL microfuge tube and dried under air flow. Afterwards, 300 μg crude protein extract was mixed with 10 nmol CTP and 8 μL assay buffer (20 mM $MgCl_2$, 50 mM Na-orthovanadate, and 0.25% CHAPS) to a final volume of 100 μL. For the coupled phytol kinase/phytyl-P kinase assay, 1 nmol of phytol instead of phytyl-P was used in the reaction. The reaction was incubated for 30 min at 30°C and terminated by freezing in liquid N_2. Isoprenyl-phosphates were isolated and analyzed using LC-MS/MS as described above.

Subcellular Localization Studies of YFP-Tagged VTE6 Protein

For subcellular localization studies, the coding sequence (without stop codon) of *VTE6* was PCR amplified from Arabidopsis cDNA with gene-specific primers (CB157/CB158; Supplemental Table 1) carrying Gateway recombination sites and recombined into the pUBC-YFP binary vector (Grefen et al., 2010) for C-terminal fusion of *VTE6* to *YFP*, driven by the Ubiquitin10 promoter. The fusion construct *YFP:VTE6* was verified by sequencing. Agrobacterium strain GV3101 was transformed with the plasmid and infiltrated into *Nicotiana benthamiana* leaves for transient expression of YFP:VTE6 fusion proteins. After 48 h, protoplasts were isolated for localization studies by a confocal laser scanning microscope (Zeiss LSM 510) as described in detail earlier (Breuers et al., 2012).

Phylogenetic and Statistical Analysis

Amino acid sequences of VTE5- and VTE6-related proteins from Arabidopsis, *Synechocystis*, *Synechococcus*, *Anabaena*, *Nostoc*, *Chlorobium*, *Pelodyction*, *Symbiobacterium thermophilum*, *Thermoplasma acidophilum*, and *Archaeoglobus fulgidus* were obtained from GenBank. Sequence alignments were done using ClustalW, and phylogenetic trees were constructed based on the neighbor-joining method with the bootstrap test with MEGA 6 (Tamura et al., 2013). A text file of the alignment is available as Supplemental Data Set 1.

Statistical analyses were performed using Microsoft Excel 2010. Values are significantly different according to Student's *t* test ($P < 0.05$). The significance of differences in genotypes between the segregating population and the expected distribution (Figure 3) was calculated according to the χ^2 test ($P < 0.01$). At least three biological replicates were used for the biochemical analyses.

Accession Numbers

Sequence data from this article can be found in the EMBL/GenBank data libraries under Arabidopsis locus identifiers At5g04490 and At1g78620, for *VTE5* and *VTE6*, respectively.

Supplemental Data

Supplemental Figure 1. Phylogenetic Relationship of COG1836-Like and Phytol Kinase Sequences.

Supplemental Figure 2. Subcellular Localization of VTE6 to the Chloroplast Envelope Membranes.

Supplemental Figure 3. Expression of VTE5 and VTE6 in *Escherichia coli*.

Supplemental Figure 4. Membrane Glycerolipids in Leaves of the *vte6* Mutant.

Supplemental Table 1. Oligonucleotides Used in This Study.

Supplemental Data Set 1. Text File of the Alignment Corresponding to the Phylogenetic Analysis in Supplemental Figure 1.

ACKNOWLEDGMENTS

We thank Stefan Hörtensteiner (University of Zürich, Switzerland) for seeds of the *pao1* mutant of Arabidopsis. We acknowledge the help of Brigitte Dresen-Scholz, Anne Bünder, and Omnia El Said (University of Bonn) during plant growth, genotyping, and cloning. This research was in part supported by the Deutsche Forschungsgemeinschaft (SFB 645, Project Z4 to P.D., WE 2231/8-2 to A.P.M.W. and M.E., and EXC 1028 to A.P.M.W.).

AUTHOR CONTRIBUTIONS

P.D., G.H. M.E., A.P.M.W., and A.D.H. designed the experiments. K.v.D., G.H., C.P., M.E., and M.A. performed the research. A.D.H. contributed new analytic and computational tools. All authors analyzed data. P.D., K.v.D., M.E., A.P.M.W., and A.D.H. wrote the article.

Received May 5, 2015; revised September 8, 2015; accepted September 17, 2015; published October 9, 2015.

REFERENCES

Bergmüller, E., Porfirova, S., and Dörmann, P. (2003). Characterization of an *Arabidopsis* mutant deficient in γ-tocopherol methyltransferase. Plant Mol. Biol. **52:** 1181–1190.

Breuers, F.K., Bräutigam, A., Geimer, S., Welzel, U.Y., Stefano, G., Renna, L., Brandizzi, F., and Weber, A.P. (2012). Dynamic remodeling of the plastid envelope membranes - A tool for chloroplast envelope in vivo localizations. Front. Plant Sci. **3:** 7.

Cahoon, E.B., Hall, S.E., Ripp, K.G., Ganzke, T.S., Hitz, W.D., and Coughlan, S.J. (2003). Metabolic redesign of vitamin E biosynthesis in plants for tocotrienol production and increased antioxidant content. Nat. Biotechnol. **21:** 1082–1087.

Cheng, Z., Sattler, S., Maeda, H., Sakuragi, Y., Bryant, D.A., and DellaPenna, D. (2003). Highly divergent methyltransferases catalyze a conserved reaction in tocopherol and plastoquinone synthesis in cyanobacteria and photosynthetic eukaryotes. Plant Cell **15:** 2343–2356.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana.* Plant J. **16:** 735–743.

Collakova, E., and DellaPenna, D. (2001). Isolation and functional analysis of homogentisate phytyltransferase from *Synechocystis* sp. PCC 6803 and Arabidopsis. Plant Physiol. **127:** 1113–1124.

Collakova, E., and DellaPenna, D. (2003). The role of homogentisate phytyltransferase and other tocopherol pathway enzymes in the regulation of tocopherol synthesis during abiotic stress. Plant Physiol. **133:** 930–940.

Dörmann, P. (2007). Functional diversity of tocochromanols in plants. Planta **225:** 269–276.

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. **8:** 978–984.

Ferro, M., Salvi, D., Riviere-Rolland, H., Vermat, T., Seigneurin-Berny, D., Grunwald, D., Garin, J., Joyard, J., and Rolland, N. (2002). Integral membrane proteins of the chloroplast envelope: identification and subcellular localization of new transporters. Proc. Natl. Acad. Sci. USA **99:** 11487–11492.

Fitzpatrick, A.H., Bhandari, J., and Crowell, D.N. (2011). Farnesol kinase is involved in farnesol metabolism, ABA signaling and flower development in Arabidopsis. Plant J. **66:** 1078–1088.

Gasulla, F., Vom Dorp, K., Dombrink, I., Zähringer, U., Gisch, N., Dörmann, P., and Bartels, D. (2013). The role of lipid metabolism in the acquisition of desiccation tolerance in *Craterostigma plantagineum*: a comparative approach. Plant J. **75:** 726–741.

Grefen, C., Donald, N., Hashimoto, K., Kudla, J., Schumacher, K., and Blatt, M.R. (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:** 355–365.

Havaux, M., Eymery, F., Porfirova, S., Rey, P., and Dörmann, P. (2005). Vitamin E protects against photoinhibition and photooxidative stress in *Arabidopsis thaliana.* Plant Cell **17:** 3451–3469.

Hunter, S.C., and Cahoon, E.B. (2007). Enhancing vitamin E in oilseeds: unraveling tocopherol and tocotrienol biosynthesis. Lipids **42:** 97–108.

Ischebeck, T., Zbierzak, A.M., Kanwischer, M., and Dörmann, P. (2006). A salvage pathway for phytol metabolism in *Arabidopsis.* J. Biol. Chem. **281:** 2470–2477.

Joo, C.N., Park, C.E., Kramer, J.K., and Kates, M. (1973). Synthesis and acid hydrolysis of monophosphate and pyrophosphate esters of phytanol and phytol. Can. J. Biochem. **51:** 1527–1536.

Keller, Y., Bouvier, F., d'Harlingue, A., and Camara, B. (1998). Metabolic compartmentation of plastid prenyllipid biosynthesis–evidence for the involvement of a multifunctional geranylgeranyl reductase. Eur. J. Biochem. **251:** 413–417.

Kleinig, H., and Liedvogel, B. (1978). Fatty acid synthesis by isolated chromoplasts from the daffodil. [^{14}C]Acetate incorporation and distribution of labelled acids. Eur. J. Biochem. **83:** 499–505.

Larson, T.R., and Graham, I.A. (2001). A novel technique for the sensitive quantifcation of acyl CoA esters from plant tissues. Plant J. **25:** 115–125.

Löbbecke, L., and Cevc, G. (1995). Effects of short-chain alcohols on the phase behavior and interdigitation of phosphatidylcholine bilayer membranes. Biochim. Biophys. Acta **1237:** 59–69.

Lippold, F., vom Dorp, K., Abraham, M., Hölzl, G., Wewer, V., Yilmaz, J.L., Lager, I., Montandon, C., Besagni, C., Kessler, F., Stymne, S., and Dörmann, P. (2012). Fatty acid phytyl ester synthesis in chloroplasts of *Arabidopsis.* Plant Cell **24:** 2001–2014.

Lynch, D.V., Chen, M., and Cahoon, E.B. (2009). Lipid signaling in Arabidopsis: no sphingosine? No problem! Trends Plant Sci. **14:** 463–466.

Maeda, H., and DellaPenna, D. (2007). Tocopherol functions in photosynthetic organisms. Curr. Opin. Plant Biol. **10:** 260–265.

Maeda, H., Song, W., Sage, T.L., and DellaPenna, D. (2006). Tocopherols play a crucial role in low-temperature adaptation and phloem loading in Arabidopsis. Plant Cell **18:** 2710–2732.

Mène-Saffrané, L., and DellaPenna, D. (2010). Biosynthesis, regulation and functions of tocochromanols in plants. Plant Physiol. Biochem. **48:** 301–309.

Mène-Saffrané, L., Jones, A.D., and DellaPenna, D. (2010). Plastochromanol-8 and tocopherols are essential lipid-soluble antioxidants during seed dessication and quiescence in Arabidopsis. Proc. Natl. Acad. Sci. USA **107:** 17815–17820.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol. Plant. **15:** 473–497.

Overbeek, R., et al. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res. **33:** 5691–5702.

Porfirova, S., Bergmüller, E., Tropf, S., Lemke, R., and Dörmann, P. (2002). Isolation of an *Arabidopsis* mutant lacking vitamin E and identification of a cyclase essential for all tocopherol biosynthesis. Proc. Natl. Acad. Sci. USA **99:** 12495–12500.

Porra, R.J., Thompson, W.A., and Kriedemann, P.E. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: Verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochim. Biophys. Acta **975:** 384–394.

Pruzinská, A., Tanner, G., Anders, I., Roca, M., and Hörtensteiner, S. (2003). Chlorophyll breakdown: pheophorbide a oxygenase is a Rieske-type iron-sulfur protein, encoded by the *accelerated cell death 1* gene. Proc. Natl. Acad. Sci. USA **100:** 15259–15264.

Pruzinská, A., Tanner, G., Aubry, S., Anders, I., Moser, S., Müller, T., Ongania, K.-H., Kräutler, B., Youn, J.-Y., Liljegren, S.J., and Hörtensteiner, S. (2005). Chlorophyll breakdown in senescent Arabidopsis leaves. Characterization of chlorophyll catabolites and of chlorophyll catabolic enzymes involved in the degreening reaction. Plant Physiol. **139:** 52–63.

Sattler, S.E., Gilliland, L.U., Magallanes-Lundback, M., Pollard, M., and DellaPenna, D. (2004). Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. Plant Cell **16:** 1419–1432.

Savidge, B., Weiss, J.D., Wong, Y.-H.H., Lassner, M.W., Mitsky, T.A., Shewmaker, C.K., Post-Beittenmiller, D., and Valentin, H.E. (2002). Isolation and characterization of homogentisate phytyltransferase genes from *Synechocystis* sp. PCC 6803 and *Arabidopsis.* Plant Physiol. **129:** 321–332.

Schelbert, S., Aubry, S., Burla, B., Agne, B., Kessler, F., Krupinska, K., and Hörtensteiner, S. (2009). Pheophytin pheophorbide hydrolase (pheophytinase) is involved in chlorophyll breakdown during leaf senescence in *Arabidopsis.* Plant Cell **21:** 767–785.

Seaver, S.M., et al. (2014). High-throughput comparison, functional annotation, and metabolic modeling of plant genomes using the PlantSEED resource. Proc. Natl. Acad. Sci. USA **111:** 9645–9650.

Soll, J., Schultz, G., Rüdiger, W., and Benz, J. (1983). Hydrogenation of geranylgeraniol : two pathways exist in spinach chloroplasts. Plant Physiol. **71:** 849–854.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. **30:** 2725–2729.

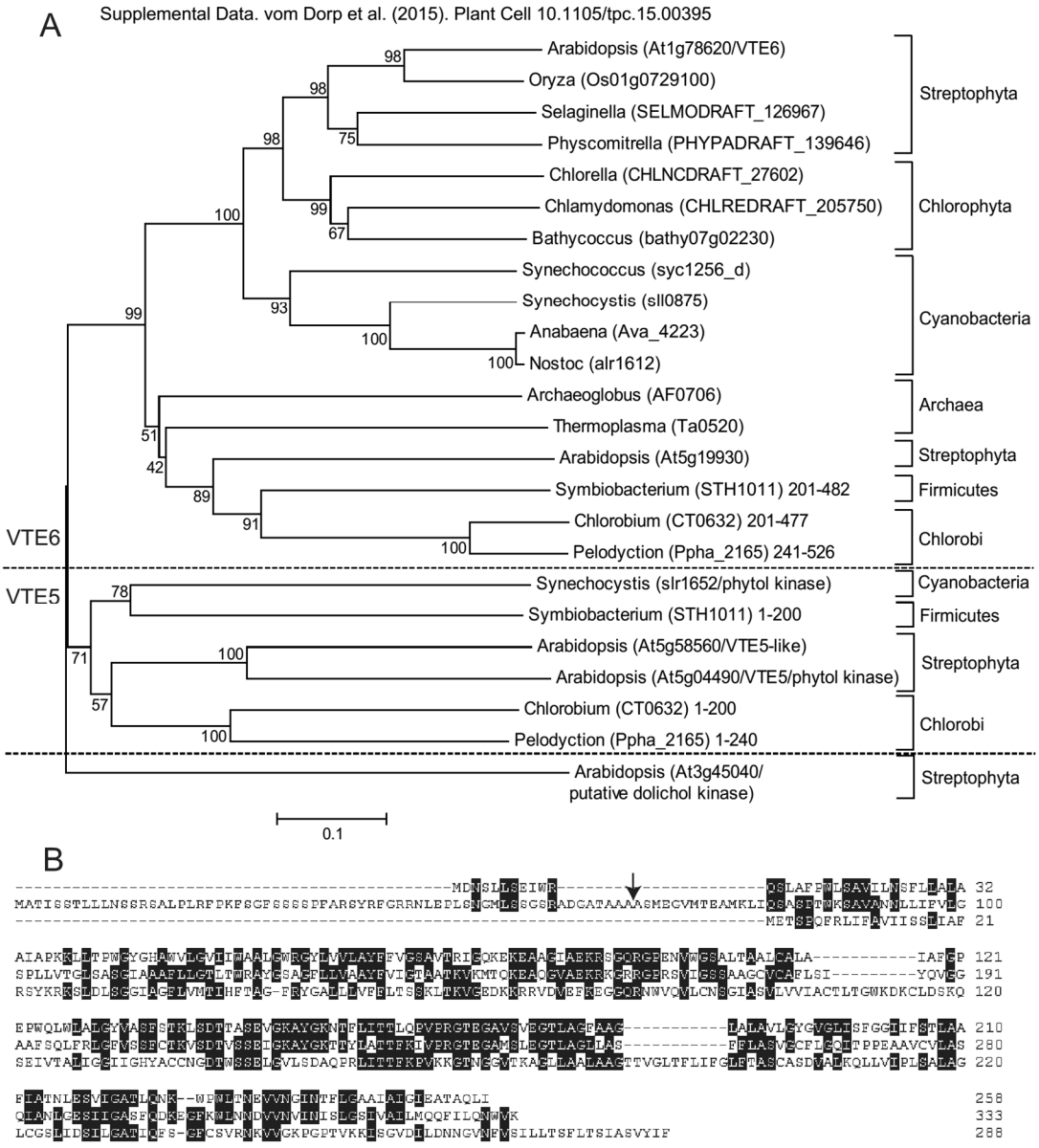
Valentin, H.E., Lincoln, K., Moshiri, F., Jensen, P.K., Qi, Q., Venkatesh, T.V., Karunanandaa, B., Baszis, S.R., Norris, S.R., Savidge, B., Gruys, K.J., and Last, R.L. (2006). The Arabidopsis *vitamin E pathway gene5-1* mutant reveals a critical role for phytol kinase in seed tocopherol biosynthesis. Plant Cell **18:** 212–224.

Welti, R., Li, W., Li, M., Sang, Y., Biesiada, H., Zhou, H.-E., Rajashekar, C.B., Williams, T.D., and Wang, X. (2002). Profiling membrane lipids in plant stress responses. Role of phospholipase D alpha in freezing-induced lipid changes in *Arabidopsis*. J. Biol. Chem. **277:** 31994–32002.

Yang, W., Cahoon, R.E., Hunter, S.C., Zhang, C., Han, J., Borgschulte, T., and Cahoon, E.B. (2011). Vitamin E biosynthesis: functional characterization of the monocot homogentisate geranylgeranyl transferase. Plant J. **65:** 206–217.

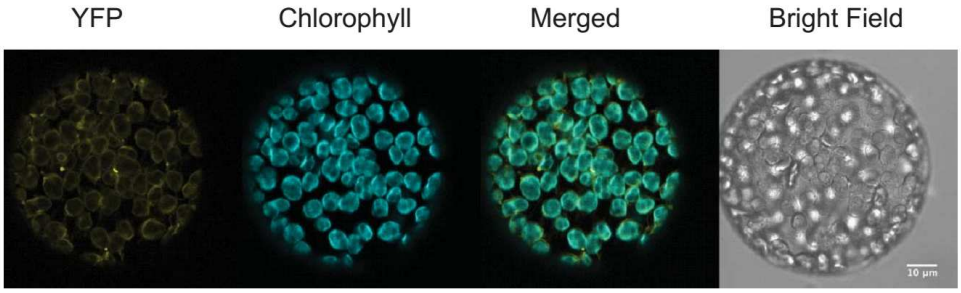
Zbierzak, A.M., Kanwischer, M., Wille, C., Vidi, P.A., Giavalisco, P., Lohmann, A., Briesen, I., Porfirova, S., Bréhélin, C., Kessler, F., and Dörmann, P. (2010). Intersection of the tocopherol and plastoquinol metabolic pathways at the plastoglobule. Biochem. J. **425:** 389–399.

Zhang, W., Liu, T., Ren, G., Hörtensteiner, S., Zhou, Y., Cahoon, E.B., and Zhang, C. (2014). Chlorophyll degradation: the tocopherol biosynthesis-related phytol hydrolase in Arabidopsis seeds is still missing. Plant Physiol. **166:** 70–79.



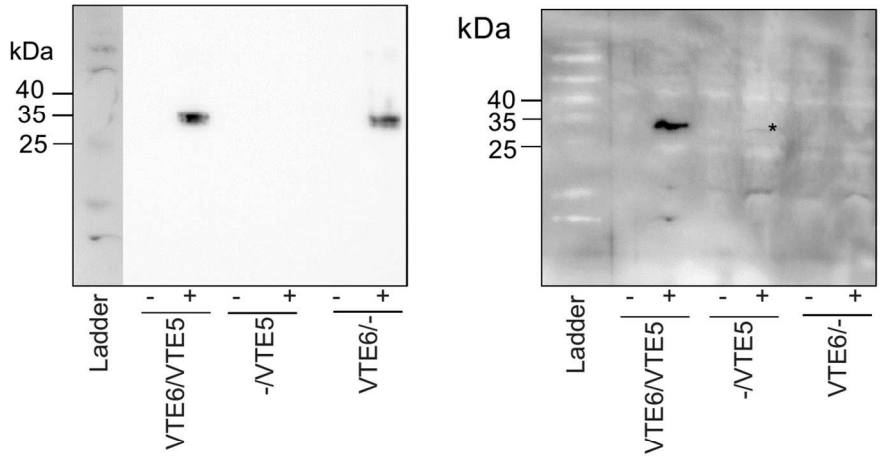
Supplemental Figure 1. Phylogenetic Relationship of COG1836-Like Sequences.
(**A**) Phylogenetic analysis of phytol kinase (VTE5) and phytyl-P kinase (VTE6) sequences from plants and bacteria using MEGA6 (Tamura et al., 2013). Protein sequences from *Arabidopsis*, *Oryza sativa*, *Physcomitrella patens*, *Selaginella moellendorfii*, *Chlorella*, *Chlamydomonas reinhardtii*, *Bathycoccus*, *Synechococcus*, *Synechocystis*, *Anabaena*, *Nostoc*, *Archeoglobus*, *Thermoplasma*, *Symbiobacterium*, *Chlorobium* and *Pelodyction* were obtained from GenBank. The phylogenetic analysis was done using the neighbor-joining-method. In the x-dimension, branch length represents evolutionary distance (number of amino acid differences per site). Bootstrap values were calculated from 1000 replicates.
(**B**) Alignment of COG1836-like sequences from *Arabidopsis* and *Synechocystis*.
The *Arabidopsis* VTE6 sequence (At1g78620, middle) was aligned with the COG1836 sequence from *Synechocystis* (sll0875, top) and At5g19930 (bottom) using ClustalW (MEGA6). The arrow indicates the predicted transit peptide cleavage site (position 65/66; ChloroP1.1).

Supplemental Data. vom Dorp et al. (2015). Plant Cell 10.1105/tpc.15.00395



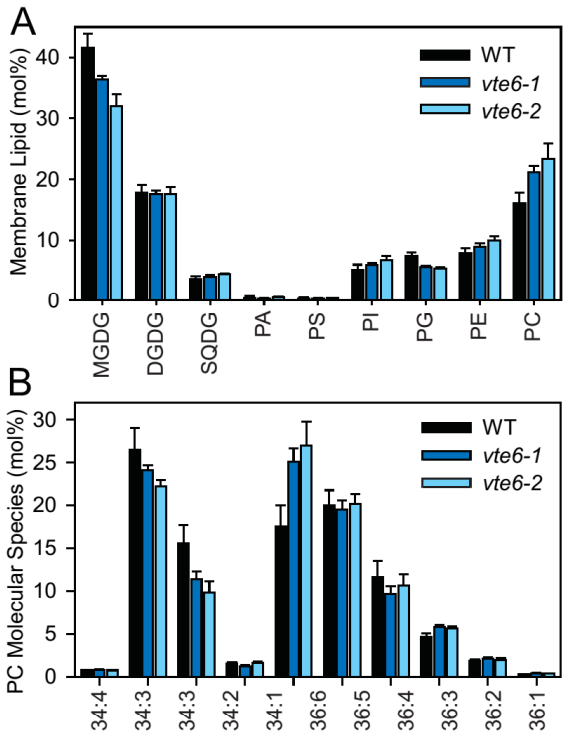
Supplemental Figure 2. Subcellular localization of VTE6.
The fusion protein of YFP:VTE6 (VTE6 C-terminally fused to yellow fluorescent protein) was transiently expressed in *N. benthamiana* leaves and fluorescence in isolated protoplasts observed by confocal microscopy.

Supplemental Data. vom Dorp et al. (2015). Plant Cell 10.1105/tpc.15.00395



Supplemental Figure 3. Heterologous expression of VTE5 and VTE6 in *E. coli*.
After ligation into pETDuet-1, His:VTE6 and/or VTE5:S fusion proteins were expressed in *E. coli* Rosetta for 4 h. Total protein extracts were isolated from cells before (-) and after (+) induction with IPTG, and expression of fusion proteins was demonstrated by immunodetection using antibodies against His- and S-tag, respectively. The asterisk indicates the weak signal of VTE5:S. Molecular masses of His:VTE6 and VTE5:S are 35 kDa.

Supplemental Data. vom Dorp et al. (2015). Plant Cell 10.1105/tpc.15.00395



Supplemental Figure 4. Membrane Glycerolipids in Leaves of the *vte6* Mutant.
(**A**) Glycerolipids are presented in mol%.
(**B**) Molecular species composition of phosphatidylcholine (PC) in mol%.
Lipids were measured in leaves extracts of WT, *vte6-1* and *vte6-2* by direct infusion mass spectrometry (Gasulla et al., 2013). Data are mean and SD of 4 measurements. *Significantly different to WT; $p < 0.05$; Student's t test. Molecular species are indicated as X:Y, where X represents the number of the carbon atoms, and Y the number of double bonds, in the tow acyl chains. DGDG, digalactosyldiacylglycerol; MGDG, monogalactosyldiacylglycerol; PA, phosphatidic acid; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PG, phosphatidyl-glycerol; PI, phosphatidylinositol; PS, phosphatidylserine; SQDG, sulfoquinovosyldiacylglycerol.

Supplemental Data. vom Dorp et al. (2015). Plant Cell 10.1105/tpc.15.00395

Supplemental Table 1: Oligonucleotides Used in this Study		
Primer	Sequence	Gene target
bn130	TCCGTTCCGTTTTCGTTTTTTAC	Transposon primer Ds5-2a G-edge
bn232	CCGGATCGTATCGGTTTTCG	Transposon primer Ds3-2a H-edge
bn233	ATGGCAACGATTTCGTCAACTC	VTE6 forward (*vte6-1*, *vte6-2*)
bn234	GTCCAGCCAATGTTCCTTCAAG	VTE6 reverse (vte6-1, vte6-2)
bn771	ACTGTACCTCCGTCAATCGCC	VTE5 forward (*vte5-2*)
bn772	AAGCTTAAGACAAGCGCGTATG	VTE5 reverse(*vte5-2*)
bn78	ATTTTGCCGATTTCGGAAC	Left border primer SALK
bn358	GCCATCCAAGCTGTTCTCTC	Forward RT-PCR primer *ACT2*
bn359	GAACCACCGATCCAGACACT	Reverse RT-PCR primer *ACT2*
bn577	GCAACGAAGGTTAAAATGACGC	Forward RT-PCR primer *VTE6*
bn578	GAAGCACCTATTATGCTCTCTC	Reverse RT-PCR primer *VTE6*
bn410	CAAACTCAGTTCCCTCCGTC	Forward RT-PCR primer *VTE5*
bn411	GTCCGTTAATAACAAGCCTTAAG	Reverse RT-PCR primer *VTE5*
bn327	CCTAGGATGGCAACGATTTCGTCAACTC	Forward VTE6 overexpression
bn328	GTCGACTTACTTGACCCAGTTCTGGAG	Reverse VTE6 overexpression
CB157	AAAAAGCAGGCTTCGAAGGAGATAGAACCATGGCAACGATTTCGTCAACTC	YFP::VTE6 expression, forward
CB158	AGAAAGCTGGGTGCTTGACCCAGTTCTGGAGTATA	YFP::VTE6 expression, reverse
CB1	AAGGATCCGATGGCAACGATTTCGTCAACTC	VTE6 expression in *E. coli*, forward
CB2	AAGCGGCCGCTTACTTGACCCAGTTCTGGAGT	VTE6 expression in *E. coli*, reverse
CB3	AACATATGATGGCAGCAACCTTACCTCTAT	VTE5 expression in *E. coli*, forward
CB4	AAGGCCGGCCCGATATCCGAAACTTAAATAAGCAGCTA	VTE5 expression in *E. coli*, reverse

Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis

Katharina vom Dorp, Georg Hölzl, Christian Plohmann, Marion Eisenhut, Marion Abraham, Andreas P.M. Weber, Andrew D. Hanson and Peter Dörmann

Plant Cell; originally published online October 9, 2015;
DOI 10.1105/tpc.15.00395

Authors' contribution to manuscript 2

C.P. performed experiments for subcellular localization, cloned constructs for the co-expression of VTE5 and VTE6 in *Escherichia coli*, and performed heterologous expression of VTE5 and VTE6 in *E. coli*

M.E. participated in drafting the manuscript, supervised experimental design for co-expression of VTE5 and VTE6 in *Escherichia coli* and performed localization studies

A.P.M.W participated in drafting the manuscript and helped with experimental design

(Draft)

MANUSCRIPT 3

amiR-vte6 Analyses Add Further Insights into the Function of VTE6 in *Arabidopsis*

amiR-vte6 Analyses Add Further Insights into the Function of VTE6 in *Arabidopsis*

Christian Plohmann, Marion Eisenhut, Peter Jahns & Andreas P.M. Weber*

Institute of Plant Biochemistry, Center of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-University, Universitätsstraße 1, D-40225 Düsseldorf

Running Head: Further Characterization of *amiR-vte6*

*Corresponding author:

Andreas P.M. Weber

Institute of Plant Biochemistry, Center of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-University, Universitätsstraße 1, D-40225 Düsseldorf

Abstract

Chlorophyll degradation by chlorophyllase, especially present during senescence, yields a large amount of phytol. Phytol can either be reutilized for chlorophyll synthesis or incorporated into tocopherol. Both synthesis pathways require a double phosphorylation of phytol to yield phytyl diphosphate, the intermediate compound which constitutes a branching point between tocopherol and chlorophyll synthesis. These phosphorylation steps are catalyzed by the action of VTE5 (phytol kinase) and VTE6 (phytyl phosphate kinase). In-depth characterization of the latter enzyme, and the corresponding gene is of importance, because it holds an exceptional position within the tocopherol synthesis pathway. *vte6-1* and *vte6-2* mutants are described to reach a lethal state during early developmental stages. The lack of tocopherols in *vte6* mutants cannot be responsible on its own for the lethality, as absence of tocopherol in other *vte* mutants does not cause such severe phenotypes. Lethality of existing *vte6* mutants renders scientists unable to study VTE6 function in all developmental stages to decipher the potential outstanding function of VTE6 in depth.

Here, early data is presented, which is based on an artificial micro RNA line for VTE6 (*amiR-vte6*), which offers the possibility to investigate VTE6 function in developmental stages which cannot be screened in existing *vte6* mutants. The phenotype of *amiR-vte6* as observed so far, was consistent with published observations about *vte6-1* and *vte6-2*. On top of supporting existing data, the investigation of *amiR-vte6* unveiled, that the pale-green phenotype which is common for *vte6-1*, *vte6-2* and *amiR-vte6* is pronounced in true leaves of young *amiR-vte6* plants in comparison to cotyledons.

The results presented here shows that *amiR-vte6* adds an important component to the toolbox for future studies about the essential function of VTE6 in *Arabidopsis.*

Introduction

More than 25 years ago α-tocopherol was found to accumulate during leaf senescence and ripening of green plant organs (Rise et al., 1989). Phytol released from chlorophyll during senescence by chlorophyllase activity (Willstätter and Stoll, 1913) was already discussed during this time to be incorporated into tocopherols (Soll et al., 1980). While tocopherols and their function have been well studied in animals (Azzi et al., 2003) and in plants (Maeda and DellaPenna, 2007; Munné-Bosch and Alegre, 2002) since then, the interconversion of released phytol into tocopherols was not deciphered completely until recently. For a long time, geranylgeranyl diphosphate (GGDP), originating from the MEP pathway (non-mevalonate pathway) was thought to be the major source of phytyl diphosphate for tocopherol synthesis. Phytol on the other hand was not largely taken into account to be a basis for phytyl diphosphate synthesis.

In 2006 it was shown by feeding experiments with phytol that *Arabidopsis* seedlings incorporate phytol into tocopherols (Ischebeck et al., 2006). In the same year Valentin and co-workers, successfully identified VTE5 (Vitamin E) in *Arabidopsis*, the phytol kinase which phosphorylates phytol originating from chlorophyll degradation to yield phytyl phosphate (Valentin et al., 2006). The analysis shows that chlorophyll derived phytol serves as an important precursor for tocopherol synthesis. Hence, the assumption made by Rise and co-workers in 1989 was shown to be right. However, phytyl phosphate has to be phosphorylated a second time to yield phytyl diphosphate which is required for tocopherol synthesis. This irreplaceable step was under debate for another ten years, because the specific phosphorylation reaction was not assigned to a specific protein or further characterized. Despite the possibility of a two-step phosphorylation by one enzyme (VTE5), another model including two independent kinases (VTE5 + X) was discussed. Using radioactive [γ-32P]CTP, Ischebeck and co-workers could show in *Arabidopsis* that the majority of radioactive labeled phosphate is incorporated into phytyl phosphate and only in parts into phytyl diphosphate, which strongly argued towards the model which consists of two individual kinases (Ischebeck et al., 2006). Moreover, VTE5 was shown to cluster in prokaryotes together with a second gene with high similarity to VTE5 itself, which added further support to the hypothesis that two independent kinases are occupied during the generation of phytyl diphosphate from phytol (Seaver et al., 2014). However, the second kinase (phytyl

phosphate kinase) was not described in literature until recently. By the identification and initial characterization of VTE6, the remaining phytyl phosphate kinase was described by vom Dorp and colleagues (vom Dorp et al., 2015).

The characterization of two *Arabidopsis vte6* null mutants disclosed a severe phenotype. *vte6* plants are dependent on sucrose supplementation and cannot be grown on soil. Mutants do not accumulate tocopherols in leaves as expected and lack chlorophyll in comparison to wild-type controls. In addition to that, homozygous *vte6* plants do not reach fertile states. Consequently, the recent study presented by vom Dorp could shed light on VTE6 function during early development in *Arabidopsis*, but data on later stages or the seed state is limited due to lethality of *vte6-1* and *vte6-2*. Seeds could not be selected for homozygosity of the mutated allele prior to germination and development of green organs. Thus, the effect of VTE6 function on the developing seed and the fate of phytol and tocopherol which originate from chlorophyll degradation during senescence during later developmental stages has to be studied in the future.

The generation of artificial microRNA (amiRNA) lines (Schwab et al., 2006) targeting *VTE6* expression in *Arabidopsis* can aid in the characterization of developmental stages which cannot be studied in *vte6-1* and *vte6-2*. Characterization of those states is crucial to understand the fate of phytol and tocopherols and their important function for seed establishment, seed longevity, germination, and development of the early seedling.

Here, early data is presented which originates from the analysis of an artificial micro RNA line (*amiR-vte6*) in which *VTE6* expression is down-regulated to a critical state for *Arabidopsis*. Still the *VTE6* expression level is sufficient to allow for studies of the complete life cycle of *Arabidopsis*. *amiR-vte6* offers the possibility to study tocopherol and phytol content in the seed, prior to germination where tocopherols were shown to possess an important function as protectant versus oxidative damage during β-oxidation and contribute to establish seed longevity (Mene-Saffrane et al., 2010; Sattler et al., 2004). In addition to that, *amiR-vte6* can aid in the characterization of later developmental stages which have not been studied yet.

Results

Co-expression Analysis of *VTE6*

Based on the atted-II platform (Obayashi et al., 2009; Obayashi et al., 2007), a list of genes which are co-expressed with VTE6 was generated to support the data presented earlier, showing that VTE6 is the proposed phytyl phosphate kinase. (Supplementary Data 1). This approach aimed at finding additional key players of tocopherol synthesis. Within this list, several interesting players of tocopherol synthesis (homogentisate phytyltransferase 1, homogentisate prenyltransferase, geranylgeranyl reductase, and geranylgeranyl pyrophosphate synthase 1) and chlorophyll biosynthesis (uroporphyrinogen decarboxylase, coproporphyrinogen III oxidase) were found.

Artificial micro RNA reduced *VTE6* expression in *Arabidopsis thaliana*

The initial lack of suitable mutants for At178620 (*VTE6*), either T-DNA lines with homozygous insertion of the T-DNA for *Arabidopsis* or *Synechocystis* knockout lines, gave rise to the generation of amiRNA lines for *Arabidopsis* in which *VTE6* expression is (at least) partially repressed (*amiR-vte6*). Using two different amiRNA targets, one suitable homozygous line was generated. Reduced *VTE6* expression could be verified by RT-PCR using three individual plants, amplifying full-length transcript based on cDNA (Figure 1).

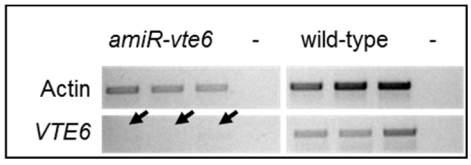


Figure 2 ***VTE6* expression levels in *amiR-vte6*.** Three individual plants for *amiR-vte6* and wild-type control were taken to investigate the expression level of *VTE6*. PCR (30 cycles) was analyzed on a 1% agarose gel. Black arrows point at feint signals for *VTE6* in *amiR-vte6*. (-) negative control, H_2O.

Actin controls were included in the same PCR study to compare expression levels between *amiR-vte6* and the wild-type control. Since actin is known to be a house-keeping gene, it is thought to show equal expression levels throughout all tissues and

both lines (*amiR-vte6* and wild-type). Although signals for actin were slightly weaker in comparison to the wild-type control, one can see that the expression of *VTE6* in *amiR-vte6* was reduced in contrast to the wild-type. PCR-amplified cDNA for *VTE6* resulted in a very feint band after 30 PCR cycles in *amiR-vte6* while the signal could be easily spotted in the wild-type control.

amiR-vte6 showed a pale-green phenotype with stunted root growth

amiR-vte6 plants were grown under ambient air conditions (380ppm CO_2, 12h light. 12h darkness) to monitor growth and morphology. *amiR-vte6* plants showed pale-green to yellow colorization of the leaves, which was pronounced on day 14 during early development (Figure 2).

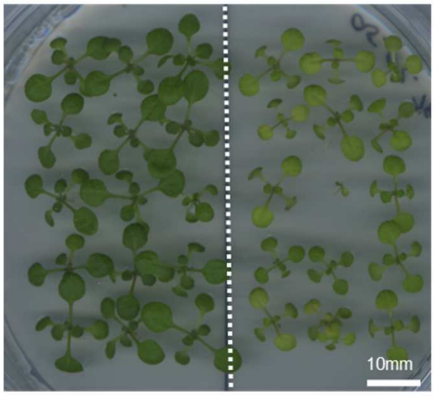


Figure 3 ***amiR-vte6*** **development.** *amiR-vte6* and wild-type plants were germinated and grown on half-strength MS media under standard ambient air conditions (380 ppm CO_2). Development of amiR-vte6 and the control is shown as monitored after 14 days of growth in 12 hours light/darkness cycle.

During initial growth studies on MS growth plates it could be observed that, in addition to the alteration in leaf colorization, root growth seemed to be affected. Growing *amiR-vte6* on vertical-placed growth plates next to wild-type plants for twelve days revealed a root growth phenotype, as expected after the initial observation was made during normal growth of *amiR-vte6*. Plants were ultimately documented on day twelve instead of day 14, because roots of the wild-type were not able to grow further down the plate due to size limitations. Root growth was measured using Fiji (Schindelin et al., 2012) in combination with the plugin "NeuronJ" (Meijering et al., 2004) using 20 individual plants per line and treatment (Figure 3).

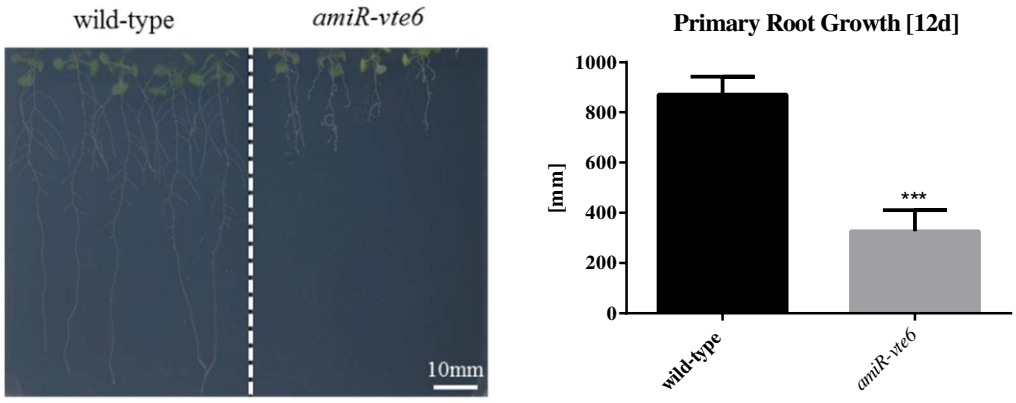


Figure 4 **Root measurements of *amiR-vte6* and wild-type *Arabidopsis*.** *amiR-vte6* and wild-type *Arabidopsis* (Col-0) were germinated and grown vertically to allow for root length measurements. Significant differences between the wild-type control and the *amiR-vte6* mutant are marked with one, two, or three asterisks (Student's t-test, * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$). 20 plants per line were taken for measurements.

Analysis of the root growth phenotype showed that root growth of the *amiR-vte6* is significantly reduced in comparison to the wild-type control. The mean root length of investigated mutants was reduced down to 40% of the wild-type control.

Early development of *amiR-vte6*

amiR-vte6 grew similar to the wild-type control until day eight. Beginning with day nine, wild-type and *amiR-vte6* plants started to emerge true leaves. While true leaves of the wild-type showed no alteration, the first true leaves of *amiR-vte6* exhibited yellow colorization. While the differences between *amiR-vte6* and the control were difficult to see for the untrained eye on day eight (due to the size of the emerging true leaves), the difference could be spotted easily within the following days. Differences in leaf colorization seemed to be pronounced in true leaves which could be visualized exceptionally well between day eleven and day twelve (Figure 4, Top). 20 day old plants did not only differ in colorization but additionally in growth. *amiR-vte6* showed stunted growth in comparison to the wild-type control. Along with that, total leaf area seemed to be reduced in *amiR-vte6* (Figure 4, Bottom).

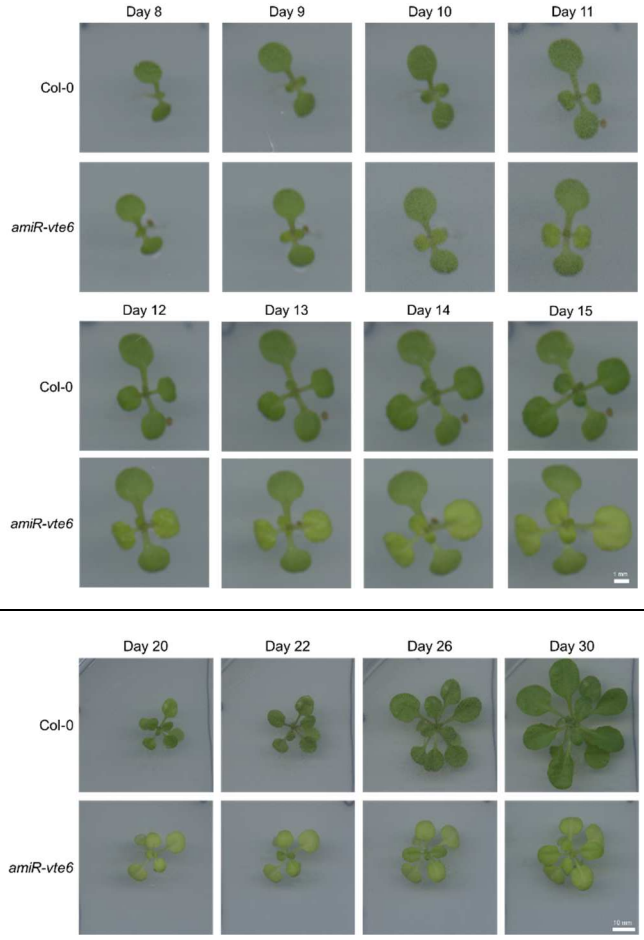
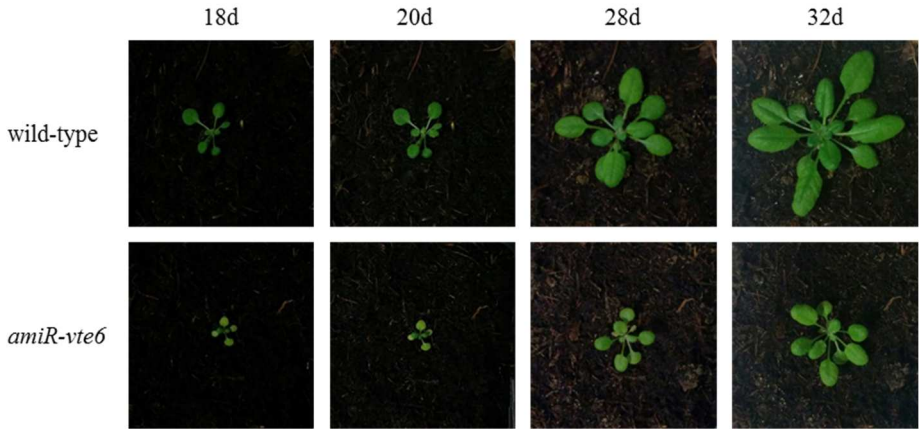


Figure 5 **Time series of *amiR-vte6* early development.** *amiR-vte6* and wild-type plants were grown on half strength MS media in a 12 hours day/night rhythm. **(Top)** Plant development was documented each day starting at day eight until day 15 (white scaling bar depicts 1 mm), and **(Bottom)** after transfer to fresh MS plates for additional 15 days (white scaling bar depicts 10 mm).

Growth of *amiR-vte6* after transfer to soil

Wild-type and *amiR-vte6* plants were transferred routinely from MS media to soil for further growth and seed propagation. Growth and leaf colorization started to resemble the wild-type state starting with day 28 (Figure 5). Noticeably, this effect was also obtained if plants were further grown on MS media (data not shown). *amiR-vte6* pigmentation started to recover in later stages. Leaf area was determined using Fiji (Schindelin et al., 2012).



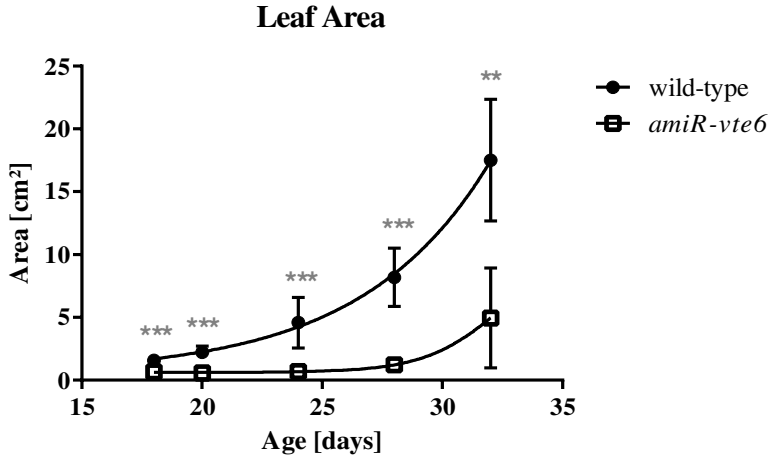


Figure 6 **Comparison of plant size and leaf colorization.** Wild-type plants and *amiR-vte6* plants were monitored after transfer to soil to document growth and leaf colorization. **(Top)** Pictures were taken for mutant and control together in one setup to guarantee identical conditions concerning light, exposure time and aperture. Distance between plants and lens was adjusted to be equal over the whole course of the experiment to ensure true size comparison. Pictures show the exact same plant for both lines in the same orientation. **(Bottom)** Measurement of leaf area of six individual plants per line. Each plant was documented under standardized conditions to allow for true area comparison and automatic measurement of the leaf area. Significant differences between the wild-type control and the *amiR-vte6* mutant are marked with one, two, or three asterisks (Student's t-test, * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$).

amiR-vte6 showed alterations in leaf pigmentation

A set of pigments in *amiR-vte6* and wild-type plants was quantified at day 14 (see Figure 4; Top), because the pale-green phenotype in *amiR-vte6* was very prominent at this developmental time point. Pigments were measured via Reverse-phase High-Performance Liquid Chromatography (RP-HPLC) and normalized by fresh weight in case of chlorophyll contents. All additional pigments were quantified and normalized based on total chlorophyll content (Figure 6). Five biological replicates containing 25 seedlings (14 day old) each, were tested for *amiR-vte6* and the wild-type.

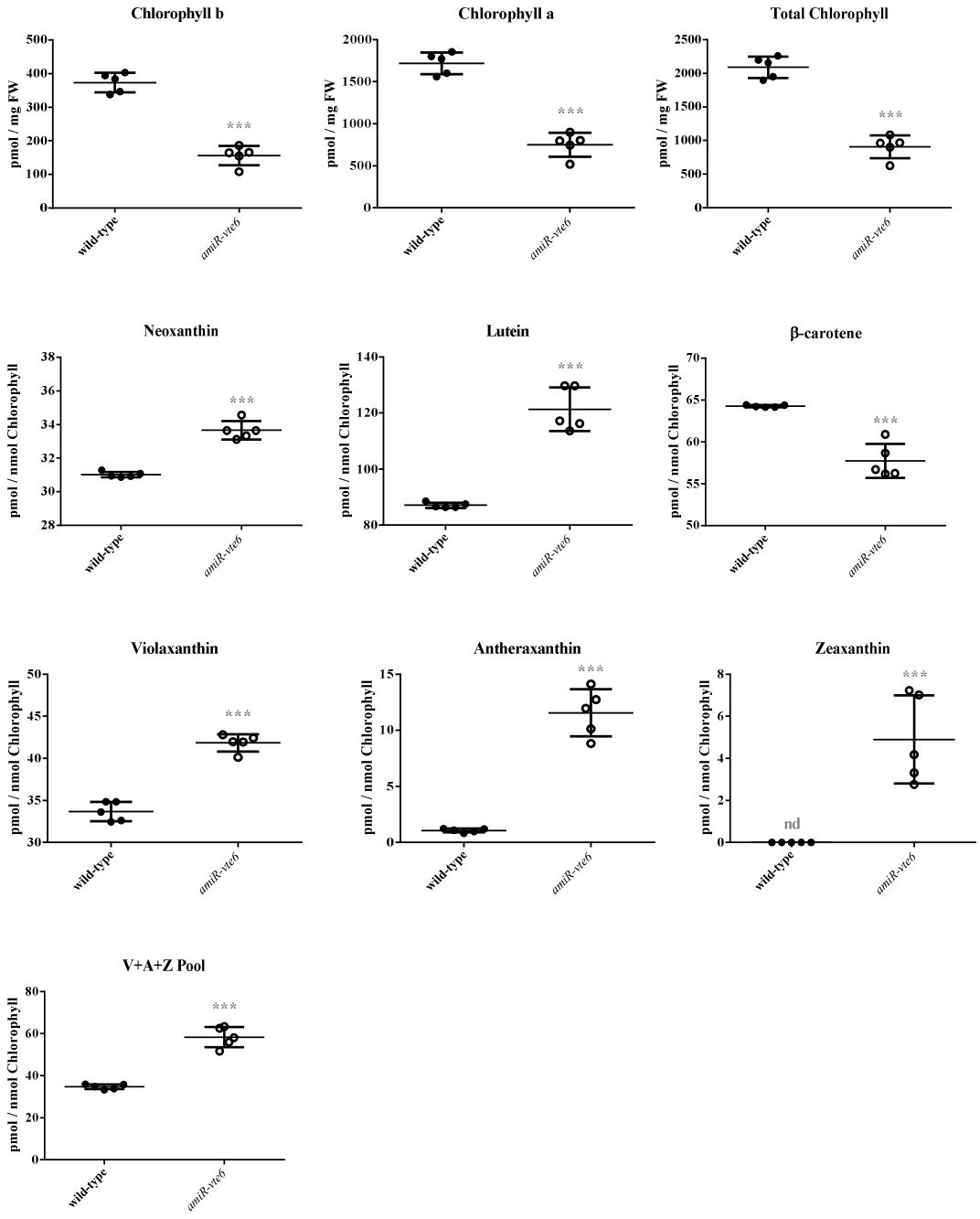


Figure 7 **Pigment quantification of *amiR-vte6* and the wild-type after 14 days.** Pigments were extracted and analyzed via reverse-phase high-performance liquid chromatography and normalized by fresh weight in case of chlorophyll and normalized by total chlorophyll for all other pigments. Significant differences between the wild-type control and the *amiR-vte6* mutant are marked with one, two, or three asterisks (Student's t-test, * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$). Five individual samples were tested per line.

Pigment analyses of the *amiR-vte6* mutant showed that the amount of several analyzed pigments was altered in comparison to the wild-type. Chlorophyll content (Chlorophyll a and b) was reduced to 44%, respectively to 42% of the wild-type. The ratio between chlorophyll a and b was not altered dramatically (but still significantly

different). The ratio between chlorophyll a and b in the wild-type control was measured with 4.6 ± 0.02, while the *amiR-vte6* mutant showed a ratio of 4.8 ± 0.03 (p<0,001). β-carotene and neoxanthin levels in *amiR-vte6* differed only slightly, whereas lutein levels were increased by 40% in the mutant.

Besides these pigments, the V+A+Z-pool (violaxanthin, antheraxanthin, and zeaxanthin) was quantified, too. Pigment analysis of the V+A+Z-pool revealed alterations if *amiR-vte6* pigmentation was compared to wild-type pigmentation. The violaxanthin level was increased in the mutant by 24% in comparison to the wild-type. Most noticeably, antheraxanthin level in the mutants was ten times higher than in the control. Last, zeaxanthin could not be detected in the wild-type control, while it could be detected in *amiR-vte6* with 5 pmol / nmol chlorophyll.

amiR-vte6 phytol content

Free phytol content of wild-type and *amiR-vte6* was measured to determine if the variation in total chlorophyll content between wild-type and mutant was mirrored in the content of free phytol, as phytol is one degradation product of chlorophyll. The results presented in figure 7 show that the relative phytol content showed no variation between *amiR-vte6* and the wild-type control.

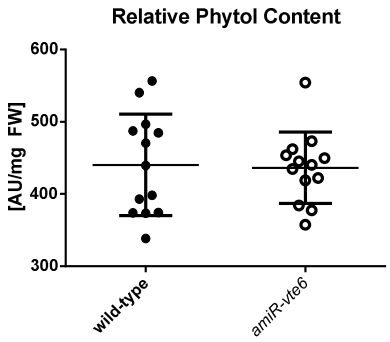


Figure 8 **Free Phytol measurement.** Free phytol was measured by GC/MS in 14 day old seedlings and normalized by fresh-weight. Significant differences between samples are marked with one, two, or three asterisks (Student's t-test, * = p<0.05; ** = p<0.01; *** = p<0.001). Twelve individual samples, consisting of a pool of 40 seedling each, were tested per line.

amiR-vte6 suffered from mild light stress

amiR-vte6 early growth was monitored under elevated light conditions to elucidate whether the mutant shows an increased response towards mild increases in light

intensity. Therefore, *amiR-vte6* was grown on half strength MS media at 150 µE and 250 µE (in contrast to standard 100 µE). To keep all additional variables as stable as possible this experiment was performed in the same plant growth cabinet like before. Variation in light intensity was achieved by placing samples at different height levels underneath the light source. Plants were monitored at day 10 and day 18 (Figure 8).

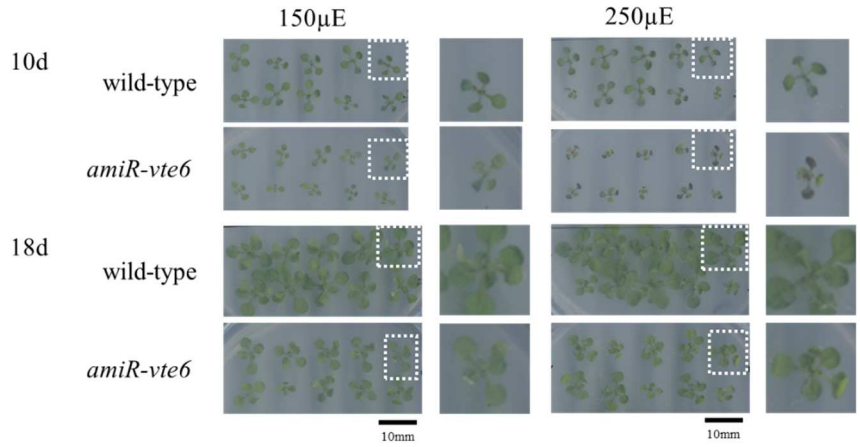


Figure 9 ***amiR-vte6*** **grown at 150µE and 250.** *amiR-vte6* was grown along with the wild-type control for 18 days on half strength MS media at different light intensities (150 µE, 250 µE). Plant development was documented at day 10 and day 18 in bottom view to visualize dark pigment accumulation in *amiR-vte6*. Individual plants marked in dashed boxes are magnified in the right part for a close-up view.

amiR-vte6 grown for ten days under elevated light conditions differed remarkably in both, size and colorization of the leaves, from the wild-type control. Mutant growth was stunted under both tested stress conditions. Moreover, *amiR-vte6* showed yellow colored leaves. In addition, *amiR-vte6* plants grown at 250 µE showed dark violet pigmentation due to anthocyanin accumulation on the bottom side of the leaf at day 10. Anthocyanin accumulation was not detected in the wild-type control of plants grown at 250 µE. *amiR-vte6* which was grown under milder light intensities (150 µE) did not show a tendency towards anthocyanin accumulation, too.

Noticeably, anthocyanin accumulation vanished until day 18 in the *amiR-vte6 plants*. Still, *amiR-vte6* differed in size and leaf colorization (top view) as shown in earlier experiments (Figure 5, Figure 6).

DISCUSSION

It was recently shown that At1g78620 codes for the proposed phytyl phosphate kinase (VTE6), which is involved in chlorophyll degradation and tocopherol synthesis (vom Dorp et al., 2015). The initial hypothesis of At1g78620 being a photorespiratory transport protein, which was derived based on data published earlier (Bordych et al., 2013) had to be rejected. Still, data on the generated At1g78620 mutant *amiR-vte6* adds up to the work presented by vom Dorp and colleagues. It offers further insights into VTE6 function in tocopherol metabolism and beyond and opens additional paths for further studies about VTE6 function in *Arabidopsis*.

VTE6 is co-expressed with genes for tocopherol synthesis

Co-expression data (Supplementary Data 1) revealed a strong correlation between *VTE6* expression and the expression of several major genes which are occupied in tocopherol and phylloquinone biosynthesis. Within these genes which are co-expressed with *VTE6*, one can find *VTE2-1* (*HPT1*, At2g18950, homogentisate phytyltransferase 1) and its paralog *VTE2-2* (*PDS2*, At3g11945, homogentisate prenyltransferase). Both enzymes catalyze reactions starting with homogentisic acid which serves as a prenyl group donor (Tian et al., 2007; Venkatesh et al., 2006). VTE2-2 catalyzes the condensation of homogentisic acid and solanesyl diphosphate to yield 2-demethylplastoquinol-9 which is further oxidized and methylated to yield plastoquinone-9 for the electron transport chain in the thylakoid membrane during light reaction of photosynthesis. *VTE2-2* characterization was performed and published under its former AGI code At3g11950 (Tian et al., 2007).

VTE2-1 catalyzes a comparable reaction by condensation of homogentisic acid with phytyl diphosphate. This reaction forms 2-methyl-6-phytylbenzoquinol which is the first committed intermediate in tocopherol biosynthesis (Collakova and DellaPenna, 2003).

Besides being synthesized through phytyl phosphate, a second way for the generation of phytyl diphosphate is reported (Keller et al., 1998), which is the reduction of geranylgeranyl diphosphate to phytyl diphosphate by the enzyme GGR (At4g38460, geranylgeranyl reductase). The corresponding gene was found in the list of genes which are co-expressed with *VTE6*, too. Likewise, the gene coding for the enzyme,

which catalyzes the preceding reaction yielding geranylgeranyl diphosphate, was detected among the co-expressed genes. *GGPS1* (At4g36810, geranylgeranyl pyrophosphate synthase 1). *GGPS1* is the last noticeable gene on the co-expression list for *VTE6*, which facilitates reaction closely connected to the tocopherol synthesis pathway around *VTE6*.

Further, Seaver and colleagues showed by comparative genomics that At1g78620 (*VTE6)* homologues form structural genomic clusters in prokaryotes together with *VTE5* (Seaver et al., 2014). *VTE5* itself has been already demonstrated to catalyze the conversion of phytol to phytyl phosphate (Ischebeck et al., 2006) yielding an important precursor for tocopherol and phylloquinone synthesis.

Co-expression data generated on the atted-II platform as demonstrated here, yielded powerful results itself, for the interpretation of the role of VTE6 as the phytyl phosphate kinase in *Arabidopsis*, as it depicts a network of genes around VTE6 which have already been demonstrated to be occupied in tocopherol synthesis or closely related pathways.

amiR-vte6 endorsed the *vte6-1* and *vte6-2* phenotype presented recently

The amiRNA approach was chosen to repress *VTE6* expression only partially, because earlier results suggested that a complete shutdown is lethal for *Arabidopsis*. Preceding studies showed that the generation of homozygous T-DNA lines (Salk_075640 and Salk_002072) could not be achieved. Further, the approach to study VTE6 function in *Synechocystis* yielded no results. No complete knock-out of the *VTE6* homologue sll0875 in *Synechocystis* could have been achieved (Supplementary Figure 2).

The phenotype of *amiR-vte6* presented here, pictures a mild version of the *vte6-1* and *vte6-2* phenotype, which was published recently by vom Dorp and co-workers. Both, the transposon lines *vte6-1* and *vte6-2*, as well as the presented amiRNA line *amiR-vte6* share a common phenotype with pale-green leaves, stunted growth in combination with weak development of the root system (vom Dorp et al., 2015).

In contrast to *vte6-1* and *vte6-2*, *amiR-vte6* could be grown on soil in addition to the growth on artificial MS media. Routinely, *amiR-vte6* was grown for at least three weeks on MS media before plants were transferred to soil. As shown, the phenotype

of *amiR-vte6* was most prominent within early development, exclusively documented while growing on MS media. Moreover, *amiR-vte6* growth is not dependent on supplementation by sucrose as shown for *vte6-1* and *vte6-2*. These differences are most likely caused by the nature of mutant lines under investigation.

vte6-1 and *vte6-2* are transposon lines with an insertions in the first exon of At1g78620, which abolish *VTE6* expression completely. Both lines are considered to be null mutants (vom Dorp et al., 2015). Expression of *VTE6* in *amiR-vte6* was modulated by the amiRNA approach (Ossowski et al., 2008; Schwab et al., 2006), which led to the partial degradation of expressed mRNA of *VTE6*. Different factors affected the effectiveness of the repression of *VTE6* expression. Utilization of the ubiquitin promoter system in case of *amiR-vte6* guaranteed for expression of the amiRNA in a wide variety of tissues in *Arabidopsis*. On the other hand the promoter strength was too weak to shut down *VTE6* expression completely. Remarkably, *VTE6* expression was not repressed completely in *amiR-vte6*, as it was shown that *VTE6* transcript can be detected in reduced amounts.

Moreover, the location of the amiRNA construct itself in the *Arabidopsis* genome after transformation via *Agrobacterium*, could have had an effect on the expression of the amiRNA construct independent from the promoter itself. Certain areas in the genome are not as frequently transcribed than others (positional inhibitory effect). The fact that only one amiRNA line was available for this study could not rule out, that additional amiRNA lines would show different (more severe) effects, yielding a lethal phenotype like the one for *vte6-1* and *vte6-2*. Reduced *VTE6* expression in *amiR-vte6* ensured that the phenotype as observed in *amiR-vte6* embodied a mild version of the phenotype as described for *vte6-1* and *vte6-2*.

While *amiR-vte6* resembled the phenotype of *vte6-1* and *vte6-2* during early development with growth retardation, leaf colorization, and root growth defects, it was shown to be different in later stages while growing on soil. Additional, it cannot be ruled out that later stages *per se* would not have suffered, because growth of *amiR-vte6* on MS media was not examined consequently in late stages in detail. Still, it is rather unlikely that the transfer from MS media to soil addressed for the changes in the phenotypic observation.

While null mutants are unable to grow without supplementation and on soil and are incapable of reproduction, *amiR-vte6* was shown to grow on soil. Most noticeably, *amiR-vte6* was not only able to grow further, but also re-established a more wild-type like phenotype in later stages. The pale-green phenotype vanished over time and the mutant line was able to reproduce. Utilization of the amiRNA approach addressed for these differences. Promoter activity of the UBQ promoter in later developmental stages of *amiR-vte6* might have been not strong enough to repress *VTE6* expression as efficiently as in earlier stages so that *amiR-vte6* grew more wild-type like in later stages due to higher VTE6 expression in general.

Moreover, the work presented here, showing that *amiR-vte6* recovers from the phenotype presented, gives room for the speculation that VTE6 function is especially crucial in distinct developmental stages of *Arabidopsis*. In this context the data presented by vom Dorp and colleagues and data about *amiR-vte6* support each other (vom Dorp et al., 2015). It was described that VTE6 function is of major importance during early development (as shown by leaf pigmentation, stunted growth and the root phenotype in early stages).

The severe impact on early development is, to a large extend, caused by the exceptional role that tocopherols play during seed germination. Tocopherols were shown to prevent lipid peroxidation caused by ROS (reactive oxygen species) during β-oxidation of fatty acids (Sattler et al., 2004). This conclusion is further supported by the fact that *vte6-1* and *vte6-2* are dependent on sucrose supplementation for germination, which is a common treatment for mutants in β-oxidation to allow for germination and initial development through supply of energy (Hayashi et al., 1998). The fact that *amiR-vte6* can germinate without external addition of sucrose in contrast to the null mutants *vte6-1* and *vte6-2,* is once again caused by the intensity by which *VTE6* expression is reduced. While the partial reduction in *VTE6* expression did not demand for additional energy supply to bypass β-oxidation of fatty acids, the complete shutdown of *VTE6* expression does make sucrose supplementation obligate to prevent oxidative damage during β-oxidation. In this context it is of major interest to determine the degree by which *amiR-vte6* is affected by oxidative damage through β-oxidation of fatty acids. Future experiments have to clarify this question.

Interestingly another common factor could be found regarding late stages of development, as both, *vte6* and *amiR-vte6* showed tendencies towards a defect in seed production. Generation of homozygous seeds of *vte6-1* and *vte6-2* is reduced down to 11% in contrast to the expected 25% (vom Dorp et al., 2015).

While seed productivity was not quantified in *amiR-vte6*, the tendency was shown that overall seed yield for *amiR-vte6* is strongly reduced (Supplementary Figure 3 and Supplementary Figure 4). *amiR-vte6* siliques under investigation showed that they were not as tightly packed as siliques of the wild-type control, pointing towards abortion of seed development in an early phase. These experiments were not performed in a large scale and need further support in the future to quantify the effect in *amiR-vte6* on seed establishment. This early data strongly suggests a critical role for VTE6 during seed establishment.

Tocopherol and precursor levels in seed and leave affected *Arabidopsis* development

VTE6 catalyzes the reaction from phytyl phosphate to phytyl diphosphate which constitutes an intermediate compound of tocopherol synthesis. It was shown that tocopherol synthesis is strongly dependent on phytyl diphosphate originating from chlorophyll degradation during senescence (Valentin et al., 2006; vom Dorp et al., 2015). However, tocopherol content in the seed or green tissues can be ruled out to be responsible for the severe phenotype shown for *vte6-1* and *vte6-2* exclusively on its own. It is known that tocopherol-free *Arabidopsis* mutants (*vte1*, *vte2*) are able to develop and reproduce while showing a milder phenotype than the described *vte6* mutants (Havaux et al., 2005; Sattler et al., 2004; vom Dorp et al., 2015). Still tocopherol content of the seed is described to impact seed longevity and reducing lipid peroxidation during germination (Sattler et al., 2004). Therefore, the lack of tocopherol is most likely only one important piece in the complex phenotype of *vte6* mutant plants.

The tocopherol content of *amiR-vte6* itself was not analyzed. The decrease in tocopherol was shown to be reduced for the *vte6-1* and *vte6-2* mutant, which argues towards a reduction of tocopherol in the *amiR-vte6* mutant as well. Preliminary results for *amiR-vte6* showed that the content of the precursor phytol was indistinguishable

from the wild-type. Decreased VTE6 expression in *amiR-vte6* was most likely causing an increase in the amount of phytyl phosphate in all tissues, due to the conversion of phytol to phytyl phosphate by VTE5. Subsequently, phytyl diphosphate concentrations must have been lowered in mutant plants. This decrease in total phytyl diphosphate in combination with the (hypothetical) increase in intermediates (phytyl phosphate) very likely had an impact on the phenotype observed for *amiR-vte6* (and *vte6-1* and *vte6-2*). Still, no toxic function of phytyl phosphate is reported in the literature so far.

Results for phytol content originated from a preliminary experiment and were by far not as detailed and professionally performed as described by vom Dorp (vom Dorp et al., 2015). Future experiments have to aim at an in-detail characterization of the level of phytyl, phytyl phosphate, and phytyl diphosphate in *amiR-vte6* in close collaboration with vom Dorp and colleagues.

Already published data for tocopherol mutants in *Arabidopsis* revealed that the variety of observed phenotypes is caused by a combination of tocopherol absence/reduction and the amount and the distinct type of the accumulating intermediates. For example, Sattler and co-workers pointed out that the accumulation of 2,3-dimethyl-6-phytyl-1,4-benzoquinone (DMPBQ) in the *vte1* mutant can compensate for the absence of tocopherol with regard to its function as antioxidant (Sattler et al., 2003; Sattler et al., 2004).

As a consequence, the lack of tocopherols *per se* does not result in such an extreme phenotype as observed in *vte6* mutant plants. In this context it is noteworthy that the lack of tocopherols in *Synechocystis* do not cause a severe effect on *Synechocystis*, as shown for a mutant defect in homogentisate phytyltransferease (VTE2) (Collakova and DellaPenna, 2001). Still, it was impossible to generate knockout mutants for the VTE6 homologue sll0875 in *Synechocystis* for the study presented here (Supplementary Data, Supplementary Figure 2).

The fact that *vte6* mutants showed the most severe phenotype among all investigated mutants of tocopherol synthesis so far, is highly interesting and gives rise to further questions about intermediates of tocopherol synthesis and their respective functions. In the context of VTE6, further studies have to aim at the characterization of potential influences of phytyl phosphate accumulation in *Arabidopsis* and its effects on growth and development.

***amiR-vte6* was severely affected by light-stress**

One component of the multifarious reaction of *Arabidopsis* upon reduction in VTE6 expression was an increased sensitivity towards elevated light intensities. It could be shown that the *amiR-vte6* mutant showed strong reactions towards (mild) increases in light intensity during early development. This observation is supported by the fact that tocopherols, as a product of phytyl diphosphate, are an important element of plant light stress response (Niyogi, 1999).

Differences in colorization between wild-type and *amiR-vte6* were visible by eye during development under mild light stress and pointed towards accumulation of anthocyanin. Anthocyanin was shown to be involved in protection from photoinhibition (Gould, 2004; Smillie and Hetherington, 1999). Even under standard growth conditions, pigment composition was altered in the *amiR-vte6* mutant with distinct increases in lutein, antheraxanthin, and zeaxanthin levels, together with an increase in the overall V+A+Z pool (Figure 6). These adaptions in pigment composition of *amiR-vte6* can be seen as a reaction to compensate for the lack of tocopherols as protectant against increased light intensities.

Both, lutein and zeaxanthin are well known for their photoprotective function during high-light stress. Lutein was shown to be an important factor in light stress response in plants (Kawabata and Takeda, 2014) as it decreases triplet chlorophyll formation in the photosystem (Dall'Osto et al., 2006).

The total increase in the V+A+Z pool in combination with the shift from violaxanthin to zeaxanthin (over antheraxanthin) in the *amiR-vte6* mutant added a second line of evidence to show that *amiR-vte6* is indeed suffering under standard light conditions. Zeaxanthin is another important component in plant photoprotection, by quenching excited singlet chlorophyll and thereby in non-photochemical quenching (Jahns and Holzwarth, 2012). The shift from violaxanthin towards zeaxanthin was already reported to be caused by light stress and helps plants to withstand periods of excessive light intensity (Havaux and Niyogi, 1999; Johnson et al., 2007).

Data as shown here, strongly points out, that *amiR-vte6* compensated a lack in tocopherols with an increase in (other) photoprotective pigments, such as shown for lutein and zeaxanthin. Interestingly, in case of zeaxanthin it was already shown that a

lack in zeaxanthin can be bypassed by accumulation of tocopherols which indicates overlapping function for tocopherols and zeaxanthin (Havaux et al., 2000).

The increase in the zeaxanthin level was accompanied by a decrease in β-carotene which can be explained by the fact that zeaxanthin synthesis can be based on the hydroxylation of β-carotene during reactions towards increased light stress of plants (Depka et al., 1998). The observed increase of zeaxanthin, or more specifically the shift from violaxanthin towards zeaxanthin was caused by a combination of the xanthophyll cycle (Eskling et al., 1997) and synthesis of zeaxanthin based on β-carotene.

Analyses of the pigment composition of *amiR-vte6* showed that the decrease in *VTE6* expression caused a complex reaction within the plant leading to alterations on the pigment level to circumvent the lack in tocopherols for photoprotective purposes.

The given data does not answer the question which of these effects were directly caused by the lack in tocopherols, phytyl diphosphate or one of the intermediate substrates. Still, besides being affected in β-oxidation during germination, light stress reaction adds another puzzle piece to the complex picture of the presented *vte6* phenotype.

Pigment composition in different leave types

Phenotypic analysis showed that the colorization phenotype of *amiR-vte6* was pronounced in true leaves in comparison to cotyledons (Figure 4). Yellow staining seemed to be exclusively found in emerging true leaves of 14 day old plants. To test, if this difference could be put into numbers via quantification of pigments, true leaves of *amiR-vte6* and the wild-type control were separated for individual pigment analysis (Supplementary Figure 1). This pilot experiment indicated that total chlorophyll reduction is indeed pronounced in true leaves of *amiR-vte6*. Regarding cotyledons the analysis pointed towards no significant difference between the mutant and the wild-type.

Analysis of the remaining set of pigments indicated that most pigments followed the trend which was discussed for whole leaf extracts earlier. Overall neoxanthin, lutein, violaxanthin, and total V+A+Z pool pigments were increased in the mutant. Still, individual analysis of pigment pools for true leaves and cotyledons indicated no trend

that alterations are exclusively pronounced in true leaves. The reduction of β-carotene which has been discussed for the whole leaf analysis could also been shown after separation of leaves.

One important fact which has to be mentioned is, that the detailed analysis for antheraxanthin and zeaxanthin did not follow the results presented above for whole leaf extracts. While antheraxanthin levels did not seem to follow any trend, zeaxanthin levels were exclusively high in cotyledons of the wild-type. In contrast to that, zeaxanthin was measured earlier to be increased in *amiR-vte6* and not in the wild-type.

Despite being measured in a very small sample size of n=3, these preliminary results indicated that the yellow colorization of the true leaves was to a large extend caused by alterations in chlorophyll levels. Alterations in the other pigments under investigation seem not to add up to this (visual) effect.

It is very important to mention that this experiment was only performed once and needs further support. In addition to this, the experimental setup needs improvement and modifications, as separation of leaves and isolation took longer than isolation of samples for whole leave extract measurements. As pigment levels are very sensitive to changing light conditions, the prolonged sampling procedure had most likely an effect on the measurement. While whole leave samples could have been isolated under growing light conditions within seconds, separation of leaves had to be performed in the laboratory. As a consequence, samples were transported from the growth chamber into the laboratory and kept in the laboratory while being in queue for isolation. These practical necessities in combination with the already prolonged sampling time could have definitely caused tremendous differences.

Still these early results as presented here can serve as a starting point for further investigations in which true leaves are isolated from cotyledons for pigment analysis. These experiments have to ensure, that light conditions are stable among all samples and that sampling is performed more efficiently to guarantee that the results are not falsified by wrong handling.

Overall, these results indicated that the phenotype as discussed was strongly pronounced in autotrophic tissues which relied on photosynthesis for energy production. As a consequence the very next experimental steps in the characterization

of *amiR-vte6* have to aim at measuring the photosynthetic capacity of young plants, to check if they show indeed impaired photosynthesis.

CONCLUSION

The generation of *amiR-vte6* offers a valuable tool to comprehend the phenotype for existing *vte6* mutants and offer the opportunity to study processes and developmental stages which cannot be investigated in already described *vte6* mutants for *Arabidopsis*. The necessity for such intensified studies has to be emphasized, because VTE6 most likely plays a crucial role in the plant. All *Arabidopsis* mutants described for tocopherol synthesis so far, were shown to be influenced by the mutation but not lethal beyond a certain developmental stage like *vte6-1* and *vte6-2*. This novelty in the analyses of tocopherol mutants in *Arabidopsis* is accompanied by the fact that *Synechocystis sp.* PCC6803 is also severely affected by mutations of VTE6, as knockout mutants for VTE6 could not be generated throughout this work, even under high selective pressure.

As VTE6 was shown to play an important role in *Arabidopsis*, it is necessary to complete the toolbox to study VTE6 function in *Arabidopsis*. *amiR-vte6* was described to show a comparable phenotype to the existing *vte6* mutants during early development. In contrast to existing *vte6* lines, *amiR-vte6* lacks the lethality of *vte6-1* and *vte6-2* in later developmental stages. This opens up the possibility for intensified studies on these developmental stages.

Still, data on *amiR-vte6* is currently limited. Future experiments have to aim at several different aspects which are important for the characterization of *amiR-vte6*. Ongoing experiments target tocopherol levels in all developmental stages of the artificial microRNA line to understand the outstanding role of VTE6 and the divergent phenotype during development. In this context the expression level of *VTE6* in the wild-type throughout a developmental course would be another prime target for further investigations. This approach can clarify if VTE6 function is indeed of outstanding importance at certain time points within development.

First evidence was discussed, arguing that the role of VTE6 cannot be limited to its participation in tocopherol synthesis alone, because it is known that tocopherol deficiency *per se* does not cause lethality in *Arabidopsis* or *Synechocystis*. This fact illustrates that intensified studies about VTE6 are inevitable. It has to be clarified if the lack in VTE6 expression causes alterations in yet not considered pathways in the plant.

Altogether, VTE6 is a highly interesting enzyme during *Arabidopsis* development and further experiments with *vte6-1* and *vte6-2* and especially *amiR-vte6* will shed light on its multifarious and important functions. With *amiR-vte6*, this work as described here, could add a powerful tool to study these functions in *Arabidopsis*.

MATERIAL AND METHODS

Arabidopsis Growth

Arabidopsis plants were grown in 12 hours light/darkness cycle and 65% humidity in Percival growth cabinets (clf plant climatics, Wertingen, Germany). Plants were either grown on half-strength MS media or soil (*Arabidopsis* substrate).

Prior to growth on MS media, seeds were surface sterilized using chlorine gas for 90 minutes. Remaining chlorine gas was allowed to evaporate for 60 minutes under sterile conditions. Growth on MS media was maintained in Petri dishes which were sealed with gaze tape. Prior to placement in growth cabinets, seeds on MS plates were incubated at 4°C in darkness for 24 hours.

***Synechocystis* Growth**

Synechocystis PCC6803 was grown on BG11 media on a shaker under constant light of 100 µE and 30 °C either under photoautotroph or heterotroph conditions with addition of 5 mM glucose to enable potential photosynthetic mutants to grow. *Synechocystis* was either cultivated in liquid BG11 media or grown on solid BG11 agar plates. Stock solutions. Agar plates were produced using 2x BG11 media mixed with 2x agar (3g/100ml). Both, BG11 stocks and 2x agar was autoclaved prior to preparation of agar plates.

Generation of amiRNA silenced *Arabidopsis* lines

Stable *Arabidopsis* amiRNA lines were generated following the approach described by Ossowski and co-workers (Ossowski et al., 2008; Schwab et al., 2006). Potential amiRNA targets were identified using the WMD3 online tool (http://wmd3.weigelworld.org/cgi-bin/webapp.cgi?page=Home;project=stdwmd) (Ossowski et al., 2008). Two individual (top rated) targets (Table 1) instead of just one were chosen, because amiRNA efficiency is was not guaranteed. Corresponding oligonucleotides suggested by the online platform were designed (Table 2).

Table 1 **Targets used for generation of amiRNA constructs to reduce expression of *VTE6*.** Targeting sequences were obtained from the online platform WMD3. Two different targets were chosen and corresponding sequence introduced in the vectors pCB1 and pCB2

	Expression Vector	*5'-3' sequence*
Target 1	pCB1	TAGTACCAAGCAAAAACGCGC
Target 2	pCB2	TATCAGTCGTAAAATACGCGC

Table 2 **Oligonucleotides utilized for generation of amiRNA constructs.** Based on the targets chosen the WMD3 platform suggested the oligonucleotides mentioned in this table to generate the final amiRNA construct. CB167-CB174 were ordered and used as suggested by the WMD3 protocol (oligonucleotides I, II, III, and IV according to the protocol). The oligonucleotides A and B were used in a modified version (gSV1 and miFB14) to allow compatibility with the Gateway® vector system.

Name	*Sequence 5'-3'*	*Note*
CB167	gaTAGTACCAAGCAAAAACGCGCtctctcttttgtattcc	I – Target 1
CB168	gaGCGCGTTTTTGCTTGGTACTAtcaaagagaatcaatga	II – Target 1
CB169	gaGCACGTTTTTGCTAGGTACTTtcacaggtcgtgatatg	III - Target 1
CB170	gaAAGTACCTAGCAAAAACGTGCtctacatatatattcct	IV - Target 1
CB171	gaTATCAGTCGTAAAATACGCGCtctctcttttgtattcc	I - Target 2
CB172	gaGCGCGTATTTTACGACTGATAtcaaagagaatcaatga	II - Target 2
CB173	gaGCACGTATTTTACCACTGATTtcacaggtcgtgatatg	III - Target 2
CB174	gaAATCAGTGGTAAAATACGTGCtctacatatatattcct	IV - Target 2
gSV1	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCACCCTGC AAGGCGATTAAGTTGG	A – Gateway
miFB14	GGACCACTTTGTACAAGAAAGCTGGGTCGCGGATAACA ATTTCACACAGG	B – Gateway

Artificial microRNA expression constructs were cloned based on the pRS300 vector template according to the published protocol (Schwab, R; http://wmd3.weigelworld.org/downloads/Cloning_of_artificial_microRNAs.pdf) Cloned amiRNA sequences were inserted into pDONR207 via recombination (BP-reaction) as described in the Gateway® manual. Constructs were verified via sequencing in pDONR207 and further recombined into the desired final vector by Gateway® LR reaction.

Constitutive expression of At1g78620 amiRNA

amiRNA sequences for target 1 and target 2 were recombined into pUB-Dest via LR reaction for constitutive expression of amiRNA in *Arabidopsis* Col-0 under the control of the ubiquitin10-promoter (pUBQ10::amiRNA1 and pUBQ10::amiRNA2). Data as presented here originates from a mutant in which the amiRNA construct consists of Target 1.

Pigment quantification

Reverse-phase High-Performance Liquid Chromatography (RP-HPLC) was utilized to measure absolute amounts of several pigments in leaf extract. RP-HPLC takes advantage of hydrophobic interaction between the hydrophobic pigment which is analyzed and a hydrophobic (stationary) phase. Based on the pigments hydrophobicity, pigments were eluted from the stationary phase at different time points (retention time). Elution of a compound was detected photometrically. Through comparison with documented retention times the respective compound which was eluted could have been identified. Further, by integration of the area below the respective peak in retention time the compound was quantified (Färber et al., 1997).

Plant material for pigment analysis was harvested and immediately frozen in liquid nitrogen until further processing. Prior to pigment extraction leaf material was ground in liquid nitrogen and aliquoted into sample tubes and weighed. Ground material was taken up in 500 µl 100% Acetone and mixed by vortexing for 10 seconds. Sample mixtures were kept at -20°C overnight. The following day the samples were centrifuged at 4 °C and 10000g for 10 minutes. The supernatant containing the pigments was filtered through Minisart filter (0.2 µm).

Filtrate was stored at -20 °C until HPLC analysis, but not longer than 24 hours. After HPLC analysis, pigments were identified by their respective retention time and quantified by using conversion factors to deduce the absolute amount based on the area generated by the peak in retention time (Table 3).

Table 3 **Retention times and conversion factors for absolute quantification of pigments.**

Pigments	*t'_R (minutes)*	*Conversion factor (area*pmol^{-1})*
Neoxanthin	3.2	3434
Violaxanthin	4.2	3832
Antheraxanthin	6.3	3045
Lutein	9.3	3167
Zeaxanthin	10.3	2764
Chlorophyll b	13.3	1227
Chlorophyll a	13.9	1389
ß-Carotene	16.8	2852

REFERENCES

Azzi, A., Gysin, R., Kempna, P., Ricciarelli, R., Villacorta, L., Visarius, T., Zingg, J.M. (2003) The role of alpha-tocopherol in preventing disease: from epidemiology to molecular events. Mol Aspects Med, **24**(6), 325-336.

Bordych, C., Eisenhut, M., Pick, T.R., Kuelahoglu, C., Weber, A.P. (2013) Co-expression analysis as tool for the discovery of transport proteins in photorespiration. Plant Biol (Stuttg), **15**(4), 686-693.

Collakova, E. and DellaPenna, D. (2001) Isolation and functional analysis of homogentisate phytyltransferase from Synechocystis sp. PCC 6803 and Arabidopsis. Plant physiology, **127**(3), 1113-1124.

Collakova, E. and DellaPenna, D. (2003) The role of homogentisate phytyltransferase and other tocopherol pathway enzymes in the regulation of tocopherol synthesis during abiotic stress. Plant physiology, **133**(2), 930-940.

Dall'Osto, L., Lico, C., Alric, J., Giuliano, G., Havaux, M., Bassi, R. (2006) Lutein is needed for efficient chlorophyll triplet quenching in the major LHCII antenna complex of higher plants and effective photoprotection in vivo under strong light. BMC Plant Biol, **6**(1), 32.

Depka, B., Jahns, P., Trebst, A. (1998) Beta-carotene to zeaxanthin conversion in the rapid turnover of the D1 protein of photosystem II. FEBS letters, **424**(3), 267-270.

Eskling, M., Arvidsson, P.O., Åkerlund, H.E. (1997) The xanthophyll cycle, its regulation and components. Physiol Plant, **100**(4), 806-816.

Färber, A., Young, A.J., Ruban, A.V., Horton, P., Jahns, P. (1997) Dynamics of Xanthophyll-Cycle Activity in Different Antenna Subcomplexes in the Photosynthetic Membranes of Higher Plants (The Relationship between Zeaxanthin Conversion and Nonphotochemical Fluorescence Quenching). Plant physiology, **115**(4), 1609-1618.

Gould, K.S. (2004) Nature's Swiss Army Knife: The Diverse Protective Roles of Anthocyanins in Leaves. J Biomed Biotechnol, **2004**(5), 314-320.

Havaux, M., Bonfils, J.P., Lutz, C., Niyogi, K.K. (2000) Photodamage of the photosynthetic apparatus and its dependence on the leaf developmental stage in the npq1 Arabidopsis mutant deficient in the xanthophyll cycle enzyme violaxanthin de-epoxidase. Plant physiology, **124**(1), 273-284.

Havaux, M., Eymery, F., Porfirova, S., Rey, P., Dormann, P. (2005) Vitamin E protects against photoinhibition and photooxidative stress in Arabidopsis thaliana. The Plant cell, **17**(12), 3451-3469.

Havaux, M. and Niyogi, K.K. (1999) The violaxanthin cycle protects plants from photooxidative damage by more than one mechanism. Proceedings of the National Academy of Sciences of the United States of America, **96**(15), 8762-8767.

Hayashi, M., Toriyama, K., Kondo, M., Nishimura, M. (1998) 2,4-Dichlorophenoxybutyric acid-resistant mutants of Arabidopsis have defects in glyoxysomal fatty acid beta-oxidation. The Plant cell, **10**(2), 183-195.

Ischebeck, T., Zbierzak, A.M., Kanwischer, M., Dormann, P. (2006) A salvage pathway for phytol metabolism in Arabidopsis. The Journal of biological chemistry, **281**(5), 2470-2477.

Jahns, P. and Holzwarth, A.R. (2012) The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. Biochimica et biophysica acta, **1817**(1), 182-193.

Johnson, M.P., Havaux, M., Triantaphylides, C., Ksas, B., Pascal, A.A., Robert, B., Davison, P.A., Ruban, A.V., Horton, P. (2007) Elevated zeaxanthin bound to oligomeric LHCII enhances the resistance of Arabidopsis to photooxidative stress by a lipid-protective, antioxidant mechanism. The Journal of biological chemistry, **282**(31), 22605-22618.

Kawabata, Y. and Takeda, S. (2014) Regulation of xanthophyll cycle pool size in response to high light irradiance in Arabidopsis. Plant Biotechnology, **31**(3), 229-240.

Keller, Y., Bouvier, F., d'Harlingue, A., Camara, B. (1998) Metabolic compartmentation of plastid prenyllipid biosynthesis--evidence for the

involvement of a multifunctional geranylgeranyl reductase. Eur J Biochem, **251**(1-2), 413-417.

Maeda, H. and DellaPenna, D. (2007) Tocopherol functions in photosynthetic organisms. Curr Opin Plant Biol, **10**(3), 260-265.

Meijering, E., Jacob, M., Sarria, J.C., Steiner, P., Hirling, H., Unser, M. (2004) Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. Cytometry A, **58**(2), 167-176.

Mene-Saffrane, L., Jones, A.D., DellaPenna, D. (2010) Plastochromanol-8 and tocopherols are essential lipid-soluble antioxidants during seed desiccation and quiescence in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America, **107**(41), 17815-17820.

Munné-Bosch, S. and Alegre, L. (2002) The Function of Tocopherols and Tocotrienols in Plants. Critical Reviews in Plant Sciences, **21**(1), 31-57.

Niyogi, K.K. (1999) PHOTOPROTECTION REVISITED: Genetic and Molecular Approaches. Annu Rev Plant Physiol Plant Mol Biol, **50**, 333-359.

Obayashi, T., Hayashi, S., Saeki, M., Ohta, H., Kinoshita, K. (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic acids research, **37**(Database issue), D987-991.

Obayashi, T., Kinoshita, K., Nakai, K., Shibaoka, M., Hayashi, S., Saeki, M., Shibata, D., Saito, K., Ohta, H. (2007) ATTED-II: a database of co-expressed genes and cis elements for identifying co-regulated gene groups in Arabidopsis. Nucleic acids research, **35**(Database issue), D863-869.

Ossowski, S., Schwab, R., Weigel, D. (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. The Plant journal : for cell and molecular biology, **53**(4), 674-690.

Rise, M., Cojocaru, M., Gottlieb, H.E., Goldschmidt, E.E. (1989) Accumulation of alpha-Tocopherol in Senescing Organs as Related to Chlorophyll Degradation. Plant physiology, **89**(4), 1028-1030.

Sattler, S.E., Cahoon, E.B., Coughlan, S.J., DellaPenna, D. (2003) Characterization of tocopherol cyclases from higher plants and cyanobacteria. Evolutionary

implications for tocopherol synthesis and function. Plant physiology, **132**(4), 2184-2195.

Sattler, S.E., Gilliland, L.U., Magallanes-Lundback, M., Pollard, M., DellaPenna, D. (2004) Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. The Plant cell, **16**(6), 1419-1432.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods, **9**(7), 676-682.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N., Weigel, D. (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. The Plant cell, **18**(5), 1121-1133.

Seaver, S.M., Gerdes, S., Frelin, O., Lerma-Ortiz, C., Bradbury, L.M., Zallot, R., Hasnain, G., Niehaus, T.D., El Yacoubi, B., Pasternak, S., Olson, R., Pusch, G., Overbeek, R., Stevens, R., de Crecy-Lagard, V., Ware, D., Hanson, A.D., Henry, C.S. (2014) High-throughput comparison, functional annotation, and metabolic modeling of plant genomes using the PlantSEED resource. Proceedings of the National Academy of Sciences of the United States of America, **111**(26), 9645-9650.

Smillie, R.M. and Hetherington, S.E. (1999) Photoabatement by anthocyanin shields photosynthetic systems from light stress. Photosynthetica, **36**(3), 451-463.

Soll, J., Kemmerling, M., Schultz, G. (1980) Tocopherol and plastoquinone synthesis in spinach chloroplasts subfractions. Arch Biochem Biophys, **204**(2), 544-550.

Tian, L., DellaPenna, D., Dixon, R.A. (2007) The pds2 mutation is a lesion in the Arabidopsis homogentisate solanesyltransferase gene involved in plastoquinone biosynthesis. Planta, **226**(4), 1067-1073.

Valentin, H.E., Lincoln, K., Moshiri, F., Jensen, P.K., Qi, Q., Venkatesh, T.V., Karunanandaa, B., Baszis, S.R., Norris, S.R., Savidge, B., Gruys, K.J., Last, R.L. (2006) The Arabidopsis vitamin E pathway gene5-1 mutant reveals a critical role for phytol kinase in seed tocopherol biosynthesis. The Plant cell, **18**(1), 212-224.

Venkatesh, T.V., Karunanandaa, B., Free, D.L., Rottnek, J.M., Baszis, S.R., Valentin, H.E. (2006) Identification and characterization of an Arabidopsis homogentisate phytyltransferase paralog. Planta, **223**(6), 1134-1144.

vom Dorp, K., Holzl, G., Plohmann, C., Eisenhut, M., Abraham, M., Weber, A.P., Hanson, A.D., Dörmann, P. (2015) Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis. The Plant cell.

Willstätter, R. and Stoll, A. (1913) Die Wirkungen der Chlorophyllase. In Untersuchungen über Chlorophyll, Springer Berlin Heidelberg: pp 172-187.

Supplementary Data

VTE6 co-expressed genes

Supplementary Data 1 **List of *VTE6* co-expressed genes.** The list is depicted as generated and ranked by the atted-II platform with genes' locus and function (if available)

Locus	*Function*
At5g04260	WCRKC thioredoxin 2
At5g19980	golgi nucleotide sugar transporter 4
At2g40600	appr-1-p processing enzyme family protein
At1g03475	Coproporphyrinogen III oxidase
At2g16510	ATPase, F0/V0 complex, subunit C protein
At1g50170	sirohydrochlorin ferrochelatase B
At5g64290	dicarboxylate transport 2.1
At3g04480	endoribonucleases
At2g15290	translocon at inner membrane of chloroplasts 21
At1g28140	
At1g18600	RHOMBOID-like protein 12
At1g11200	Protein of unknown function (DUF300)
At1g16560	Per1-like family protein
At3g25540	TRAM, LAG1 and CLN8 (TLC) lipid-sensing domain containing protein
At4g36810	geranylgeranyl pyrophosphate synthase 1
At5g16660	
At1g63970	isoprenoid F
At1g10830	15-cis-zeta-carotene isomerase
At2g36360	Galactose oxidase/kelch repeat superfamily protein
At1g33780	Protein of unknown function (DUF179)
At4g35360	Uncharacterised conserved protein (UCP030210)
At1g12840	vacuolar ATP synthase subunit C (VATC) / V-ATPase C subunit / vacuolar proton pump C subunit (DET3)
At4g27390	
At2g02590	
At1g80510	Transmembrane amino acid transporter family protein
At1g02475	Polyketide cyclase/dehydrase and lipid transport superfamily protein
At5g45390	CLP protease P4
At5g52540	Protein of unknown function (DUF819)
At3g01360	Family of unknown function (DUF716)
At3g49390	CTC-interacting domain 10
At1g32070	nuclear shuttle interacting
At2g41090	Calcium-binding EF-hand family protein
At4g38460	geranylgeranyl reductase
At2g20270	Thioredoxin superfamily protein
At5g10870	chorismate mutase 2
At3g19490	sodium:hydrogen antiporter 1
At5g02410	DIE2/ALG10 family
At2g37920	copper ion transmembrane transporters

At2g44870	
At4g14400	ankyrin repeat family protein
At4g18230	
At5g19630	alpha/beta-Hydrolases superfamily protein
At2g40490	Uroporphyrinogen decarboxylase
At1g74640	alpha/beta-Hydrolases superfamily protein
At2g43090	Aconitase/3-isopropylmalate dehydratase protein
At1g29390	cold regulated 314 thylakoid membrane 2
At3g55250	
At3g50920	Phosphatidic acid phosphatase (PAP2) family protein
At4g26070	MAP kinase/ ERK kinase 1
At2g25610	ATPase, F0/V0 complex, subunit C protein
At4g24460	CRT (chloroquine-resistance transporter)-like transporter 2
At2g14560	Protein of unknown function (DUF567)
At1g29040	
At2g39080	NAD(P)-binding Rossmann-fold superfamily protein
At2g39840	type one serine/threonine protein phosphatase 4
At3g58730	vacuolar ATP synthase subunit D (VATD) / V-ATPase D subunit / vacuolar proton pump D subunit (VATPD)
At1g05940	cationic amino acid transporter 9
At3g11945	homogentisate prenyltransferase
At5g65840	Thioredoxin superfamily protein
At3g32930	
At5g58560	Phosphatidate cytidylyltransferase family protein
At3g15360	thioredoxin M-type 4
At5g17560	BolA-like family protein
At1g03430	histidine-containing phosphotransfer factor 5
At3g06050	peroxiredoxin IIF
At3g04970	DHHC-type zinc finger family protein
At5g43280	delta(3,5),delta(2,4)-dienoyl-CoA isomerase 1
At2g18950	homogentisate phytyltransferase 1
At1g56190	Phosphoglycerate kinase family protein
At5g12130	integral membrane TerC family protein
At5g08410	ferredoxin/thioredoxin reductase subunit A (variable subunit) 2
At5g60750	CAAX amino terminal protease family protein
At2g02400	NAD(P)-binding Rossmann-fold superfamily protein
At5g06130	chaperone protein dnaJ-related
At4g34480	O-Glycosyl hydrolases family 17 protein
At5g52110	Protein of unknown function (DUF2930)
At1g68000	phosphatidylinositol synthase 1
At5g59500	protein C-terminal S-isoprenylcysteine carboxyl O-methyltransferases
At3g19460	Reticulon family protein
At3g22231	pathogen and circadian controlled 1
At1g22700	Tetratricopeptide repeat (TPR)-like superfamily protein
At1g03680	thioredoxin M-type 1

At3g44620	protein tyrosine phosphatases;protein tyrosine phosphatases
At3g57280	Transmembrane proteins 14C
At1g74440	Protein of unknown function (DUF962)
At4g22830	
At5g11810	
At5g19250	Glycoprotein membrane precursor GPI-anchored
At2g32520	alpha/beta-Hydrolases superfamily protein
At5g38660	acclimation of photosynthesis to environment
At4g34720	ATPase, F0/V0 complex, subunit C protein
At4g33350	Tic22-like family protein
At5g37480	
At5g54610	ankyrin
At1g27930	Protein of unknown function (DUF579)
At5g03350	Legume lectin family protein
At4g10000	Thioredoxin family protein
At4g36750	Quinone reductase family protein
At4g09650	ATP synthase delta-subunit gene
At1g55910	zinc transporter 11 precursor
At1g70330	equilibrative nucleotide transporter 1
At2g39290	phosphatidylglycerolphosphate synthase 1
At1g10500	chloroplast-localized ISCA-like protein
At5g21070	
At2g04700	ferredoxin thioredoxin reductase catalytic beta chain family protein
At1g66670	CLP protease proteolytic subunit 3
At5g08280	hydroxymethylbilane synthase
At3g06470	GNS1/SUR4 membrane protein family
At3g25040	endoplasmic reticulum retention defective 2B
At2g46170	Reticulon family protein
At4g32915	
At5g27830	
At4g38250	Transmembrane amino acid transporter family protein
At4g26550	Got1/Sft2-like vescicle transport protein family
At5g39250	F-box family protein
At5g43830	Aluminium induced protein with YGL and LRDR motifs
At1g30910	Molybdenum cofactor sulfurase family protein
At1g27385	
At2g25310	Protein of unknown function (DUF2012)
At3g23400	Plastid-lipid associated protein PAP / fibrillin family protein
At4g29120	6-phosphogluconate dehydrogenase family protein
At3g29350	histidine-containing phosphotransmitter 2
At5g67560	ADP-ribosylation factor-like A1D
At2g01110	Sec-independent periplasmic protein translocase
At4g01940	NFU domain protein 1
At4g31530	NAD(P)-binding Rossmann-fold superfamily protein
At1g50320	thioredoxin X

At2g20890	photosystem II reaction center PSB29 protein
At1g21270	wall-associated kinase 2
At2g16660	Major facilitator superfamily protein
At1g56500	haloacid dehalogenase-like hydrolase family protein
At4g33460	ABC transporter family protein
At1g15140	FAD/NAD(P)-binding oxidoreductase
At1g79870	D-isomer specific 2-hydroxyacid dehydrogenase family protein
At5g03080	Phosphatidic acid phosphatase (PAP2) family protein
At3g25805	
At5g18520	Lung seven transmembrane receptor family protein
At4g34950	Major facilitator superfamily protein
At3g26710	cofactor assembly of complex C
At3g51260	20S proteasome alpha subunit PAD1
At5g58970	uncoupling protein 2
At2g45630	D-isomer specific 2-hydroxyacid dehydrogenase family protein
At1g49750	Leucine-rich repeat (LRR) family protein
At5g10760	Eukaryotic aspartyl protease family protein
At1g74520	HVA22 homologue A
At4g01900	GLNB1 homolog
At2g45010	PLAC8 family protein
At5g25050	Major facilitator superfamily protein
At4g28770	Tetraspanin family protein
At5g55710	
At3g58460	RHOMBOID-like protein 15
At1g02560	nuclear encoded CLP protease 5
At3g27890	NADPH:quinone oxidoreductase
At5g67590	NADH-ubiquinone oxidoreductase-related
At1g44920	
At1g21250	cell wall-associated kinase
At3g04790	Ribose 5-phosphate isomerase, type A protein
At4g25370	Double Clp-N motif protein
At3g43520	Transmembrane proteins 14C
At1g67280	Glyoxalase/Bleomycin resistance protein/Dioxygenase superfamily protein
At2g42130	Plastid-lipid associated protein PAP / fibrillin family protein
At3g02730	thioredoxin F-type 1
At3g09250	Nuclear transport factor 2 (NTF2) family protein
At1g71090	Auxin efflux carrier family protein
At1g63110	GPI transamidase subunit PIG-U
At1g16880	uridylyltransferase-related
At4g25130	peptide met sulfoxide reductase 4
At5g61540	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
At3g51520	diacylglycerol acyltransferase family
At2g45790	phosphomannomutase
At1g13750	Purple acid phosphatases superfamily protein
At4g31600	UDP-N-acetylglucosamine (UAA) transporter family

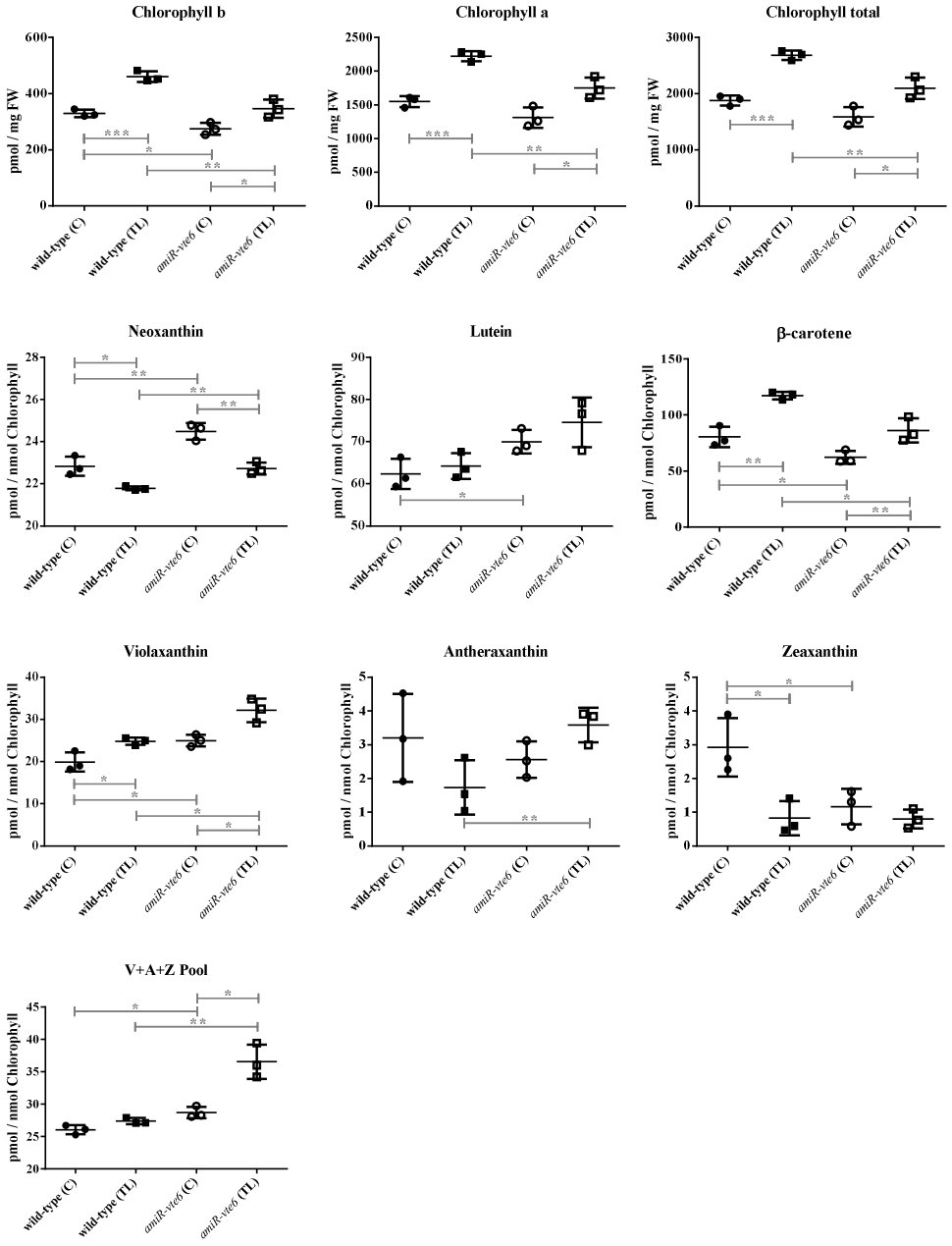
At2g47760	asparagine-linked glycosylation 3
At5g07370	inositol polyphosphate kinase 2 alpha
At1g09130	ATP-dependent caseinolytic (Clp) protease/crotonase family protein
At2g45740	peroxin 11D
At3g10060	FKBP-like peptidyl-prolyl cis-trans isomerase family protein
At3g04090	small and basic intrinsic protein 1A
At2g44210	Protein of Unknown Function (DUF239)
At3g22425	imidazoleglycerol-phosphate dehydratase
At2g38270	CAX-interacting protein 2
At1g14310	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
At2g27290	Protein of unknown function (DUF1279)
At4g26510	uridine kinase-like 4
At4g02530	chloroplast thylakoid lumen protein
At1g80600	HOPW1-1-interacting 1
At5g38630	cytochrome B561-1
At1g64850	Calcium-binding EF hand family protein
At4g34190	stress enhanced protein 1
At1g69390	homologue of bacterial MinE 1
At5g45550	Mob1/phocein family protein
At3g21220	MAP kinase 5
At3g15110	
At5g59250	Major facilitator superfamily protein
At3g60440	Phosphoglycerate mutase family protein
At1g24040	Acyl-CoA N-acyltransferases (NAT) superfamily protein
At1g12410	CLP protease proteolytic subunit 2
At3g12800	short-chain dehydrogenase-reductase B
At1g73655	FKBP-like peptidyl-prolyl cis-trans isomerase family protein
At1g43560	thioredoxin Y2
At5g19370	rhodanese-like domain-containing protein / PPIC-type PPIASE domain-containing protein
At5g40810	Cytochrome C1 family
At5g24210	alpha/beta-Hydrolases superfamily protein
At3g04210	Disease resistance protein (TIR-NBS class)
At3g07800	Thymidine kinase
At2g20690	lumazine-binding family protein
At5g08160	serine/threonine protein kinase 3
At1g31300	TRAM, LAG1 and CLN8 (TLC) lipid-sensing domain containing protein
At2g24860	DnaJ/Hsp40 cysteine-rich domain superfamily protein
At4g12060	Double Clp-N motif protein
At5g35100	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein
At1g20050	C-8,7 sterol isomerase
At2g25830	YebC-related
At3g54300	vesicle-associated membrane protein 727
At5g63870	serine/threonine phosphatase 7
At1g08640	Chloroplast J-like domain 1
At2g20930	SNARE-like superfamily protein

At1g09330	
At5g52970	thylakoid lumen 15.0 kDa protein
At3g56010	
At4g11010	nucleoside diphosphate kinase 3
At1g17610	Disease resistance protein (TIR-NBS class)
At1g32500	non-intrinsic ABC protein 6
At3g24160	putative type 1 membrane protein
At1g06470	Nucleotide/sugar transporter family protein
At1g22750	
At4g14870	secE/sec61-gamma protein transport protein
At3g27925	DegP protease 1
At1g25520	Uncharacterized protein family (UPF0016)
At4g28030	Acyl-CoA N-acyltransferases (NAT) superfamily protein
At1g62430	CDP-diacylglycerol synthase 1
At1g12990	beta-1,4-N-acetylglucosaminyltransferase family protein
At5g59300	ubiquitin carrier protein 7
At1g17120	cationic amino acid transporter 8
At3g58140	phenylalanyl-tRNA synthetase class IIc family protein
At4g08280	Thioredoxin superfamily protein
At3g45050	
At3g57240	beta-1,3-glucanase 3
At1g31130	
At1g19740	ATP-dependent protease La (LON) domain protein
At2g17033	pentatricopeptide (PPR) repeat-containing protein
At1g78560	Sodium Bile acid symporter family
At3g12780	phosphoglycerate kinase 1
At2g23940	Protein of unknown function (DUF788)
At5g55450	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
At5g55630	Outward rectifying potassium channel protein
At3g18060	transducin family protein / WD-40 repeat family protein
At1g32080	membrane protein, putative
At3g07950	rhomboid protein-related
At4g37040	methionine aminopeptidase 1D
At4g35000	ascorbate peroxidase 3
At4g32390	Nucleotide-sugar transporter family protein
At4g14800	20S proteasome beta subunit D2
At3g24760	Galactose oxidase/kelch repeat superfamily protein
At4g17570	GATA transcription factor 26
At3g18430	Calcium-binding EF-hand family protein
At5g12040	Nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase family protein
At2g26260	3beta-hydroxysteroid-dehydrogenase/decarboxylase isoform 2
At2g36300	Integral membrane Yip1 family protein
At2g26900	Sodium Bile acid symporter family
At2g13100	Major facilitator superfamily protein
At2g46820	photosystem I P subunit

At5g02040	prenylated RAB acceptor 1.A1
At4g25570	Cytochrome b561/ferric reductase transmembrane protein family
At4g09580	SNARE associated Golgi protein family
At1g22940	thiamin biosynthesis protein, putative
At5g16150	plastidic GLC translocator
At3g05000	Transport protein particle (TRAPP) component
At1g66970	SHV3-like 2
At3g13235	ubiquitin family protein
At3g46820	type one serine/threonine protein phosphatase 5
At5g41210	glutathione S-transferase THETA 1
At3g23600	alpha/beta-Hydrolases superfamily protein
At4g25840	glycerol-3-phosphatase 1
At4g32260	ATPase, F0 complex, subunit B/B', bacterial/chloroplast
At2g31440	
At3g52760	Integral membrane Yip1 family protein
At2g26670	Plant haem oxygenase (decyclizing) family protein
At2g04450	nudix hydrolase homolog 6
At2g37220	RNA-binding (RRM/RBD/RNP motifs) family protein
At3g13410	
At4g16060	
At1g80500	SNARE-like superfamily protein
At1g04690	potassium channel beta subunit 1
At3g55040	glutathione transferase lambda 2
At5g06290	2-cysteine peroxiredoxin B
At3g10920	manganese superoxide dismutase 1
At3g05280	Integral membrane Yip1 family protein
At2g35490	Plastid-lipid associated protein PAP / fibrillin family protein
At1g76960	
At5g11060	KNOTTED1-like homeobox gene 4
At4g12650	Endomembrane protein 70 protein family
At3g11630	Thioredoxin superfamily protein
At5g12860	dicarboxylate transporter 1
At2g32380	Transmembrane protein 97, predicted
At4g19185	nodulin MtN21 /EamA-like transporter family protein
At1g58080	ATP phosphoribosyl transferase 1
At2g23200	Protein kinase superfamily protein
At5g59770	Protein-tyrosine phosphatase-like, PTPLA
At5g42420	Nucleotide-sugar transporter family protein

Dissection of leaf material for pigment analysis

Earlier results suggested that the alteration in leaf pigmentation is more pronounced in emerging true leaves than in cotyledons (Figure 4). Therefore leaves of 14 day old seedlings (four leaves, with two true leaves, which were distinguishable from cotyledons) were separated to measure pigments for cotyledons and true leaves individually (Supplementary Figure 1).



Supplementary Figure 1 **Pigment quantification of *amiR-vte6* in isolated cotyledons and isolated true leaves after 14 days.** Pigments were extracted and analyzed via reverse-phase high-performance liquid chromatography and normalized by fresh weight in case of chlorophyll and normalized by total chlorophyll for all other pigments. Significant differences between the cotyledons (CO) and true leaves (TL) are marked with one, two, or three asterisks (Student's t-test, * = $p<0.05$; ** = $p<0.01$; *** =

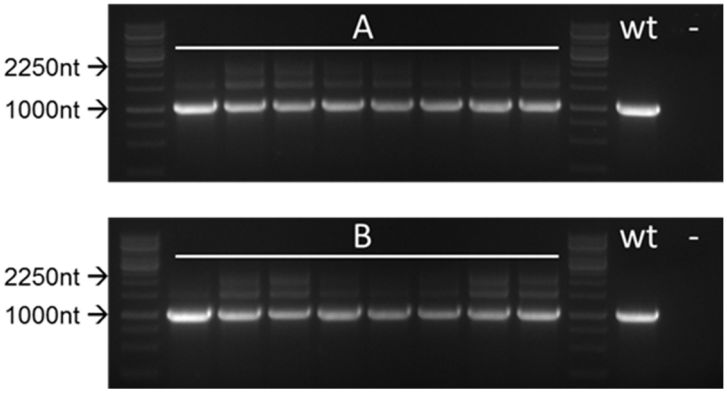
$p<0.001$). Three individual samples were tested per line and leave type. For each replicate, the leaves from 40 individual 14 day old plants were separated and analyzed.

Regarding the chlorophyll results, one can see that for both, wild-type control and *amiR-vte6* the content for chlorophyll a, chlorophyll b, and total chlorophyll differed between true leaves and cotyledons. Noticeably, if cotyledons of amiR-vte6 and the control were compared, no significant difference was obtained for chlorophyll a and total chlorophyll. However, the data showed a significant reduction in chlorophyll a and total chlorophyll in true leaves of the mutant.

This trend was not visible for neoxanthin, β-carotene, and violaxanthin. Regarding these pigments the alteration between mutant and control was not restricted exclusively to true leaves, as significant increases (neoxanthin and violaxanthin), respectively a significant decreases (β-carotene) was indicated for both tissues under investigation. Analysis of antheraxanthin and zeaxanthin revealed no trend as the obtained results were diffuse with high variation between individual samples.

Generation of a sll0875 knockout mutant in *Synechocystis* was impossible

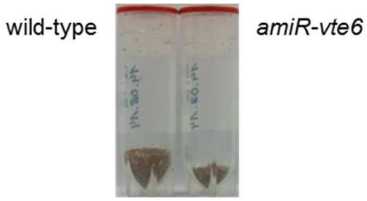
Generation of knockout lines for the *VTE6* homologue sll0875 in *Synechocystis* sp. PCC6803 was impossible to achieve because the knocked out allele did not fully segregate. Even under strong selective pressure through addition of twice the normal amount of kanamycin to the growth media, a wild-type allele persists within the genome. All tested lines exhibited a wild-type specific signal if screening for Kan^R insertion in sll0875 which was introduced to interrupt the coding gene sequence (Supplementary Figure 2).



Supplementary Figure 2 **Segregation analysis of *Synechocystis sp*. PCC6803 transformants.** Full-length sll0875 was amplified using PCR to verify if Kan^R cassette is inserted into the sll0875 coding region. Positive insertion resulted in a fragment of approximately 2250 bp, while the wild-type fragment without insertion is approximately 1000 bp long. **(A)** Kan^R insertion in the same orientation like the interrupted sll0875, **(B)** Kan^R insertion points in the opposite direction like sll0875 coding region. Intermediate bands between 1000 bp and 2250 bp are artificial and unspecific and occurred frequently in our analyses.

Seed yield for *amiR-vte6*

During routine propagation of new seed batches for *amiR-vte6* it was detected that the amount of seeds harvested from the mutant line was not reflecting the corresponding amount of a simultaneously grown wild-type *Arabidopsis* wild-type batch. Preliminary results indicated that the amount of seeds per plant might be reduced in comparison with the wild-type control (Supplementary Figure 3).

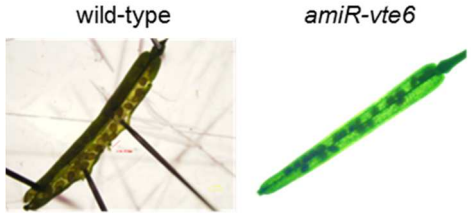


Supplementary Figure 3 **Seed yield of eight individual plants.** Eight plants for wild-type and *amiR-vte6* were grown simultaneously to harvest produced seeds. Seeds of all individual plants were harvested and pooled to compare the amount of seeds for both lines. Seeds of the wild-type batch are shown on the right, seeds of *amiR-vte6* are shown on the left side of the picture.

Since this could not be tested systematically due to time restrictions, a very rough approach was chosen to get at least an insight if it would be worth testing in a larger approach.

Siliques of simultaneously grown mutant lines and the wild-type were checked for their absolute seed content. This was either been done by opening the siliques or by

using intense light to shine through the siliques to visualize the single seeds. Supplementary Figure 4 shows a representative picture of several investigated siliques of both, the wild-type control and *amiR-vte6*.



Supplementary Figure 4 **Seed content in siliques of *amiR-vte6* and wild-type *Arabidopsis* plants.** Siliques of wild-type *Arabidopsis* were opened before they reached senescence to visualized seed content **(left)**. In case of *amiR-vte6*, siliques were placed below a binocular and light was applied from underneath to make siliques translucent. Seeds inside the silique are visible as dark spots **(right)**.

Investigation of the seed yield of *amiR-vte6* pointed towards reduced amounts or ripening seeds, as indicated by the analysis of single siliques (Supplementary Figure 4). While siliques of the wild-type control were tightly packed with seeds, *amiR-vte6* siliques harbored less seeds and were not as tightly packed as the wild-type control. Several free spots between single seeds were detected.

Authors' contribution to manuscript 3

C.P. performed all experiments, performed data analyses and wrote the manuscript

M.E. supervised experimental design, helped with data analyses, provided crucial information about the work with *Synechocystis*, and took part in isolation of leaf samples for pigment analysis.

P.J. Helped with experimental design. Pigment Analyses was exclusively performed in his laboratory

A.P.M.W supervised the project and helped in experimental design

CONCLUDING REMARKS

This PhD thesis started with the aim to discover new transport proteins, which are occupied in photorespiration. Deciphering the complete photorespiratory cycle in plants, is one prerequisite for modulating it in future approaches. This ultimately aims at improving agricultural yield. The reason, why an "improved" photorespiratory cycle can help in increasing plants' yield, can be found in the high energy consumption of the photorespiratory cycle, which is needed to recycle toxic compounds. Bypassing certain steps of this photorespiratory cycle can save energy, which can be used by the plant to improve its biomass.

Since forward genetic approaches were not sufficient to find these missing transport proteins, co-expression analysis was used to find putative targets for further work. These potential photorespiratory targets were not yet characterized as genes coding for transport proteins in the plant photorespiratory cycle. The first publication in this thesis showed, how the approach was carried out.

The genes, which were shown to be co-expressed with known photorespiratory genes, were investigated for their potential role in photorespiration. However, none of the tested genes could be shown to be occupied in photorespiration.

Still, Myouga and co-workers described a pale-green phenotype (Myouga et al., 2010) for the candidate gene At1g78620. In-depth characterization of an amiRNA line showed that a lack in At1g78620 expression leads to a pale-green phenotype like observed for photorespiratory mutants (Timm and Bauwe, 2013). This phenotype could not been proven to be a real photorespiratory phenotype as plants were unaffected by elevated CO_2 levels, which are well known to suppress photorespiratory phenotypes.

Ongoing work could show that At1g78620 is VTE6 the missing phytyl phosphate kinase (Vom Dorp et al., 2015), which is occupied in tocopherol synthesis.

The fact that At1g78620 is not occupied in photorespiration, emphasizes that results from co-expression analysis can aid in finding hypotheses about genes' function, but

never guarantee this function. The question arises, why genes of tocopherol synthesis cluster together with known photorespiratory genes during co-expression analysis.

A possible explanation for that is, that photorespiration and tocopherol synthesis are both interconnected through senescence in plants. While photorespiratory conditions can cause senescence (Zhu et al., 2005), tocopherol synthesis can be based on the chlorophyll degradation product phytol, which accumulates during senescence. Therefore, an intertwined expression of both, photorespiratory genes and genes of tocopherol synthesis, makes sense.

Still, this is a very basic explanation. Future work with *vte6* mutants has to aim at the characterization of this interaction between photorespiration and tocopherol synthesis. The generation of the amiRNA line *amiR-vte6* for this work, provides an essential tool to study this interconnection.

REFERENCES

Myouga, F., Akiyama, K., Motohashi, R., Kuromori, T., Ito, T., Iizumi, H., Ryusui, R., Sakurai, T., Shinozaki, K. (2010) The Chloroplast Function Database: a large-scale collection of Arabidopsis Ds/Spm- or T-DNA-tagged homozygous lines for nuclear-encoded chloroplast proteins, and their systematic phenotype analysis. The Plant journal : for cell and molecular biology, **61**(3), 529-542.

Timm, S. and Bauwe, H. (2013) The variety of photorespiratory phenotypes - employing the current status for future research directions on photorespiration. Plant Biol (Stuttg), **15**(4), 737-747.

Vom Dorp, K., Holzl, G., Plohmann, C., Eisenhut, M., Abraham, M., Weber, A.P., Hanson, A.D., Dormann, P. (2015) Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis. The Plant cell.

Zhu, Y.-R., Lü, X.-Y., Wang, S.-F., Wang, N.-N., Wang, Y. (2005) The relationship between photorespiration and senescence of S. polyrrhiza P143. Plant Science, **168**(6), 1625-1632.

ADDENDUM

All materials and methods used during the work presented in this thesis which are not described in detail in any of the manuscripts are described in the following part

Consumables and chemicals

All consumables used in this study were purchased from Sarstedt AG & Co (Nümbrecht, Germany), Carl Roth (Karlsruhe, Germany), and VWR (Darmstadt, Germany) or (if bought elsewhere) it is stated in the text.

Chemicals were ordered from Sigma-Aldrich (Steinheim, Germany) or Carl Roth (Karlsruhe, Germany), all exceptions are labeled with the corresponding manufacturer.

Bioinformatics

Co-expression analysis

Co-expression analysis was carried out on the atted-II platform (Obayashi et al., 2009; Obayashi and Kinoshita, 2010; Obayashi et al., 2007; Obayashi et al., 2011). Analysis of co-expressed genes for *VTE6* was performed the same way using the standard parameters like previously for the pilot experiment (Bordych et al., 2013).

Arabidopsis accession numbers, DNA and protein sequences

DNA and protein sequences used for cloning purposes or *in silico* experiments were taken from different online resources. In case of *Arabidopsis thaliana* sequences were obtained from the Aramemnon database. *Synechocystis sp.* PCC 6803 coding sequence for the *VTE6* homolog sll0875 was retrieved from the CyanoBase database (http://genome.microbedb.jp/cyanobase/Synechocystis) (Nakamura et al., 1998).

General methods

Cloning procedure

Standard cloning procedures were used to clone vector constructs for this work, either conventional cloning based on restriction digest or cloning based on homologous recombination utilizing the Gateway® system.

Oligonucleotide primers used for both cloning systems were ordered at Sigma-Aldrich (Steinheim, Germany).

Vector backbones

All vectors used as backbone for cloning procedures are listed in Supplementary Material and Methods 1 along with common features of the vector system.

Supplementary Material and Methods 1. **Vector backbones used within this thesis.** All vector backbones used within this thesis are described along with their function and resistance markers.

Vector	*Function/Features*	*Resistance(s)*	*Reference*
pJet 1.2	Subcloning	Amp^R	Fermentas
pDONR207	Subcloning	$Gent^R$	Invitrogen
pUT-Kan	Plant Expression; Ubiquitin10-Promoter	Kan^R	Prof. K. Schumacher, Universität Heidelberg
pUBN-YFP	n-terminal fusion of YFP for localization studies	$Spec^R$	(Grefen et al., 2010)
pRS300	Backbone amiRNA construction	Amp^R	(Schwab et al., 2006)

Escherichia coli strains

Different *Escherichia coli* strains were used in this thesis. DH5α, XL1-BLUE, and Mach1 cells were used for standard cloning procedures. Amplification of Gateway® basic vectors was performed in 2T1 cells. All *E. coli* strains along with their respective features and functions are consolidated in Supplementary Material and Methods 2.

Supplementary Material and Methods 2. ***Escherichia coli*** **strains.** All strains mentioned were used for cloning application (conventional via restriction digest or via homologues recombination/Gateway®)

Strain	*Genotype*	*Function*	*Reference*
DH5α	F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK-mK+), λ–	Cloning	Invitrogen/Life Technologies (Carlsbad, United States of America
XL1-BLUE	endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB+ lacIq Δ(lacZ)M15] hsdR17(rK- mK+)	Cloning	Stratagene (LaJolla, United States of America)
Mach1	ΔrecA1398 endA1 tonA Φ80ΔlacM15 ΔlacX74 hsdR(rK-mK+)	Cloning	Invitrogen/Life Technologies (Carlsbad, United States of America
One-shoot®-2T1^R	F-mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araΔ139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG fhuA::IS2	Gateway ® Cloning	Invitrogen/Life Technologies (Carlsbad, United States of America

Standard cloning PCR

Phusion polymerase (Thermo Scientific) based PCR was performed if amplified DNA fragments were applied to latter cloning approaches to decrease the probability of single point mutations. PCR sample composition and the corresponding PCR program is depicted below (Supplementary Material and Methods 3 and 4)

Supplementary Material and Methods 3. **Standard PCR scheme for Phusion PCR**. In most cases 5X Phusion HF buffer was used. In some rare cases when initial PCR amplification failed and the sequence to be amplified was considered to show a high GC content, an alternative buffer (5X GC buffer) was used instead.

Component	*Volume*
5x Phusion HF/GC buffer	10.0 µl
Phusion polymerase	0.5 µl
Primer 1 [10 µM]	2.5 µl
Primer 2 [10 µM]	2.5 µl
dNTPs (10 mM each]	1.0 µl
Template DNA	1.0 µl
H_2O	ad. 50.0 µl

Supplementary Material and Methods 4. **PCR program for phusion PCR.** Primer annealing temperature (X) depended on the oligonucleotide composition and was calculated using the calculation algorithm provided by Thermo Scientific (http://www.thermoscientificbio.com/webtools/tmc/)

Duration	*Temperature*
0'30"	98 °C
0'10"	98 °C
0'10"	X °C
0'15"/kb	72 °C
5'00"	72 °C
∞	20 °C

Ligation into vector backbone

Ligation was used to introduce DNA fragments into vector backbones (either for subcloning into pJet1.2 or into the final vector). All ligation steps were carried out according to the manual provided by the manufacturer (Thermo Scientific, CloneJET PCR Cloning Kit).

Standard restriction digest

Restriction digests performed for conventional cloning approaches in this thesis, were performed with buffers and restriction enzymes provided by New England Biolabs (NEB). 2 µg of vector DNA was taken up in 2.0 µl of suitable (10x) restriction buffer (depending on used restriction enzymes) and samples were filled up to 18.0 µl respectively 19.0 µl (depending on the number of restriction enzymes). After addition of 1.0 µl of desired restriction enzymes, samples were incubated at appropriate temperatures (mostly 37°C, in rare cases 25°C or 65 °C) for at least 2 hours. Agarose gel electrophoresis was performed with digested DNA samples to monitor cutting efficiency and/or to isolate DNA fragments resulting from the enzymatic digest.

E. coli transformation

Competent *E. coli* cells were transformed with desired vector DNA using the heat shock transformation protocol. Competent cells were thawed gently on ice and the DNA of choice (whole ligation sample or 0.1 µl of vector DNA) was added. Samples were gently mixed and incubated on ice for up to 30 minutes. In most cases incubation for 10 minutes was sufficient. Uptake of vector DNA by *E. coli* was initiated by heat

shock treatment. Therefore, samples were transferred to 42°C (water bath) and incubated for exactly 45 seconds. After heat-shock samples were placed back on ice for 2 minutes to regenerate. After addition of 500 µl LB medium (Supplementary Material and Methods 5) samples were incubated on a shaker at 37°C for at least 1 hour. Samples were centrifuged down (1500g for 3 minutes) to concentrate cells and discard excessive amounts of LB media. Cells were re-suspended in a small volume of fresh LB media and streaked out on LB plates containing appropriate antibiotics (Supplementary Material and Methods 6)

E. coli growth conditions

E. coli cells were either grown on solid LB media for isolation of single colonies or in liquid LB media (Supplementary Material and Methods 5) for amplification of single clones. Cells on LB plates were grown in a chamber running at 37 °C overnight (approximately 10-12 hours) in presence of appropriate antibiotics (Supplementary Material and Methods 6). Liquid cultures were started via inoculation of single colonies into sample tubes containing 5 ml liquid LB media in combination with desired antibiotics and grown overnight on a shaker at 100 rpm and 37°C.

Supplementary Material and Methods 5. **Standard LB media** for *E. coli* cultivation

Compound	*Amount*
Bacto-Tryptone	200 µg/ml
Yeast extract	25 µg/ml
Sodium chloride	30 µg/ml
Bacto Agar (for LB Plates)	100 µg/ml
H_2O	ad. 1000ml

Supplementary Material and Methods 6. **Antibiotics for *E. coli* selection.** Antibiotics were used in the following concentrations for selective growth of *E. coli*

Antibiotic	*Final concentration*
Ampicillin (Amp)	200 µg/ml
Gentamycin (Gent)	25 µg/ml
Kanamycin (Kan)	30 µg/ml
Spectinomycin (Spec)	100 µg/ml
Streptomycin (Strep)	300 µg/ml

Plasmid Isolation from *E. coli*

Plasmid DNA was isolated from *E. coli* liquid cultures utilizing a modified alkaline lysis protocol (Birnboim and Doly, 1979), which is described in the following paragraph. Roughly 3 ml *E. coli* suspension from overnight cultures were spun down in 1.5 ml sample tubes at maximum speed and room temperature in two consecutive rounds of centrifugation. Supernatant (LB-medium depleted of cells) was discarded in between centrifugation steps. Cells were re-suspended in 250 µl re-suspension buffer (25 mM Tris-HCl , 10mM EDTA, pH 8.0) and mixed with 1V of lysis buffer (0.2 N NaOH, 1.0% SDS). Mixture was neutralized by the addition of 350 µl neutralization buffer (5M KOAc, pH 4.8). Sample was mixed gently by inversion and set on ice for 5 minutes prior to centrifugation at maximum speed for 10 minutes. After centrifugation the supernatant was transferred into a fresh sample tube and covered with 1V 100% isopropanol and inverted several times. Samples were subsequently centrifuged for 5 minutes at maximum speed to precipitate nucleic acids. After removal of the supernatant, the DNA pellet was washed twice with 500 µl 70% (v/v) ethanol. Intermediate centrifugation steps were carried out at maximum speed for 3 minutes. After removal of ethanol, the precipitated DNA was air-dried in the sample tube bottom side-up for 10 minutes on a paper tissue and re-suspended in 50 µl Milli-Q water containing 20 µg/ml DNase-free RNase A.

In case vector DNA was used for ligation into pJet 2.1 or sent in for sequencing, vector DNA was further purified using the Promega Wizard® PCR and Gel Clean Up system which was used according to the manufacturer's manual.

Vector DNA was stored at -20°C until further use or at 4°C if used within the next day for further applications.

DNA Agarose Gel

Agarose gels for gel electrophoresis were used with agarose concentrations between 0.5% and 4.0% depending on the size of the DNA fragment which should be visualized or isolated. Gels up to 2% agarose concentration were produced using the following standard protocol.

The desired amount of agarose was mixed with the according volume of TAE-buffer (Supplementary Material and Methods 7). Agarose was melted in the microwave at maximum power for about two minutes while frequently mixing, until no agarose residues were visible in the gel. The gel was placed on a stirrer to cool down and 0.01% (v/v) ethidium bromide was added while stirring. Gels were poured into casting trays and fitted with a comb to yield the desired amount and size of pockets.

For agarose gels exceeding agarose concentrations of 2% agarose was pre-incubated with the appropriate volume of TAE buffer (Supplementary Material and Methods 7) for at least 30 minutes, while constantly stirring. The rest of the procedure was carried out according to the protocol described for standard concentrations of agarose.

Supplementary Material and Methods 7. **50x TAE buffer for agarose gel electrophoresis**

Component	*Amount/Volume*
TRIS-HCl	242.0 g
Glacial acetic acid	57.1 ml
0.5 M EDTA	100 ml
H_2O	ad. 1000 ml

Once the gel was solid, the comb was removed gently and gels were placed into agarose gel running chambers with appropriate volumes of 1x TAE running buffer. DNA samples were loaded into pockets along with a suitable DNA marker for size comparison in at least one additional pocket. Gels ran according to their size at 120-150V for 15-30 minutes. DNA separated in the agarose gel was visualized using UV light and documentation pictures were stored digitally.

In case distinct DNA bands (e.g. PCR products) should be separated by the gel for isolation, the gel was place on a UV table generating weak UV light to cut out the respective DNA band using a scalpel. Usage of weak UV lights reduced the possibility of generating point mutations in the PCR fragment.

Plants

Seed sterilization

Arabidopsis seeds were surface sterilized using chlorine gas. Seeds were split up in small amounts in reaction tubes (to ensure that all seeds are in contact with air) and placed with open lid in a desiccator. Chlorine gas was produces using 100 ml of sodium hypochlorite mixed with 3 ml hydrochloric acid (fuming, 37%). Desiccator was kept closed under gas production for 90 minutes. Remaining gas in the reaction tube was evaporated under the clean bench for at least 30 minutes afterwards.

Arabidopsis growth

Surface sterilized seeds were spotted on half-strength Murashige and Skoog media (MS, Supplementary Material and Methods 8) (Duchefa, Harleem, Netherlands) in petri dishes if not stated otherwise.

Supplementary Material and Methods 8. **Standard half-strength Murashige and Skoog (MS) media**

Component	*Amount/Volume*
MS basal salts	2.2g
0.5 M MES-KOH (pH7.5)	10 ml
Plant agar	8 g
H_2O	ad. 1000 ml

Petri dishes were sealed with patch-tape and incubated for 24 hours in darkness at 4°C for stratification. Afterwards plates were incubated in plant growth cabinets (CLF Plant Climatics, Wertingen, Germany) under standard conditions (12 hours light/darkness cycle, 22°C/18°C) either with ambient CO_2 conditions or elevated CO_2-supply (3000 ppm). 14 days after germination plants were transferred to soil if needed and further grown in growth cabinets under the same conditions as stated above.

Subcellular localization studies – Cloning reporter construct

Subcellular localization studies were performed in tobacco plants (*Nicotiana benthamiana*), which were transiently transformed with a reporter construct for At1g78620. At1g78620 (without stop-codon) was amplified with a gene-specific primer pair (Supplementary Material and Methods 9) carrying Gateway®

recombination sites based on Col-0 cDNA and recombined into pDONR207 via BP reaction. Sequence of At1g78620 was verified via sequencing and the correct sequence was transferred into pUBC-YFP to fuse At1g78620 (Supplementary Material and Methods 10) with the yellow fluorescent protein for localization studies via LR reaction.

Supplementary Material and Methods 9. **Oligonucleotides for fusion of At1g78620 with YFP**

Name	*5'-3' Sequence*	*Description*
CB157	AAAAAGCAGGCTTCGAAGGAGATAGAACCATGGCAACGATTTCGTCAACTC	At1g78620 forward Gateway
CB158	AGAAAGCTGGGTGCTTGACCCAGTTCTGGAGTATA	At1g78620(-Stop) reverse Gateway

Supplementary Material and Methods 10. **Vector for subcellular localization studies**

Name	*Vector backbone*	*Insertion*	*Description*
pCB5	pUBC-YFP	pUBQ10::At1g78620(-Stop)::YFP	YFP-Fusion

DNA isolation from *Arabidopsis* leaves

The following protocol was adapted from the Book "Arabidopsis - A Laboratory Manual" by D. Weigel and J. Glazebrook and used for quick extraction of DNA from *Arabidopsis* leaves which is sufficient for downstream PCR protocols (Weigel and Glazebrook, 2002).

One *Arabidopsis* leaf was harvested into 2 ml sample tubes and shock-frozen in liquid nitrogen. Material was ground with a glass bead in a bead-mill and subsequently mixed with 400 µl extraction buffer (Supplementary Material and Methods 11). Non-soluble parts were precipitated via centrifugation at full speed for 5 minutes at room temperature. The supernatant was subjected to a second identical centrifugation step to extract remaining non-soluble parts. The supernatant of this step was transferred into a fresh 1.5 ml sample tube and mixed with 1V isopropanol. DNA precipitation was facilitated via centrifugation at full speed for 5 minutes at room temperature. DNA pellet was washed twice with 1 ml ethanol (70%) and air-dried for 10 minutes. DNA

was re-suspended in 50 µl Milli-Q water and stored at -20°C until further processing in PCR reactions.

Supplementary Material and Methods 11. **Arabidopsis DNA Extraction buffer for quick DNA isolation**

Component	*Concentration*
Tris-HCl (pH 7.5)	200 mM
Sodium chloride	250 mM
EDTA	25 mM
SDS	0.5 %

Verification of T-DNA lines

Putative T-DNA lines were screened for presence of the T-DNA to verify if plants are wild type or hetero-, respectively homozygous for the T-DNA insertion. Via standard PCR (Supplementary Material and Methods 12 and 13) with gene-specific oligonucleotides (Supplementary Material and Methods 14) and T-DNA-specific oligonucleotides (Supplementary Material and Methods 15) presence of T-DNA in the desired location was verified according to the protocol suggested by the Salk Institute Genomic Analysis Laboratory (http://signal.salk.edu/tdnaprimers.2.html).

Supplementary Material and Methods 12. **PCR mix for standard T-DNA-screening reaction.** The term "self-made" describing the taq polymerase, indicates that taq polymerase used for screening purposes was synthesized in our lab by heterologous expression in *E. coli*

Component	*Volume*
5x polymerase buffer	4.0 µl
"self-made" Taq	0.5 µl
Primer 1 (gene-specific) [10 µM]	0.5 µl
Primer 2 (gene-/T-DNA-specific) [10 µM]	0.5 µl
dNTPs (10 mM each]	0.5 µl
Template DNA	1.0 µl
H_2O	ad. 20.0 µl

Supplementary Material and Methods 13. **Standard PCR program for T-DNA-screening PCR.** Denaturation, primer annealing and elongation phase ran for 35 cycles

Duration	*Temperature*
2'00"	94 °C
0'30"	94 °C
0'30"	58 °C
1'00"	72 °C
2'00"	72 °C
∞	20 °C

Supplementary Material and Methods 14. **Oligonucleotide pairs for T-DNA verification.** All gene-specific oligonucleotides used for T-DNA verification are described along with the corresponding gene locus and T-DNA line

Name	*T-DNA Line*	*T-DNA Locus*	*5'-3' Sequence*
CB213	SALK_147856	At1g14270	TAACGAAACTGTCTGCCATCC
CB214			GCATCTTCCTTGCTTGTCAAG
CB211	SAIL_1144G11	At1g22850	TCAACCCGTACAGGTAATTGC
CB212			TGCTTCAATCCAATCTTCGTC
CB126	SALK_103897	At1g28140	CTGTGAGAGATGCGAATAGGG
CB127			CTCTGCATTACGCAAATTTGG
CB128	SALK_136302	At1g28140	TTCTTGCTCGAGCTTTTCAAC
CB129			TCGATGGAACCGTTTACAAAG
CB207	SALK _020108	At1g44920	CGCCAACGTGAATATTGAGG
CB208			ACTTCCCATTGTCACCAACAC
CB200	GABI_478F05	At1g49380	ATGAGCAATTTGGGACAACAG
CB201			TTGTATTGGCCTTAACGTTGC
CB215	GABI_787F04	At1g61740	TGCCCACACAAGTTCTATGTG
CB216			AATCCATGGAAACCTCTTTGG
CB130	SALK_002072	At1g78620	AATACATGACCTTCGTGGCAG
CB131			ACAATATCACAGGGAACAGCG
CB132	SALK_075640	At1g78620	TCGTATCTTATTGCGGACCAC
CB133			CAGATTGCCAATCTTGGAGAG
CB134	GABI_108A10	At2g04360	ATAAGGGAAAAGAAGAGGCCC
CB135			TATTTGAAGATTGCGTACGGG
CB194	SALK_143171	At2g42770	GGAGTGAGGATGAGAAGAGCC
CB195			GCTTGACGAGGTTTGTCTCTG
CB204	SALK_014056	At2g42770	TTGGTACCACGCATAAGAACC
CB205			TTATTTATTAACCACCGCGCC
CB153	GABI_132G05	At3g02690	CGCTGCTTCTTCTTCCTCTTC
CB154			AACCATGACAGTGCCAATAGC
CB155	GABI_349G09	A13g02690	CCAGTTGCAATGAGGAGAAAG
CB156			CTTGTCTGCATGTTGGTGATG
CB181	SALK_030049	At3g26085	TTTGGTTTCAAAGCTTGTTGC
CB182			CTAGTGAAGCAGCAAACCGAC
CB183	SALK_113749	At3g48200	AATCTGTACTCCATTGCGGTG
CB184			TCCGCATGATCTTAGATACCG
CB185	SALK_052004	At3g48200	TCTCGCCTTACCTTCAATCTG
CB186			ATCTGTTTCAACCATGGCTTG

CB217	GABI_216E06	At3g50685	ATAATCCCTGGCTAAATTGCG
CB218			TCAGGCGGCTATATCTGAATG
CB136	SALK_131730	At3g61870	ACCAGATTTCATGTCCACTCG
CB137			CCCATCACAATTGGGTGTATC
CB138	SALK_131721	At3g61870	TACTTGTCGAACAACCGTTCC
CB139			GTTGGATCTGAAGCTCAGTCG
CB198	SALK_106932	At4g19450	GAACATTCTTTGTCGCTGCTC
CB199			TGGTCGTTTAATTCAGGCAAC
CB144	SALK_015620	At4g19450	TGATTCTGTGCTGCGTCATAC
CB145			ATTGGTACCATGCTGATCTCG
CB149	SALK_010361	At5g22330	TGCATCTTATCCCCAGAAAAG
CB150			GCTGAAGAAGAAGCTCCATTG
CB149	SALK_010362	At5g22330	TGCATCTTATCCCCAGAAAAG
CB150			GCTGAAGAAGAAGCTCCATTG
CB140	SALK_060595	At5g27290	CTTTCCAAGCAACTATTCCCC
CB141			TGGAATCTTTCCAAAGCAATG
CB142	SALK_077727	At5g27290	AGGAACGTGTTGTTCCTACCC
CB143			ACGGAATATGCTCATGAAACG
CB165	GABI_311D02	At5g59250	CTTTATGAAGGAAGAAGGCGG
CB166			AGAGACTCGAGTTGCATCAGC
CB179	SAIL_1253A02	At5g59250	TGGTCTTGAACAGTCACCTGC
CB180			CAAAGCTTTCAAATTAGGCCC
CB209	SALK_023292	At5g67370	TCAGGAACCGGTGATAACTTG
CB210			GAATCAGATCGAGCAATGCTC

Supplementary Material and Methods 15. **T-DNA specific oligonucleotides used in this work.** According to the T-DNA insertion line one of the mentioned oligonucleotides was used to amplify T-DNA specific sequences

Name		*5'-3' sequence*	*Purpose*
CB112	LBb1.3	ATTTTGCCGATTTCGGAAC	SALK
CB113	Lba1	TGGTTCACGTAGTGGGCCATCG	SALK
CB151	GABI-LB	ATATTGACCATCATACTCATTGC	GABI-KAT
CB152	GABI-RB	GTGGATTGATGTGATATCTCC	GABI-KAT
CB187	SAIL LB1	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC	SAIL

PCR reactions were subjected to an agarose gel electrophoresis with 1% Agarose in TAE buffer and 0.1% ethidium bromide if not stated otherwise. Homozygous T-DNA lines yield no positive reaction in wild type-specific PCR but in T-DNA-specific PCR (wild type lines: vice versa). In case of a heterozygous insertion of the T-DNA, both PCR reactions show a signal.

Transformation of *Agrobacterium tumefaciens* – Freeze-Thaw method

Agrobacterium tumefaciens (strain GV2260) was transformed using the freeze-thaw protocol (Weigel and Glazebrook, 2006). The protocol was modified slightly and is therefore described in the following.

GV2260 cells were gently thawed on ice. 5µl of the desired vector was added. Total amount of plasmid DNA used to transform could not be calculated because vector DNA was isolated using the alkaline lysis protocol without further cleanup by the Promega PCR clean Kit. Cells were incubated with vector DNA for 5 minutes on ice and subsequently frozen in liquid nitrogen. Samples were kept in liquid nitrogen for 5 minutes and then thawed at 37°C for additional 5 minutes. Cells were kept for 2 minutes on ice to allow for regeneration from temperature treatment and then supplemented with 1 ml of LB media. Samples were put in an incubation cabinet for agrobacteria running a shaker at 120 rpm and temperature of 30°C for at least 4 hours. Afterwards cells were concentrated via centrifugation (1500 rpm for 3 minutes) and excessive amounts of LB media were discarded. Cells were spread out on LB plates in two different concentrations (10% and 90% of total volume of cells) with the appropriate antibiotics (Supplementary Material and Methods 16) and kept at 30°C for at least 48 hours for selective growth.

Positive colonies were streaked out again on LB plates with the selective antibiotics for screening via colony PCR.

Supplementary Material and Methods 16. **Antibiotics for *A. tumefaciens* selection.** Antibiotics were used with the following concentrations for selective growth of *A. tumefaciens*

Antibiotic	*Final concentration*
Ampicillin (Amp)	200 µg/ml
Gentamycin (Gent)	25 µg/ml
Kanamycin (Kan)	50 µg/ml
Spectinomycin (Spec)	100 µg/ml
Rifampicin (Rif)	150 µg/ml

Plant transformation

A. thaliana Col-0 plants were transformed via *Agrobacterium tumefaciens* using the floral-dip method described in detail in the literature (Bernhardt et al., 2012). All transformation experiments were carried out using the *A. tumefa*ciens strain GV2260.

Transient transformation of *Nicotiana benthamiana*

Tobacco (*Nicotiana benthamiana*) was transiently transformed with reporter constructs via injection of agrobacteria carrying the respective vector construct. Three week old tobacco plants were injected with agrobacteria using a syringe without needle on the abaxial side of the lower leaves. Therefore single overnight colonies (ONC) of transformed and screened agrobacteria were transferred to liquid LB media with appropriate antibiotics and grown overnight at 30°C on a shaker. The next day, cells were harvested by centrifugation at 2000g and re-suspended in infiltration buffer (Supplementary Material and Methods 17) to an OD_{600} of 1.0.

Supplementary Material and Methods 17. **Infiltration buffer.** Buffer was freshly prepared from stocks right before use.

Component	
MES-KOH, pH 5.6	50 mM
KH_2PO_4	2 mM
Acetosyringone	0.2 mM
Glucose	0.5% (w/v)
NH_4Cl	0.1% (w/v)
$MgSO_4 \times 7\ H_2O$	0.03% (w/v)
KCl	0.015% (w/v)
$CaCl_2$	0.001% (w/v)
$FeSO_4 \times 7\ H_2O$	0.00025% (w/v)

Following inoculation tobacco plants were grown further at 12h day/night rhythm at 24°C for at least 24 hours until expression of transiently expressed constructs was verified by microscopy.

Root growth measurements

Seeds of *Arabidopsis* wild type plants and mutants were sterilized and spotted together on half-strength MS media on big, quadric plates in the upper fifth on one line. After stratification, plates were incubated vertical in a 12 hours day/night cycle. Root growth was monitored periodically and documented by scanning and stored digitally. Root elongation was measured using Fiji (Schindelin et al., 2012) “Axon Elongation” plugin.

Leaf Area

Leaf area was measured after transfer of plants from half-strength MS media to soil. Seeds were sterilized, spotted on plates and incubated under different CO_2 conditions in a 12 hours day/night cycle. Having first four true leaves emerged, plants were transferred to soil with one plant per pot and further grown under respective conditions. Plants were periodically photographed for documentation and measurements. Measurement of total leaf area per plant was analyzed using Fiji (Schindelin et al., 2012).

Quick Chlorophyll quantification

50 mg of ground and frozen plant material was mixed with 1 ml of acetone (100%) for 10 seconds on a vortexer. Mixture was incubated at room temperature for 10 minutes and subsequently centrifuged at full-speed for 3 minutes at room temperature. 400 µl of the supernatant were transferred into a new reaction tube and combined with 100µl of distilled water. Chlorophyll quantification was carried out in 96 well plates using 200 µl of the prepared chlorophyll solution. OD was measured at 646 nm, 663 nm, and 750 nm.

Total chlorophyll content [µg/ml] was calculated using the following formula:

$$Chlorophyll_{well} = 17.76 * (OD_{646} - OD_{750}) + 7.34 * (OD_{663} - OD_{750})$$

Calculation of chlorophyll content in the sample [µg/mg FW*] was derived using the following formula:

$$Chlorophyll_{total} = \frac{Chlorophyll_{well}}{FW * 0.8}$$

*FW = fresh-weight

RNA Extraction

Two different standard protocols were used according to the desired downstream utilization of isolated RNA. Both protocols are briefly described in the following paragraphs.

1. Citric acid Method

Quick RNA isolation for cDNA synthesis was performed following the protocol published by Oñate-Sánchez and Vicente-Carbajosa which was modified as described in the following paragraph (Onate-Sanchez and Vicente-Carbajosa, 2008).

Arabidopsis leaf tissue was frozen and ground in liquid nitrogen. 50 mg of frozen material was mixed with 600 µl cell lysis solution (Supplementary Material and Methods 18). Samples were kept at room temperature until ground material was thawed (approximately three minutes).

Samples were covered with 250 µl protein-DNA precipitation buffer (Supplementary Material and Methods 19), mixed and incubated on ice for at least 10 minutes. Samples were centrifuged at full speed and 4 °C. 500 µl of the supernatant was transferred into new reaction tubes and gently mixed with 1V isopropanol. After centrifugation at full speed and 4 °C the supernatant was discarded and the pellet was washed twice with 500 µl 70% ethanol. Pellets were dried in the reaction tube (upside-down) for approximately 10 minutes until ethanol evaporated. RNA was re-solubilized in 25 µl Milli-Q water and subjected to DNase treatment (RQ1 RNase-free DNase; Promega).

3 µl of 10x RQ1 buffer and 2 units of RQ1 RNase-free DNase were added to the RNA solution and incubated at 37 °C for 30 minutes. After DNase treatment RNA was precipitated again using ammonium acetate. Therefore, 30 µl of DNase-treated RNA was filled up to 100 µl with Milli-Q water and covered with 50 µl of ammonium acetate (7.5M) and 400 µl 100% ethanol. After centrifugation for 30 minutes supernatant was discarded and the pellet was washed twice as described before with 70% ethanol, dried and re-suspended in 25 µl Milli-Q water. Portions of the RNA samples were directly subjected to cDNA synthesis and aliquots of RNA were stored at -80 °C.

Supplementary Material and Methods 18. **Cell lysis solution**

Component	*Concentration/Volume*
Citric acid	132 mM
Sodium citrate	68 mM
EDTA	1 mM
N-laurosylsarcosine	2.0% (w/v)

Supplementary Material and Methods 19. Protein-DNA precipitation solution

Component	*Concentration/Volume*
Sodium chloride	4 M
Citric acid	32 mM
Sodium citrate	16 mM

2. RNase-All Method

The original RNase-All protocol (Chomczynski and Sacchi, 1987) was modified and performed as stated in the following paragraph.

Plant tissue was collected and frozen in liquid nitrogen. Samples were ground in liquid nitrogen to a fine powder and covered with 1 ml RNase ALL working solution (Supplementary Material and Methods 20 and 21) per 100 mg plant powder and stirred while still frozen. Once samples were thawed (approximately after 20 minutes), material was transferred into fresh sample tubes. Protein extraction from the sample was performed using 300 µl chloroform:isoamylalcohol (24:1). After addition, samples were vortexed and centrifuged at 10000g for 10 minutes at 4°C. The aqueous phase was transferred into fresh sample tubes and mixed with 700 µl phenol. Another portion of 300 µl chloroform:isoamylalcohol (24:1) was added, mixed and centrifuged like before. Upper phase after centrifugation was subjected to RNA precipitation.

RNA precipitation was performed using 1/20V 1N acetic acid while mixing gently by inverting the sample tube. After addition of 1V 100% ethanol, samples were mixed and kept at -20°C for 2 hours to allow for precipitation. Samples were centrifuged for 30 minutes at maximum speed at 4 °C. After centrifugation RNA was hardly visible as a transparent pellet. Pellet was washed twice with 1ml ethanol (80%) and subsequent centrifuged at full speed and 4 °C for 10 minutes. RNA was air-dried with the sample tube bottom side up for approximately 10 minutes and resolved in 500 µl Milli-Q water.

RNA was subjected to a second precipitation step. Therefore, samples were heated up to 50 °C for 5 minutes and centrifuged for 1 minute at full speed and 4 °C, to precipitate remaining contaminants. The supernatant was transferred into a fresh sample tube and mixed with 1V of 5M lithium chloride and incubated overnight at -20°C. After centrifugation at full speed and 4 °C for 10 minutes, the supernatant was discarded and the RNA pellet was washed twice according to the previous washing treatment. RNA was solubilized in 50 µl Milli-Q water at 50°C for ten minutes.

Supplementary Material and Methods 20. **RNase ALL Stock solution.** pH was adjusted to 7.0 with 1N NaOH

Component	*Concentration/Volume*
Guanidin isothiocyanate	4 M
Sodium acetate	25 mM
N-laurosylsarcosine	0.5% (w/v)

Supplementary Material and Methods 21. **RNase Working solution.** Working solution was freshly prepared from stock solutions prior to isolation

Component	*Volume*
RNase ALL Stock Solution	1V
Water-saturated phenol	1V
2-Mercaptoethanol	0.007V

RNA from both extraction protocols was quantified using the peqlab NanoDrop™ and stored at -80°C until further applications.

DNase treatment

RNA isolated by the RNase All method used for latter cDNA synthesis was further deployed of contaminating DNA through DNase treatment with RQ1 RNase-free DNase (Promega) according to the following protocol (Supplementary Material and Methods 22).

Supplementary Material and Methods 22. **Standard DNA digest of RNA samples**

Component	*Amount/Volume*
RNA	2 µg / max. 16.0 µl
RQ1 buffer (10x)	2.0 µl
RQ1 DNase (1U/µl)	2.0 µl
Milli-Q water	ad. 20.0 µl

Prepared samples were mixed gently and incubated at 37 °C for 30 minutes in a heating block. DNase treatment was stopped by addition of 1 µl of DNase Stop Solution and heating up to 65 °C for 10 minutes according to the manufacturer's manual. DNase-treated RNA was checked via RNA gel electrophoresis and directly subjected to cDNA synthesis afterwards or stored at -80 °C.

RNA gel electrophoresis

RNA isolated with protocols described above were subjected to gel electrophoresis using a 1.5 % formaldehyde RNA gel. 100 ml gel were produced according to the following protocol.

1.5 g agarose was boiled in 72 ml of Milli-Q water via microwaving and 10 ml of 10x MEN buffer (Supplementary Material and Methods 23) was added. Once the mixture was cooled down 18 ml of formaldehyde were added and the solution was stirred briefly and poured into gel casts.

5 µg of RNA were taken up in a total volume of 10 µl Milli-Q water and mixed with 10 µl of loading buffer (Supplementary Material and Methods 24). RNA in loading buffer was heated up to 60 °C for 10 minutes and subsequently cooled down on ice for at least 5 minutes. After the addition of 1 µl ethidium bromide with a total concentration of 1 ng/ml, samples were loaded onto the formaldehyde gel. Gels ran up to two hours at 60 V. RNA was visualized using UV light subsequently after the gel run.

Supplementary Material and Methods 23. **10x MEN buffer**

Component	*Concentration*
MOPS	2 M
Sodium acetate	500 mM
EDTA	50 mM

Supplementary Material and Methods 24. **RNA loading buffer composition for 10µl RNA samples**

Component	*Volume*
10x MEN buffer	1.0 µl
Formaldehyde (37%)	3.0 µl
Formamide (100%)	5.0 µl
Glycerol (60%)	1.0 µl

cDNA synthesis

cDNA synthesis was carried out using RNA isolated following the extraction protocols described earlier and treated with DNase. cDNA used for cloning procedures were synthesized on the basis of RNA isolated utilizing the quicker citric acid-based protocol. cDNA for semiquantitative RT-PCRs was generated using RNA which was isolated by the RNase ALL protocol, which is suitable for quantitative isolation of RNA. cDNA synthesis was performed using the SuperScript II reverse transcriptase manufactured by Invitrogen™ according to the following protocol and sample composition (Supplementary Material and Methods 25)

Supplementary Material and Methods 25. **cDNA synthesis basic sample preparation**

Component	*Volume*
DNase-treated RNA (~2µg)	12.0 µl
oligo-dT$_{12-18}$ Primer (500 µg/ml)	1.0 µl
dNTP-Mix (10mM each)	1.0 µl

All following steps containing temperature adjustments were performed in a PCR-cycler to ensure constant conditions. Samples were heated up to 65 °C for 5 minutes and cooled down on ice subsequently. After the addition of 4.0 µl of First-Strand-buffer (5x) and 2.0 µl DTT (0.1 M), the sample was gently mix and incubated at 42 °C for primer annealing. cDNA synthesis was started by the addition of 1.0 µl SuperScript II reverse transcriptase and further incubation at 42 °C for 1 hour. Reverse transcriptase was inactivated by heat (70 °C) for 15 minutes. cDNA was stored at -20 °C until further use.

Semiquantitative RT-PCR

Expression level of target mRNAs was quantified relative to the expression of the housekeeping gene actin. Two different PCR mixes were mixed to amplify actin coding mRNA and the mRNA of the gene of interest on the basis of cDNA. Standard PCRs were set up according to the following scheme (Supplementary Material and Methods 26 and 27). The total number of cycles (denaturation, primer annealing and extension) varied, because samples of each PCR batch were analyzed on an agarose gel after 25, 30, and 35 cycles to prevent oversaturation on the documentation gel which would have rendered quantification impossible.

Supplementary Material and Methods 26. **Pipetting scheme for RT-PCR**

Component	*Volume*
5x polymerase buffer	4.0 µl
GO-Taq Polymerase	0.05 µl
Primer 1 [10µM]	0.5 µl
Primer 2 [10µM]	0.5 µl
dNTPs (10mM each]	0.5 µl
cDNA (equal concentration per sample)	1.0 µl
H_2O	ad. 20.0 µl

Supplementary Material and Methods 27. **PCR program for RT-PCR**

Duration	*Temperature*
2'00"	94 °C
0'30"	94 °C
0'30"	58 °C
1'00"	72 °C
2'00"	72 °C
∞	20 °C

PCR samples were subjected to an agarose gel (1%) and visualized using UV-light.

Synechocystis growth conditions

Synechocystis PCC6803 was grown on BG11 media (Supplementary Material and Methods 28-30) on a shaker under constant light of 100 µE and 30 °C either under photoautotroph or heterotroph conditions with addition of 5 mM glucose to enable potential photosynthetic mutant lines to grow. *Synechocystis* was either cultivated in liquid BG11 media or grown on solid BG11 agar plates. Stock solutions were prepared and maintained at 4°C in darkness. Agar plates were produced using 2x BG11 media mixed with 2x Agar (3g/100 ml). Both, BG11 stocks and 2x agar was autoclaved prior to preparation of agar plates.

Supplementary Material and Methods 28. **BG11 media.**

Compound	*Amount/Volume*
MilliQ water	970.0 ml
Macronutrients (100x)	10.0 ml
Trace metals (1000x)	1.0 ml
1M Hepes (pH7.7, NaOH)	20.0 ml
175 mM K_2HPO_4	1.0 ml
Fe-NH_4-citrate (6 mg/ml)	1.0 ml
189 mM Na_2CO_3	1.0 ml
$Na_2S_2O_3$ x 5 H_2O	3.0 g

Supplementary Material and Methods 29. **Macronutrients (100x) stock solution.**

Compound	*Amount/Volume*
$NaNO_3$	74.8 g
$MgSO_4$ x 7 H_2O	3.75 g
$CaCl_2$ x 2 H_2O	1.8 g
Citric acid (monohydrate)	0.325 g
Na_2-EDTA	0.05 g
H_2O	ad. 500 ml

Supplementary Material and Methods 30. **Trace metal (1000x) stock solution.**

Compound	*Amount/Volume*
H_3BO_3	2.86 g
$MnCl_2$ x 4 H_2O	1.81 g
Na_2MoO_4 x 2 H_2O	0.391 g
$CuSO_4$ x 5 H_2O	0.079 g
$Co(NO_3)_2$ x 6 H_2O	0.05 g
H_2O	ad. 500 ml

Synechocystis Knockout construct cloning

Knockout mutants for *Synechocystis* were generated by the insertion of a Kanamycin ($Kan^{R)}$ cassette into the gene of interest by homologues recombination. Cloning procedure followed standard protocols utilizing restriction-digest based approaches. The gene of interest was amplified using gene-specific primers on basis of *Synechocystis* DNA (kindly provided by Dr. Marion Eisenhut) introducing restriction sites to the 5'- and the 3'-end of the amplified sequence. The PCR fragment was subjected to an agarose gel electrophoresis (1% agarose) and the respective band was isolated. Corresponding DNA was extracted with the Promega SV Wizard DNA extraction kit. The DNA fragment was ligated into pJET1.2, amplified and selected in *E. coli* Mach1 cells and the combined vector was isolated following the miniprep

isolation protocol. Presence of the insert and orientation was verified by restriction enzyme digest and sequence was checked by sequencing. Vector product was digested with a blunt-end restriction enzyme to insert Kan^R cassette via standard blunt-end ligation into the gene-coding DNA within the vector.

Two different constructs were selected having both orientations of the Kan^R cassette in the gene of interest. *Synechocystis* was transformed with the vector product as described in the following paragraph.

Transformation of *Synechocystis* PCC6803

Synechocystis PCC6803 wild type strain or mutant line were grown in liquid BG11 under standard conditions. Prior to transformation, 2 ml cell suspension were harvested, centrifuged and re-suspended in 200 µl of fresh BG11 media. 2 µl of the isolated vector was added to the cell suspension and cells were incubated overnight on the shaker in darkness under standard temperature conditions. The following day cells were streaked out on BG11 media without additives and incubated in the shaker. Plates were covered with tissue to prevent excessive amounts of light, until the next day. Once cells were acclimatized to the light conditions, the cover was removed. Cells grew up to 5 days under normal conditions until green colonies emerged. BG11-agar was gently lifted in the plate and 1/5 concentration of kanamycin (10 µg/ml) was added underneath the agar and distributed by slight rotations. Once kanamycin was added, the agar was gently put back into the plate without introducing any air bubble underneath the agar, because this would have caused spots on the plate which are not in contact with kanamycin and finally led to false-positive colonies.

Plates were sealed and further incubated under standard conditions. Positive colonies stayed green, since they were resistant to kanamycin and could be transferred onto another BG11 plate harboring 1/5 of the final kanamycin concentration. Single colonies were taken up from the plate using custom-cut sterile paper pieces and injected into the fresh BG11 agar. After transfer, plates were once again covered with tissue and incubated under standard conditions until the next morning. The next day the tissue cover was removed and cells were grown until they reached the surface of the agar and formed dense cell formations on the agar surface. Colonies which were able to grow efficiently on 1/5 concentration of kanamycin were once again transferred

to fresh BG11 media with the final concentration of kanamycin (50µg/ml) and injected into the agar. The first night plates were covered and subsequently put under constant light standard conditions. All lines which tolerated the increased amount of kanamycin were streaked out on BG11 with 1/1 concentration of kanamycin and grown for 12 hours under reduced light, followed by standard conditions. In case selective pressure should have been increased, cells were transferred once again to plates containing twice the amount of kanamycin (100µg/ml).

Once cells were green and dense enough, a sample was taken for DNA isolation to verify the genotype.

DNA isolation *Synechocystis*

Synechocystis cells for DNA isolation were streaked out on BG11 and grown for 2 days or until cells were dense and green. Small amounts of cell material were re-suspended in 200 µl $T_{10}E_1$ buffer. Samples were mixed with 200 µl water-saturated phenol and heated up to 65 °C for 10 minutes. Samples were mixed every 2 minutes while being heated up. After heating, samples were cooled down on ice to reach room temperature. Cooled samples were mixed with 200 µl chloroform and centrifuged at full speed and 4 °C for 5 minutes. After centrifugation the upper aqueous phase was transferred into a clean reaction tube and washed twice with 200 µl chloroform followed by centrifugation at full speed and 4°C for 5 minutes. DNA solutions were kept on ice if subsequently subjected to further PCR analysis or stored at -20 °C.

***Synechocystis* genotyping**

Potential mutant lines of *Synechocystis* were screened via PCR analysis. Therefore the gene under investigation was amplified using gene-specific oligonucleotides, yielding full-length DNA. Standard PCR mix (Supplementary Material and Methods 12) was used in combination with a modified standard PCR program (Supplementary Material and Methods 31) with an increased elongation time to allow for amplification of the gene of interest with the Kan^R insertion. PCR program ran 35 cycles to guarantee dense signals after separation in a gel electrophoresis.

Supplementary Material and Methods 31. **Standard PCR program**. PCR program for mutant screening of *Synechocystis* PCC 6803 with Kan^R cassette insertion. Program ran 35 cycles.

Duration	*Temperature*
2'00"	94 °C
0'30"	94 °C
0'30"	58 °C
2'30"	72 °C
2'00"	72 °C
∞	20 °C

Wild type samples were included in the analysis to have proper negative controls available. In case of wild type samples, it was expected to see a band corresponding to the size of the gene of interest. In case of fully segregated mutated lines the size of the corresponding PCR fragment was shifted up by approximately 1.3 kb which resembles the size of the Kan^R cassette. Non-fully segregating lines were expected to show an intermediate patterning, yielding both fragments on the agarose gel.

However, under the conditions applied in this thesis, non-fully segregating lines showed an unexpected patterning with three bands. Both, distinct signals for the mutated version and the wild type version of the gene of interest were detectable in agarose gel electrophoresis, along with a third, distinct signal in between both formerly mentioned signals. This third band has neither been characterized by sequencing yet, nor can a suitable explanation been given. However, this phenomenon was observed several times in the lab during independent experiments (personal communication, Dr. Marion Eisenhut, University of Düsseldorf).

REFERENCES

Bernhardt, K., Vigelius, S.K., Wiese, J., Linka, N., Weber, A.P. (2012) Agrobacterium-mediated Arabidopsis thaliana transformation: an overview of T-DNA binary vectors, floral dip and screening for homozygous lines. Endocytobiosis & Cell Research, **22**.

Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic acids research, **7**(6), 1513-1523.

Bordych, C., Eisenhut, M., Pick, T.R., Kuelahoglu, C., Weber, A.P. (2013) Co-expression analysis as tool for the discovery of transport proteins in photorespiration. Plant Biol (Stuttg), **15**(4), 686-693.

Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem, **162**(1), 156-159.

Grefen, C., Donald, N., Hashimoto, K., Kudla, J., Schumacher, K., Blatt, M.R. (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. The Plant journal : for cell and molecular biology, **64**(2), 355-365.

Nakamura, Y., Kaneko, T., Hirosawa, M., Miyajima, N., Tabata, S. (1998) CyanoBase, a www database containing the complete nucleotide sequence of the genome of Synechocystis sp. strain PCC6803. Nucleic acids research, **26**(1), 63-67.

Obayashi, T., Kinoshita, K., Nakai, K., Shibaoka, M., Hayashi, S., Saeki, M., Shibata, D., Saito, K., Ohta, H. (2007) ATTED-II: a database of co-expressed genes and cis elements for identifying co-regulated gene groups in Arabidopsis. Nucleic acids research, **35**(Database issue), D863-869.

Obayashi, T., Hayashi, S., Saeki, M., Ohta, H., Kinoshita, K. (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic acids research, **37**(Database issue), D987-991.

Obayashi, T. and Kinoshita, K. (2010) Coexpression landscape in ATTED-II: usage of gene list and gene network for various types of pathways. Journal of plant research, **123**(3), 311-319.

Obayashi, T., Nishida, K., Kasahara, K., Kinoshita, K. (2011) ATTED-II updates: condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. Plant & cell physiology, **52**(2), 213-219.

Onate-Sanchez, L. and Vicente-Carbajosa, J. (2008) DNA-free RNA isolation protocols for Arabidopsis thaliana, including seeds and siliques. BMC Res Notes, **1**, 93.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods, **9**(7), 676-682.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N., Weigel, D. (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. The Plant cell, **18**(5), 1121-1133.

Weigel, D. and Glazebrook, J. (2002) Arabidopsis. A Laboratory Manual, **165**.

Weigel, D. and Glazebrook, J. (2006) Transformation of agrobacterium using the freeze-thaw method. CSH Protoc, **2006**(7).

PUBLICATIONS

The following manuscripts were published while working for the thesis presented here:

Bordych, C., Eisenhut, M., Pick, T.R., Kuelahoglu, C., Weber, A.P. (2013) Co-expression analysis as tool for the discovery of transport proteins in photorespiration. Plant Biol (Stuttg), **15**(4), 686-693.

Eisenhut, M., Pick, T.R., Bordych, C., Weber, A.P. (2013) Towards closing the remaining gaps in photorespiration--the essential but unexplored role of transport proteins. Plant Biol (Stuttg), **15**(4), 676-685.

vom Dorp, K., Holzl, G., Plohmann, C., Eisenhut, M., Abraham, M., Weber, A.P., Hanson, A.D., Dörmann, P. (2015) Remobilization of Phytol from Chlorophyll Degradation Is Essential for Tocopherol Synthesis and Growth of Arabidopsis. The Plant cell.

ACKNOWLEDGEMENTS

Prof. Dr. Andreas Weber für die Möglichkeit meine Doktorarbeit in seinem Labor anzufertigen, stetige Diskussionsbereitschaft, Hilfsbereitschaft und Betreuung und zuletzt natürlich auch für die Finanzierung

Apl. Prof. Dr. Peter Jahns für die unkomplizierte und schnelle Übernahme des Korreferats als zweiter Gutachter, für Hilfe bei der Planung meiner Experimente und die Möglichkeit, dass ich Experimente in seinem Labor durchführen durfte

Dr. Marion Eisenhut für die Betreuung meiner Arbeit und Hilfestellung wann immer nötig, zwei zusätzliche Hände als ich nur eine hatte, einen Intensivkurs in „Cyanos“ und viele Oster- und Weihnachtsüberraschungen.

Dr. Marc Linka für die Betreuung in der ersten ganz heißen Phase

Dr. Nicole Linka für ein stetig offenes Ohr, egal worum es ging und für Schreibtisch-Asyl am Ende

Dr. Andrea Bräutigam für die Bereitschaft immer wieder in Statistikfragen auszuhelfen, BBQ-Anregungen und die stetige Erinnerung an die Tikki-Bar

Dr. Thea Pick für Hilfe und ein offenes Ohr, Organisation bei allem was im Labor anfällt, Schreibtisch-Asyl am Anfang und am Ende

Dr. Tabea Mettler-Altmann für Einsatzbereitschaft, als ich noch mal eben Phytol messen wollte

Dr. Fabio Facchinelli für Diskussion, All-Ins und Kaffeemahlen

Samantha Kurz für eine helfende Hand wann immer ich eine benötigt habe, Ausflüge zum schwarzen Loch, Leiter halten, wenn die Hose nicht hält, Mittagessen und viel Spaß

Nadine Hocken für ein zusätzliches Ohr im R-Kurs im Tausch gegen meine Augen, Kaffee, Schokowaffeln und Käse, Ermunterung und eine Menge Spaß

Lisa Leson für S1 auch samstags - nur der FC!

Katrin Weber für Wasser, Milch, Hund, neue Schimpfwortkombinationen und Hilfe bei der Massenspektrometrie

Elisabeth Klemp für geduldige Hilfe bei der Auswertung der Massenspektrometrie

Maria Graf für Lachen in zehn Oktaven und Pigmentanalysen

Allen **Autoren** und **Co-Autoren**, die an der Anfertigung der Publikationen beteiligt waren

Allen **Gärtnern** für die Betreuung der Pflanzen

Alle, die ich bis hierher vergessen habe

Meinen Eltern für stetiges Interesse an meiner Arbeit auch wenn sie immer noch kaum bis gar nichts verstehen

Zu allerletzt **Katharina** für die Liebe, das Rücken-frei-halten, stetiges Interesse an meinem Laboralltag und überhaupt