

The evolution of  
B-family GATA transcription factors  
in the plant lineage

Peter Michael Schröder

Vollständiger Abdruck der von der TUM School of Life Sciences der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Ralph Hückelhoven

Prüfer der Dissertation:

1. Prof. Dr. Claus Schwechheimer
2. Prof. Dr. Kay H. Schneitz

Die Dissertation wurde am 02.06.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Life Sciences am 13.07.2023 angenommen.

## Abstract

GATAs are highly conserved transcription factors in animals, plants, and fungi. Their distinguishing feature is a zinc-finger DNA-binding domain that recognizes the core sequence W-G-A-T-A-R. Plant GATAs can be subdivided into the four classes A – D based on the sequence of their zinc-finger domains. Members of class B can be subdivided based on the existence of two domains of unknown biochemical functions, the HAN- and LLM-domains. Previous research, mainly from our group, has identified the role of LLM-domain B-GATAs in the regulation of many physiological processes in *Arabidopsis thaliana* e.g., chlorophyll synthesis, chloroplast division and stomatal development. While angiosperms only encode B-GATAs with one of the two domains, B-GATAs from the bryophytes *Marchantia polymorpha* and *Physcomitrium patens* contain both HAN- and LLM-domains.

The aim of my project was to investigate the evolution of B-GATAs in the plant lineage. To this end, I performed comprehensive phylogenetic analyses, suggesting that the HAN-domain was acquired by plants during terrestrialisation. In order to analyse the functions of B-GATAs in *M. polymorpha* and *P. patens*, I generated *gata* mutants in both species through CRISPR/Cas9-mediated genome editing. Since previous efforts to generate an *Arabidopsis thaliana gata hexuple* mutant lacking all six LLM B-GATAs failed, I also generated this mutant through CRISPR/Cas9-mediated genome editing.

I characterized the *M. polymorpha GATA* mutants and found a constitutive high-light stress response. Subsequent transcriptomic analyses confirmed this phenotype in the new *A. thaliana gata hex* mutant. Based on these results, I concluded that B-GATAs have a protective role in high-light stress and that this function might be evolutionarily conserved.

## Zusammenfassung

GATAs sind hochkonservierte Transkriptionsfaktoren in Tieren, Pflanzen und Pilzen. Ihr charakteristisches Merkmal ist eine Zinkfinger DNA-Bindedomäne, die die Kernsequenz W-G-A-T-A-R erkennt. GATAs aus Pflanzen können basierend auf der Sequenz ihrer Zinkfinger DNA-Bindedomäne in die vier Klassen A – D eingeteilt werden. Mitglieder der Klasse B können zudem durch das Auftreten zweier Domänen mit unbekanntem biochemischen Funktionen, der HAN- und der LLM-Domäne, weiter unterteilt werden. Vorgegangene Forschung, hauptsächlich aus unserem Labor, hat die Rolle der LLM-Domänen B-GATAs in *Arabidopsis thaliana* bei der Regulation vieler physiologischer Prozesse identifiziert, z.B. der Chlorophyllsynthese, der Chloroplastenteilung und der Entwicklung von Stomata. Während Angiospermen nur B-GATAs mit je einer der beiden Domänen kodieren, enthalten B-GATAs der Bryophyten *Marchantia polymorpha* und *Physcomitrium patens* sowohl HAN- als auch LLM-Domänen.

Das Ziel meines Projekts war es, die Evolution der B-GATAs in der Entwicklungsgeschichte der Pflanzen zu untersuchen. Zu diesem Zweck führte ich umfassende phylogenetische Untersuchungen durch, die nahelegen, dass Pflanzen die HAN-Domäne während der Terrestrialisierung gewannen. Um die Funktionen der B-GATAs aus *M. polymorpha* und *P. patens* zu analysieren, erzeugte ich *gata* Mutanten in beiden Spezies durch CRISPR/Cas9-vermittelte Genomeditierung. Da vorherige Versuche, eine *A. thaliana gata hexuple* Mutante zu erzeugen, bei der alle sechs LLM B-GATAs mutiert sind, gescheitert waren, erzeugte ich erstmals erfolgreich auch diese Mutante durch CRISPR/Cas9-vermittelte Genomeditierung.

Ich charakterisierte die *M. polymorpha GATA* Mutanten und fand eine konstitutive Starklichtantwort. Anschließende transkriptomische Analysen bestätigten diesen Phänotyp in der neuen *A. thaliana gata hex* Mutante. Basierend auf diesen Ergebnissen komme ich zu dem Schluss, dass B-GATAs eine schützende Rolle bei Starklicht-Stress haben und dass diese Funktion möglicherweise evolutionär konserviert sein könnte.

## Acknowledgements

I would like to express my gratitude and appreciation to all those who have supported me throughout the journey of my PhD during the last five years. Without their encouragement, guidance, and contributions, this accomplishment would not have been possible.

First and foremost, I would like to thank my supervisor, Prof. Dr. Claus Schwechheimer, for giving me the opportunity to work on this exciting project. His support and guidance were instrumental in shaping my research.

I would like to extend my thanks to my mentor, Prof. Dr. Miltos Tsiantis, and my second examiner, Prof. Dr. Kay Schneitz, for their constructive criticism and valuable feedback during our TAC-meetings, as well as to the chair of my examining committee, Prof. Dr. Ralph Hückelhoven.

I am very thankful for my collaborators Dr. Nora Gutsche and Prof. Dr. Sabine Zachgo, well as Dr. Mauricio Lopez-Obando and Prof. Dr. Eva Sundberg, for hosting me and teaching me the basics of *Marchantia polymorpha* and *Physcomitrium patens* research. I would also like to thank Dr. Barbro Winkler, Dr. Boris Hedtke, Prof. Dr. Bernhard Grimm, Dr. Timo Engelsdorf, Dr. Ondřej Novák, and Dr. Oguz Top. Your contributions helped to make this project possible.

Many thanks to my colleagues and friends Alina, Lana, Carlos, Jan, and Dario for the fruitful discussions, climbed mountains and all the fun we had in the lab. Special thanks to Bang-Yu, who assisted me with the revisions for the paper. Thanks to all members of the chair for your support in everyday lab life and the great atmosphere.

I would like to thank Barbara for her superb technical assistance, as well as Petra and Dani for their help with all administrative tasks.

I am grateful to my family for their unwavering support, trust, and belief, not only throughout the last five years but throughout my entire life.

Finally, I would like to express my gratitude to Felicitas, who was always there to listen, to support and to encourage me.

# Table of Contents

Abstract.....	i
Zusammenfassung.....	ii
Acknowledgements.....	iii
<b>1 Introduction.....</b>	<b>1</b>
1.1 GATA transcription factors in plants.....	1
1.1.1 The regulation and functions of B-GATAs in angiosperms.....	2
1.1.1.1 LLM-domain B-GATAs regulate chlorophyll biosynthesis.....	4
1.1.1.2 LLM-domain containing B-GATAs promote stomatal development.....	7
1.1.1.3 LLM-domain containing B-GATAs control cytokinin-regulated development.....	7
1.2 Plant evolution.....	10
1.3 <i>Arabidopsis thaliana</i> .....	11
1.3.1 Role of <i>A. thaliana</i> as a model organism.....	11
1.3.2 Morphology and life cycle of <i>A. thaliana</i> .....	11
1.4 <i>Marchantia polymorpha</i> .....	12
1.4.1 Role of <i>M. polymorpha</i> as a model organism.....	12
1.4.2 Morphology and life cycle of <i>M. polymorpha</i> .....	12
1.5 <i>Physcomitrium patens</i> .....	14
1.5.1 Role of <i>P. patens</i> as a model organism.....	14
1.5.2 Morphology and life cycle of <i>P. patens</i> .....	14
1.6 Plant cell walls.....	15
1.7 Photosynthesis.....	16
1.7.1 Light-dependent reactions of photosynthesis.....	16
1.7.2 Carbon reactions of photosynthesis.....	18
1.8 Excess light energy causes photooxidative damage.....	20
1.8.1 EARLY LIGHT INDUCED PROTEINS respond to high-light stress.....	20
1.8.2 Carotenoids are required for photoprotection in plants.....	21
1.9 Aims of my project.....	23
1.10 Publications created as part of the doctorate.....	23
<b>2 Results.....</b>	<b>24</b>
2.1 GATA evolution in land plants.....	24
2.2 Novel functions for B-GATAs of <i>Arabidopsis thaliana</i> .....	27
2.2.1 Generation of a <i>A. thaliana gata hexuple</i> mutant using CRISPR/Cas9.....	27
2.2.2 Deficiency in chlorophyll accumulation is a phenotype of the <i>gata hex</i> mutant.....	28
2.2.3 Decreasing stem instability is a novel phenotype of <i>A. thaliana</i> higher order <i>gata</i> mutants.....	29
2.2.4 Mutants and overexpressors of <i>A. thaliana B-GATAs</i> show altered cytokinin contents.....	32

2.3	<i>M. polymorpha</i> MpB-GATA1 has developmental- and physiological functions.....	33
2.3.1	Increased gemma cup density and reduced plant size are phenotypes of <i>Mpb-gata1</i> mutants .....	33
2.3.2	Reduced gemma cup formation is a phenotype of MpB-GATA1 overexpressors....	35
2.4	Complementation lines of <i>Mpb-gata1</i> mutants lacking HAN- or LLM-domains did not give insights in domain functions .....	36
2.4.1	Transcriptomic analysis of MpB- <i>gata1</i> suggests role in regulation of high-light stress .....	37
2.4.2	Differential regulation of the tetrapyrrole biosynthesis pathway in <i>Mpb-gata1</i> mutants.....	39
2.4.3	Differential regulation of the Calvin-Benson cycle in <i>Mpb-gata1</i> mutants.....	40
2.5	<i>Physcomitrium patens</i> B-GATAs are involved in protonema formation .....	41
2.5.1	Deficiency in protonema formation is a phenotype of <i>P. patens gata</i> mutants .....	41
2.5.2	Protonema defects of <i>ppgata</i> mutants might be related to cytokinin .....	42
2.6	B-GATAs are involved in high-light stress response .....	43
2.6.1	<i>Mpb-gata1</i> mutants maintain high-light response .....	43
2.6.2	<i>Marchantia Mpb-gata1</i> exhibit constitutive a high-light stress response.....	45
2.6.3	High-light stress induces increased air pore size in <i>M. polymorpha</i> genotypes.....	47
2.6.4	<i>Arabidopsis thaliana</i> LLM B-GATA <i>gata hex</i> maintains high-light stress response .	48
2.6.5	The <i>A. thaliana</i> LLM B-GATA <i>gata hex</i> exhibits a constitutive high-light stress phenotype .....	49
2.6.6	The role of B-GATAs in high-light stress is conserved between <i>M. polymorpha</i> and <i>A. thaliana</i> .....	50
2.6.7	Genes involved in carbon fixation are similarly regulated in <i>Mpb-gata1</i> mutants and in response to high-light .....	53
2.6.8	Transcriptomic analysis yields differences in cytokinin signalling of <i>gata</i> mutants in <i>M. polymorpha</i> and <i>A. thaliana</i> .....	54
2.6.9	The tetrapyrrole biosynthesis pathway is differentially regulated in high-light stressed BoGa and not-stressed <i>Mpb-gata1</i> mutants .....	56
2.6.10	Quantification of chlorophyll intermediates reveals no differences in <i>Mpb-gata1</i> mutants .....	56
2.6.11	Differential expression of the carotenoid biosynthesis pathway .....	58
2.6.12	Quantification of carotenoids and xanthophylls.....	58
2.6.13	The tetrapyrrole biosynthesis pathway is differentially regulated in <i>A. thaliana gata hex</i> and in response to high-light.....	60
2.6.14	Quantification of chlorophyll intermediates confirms role of <i>A. thaliana</i> B-GATAs in greening .....	60
2.6.15	Differential expression of the carotenoid biosynthesis pathway in <i>A. thaliana gata hex</i> .....	62
2.6.16	Quantification of carotenoids suggests impairment of the xanthophyll cycle in <i>gata hex</i> .....	63
2.6.17	Long term high-light stress did not reveal GATA effect in <i>M. polymorpha</i> .....	63

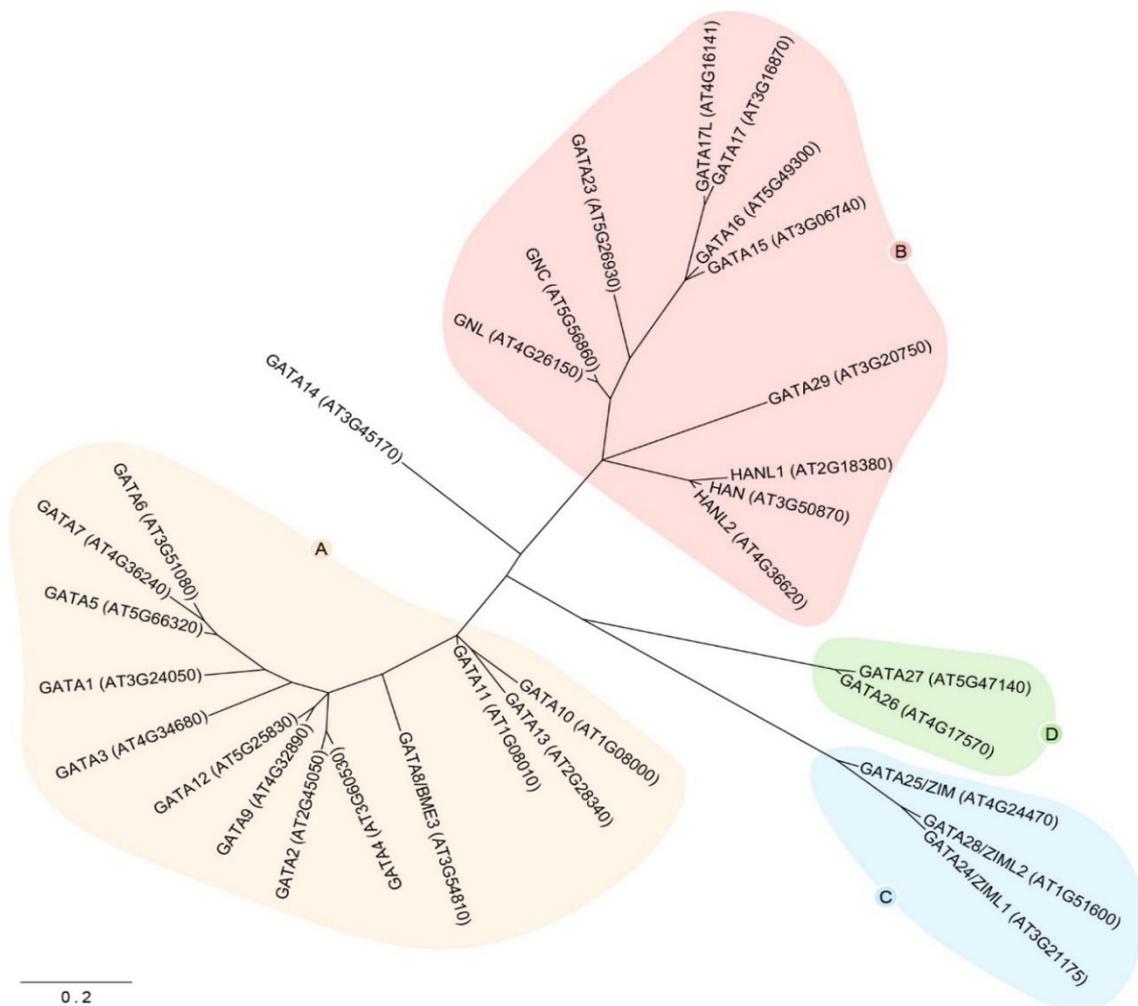
2.6.18	<i>A. thaliana</i> GNLox is deficient in anthocyanin biosynthesis .....	64
2.6.19	<i>A. thaliana gata hex</i> accumulates less ROS compared to Col-0 after exposure to high light .....	65
<b>3</b>	<b>Discussion.....</b>	<b>67</b>
3.1	Bryophytes contain B-GATAs with both HAN- and LLM-domains .....	67
3.2	B-GATAs in <i>P. patens</i> regulate protonema formation independently of cytokinin ...	68
3.3	The single B-GATA ortholog of <i>M. polymorpha</i> has physiological and developmental functions.....	69
3.4	Gemma cup formation is regulated by MpB-GATA1 .....	69
3.5	The functions of HAN- and LLM-domain remain unknown in <i>M. polymorpha</i> MpB-GATA1 .....	70
3.6	Gene expression is not sufficient to explain altered CK levels in GATA mutants and overexpressors of <i>M. polymorpha</i> and <i>A. thaliana</i> .....	71
3.7	<i>gata</i> mutants exhibit high-light stress response in low-light conditions.....	71
3.8	B-GATA factors are required to repress high-light stress responses in <i>M. polymorpha</i> and <i>A. thaliana</i> .....	72
3.9	B-GATAs regulate tetrapyrrole biosynthesis in <i>A. thaliana</i> and <i>M. polymorpha</i> .....	74
3.9.1	The role of B-GATAs in regulation of chlorophyll biosynthesis is conserved.....	74
3.9.2	MpB-GATA1 represses the high-light response of the tetrapyrrole biosynthesis pathway in <i>M. polymorpha</i> .....	74
3.9.3	<i>A. thaliana gata hex</i> did not copy the high-light response of the tetrapyrrole biosynthesis pathway .....	75
3.10	B-GATAs are not involved in the high-light response of the carotenoid biosynthesis pathway.....	76
3.10.1	The xanthophyll cycle functions independently of MpB-GATA1.....	76
3.10.2	Regulation of the carotenoid biosynthesis pathway is a conserved function of LLM B-GATAs .....	77
3.11	Loss of MpB-GATA1 affects carbon fixation .....	79
3.12	The Calvin-Benson cycle is high-light regulated in <i>M. polymorpha</i> but not in <i>A. thaliana</i> .....	79
3.13	GATAs influence flavonoid accumulation differently in <i>M. polymorpha</i> and <i>A. thaliana</i> .....	79
3.14	Loss of B-GATAs might prime <i>A. thaliana</i> for high-light stress.....	80
3.15	Summary and outlook .....	81
<b>4</b>	<b>Materials and Methods .....</b>	<b>83</b>
4.1	Biological material .....	83
4.2	Plant culture .....	83
4.3	Phylogeny construction .....	83
4.4	Thallus size quantification .....	83
4.5	Air pore size quantification .....	84
4.6	Molecular cloning and CRISPR/Cas9-mediated genome editing .....	84

4.7	Quantitative real-time PCR.....	85
4.8	RNA sequencing .....	85
4.8.1	Identification of orthologs.....	86
4.8.2	GO enrichment analysis .....	86
4.8.3	Accession Numbers .....	86
4.9	Chlorophyll quantification .....	86
4.10	Detection of superoxide with NBT staining .....	87
4.11	gRNAs used for CRISPR/Cas9-mediated genome editing.....	87
4.11.1	<i>M. polymorpha</i> gRNAs .....	87
4.11.2	<i>P. patens</i> gRNAs.....	87
4.11.3	<i>A. thaliana</i> gRNAs.....	87
4.12	Primers used in this study .....	88
4.12.1	<i>Mpb-gata1</i> genotyping primers.....	88
4.12.2	<i>P. patens</i> genotyping primers.....	88
4.12.3	<i>A. thaliana</i> SALK primers .....	88
4.12.4	<i>A. thaliana gata hex</i> genotyping primers.....	88
4.13	qRT-PCR primers.....	89
4.14	Cloning of <i>MpB-GATA1</i> .....	89
<b>5</b>	<b>Supplemental materials.....</b>	<b>90</b>
5.1	Supplemental Data.....	90
5.2	Supplemental Figures .....	93
5.3	Supplemental Tables.....	97
<b>6</b>	<b>Literature .....</b>	<b>101</b>

# 1 Introduction

## 1.1 GATA transcription factors in plants

GATA transcription factors are a class of highly conserved transcriptional regulators among all animals, plants, and fungi. Their DNA binding domain is a type IV zinc finger usually in the form  $CX_2CX_{17-20}CX_2C$ . While some GATAs from metazoa and fungi contain several zinc finger domains, plant GATAs encode only one zinc finger domain per protein. The genome of *Arabidopsis thaliana* encodes 30 GATA transcription factors (Figure 1.1). Based on the sequence of their zinc finger domains and the presence of additional domains, they can be subdivided in the families A-D according to a previously proposed classification for GATAs from *A. thaliana* and *Oryza sativa* (Reyes, Muro-Pastor et al. 2004). GATA14 was previously considered a member of class A, but my analysis based on the sequence of the zinc-finger domain placed it outside any of the defined classes. Even though GATA14 did not cluster with class-A GATAs, a highly conserved motif close to the zinc finger suggests that it is phylogenetically related to it. The zinc finger of *A. thaliana* GATAs has the form  $CX_2CX_{18}CX_2C$ , with the exception of GATA29 from class B, which has  $CX_4CX_{18}CX_2C$ , and the members of class C, GATA24, GATA25, and GATA28, which encode zinc fingers with the form  $CX_2CX_{20}CX_2C$ .



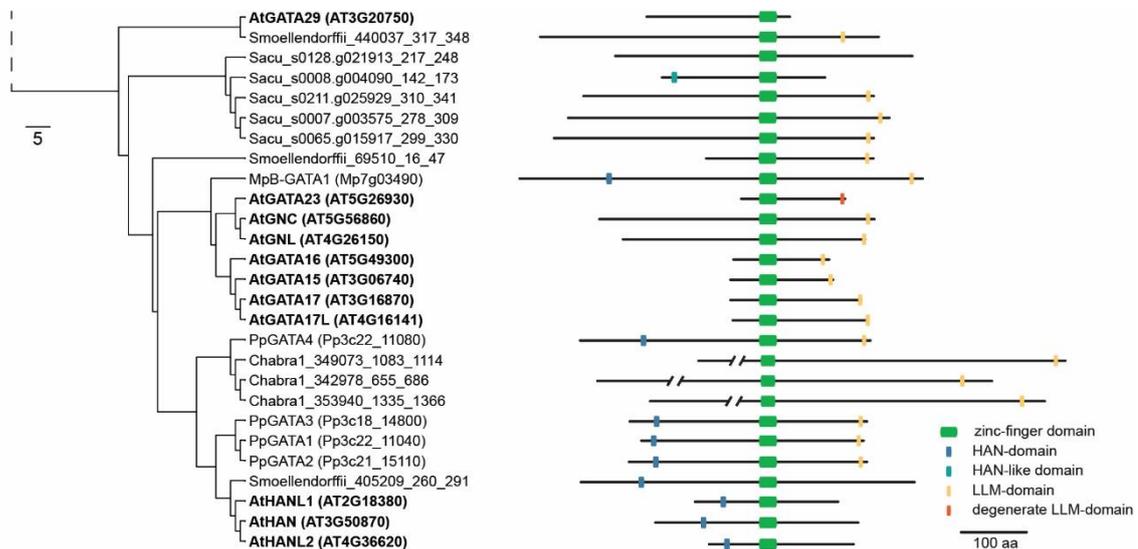
**Figure 1.1: Phylogeny of all 30 GATAs from *A. thaliana*.** Zinc finger sequences were aligned using MUSCLE (Edgar 2004) and phylogeny construction was carried out with PhyML (Guindon, Dufayard et al. 2010) via ATGC (<http://www.atgc-montpellier.fr/phyml/>). The Dayhoff model was chosen using automatic model selection by SMS based on the Bayesian Information Criterion and used for phylogeny construction (Lefort, Longueville et al. 2017). The scale bar represents a distance of 0.2 amino acid substitutions.

### 1.1.1 The regulation and functions of B-GATAs in angiosperms

Research on GATA transcription factors in plants is almost exclusively focused on class-B GATAs and a major research topic in my host laboratory. Members of this GATA class can be further separated by the presence of either an LLM- or HAN-domain, which is located at the C- or N-terminus, respectively. The HAN- and LLM-domains only appear in B-GATAs of plants and have not been found in any other proteins.

Since the sequencing of the *Arabidopsis thaliana* genome in 2000, the number of available plant genomes has expanded drastically due to the ever-sinking costs of current generation DNA sequencing technologies (Li and Harkess 2018, Paajanen, Kettleborough et al. 2019). The public availability of many new plant genome sequences in comparative genomics platforms like PLAZA (bioinformatics.psb.ugent.be/plaza; Van Bel, Diels et al. 2018) and Phytozome (phytozome.jgi.doe.gov; Goodstein, Shu et al. 2012), enabled comparative analyses in ways that were not possible only a decade ago. While plant genome sequencing efforts at first mainly focused on crop species, in recent years more and more early diverging land plant species were covered. I made use of these new resources for the construction of comprehensive phylogenies in order to investigate the origin of B-GATAs in the green lineage.

My phylogenetic analyses revealed that some species such as *Marchantia polymorpha* and *Physcomitrium patens* contain both domains on the same protein, a constellation unique to bryophytes (Figure 1.2). So far, the biochemical functions of both domains are unknown, but previous research from my host laboratory found that the LLM-domain is required for the complementation of the chlorophyll accumulation defect in the *gnc gnl* double mutant (Behringer, Bastakis et al. 2014). One study reported physical interaction of the HAN- and LLM-domain in *A. thaliana*, but this finding was dismissed when the results could not be replicated using the original materials (Zhang, Zhou et al. 2013). Understanding the functions of HAN- and LLM-domain was part of the motivation to start my project.

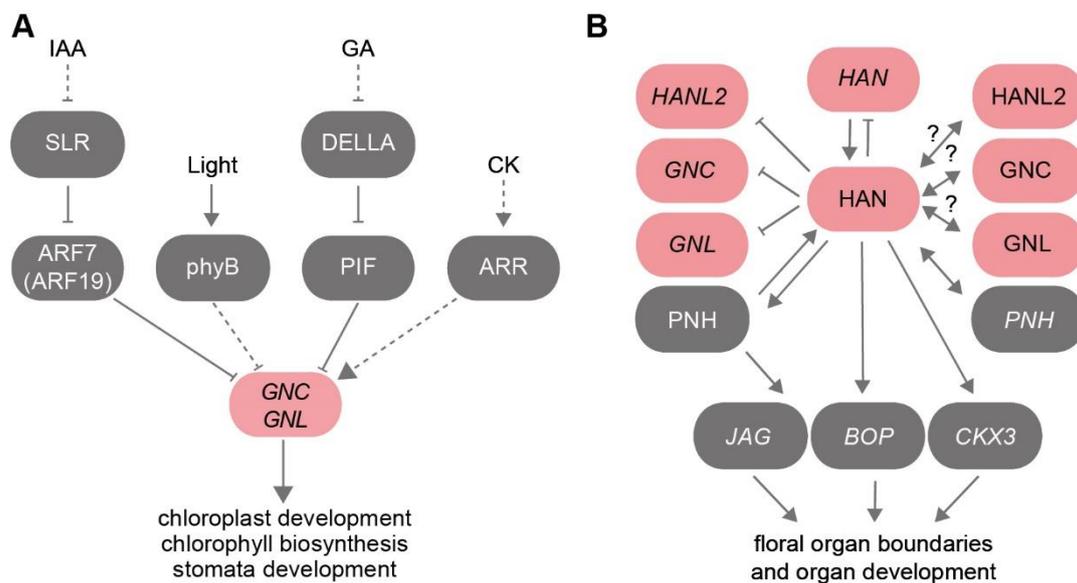


**Figure 1.2: Bryophytes contain B-GATAs with HAN- and LLM-domains.** Phylogenetic tree and schematic representation of the B-GATA factors from *Selaginella moellendorffii* (Sm), *Chara braunii* (CHABR), *Marchantia polymorpha* (Mp), *Salvinia cucullata* (Sc), *Physcomitrium patens* (Pp) and *Arabidopsis thaliana* (At). The lengths of the B-GATAs are drawn to scale and only the B-GATA zinc-finger domains were aligned. The N-terminally located HAN-domain and the C-terminally located LLM-domain or the degenerate variant LLM-domain are shown. *A. thaliana* B-GATAs is shown in bold. The scale bar represents a distance of five amino acid substitutions.

We recently published a review article on the biology, phylogeny and phylogenomics of plant GATA factors, which summarizes the present knowledge in great detail (Schwechheimer, Schröder et al. 2022). Thus, I will only provide a brief summary of the roles of B-GATAs important within the context of my work.

The LLM-domain containing B-GATAs comprise six members in *A. thaliana* and are characterized based on the existence of the name giving leucine-leucine-methionine motif at the C-termini of the proteins (Reyes, Muro-Pastor et al. 2004). Their most prominent members are *GNC* (*GATA*, *NITRATE INDUCIBLE*, *CARBON METABOLISM-INVOLVED*) and its paralog *GNL* (*GNC-LIKE*; Bi, Zhang et al. 2005, Naito, Kiba et al. 2007). Most LLM-domain B-GATAs are induced by cytokinin, however only *GNL* is known to be regulated by light (Figure 1.3; Naito, Kiba et al. 2007, Ranftl, Bastakis et al. 2016).

In our laboratory, *GNC* and *GNL* were first described as downstream targets of the gibberellin (GA) signalling pathway and PHYTOCHROME INTERACTION FACTORS (PIFs) (Richter, Behringer et al. 2010). Later, the involvement of *GNC* and *GNL* in the auxin signalling pathway via the AUXIN RESPONSE FACTORS ARF2 and ARF7 was unravelled (Richter, Behringer et al. 2013). *gnc* mutants are pale green and this phenotype is further enhanced in higher order mutants of LLM B-GATAs, suggesting at least partial functional redundancy of the transcription factors (Behringer, Bastakis et al. 2014, Ranftl, Bastakis et al. 2016). Transcriptomic analyses revealed that this observation can be explained by misregulation of many genes of the chlorophyll biosynthesis pathway, as well as of genes involved in chloroplast division (Naito, Kiba et al. 2007, Chiang, Zubo et al. 2012, Bastakis, Hedtke et al. 2018).



**Figure 1.3: Selected biological functions of *A. thaliana* B-GATAs.** **A** Biological functions of LLM B-GATAs. **B** Biological functions of HAN B-GATAs. Genes that are transcriptionally regulated are shown in italics. Arrows indicate activation, blunt arrows indicate inhibition. Dashed lines represent indirect activation or repression. Abbreviations: ARF, AUXIN RESPONSE FACTOR; ARR, ARABIDOPSIS RESPONSE REGULATOR; BOP, BLADE-ON-PETIOLE; CK, cytokinin; CKX3, CYTOKININ OXIDASE3; GA, gibberellin; IAA, auxin (indole-3-acetic acid); JAG, JAGGED; phyB; phytochrome B; PIF, PHYTOCHROME-INTERACTING FACTOR; PNH, PINHEAD; SLR, SOLITARY ROOT. Adapted from Schwechheimer, Schröder et al. 2022.

Apart from *GNC* and *GNL*, the family in *A. thaliana* contains four additional proteins with shorter N-termini named *GATA15*, *GATA16*, *GATA17*, and *GATA17-LIKE* (*GATA17L*). Our laboratory generated a comprehensive suite of mutants and transgenic lines for LLM-domain containing B-GATAs, and in many cases we identified opposing phenotypes for mutants and overexpressors (Richter, Behringer et al. 2010, Behringer, Bastakis et al. 2014, Klermund, Ranftl et al. 2016, Ranftl, Bastakis et al. 2016, Bastakis, Hedtke et al. 2018). This shows that phenotypes of *GATA* overexpressors can provide physiologically relevant insights into *GATA* factor function.

The three HAN-domain B-GATAs *HAN* (*GATA18*), *HANL1* (*HAN-LIKE1*; *GATA20*) and *HANL2* (*GATA19*) were originally described as regulators of floral development (Zhao, Medrano et al. 2004, Zhang, Zhou et al. 2013) and thus named “*hanaba taranu*” (from jap. “not enough floral leaves”). In contrast to the LLM B-GATAs, essentially nothing is known about their upstream

regulation, but transcriptomic data suggests that HAN negatively regulates its own expression, as well as the expression of *HANL2*, *GNC*, and *GNL* (Figure 1.3; Zhang, Zhou et al. 2013). In *han* mutants, floral organ identity and embryonic development are strongly affected. *HAN* is expressed in the region between the shoot meristem and newly initiated organs and the respective mutants exhibit reduced numbers of floral organs and fused sepals, as well as flattened shoot meristems (Zhao, Medrano et al. 2004). *HAN* interacts with the *CLAVATA* meristem regulators *CLV1*, *CLV2* and *CLV3* and there are indications that *HAN* controls expression of *WUSCHEL* (*WUS*), a regulator of meristem size. Furthermore, *HAN* modulates cytokinin homeostasis by inducing the expression of the *CYTOKININ OXIDASE* (*CKX3*) (Ding, Yan et al. 2015). *HAN* also acts as a growth repressor since overexpressors display altered cell division patterns in shoots and flowers (Zhao, Medrano et al. 2004). *HAN* activates the expression of *PINHEAD* (*PNH*) during floral development and interacts with *PNH* to induce expression of *JAGGED* (*JAG*) and *BLADE-ON-PETIOLE* (*BOP*) (Ding, Yan et al. 2015). Additionally, double mutants of *HAN* and *ANGUSTIFOLIA3/GRF-INTERACTING FACTOR1* (*AN3/GIF1*) show ectopic root formation in cotyledons (Kanei, Horiguchi et al. 2012). *han* mutants show altered auxin distribution in the embryo resulting in defects in the proembryo boundary and embryonic root development, likely due to lacking cell divisions required for the formation of the quiescent center (Nawy, Bayer et al. 2010). Furthermore, *han* mutants show defects in cotyledon initiation with plants developing up to four cotyledons (Figure 1.3; Zhao, Medrano et al. 2004).

#### 1.1.1.1 LLM-domain B-GATAs regulate chlorophyll biosynthesis

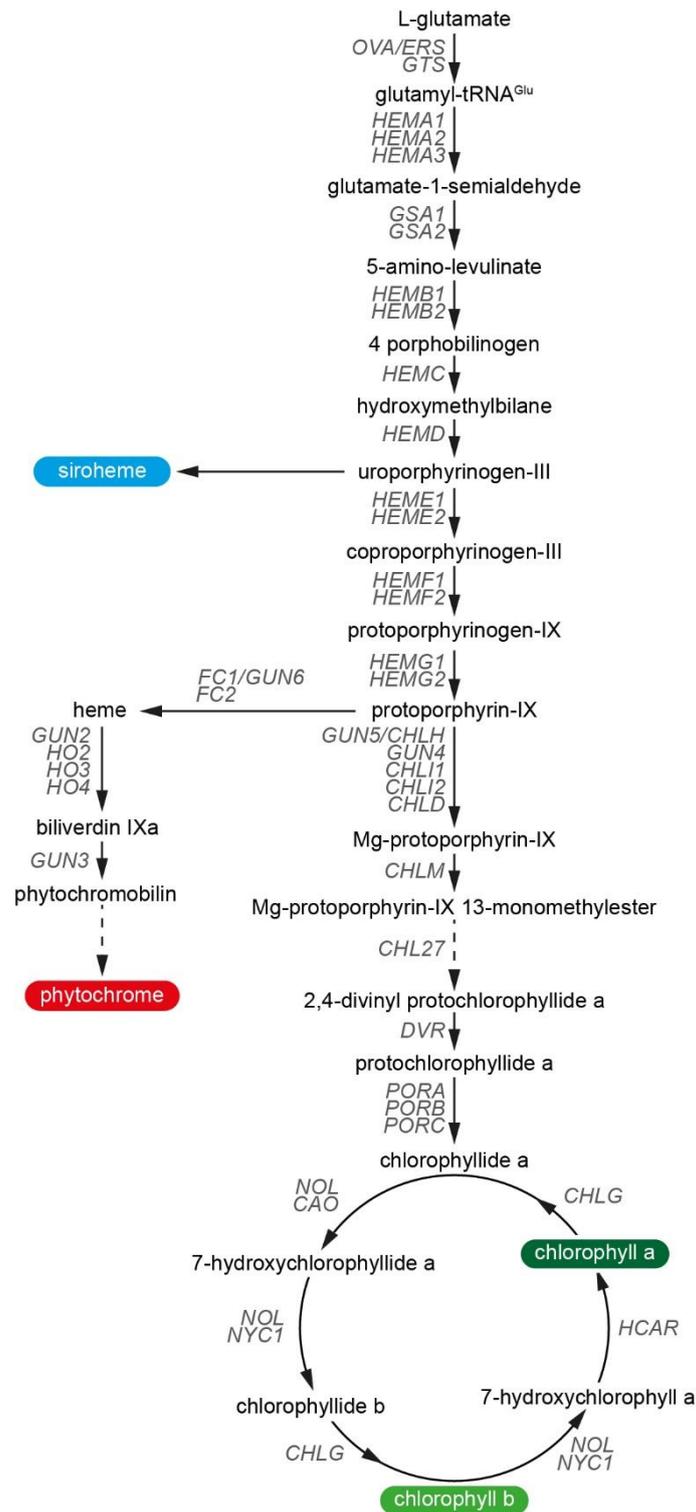
Chlorophylls are the major photosynthetically active pigments in plants and defective chlorophyll accumulation is an important phenotype of *b-gata* mutants from *A. thaliana*. Land plants produce chlorophyll *a* and chlorophyll *b*, and the latter is produced from the former by a conversion of a methyl to a formyl side chain (Tanaka and Tanaka 2006). Chemically, chlorophylls are tetrapyrroles and bind  $Mg^{2+}$  ions in the centre of their four pyrrole rings.

Chlorophylls are products of a branch of the tetrapyrrole biosynthesis pathway, and their production takes place in chloroplasts (Tripathy and Pattanayak 2012). The tetrapyrrole pathway starts with the synthesis of 5-amino-levulinate (ALA) from L-glutamate and the first cyclic intermediate of the pathway is uroporphyrinogen-III (Uro-III), which is synthesized through the condensation of eight ALA molecules (Figure 1.4). Uro-III undergoes an oxidative conversion, yielding protoporphyrin-IX (Proto-IX). Here, the tetrapyrrole biosynthesis branches off into the heme pathway and the chlorophyll pathway. The ferrochelatase-catalysed insertion of  $Fe^{2+}$  in the pyrrole ring of Proto-IX results in the formation of heme, which is a precursor for the biosynthesis of phytochromobilin, the chromophore of phytochromes. The insertion of  $Mg^{2+}$  into protoporphyrin-IX is catalysed by a magnesium chelatase (MgCH) using ATP and leads to the formation of Mg-protoporphyrin-IX (MgP) and later, chlorophyll. The three MgCH subunits are encoded by the genes *GENOMES UNCOUPLES 5* (*GUN5*), *CHELATASE D* (*CHLD*), *CHLI1* and *CHLI2* (Tanaka and Tanaka 2006). MgP is converted to Mg-protoporphyrin-IX 13-monomethylester (MME) via a Mg-protoporphyrin-IX methyltransferase (CHLM) and subsequently to 2,4-divinyl protochlorophyllide and protochlorophyllide through a Mg-protoporphyrin-IX monomethylester cyclase (CHL27) and a 3,8-divinyl protochlorophyllide a 8-vinyl reductase (DVR). Chlorophyllide *a* synthesis is catalysed by the protochlorophyllide oxidoreductases A-C (PORA-C), and chlorophyllide *a* is either directly converted to chlorophyll *a* by the chlorophyll synthase CHLG, or converted to 7-hydroxychlorophyllide *a*, chlorophyllide *b* and then to chlorophyll *b* through the activities of the chlorophyll *b* reductases NON-YELLOW COLORING1 (NYC1) and NYC1-LIKE/NOL, the chlorophyllide *a* oxygenase CAO, and CHLG (Tripathy and Pattanayak 2012).

Mutants of LLM-domain B-GATAs in *A. thaliana* exhibit a striking greening defect that can be explained through impaired transcriptional activation of genes involved in chlorophyll

biosynthesis and chloroplast development (Bastakis, Hedtke et al. 2018, Zubo, Blakley et al. 2018). Additionally, in ChIP-qPCR experiments, *GUN4*, *DVR*, *GUN2*, the sigma factors *SIG2* and *SIG6*, as well as the MgCH-subunits *GUN5* and *CHLD* were shown to be direct targets of GNL (Bastakis, Hedtke et al. 2018). Genetic experiments further revealed that overexpression of GNL was sufficient to suppress the pale green phenotype of *gun5* and *chli* mutants, whereas overexpression of *DVR* suppressed the greening defect of the *gnc gnl* mutant, suggesting that DVR abundance is the limiting factor for chlorophyll accumulation in the mutant (Bastakis, Hedtke et al. 2018). The heme biosynthesis pathway is also affected in LLM B-GATA mutants, leading to reduced phytylchromobilin synthesis, which ultimately results in decreased phytochrome function.

While LLM-domain B-GATA mutants are pale green, *A. thaliana* *GNC*- or *GNL* overexpressors (*GNCox* and *GNLox*, respectively) accumulate more chlorophyll and thus appear darker compared to the wildtype (Richter, Behringer et al. 2010, Chiang, Zubo et al. 2012). Conversely, many genes involved in chlorophyll biosynthesis are upregulated in *GNCox* (Bastakis, Hedtke et al. 2018).



**Figure 1.4: The tetrapyrrole biosynthesis pathway in *A. thaliana*.** Intermediates are indicated by black text; enzymes are indicated by grey italic text; enzymatic reactions are shown as black arrows; dashed arrows indicate additional steps not shown here. The four main products are shown in coloured boxes. Adapted from Bastakis, Hedtke et al. 2018.

In many plant species, the pair of GOLDEN-LIKE (GLK) transcription factors *GLK1* and *GLK2* are required for chloroplast formation and *glk1 glk2* double mutants in *Arabidopsis thaliana* exhibit lower amounts of chlorophyll biosynthesis intermediates compared to the wild type, as well as downregulation of genes encoding light harvesting complexes and enzymes required for chlorophyll biosynthesis (Waters, Wang et al. 2009), reminiscent of the mutants of LLM B-GATAs. *GNC* and *GNL* likely act independently of *GLK1* and *GLK2*, since the differentially

expressed genes in *gnc gnl* and *glk1 glk2* are added up in the *gnc gnl glk1 glk2* quadruple mutant (Bastakis, Hedtke et al. 2018, Zubo, Blakley et al. 2018).

#### 1.1.1.2 LLM-domain containing B-GATAs promote stomatal development

Stomata are pores in the epidermis of above-ground plant organs like hypocotyls, cotyledons, stems, leaves and floral organs that plants use to control the exchange of oxygen, CO<sub>2</sub>, and water with their environment (Casson and Hetherington 2010). In the course of my project, I uncovered the involvement of the single *B-GATA* ortholog of *M. polymorpha* in the regulation of air pore formation, structures analogous to stomata. In contrast to the air pores of *M. polymorpha*, stomata can be actively opened and closed in order to adjust to abiotic conditions, such as heat and drought. Stomatal development in *A. thaliana* is regulated by several genes, and the LLM-domain B-GATAs *GNC*, *GNL*, *GATA17*, and *GATA17L* act upstream of the key regulators *SPEECHLESS* (*SPCH*), *MUTE*, *SCREAM* and downstream of several light-signalling pathways. LLM B-GATAs induce stomata formation in hypocotyls and, less prominently, in cotyledons (Klermund, Ranftl et al. 2016). In a *gata quadruple* mutant, the number of stomata is decreased compared to the wild type, while *GNC* overexpressors show a strong increase in stomata number. *GATA* function on stomata formation can be induced by different light treatments or by overexpression of *GNC* in dark grown *pif* quadruple mutants, suggesting that *GATAs* act downstream from *PIFs* or with the *PIFs* themselves (Klermund, Ranftl et al. 2016).

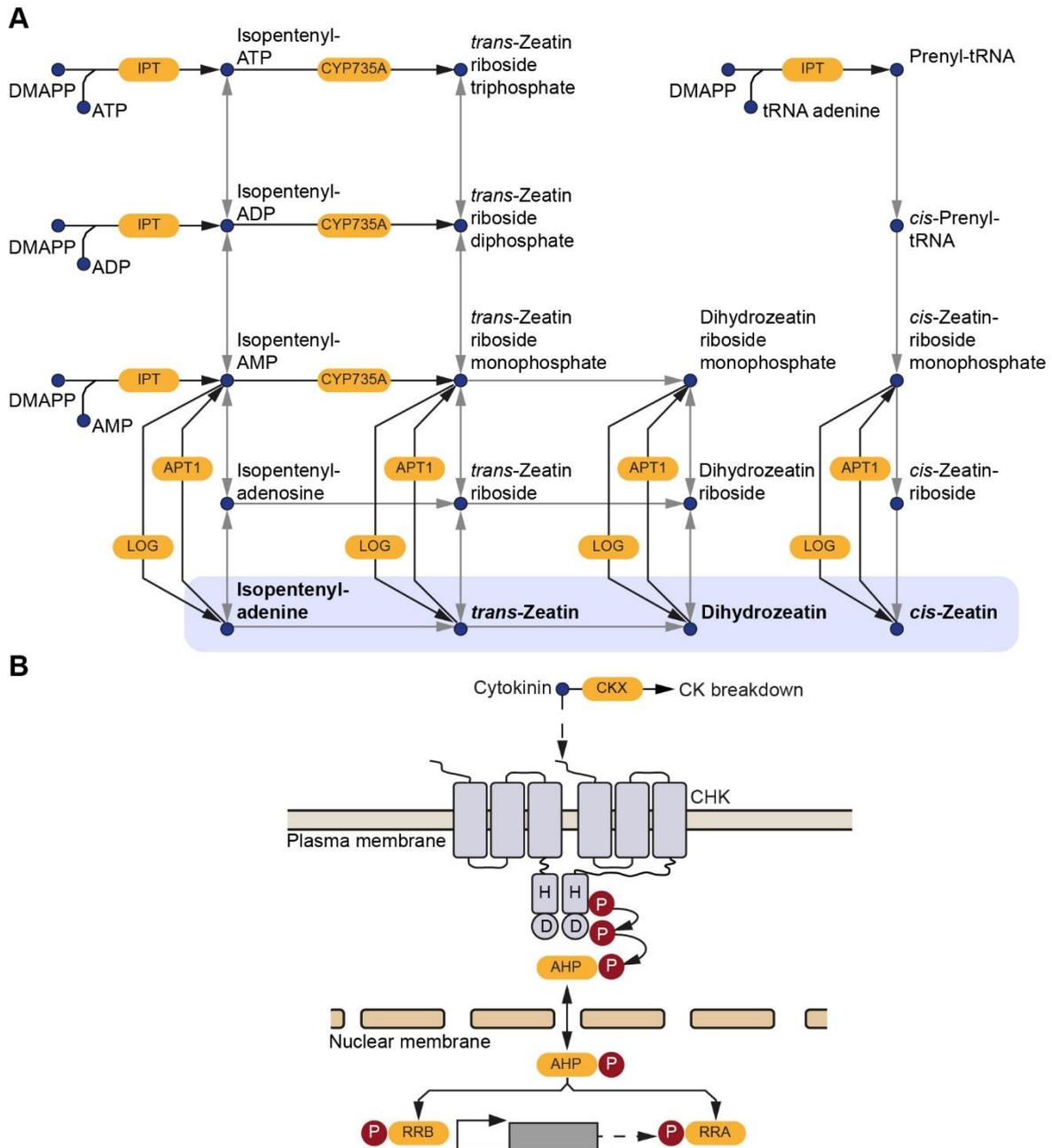
#### 1.1.1.3 LLM-domain containing B-GATAs control cytokinin-regulated development

*Arabidopsis thaliana* LLM B-GATAs *GNC*, *GNL*, *GATA16*, *GATA17*, and *GATA17L* are induced by cytokinin (CK) (Ranftl, Bastakis et al. 2016, Schwechheimer, Schröder et al. 2022). In fact, *GNL/CGA1* was originally identified based on its strong CK induced gene expression and thus named *CYTOKININ-RESPONSIVE GATA FACTOR1* (*CGA1*) (Naito, Kiba et al. 2007), before later being independently described as a regulator of floral homeotic gene action and named *GNC-LIKE* (*GNL*) (Mara 2008) due to its similarity to the previously characterized *GNC* (Bi, Zhang et al. 2005). As will be seen later, some phenotypes of the *gata* mutants I generated in the course of my project could be connected to CK, which is why I will present an overview of CK biosynthesis and signalling here.

CKs are *N*<sup>6</sup> substituted adenine derivatives with isoprenoid side chains (Kieber and Schaller 2014). The naturally occurring CKs are *cis*- and *trans*-zeatin (*cZ* and *tZ*, respectively), dihydrozeatin (DHZ), zeatin riboside, zeatin ribotide, 6-benzylamino purine (BA), thiazuron, and *N*<sup>6</sup>-( $\Delta^2$ -isopentenyl) adenine (IP) (Kieber and Schaller 2014). At least some of them are present in most eukaryotes and bacteria (Persson, Esberg et al. 1994). In plants, zeatin is the most highly abundant CK and its *trans* form is active in almost all plant species, while *cis*-zeatin is only active in some (Gajdošová, Spíchal et al. 2011). Active CKs are usually free bases, while inactive CKs can be found as ribosides and ribotides (Hothorn, Dabi et al. 2011, Lomin, Krivosheev et al. 2015).

CKs are important plant hormones and involved in many developmental processes. They often act together with other plant hormones, mainly auxin, and regulate cell division and differentiation (Sakakibara 2006). CKs promote cell proliferation and meristem activity in the shoot, while inhibiting root growth by regulating root branching and cell differentiation in the root apical meristem (Werner, Motyka et al. 2001, Moubayidin, Perilli et al. 2010, Ioio, Galinha et al. 2012, Chang, Ramireddy et al. 2015, Kieber and Schaller 2018). They also play a role in controlling root architecture by preventing the initiation of lateral roots and elongation of primary roots. CKs affect the functioning of roots by regulating the expression of genes involved in nutrient transport and uptake (Argueso, Ferreira et al. 2009, Werner, Nehnevajova et al. 2010, Kieber and Schaller 2018). Additionally, CKs work in concert with auxin to regulate the development of the vascular system. Through a pair of mutually inhibitory interactions, CKs

promote the development of phloem while auxin regulates the development of xylem (Mähönen, Bishopp et al. 2006, Bishopp, El-Showk et al. 2011, De Rybel, Adibi et al. 2014, Kieber and Schaller 2018).



**Figure 1.5: Cytokinin biosynthesis and signalling in plants. A** CK biosynthesis. **B** CK signalling. Enzymes are shown in yellow boxes, reactions as arrows and metabolites as blue dots. Active CKs are shown in bold type. Black arrows indicate enzymatic reactions with known genes, grey arrows indicate enzymatic reactions with unidentified genes. DMAPP, dimethylallyl diphosphate; IPT, isopentenyl transferase; LOG, cytokinin nucleoside 5'-monophosphate phosphoribohydrolase LONELY GUY; APT1, adenine phosphoribosyl transferase; CKX, cytokinin oxidase; CHK, CHASE domain-containing histidine kinase receptor; AHP, authentic histidine phosphotransferase; RRA, type-A response regulator; RRB, type-B response regulator; P, phosphate.

CK biosynthesis is a complex process consisting of several enzymatic reactions and mediated by regulatory mechanisms (Figure 1.5). One of the key enzymes in CK biosynthesis is an isopentenyl transferase (IPT), which catalyses the addition of a prenyl-group of dimethylallyl diphosphate (DMAPP) to the  $N^6$  position of AMP/ADP/ATP (Sakakibara 2006, Kieber and Schaller 2018). In order to convert the resulting iP ribotides to *tZ*-type CKs, the cytochrome P450 enzymes CYP735A1 and CYP735A2 are required to hydroxylate the isoprenoid side chain (Takei, Yamaya et al. 2004). It is not fully understood how *cZ*-type CKs are formed, but

research suggests that tRNA-IPTs are required to catalyse the addition of a prenyl group to the  $N^6$  position of adenine on tRNA (Kasahara, Takei et al. 2004). Finally, the CK nucleoside 5'-monophosphate phosphoribohydrolase LONELY GUY (LOG) generates physiologically active CKs from CK ribosides through a single enzymatic reaction (Kurakawa, Ueda et al. 2007, Kuroha, Tokunaga et al. 2009). The back-conversion of active CKs to CK ribotides is catalysed by the adenine phosphoribosyl transferase APT1 (Witte and Herde 2020).

Active CK levels in plants can be modulated via conjugation to sugars or catalytic breakdown. Glycosylated CKs cannot bind to the histidine kinase (HK) CK receptors and are not active in bioassays (Spíchal, Rakova et al. 2004). O-glycosylation is catalysed by glycosyltransferases and is reversible via  $\beta$ -glucosidases, while N-glycosylation is probably irreversible (Kieber and Schaller 2018). The breakdown of CK is mediated by cytokinin oxidases (CKX) through irreversible cleavage (Werner, Köllmer et al. 2006) of both active and inactive, riboside forms, of CKs. Notably, tZ- and IP-types can be cleaved by CKX, while DHZ and synthetic CKs like kinetin or 6-BA are resistant to CKX-mediated cleavage (Galuszka, Popelková et al. 2007, Zalabák, Galuszka et al. 2014, Kieber and Schaller 2018). The genes involved in CK biosynthesis and breakdown are modulated by internal and external factors, e.g. LOG and CKX genes exhibit tissue- and time specific expression patterns (Kowalska, Galuszka et al. 2010, Chickarmane, Gordon et al. 2012). Additionally, expression of genes encoding IPT and CKX is regulated by the availability of nitrate and phosphate (Argueso, Ferreira et al. 2009).

The CK signalling pathway consists of a phosphorelay system with an internal feedback loop. In short, CK is perceived by membrane-localized CHASE domain-containing histidine kinase receptors (CHKs), leading to their autophosphorylation. In turn, the CHKs phosphorylate the authentic histidine phosphotransferases (AHPs), which, after translocation to the nucleus, transfer the phosphate to the type-B response regulators (type-B RRs). The ARR gene family comprises four groups, but only type-A ARRs and type-B ARRs are of interest here. In Arabidopsis, there are ten type-A ARRs (ARR3–ARR9, ARR15–ARR17) and 11 type-B ARRs (ARR1, ARR2, ARR10–ARR14, ARR18–ARR21) (To and Kieber 2008), while *Marchantia polymorpha* only encodes one type-A ARR (MpRRA) and one type-B ARR (MpRRB), each (Aki, Mikami et al. 2019). Type-B ARRs function as transcriptional regulators and among their target genes are type-A ARRs (Hwang and Sheen 2001, Sakai, Honma et al. 2001, Taniguchi, Sasaki et al. 2007). Type-A ARRs compete with type B-ARRs for phosphorylation and thus negatively regulate CK signalling (To, Haberer et al. 2004, To, Deruere et al. 2007).

## 1.2 Plant evolution

The water-to-land transition marks one of the milestones in plant evolution. In their conquest of terrestrial habitats, the algal ancestors of extant embryophytes (land plants) faced a broad range of new environmental stimuli and abiotic stresses (Fürst-Jansen, de Vries et al. 2020). Unlike in freshwater environments, the availability of water is a limiting factor for life on land. This required the organisms to evolve water conducting structures and strategies for desiccation tolerance in times of limited water accessibility (Oliver, Tuba et al. 2000). Additionally, changes in ambient temperature are more rapid and harsh in air compared to water, which requires a higher tolerance to ever changing temperatures as well as fast adaptation responses. Next to temperature, high light irradiances are a major challenge for terrestrial photosynthetic organisms. Air does not filter light to the same degree as water, leading to spectral differences which the early land plants were subjected to. UV-B is largely absorbed by water, suggesting that the first land plants had to develop new mechanisms in order to cope with its DNA-damaging capacities (Han, Chang et al. 2019). During daytime, the light intensity fluctuates drastically and can easily reach intensities capable of causing damage to the photosynthetic machinery. These harsh conditions in the new terrestrial environments acted as the selective pressures driving plant evolution.

Even though multiple algal lineages independently adapted a terrestrial lifestyle (Raven and Edwards 2013), all extant embryophytes are derived from a single clade within the Streptophyta, the monophyletic group including streptophyte algae (or charophytes) and embryophytes (Raven and Edwards 2013, Wickett, Mirarab et al. 2014, de Vries and Archibald 2018). Embryophytes can be subdivided into plants containing internal water-conducting structures known as vasculature, thus named vascular plants or tracheophytes, and plants lacking these structures, the bryophytes. Among the vascular plants are lycophytes, ferns, gymnosperms, and angiosperms (flowering plants). The bryophytes, comprising hornworts, mosses and liverworts, have been repeatedly proposed to form a monophyletic group based on recent phylogenetic analyses facilitated through the availability of many newly sequenced plant genomes (Puttick, Morris et al. 2018, de Sousa, Foster et al. 2019, Harris, Harrison et al. 2020, Li, Nishiyama et al. 2020, Rich and Delaux 2020). Based on these results, the current hypothesis claims that bryophytes and vascular plants have been separated by over 400 million years of distinct evolution.

Even though it is unknown whether angiosperms or bryophytes represent the ancestral roles of B-GATAs from the last common ancestor of embryophytes, investigating the functions of B-GATAs in members of both angiosperms and bryophytes can help in understanding plant adaptations to different ecological niches and the genetic mechanisms underlying biological processes. In the course of my project, I found that the single B-GATA from *M. polymorpha* is involved in the regulation of air pore formation, a function possibly related to the role of *A. thaliana* LLM B-GATAs in stomata formation (Klermund, Ranftl et al. 2016). Additionally, I found a conserved role for B-GATAs in the repression of high-light stress, an important trait in the context of plant evolution and terrestrialisation.

## **1.3 *Arabidopsis thaliana***

### **1.3.1 Role of *A. thaliana* as a model organism**

*Arabidopsis thaliana* is an angiosperm of the Brassica family. Whereas other members of the Brassicaceae such as mustard and cabbages are important crop plants, *A. thaliana* is of no economic value (Meyerowitz 1987). Despite its unassuming habitus, *A. thaliana* was early proposed as “the plant *Drosophila*” (Titova 1935) and has had a long history as a model organism in plant biology, its use in classical genetic studies dating back to the mid-20<sup>th</sup> century (Laibach 1943). Due to its relatively small, diploid genome, short life cycle, and easy handling, *A. thaliana* is one of the most used models in plant biology, and commonly used ecotypes are Landsberg *erecta* (Ler) and Columbia (Col-0; (Woodward and Bartel 2018). *A. thaliana* was the first plant to have its genome sequenced, thus laying the foundation for modern plant genetics (The Arabidopsis Genome Initiative 2000). The large collection of publicly available T-DNA insertion lines has greatly accelerated research on *A. thaliana*, allowing for large-scale genetic screens and functional analyses of single genes (Alonso, Stepanova et al. 2003).

### **1.3.2 Morphology and life cycle of *A. thaliana***

*A. thaliana* is an herbaceous plant, growing 30-40 cm tall (Meyerowitz 1987). Its leaves are arranged in a rosette at the bottom of the stem carrying the inflorescences. Typically for vascular plants, the dominant phase of the life cycle of *A. thaliana* is the sporophyte, whereas the gametophyte is reduced to a few cells. *A. thaliana* is monoecious and adult plants grow small, white flowers with four petals, four sepals, six stamens and two carpels. The flowers usually self-fertilize but can be crossed easily. After pollination, the pollen tube grows within the pistil towards the ovule in order to deliver the two haploid male gametes known as sperm cells to the haploid female gametes. In the ovary, one of the sperm cells fertilizes the egg cell, resulting in the formation of a diploid zygote. The other sperm cell fuses with the two polar nuclei in order to form the triploid endosperm, a nurturing tissue for the growing embryo. After fertilization, the fruit, a silique, is formed and contains 30-60 seeds (Meyerowitz 1987). As an angiosperm, *A. thaliana* has a sporophyte dominant life cycle. The life cycle takes six weeks to complete from seed to adult plants carrying mature seeds (Woodward and Bartel 2018).

## 1.4 *Marchantia polymorpha*

### 1.4.1 Role of *M. polymorpha* as a model organism

*Marchantia polymorpha*, a common liverwort, is currently experiencing a renaissance as a model organism in plant biology. The oldest surviving illustration of *M. polymorpha* is from the 15<sup>th</sup> century (Pächt 1950, Blunt and Raphael 1994) and the species has been used in research for over 200 years, with the earliest scientific studies dating back to 18<sup>th</sup> century (Bowman 2015, Bowman 2022). In recent years, *M. polymorpha* has been rediscovered by biologists and is now routinely used in many molecular plant laboratories. Due to its phylogenetic position as bryophyte, it provides the opportunity to study genes or gene families from a different perspective than the angiosperm *A. thaliana*. The well annotated genome, easy handling in the laboratory and availability of established protocols for transformation make it the ideal subject in comparative studies using both bryophytes like *M. polymorpha* and angiosperms like *A. thaliana* to investigate gene function across over 450 million years of independent evolution. Being a bryophyte with a haploid-dominant life cycle means *M. polymorpha* is a favorable subject for genome editing (Sugano, Shirakawa et al. 2014).

### 1.4.2 Morphology and life cycle of *M. polymorpha*

Typical for bryophytes, *M. polymorpha* has a gametophyte dominant life cycle in contrast to angiosperms with their sporophyte dominant life cycle. In recent years, many publications have described the life cycle and morphology of *M. polymorpha* in great detail (Shimamura 2015, Bowman, Araki et al. 2016, Kato, Yasui et al. 2020, Yamaoka, Inoue et al. 2021, Bowman 2022, Bowman, Arteaga-Vazquez et al. 2022), which is why I will only present a short overview here.

The haploid gametophore starts out as a unicellular spore, that, through germination, gives rise to filamentous protonema tissue, which continues to grow into the mature thallus, the main plant body. *M. polymorpha* is a dioecious plant and the thalli of male and female plants are indistinguishable in their vegetative form. In the right environmental conditions, however, induction of sexual reproduction occurs and antheridiophores and archegoniophores are formed by male and female plants, respectively. These umbrella-like structures carry the antheridia and archegonia and allow easy identification of sex. Through contact with water, flagellated sperm cells are released from the antheridia located at the dorsal surface of the antheridiophores of male plants. The sperm cells need water in order to swim to the archegonia located on the ventral side of the archegoniophores of female plants. After fertilization of the egg cell, the diploid sporophyte develops in the archegonium, and haploid spores are formed through meiotic cell divisions. After sporogenesis, the sporophytic capsule dries and the spores are released (Shimamura 2015).

The thallus of *M. polymorpha* displays a three-layer structure and clear dorsoventrality. The dorsal surface houses gemma cups and air pores and is followed by a thin layer of assimilatory tissue with so-called air chambers and mesophyll cells. The inner part of the thallus houses the storage tissue, containing parenchymatous cells and oil cells with oil bodies. Oil bodies are structures only found in liverworts and contain a plethora of secondary metabolites, many of which have not been identified in other organisms (Chen, Ludwiczuk et al. 2018). The ventral surface of the thallus gives rise to scales and unicellular rhizoids. Rhizoids are either smooth or pegged, with peg-like cell wall thickenings (Kny 1890, Kamerling 1897, Schiffner 1909). The former usually anchor the thallus to the substrate and are involved in nutrient uptake, while the latter function as external water conducting tissue, since *M. polymorpha*, as a bryophyte, does not have vasculature (McConaha 1941, Duckett, Ligrone et al. 2014).

The gemma cups found on the dorsal surface of the thallus are cup-like structures that usually appear on each bifurcation of the mature thallus and continuously produce gemmae, disk-like

units of asexual propagation produced by both male and female individuals (Shimamura 2015, Kato, Yasui et al. 2020). Gemma cup development in *M. polymorpha* is dependent on the R2R3-MYB transcription factor GEMMA CUP-ASSOCIATED MYB1 (Yasui, Tsukamoto et al. 2019), an ortholog of the *A. thaliana* REGULATOR OF AXILLARY MERISTEM (RAX), which regulates axillary meristem formation (Keller, Abbott et al. 2006). The corresponding gene MpGCAM1 is expressed in developing gemma cups and *gcam1* mutants lack gemma cups. Recent research could show that cytokinin (CK) is required for gemma cup formation in *M. polymorpha* (Aki, Mikami et al. 2019, Aki, Nishihama et al. 2019).

Knockouts of the cytokinin response regulators MpRRA and MpRRB in *M. polymorpha* result in the total absence of gemma cups (Aki, Mikami et al. 2019). Overexpression of the cytokinin inactivating cytokinin oxidase CKX2 in *Marchantia* leads to reduced endogenous cytokinin levels and a strong decrease in the number of gemma cups. Interestingly, overexpressors of MpCKX2 also display an irregular distribution of air pores and mutants of MpRRB showed air pores with irregular, non-circular shapes, while mutants of MpRRA do not exhibit any phenotypes related to air pores (Aki, Nishihama et al. 2019).

*Marchantia* does not have stomata for gas exchange, but analogous air pores are present on the dorsal epidermis. In contrast to stomata, the air pores cannot actively open and close, but remain constantly open. Air pores facilitate gas exchange for the underlying air chambers containing assimilatory filaments called chlorenchyma (Green and Snelgar 1982). The chlorenchyma contains cells with many chloroplasts and is the main photosynthetically active tissue of the plant. Air pores are derived from intercellular spaces, and little is known about the genetic basis of their development. The E3 ubiquitin ligase NOPPERABO1 is required for their formation in *M. polymorpha*. MpNOP1 encodes a plasma membrane localized Plant U-box (PUB) E3 ubiquitin ligase with tandem ARMADILLO (ARM) repeats. Mutants of *NOP1* exhibit total absence of both air pores and air chambers, even though the periclinal and anticlinal cell divisions in the apical region generating protodermal and subprotodermal cell layers occur normally (Ishizaki, Mizutani et al. 2013).

## **1.5 *Physcomitrium patens***

### **1.5.1 Role of *P. patens* as a model organism**

The use of *P. patens* in genetic studies is not new, in fact it has been used in early genetic experiments (Cove 2005). Due to its potential for highly efficient homologous recombination (Schaefer and Zrýd 1997, Kamisugi, Schlink et al. 2006), it was the first non-angiosperm chosen for genome sequencing in 2008 (Rensing, Lang et al. 2008). Like *M. polymorpha*, its phylogenetic position and the available resources make it an interesting candidate for comparative studies.

### **1.5.2 Morphology and life cycle of *P. patens***

*P. patens* is a moss and, thus, belongs to the bryophytes. The life cycle, biology and systematic classification of *P. patens* have been extensively described in several publications (Cove 2005, Rensing, Goffinet et al. 2020). Like *M. polymorpha*, the dominant phase of its life cycle is the haploid gametophyte. The life cycle starts with the germination of a haploid spore, giving rise to the tip-growing, filamentous protonema. Protonema is the multicellular, fast-growing, juvenile phase of the gametophyte and is composed of chloronema and caulonema cells. Chloronema cells are the first to emerge after spore germination and contain many chloroplasts. They divide every 22 - 26h and, in contrast to caulonema, the newly formed cell plate is perpendicular to the side walls (Schween, Gorr et al. 2003). After several days, the apical cell transitions into a caulonemal cell with fewer chloroplasts in which the new cell walls are obliquely oriented (Reski 1998). Caulonema cells divide every 6-8h and form buds, specialized branch initials giving rise to the foliate gametophores (Schween, Gorr et al. 2003). The gametophyte of *P. patens* consists of a stem and leaf-like phyllids that, unlike true leaves, lack vasculature. In contrast to *M. polymorpha*, *P. patens* is monoecious and thus carries both archegonia and antheridia on the same gametophore. The male gametes, spermatozoids, are biflagellate and depend on water to reach the egg cell within the archegonium. After fertilization, the diploid zygote develops into the sporophyte, consisting of a stomata-bearing capsule in which the haploid spores are formed through meiotic cell divisions. When the mature capsule opens, the unicellular spores are released (Cove 2005, Rensing, Goffinet et al. 2020).

## 1.6 Plant cell walls

In the course of my project, I discovered a novel phenotype of *b-gata* mutants from *A. thaliana* related to stem stability. This phenotype was reminiscent of plants defective in primary- or secondary cell wall formation, which is why I will give a short overview of plant cell walls here.

The walls of plant cells largely determine the shape of cells and eventually of the plants themselves. Most plant cells are surrounded by primary cell walls outside of the plasma membrane and adjacent cells are glued together by the middle lamella. The primary cell wall is mainly made of cellulose, consisting of  $\beta$ -1,4-linked glucose molecules. Cellulose forms microfibrils that act as the main load-bearing cell wall component. The cellulose microfibrils are cross-linked by glycans and thus form a mesh as the basic structure of the primary cell wall (Taylor-Teeple, Lin et al. 2015).

A second network within the primary cell wall is formed by pectins, a heterogeneous mixture of polysaccharides rich in D-galacturonic acid. The major components of pectins are homogalacturonan, a linear chain of galacturonic acid molecules, and rhamno-galacturonan, a linear chain of alternating galacturonic acid and rhamnose molecules (Jones, Ougham et al. 2012). Usually, homogalacturonan is highly methylesterified after its biosynthesis, and demethylesterification is achieved by pectin demethylesterases (PME) *in muro* (Micheli 2001, Hongo, Sato et al. 2012). The demethylesterification allows homogalacturonan to form intermolecular  $\text{Ca}^{2+}$  bonds, resulting in the formation of a firm gel, or to become more vulnerable to breakdown by pectin-degrading enzymes like endopolygalacturonases, and thus affecting the mechanical properties of the primary cell wall (Pelloux, Rusterucci et al. 2007, Wolf, Mouille et al. 2009, Hongo, Sato et al. 2012). PMEs can be subdivided in two subfamilies based on the existence of a characteristic N-terminal PRO region that is present in members of type II/group 1 but absent in members of type I/group 2 (Micheli 2001). *A. thaliana* PME35 is a type I/group 2 PME that is upregulated by application of weight to the stem of adult plants and downregulated in stems placed horizontally (Yokoyama and Nishitani 2006, Koizumi, Yokoyama et al. 2009), suggesting that PME35 is involved in strengthening of the *A. thaliana* stem. *pme35* mutants are deficient in demethylesterification of methylesterified homogalacturonans in the primary cell wall, while the lignified cell walls in the interfascicular tissue and vasculature are not affected, leading to strong instability of the stem of adult plants (Hongo, Sato et al. 2012).

In order to allow for differentiation, specialized cells develop secondary cell walls required for water transport, mechanical support and as barrier against invading pathogens (Taylor-Teeple, Lin et al. 2015). Secondary cell wall composition varies among species, but mainly consists of cellulose, hemicellulose, and lignin. The term hemicellulose refers to a range of polysaccharides including xylans, glucans, and mannans involved in strengthening the cell wall through their interactions with cellulose and, in some cases, with lignin (Scheller and Ulvskov 2010).

Lignin forms a complex network of aromatic phenylpropanoids in the secondary cell wall and is required for increased stability and waterproofing e.g., in the vascular tissue (Neutelings 2011). In *A. thaliana*, the NAC transcription factors NAC SECONDARY WALL THICKENING PROMOTING FACTOR1 (NST1) and NST3 regulate the formation of secondary cell walls in inflorescence stems. The *nst1 nst3* double mutants are unable to stand upright and their stems are easily broken due to the complete absence of cellulose and lignin in stems, except for vascular vessels (Mitsuda, Iwase et al. 2007).

## 1.7 Photosynthesis

Plants are photoautotrophic organisms that rely on light to fuel their biosynthetic processes. In oxygenic photosynthesis, water and atmospheric carbon dioxide are utilized to generate sugars, while oxygen is produced as a by-product through the oxidation of water. Photosynthesis can be divided in two parts, the light-dependent and carbon reactions (formerly “light-independent reactions”). The light-dependent reactions take place in the thylakoid membranes of chloroplasts and require the photosystems I and II and the electron transport chain for the generation of ATP and NADPH, while the carbon reactions take place in the chloroplast stroma and result in the biosynthesis of carbohydrates through carbon fixation (Figure 1.6). In the course of my project, I uncovered regulation of photosynthesis-associated genes in *B-GATA* mutants, thus I will shortly introduce the light-dependent and carbon reactions in the following.

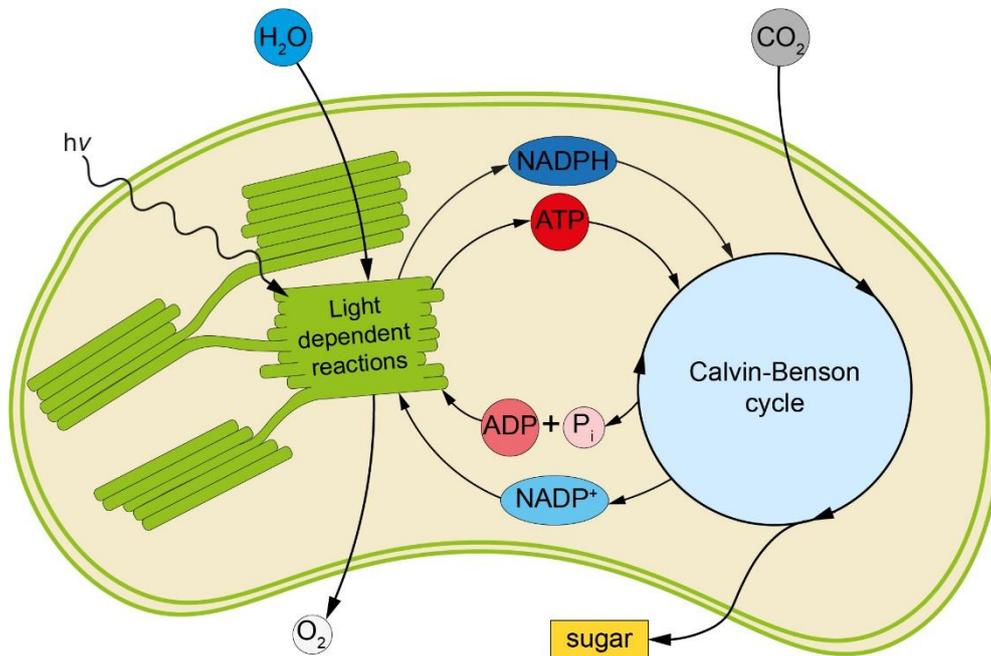
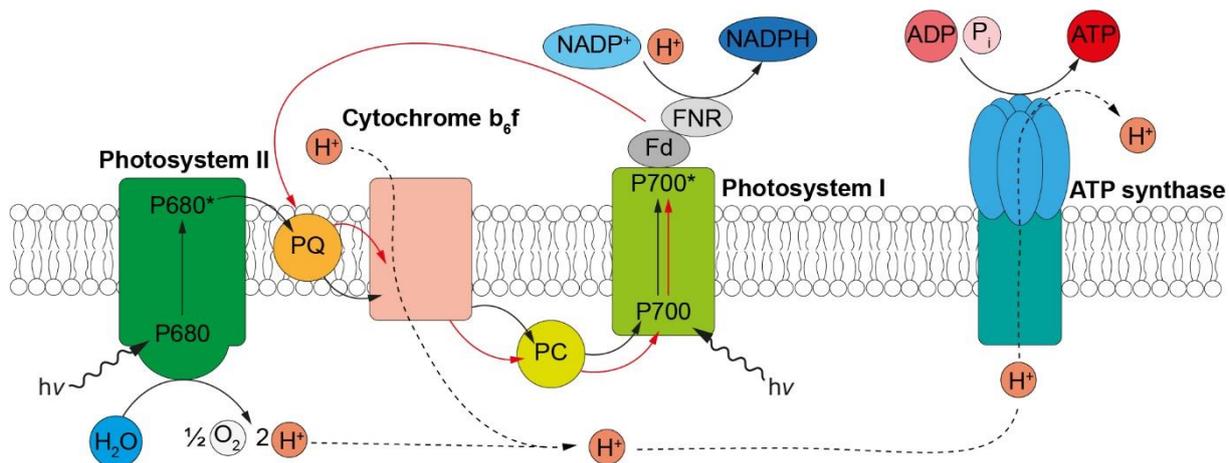


Figure 1.6: Overview of photosynthesis in the chloroplasts of plants. Adapted from Jones, Ougham et al. 2012.

### 1.7.1 Light-dependent reactions of photosynthesis

Photosystems consist of the reaction centre complex composed of several proteins and the light-harvesting complex, antenna-like structures required to bundle incoming light energy and to transfer it to the reaction centre. Light-harvesting complexes and reaction centres bind chlorophyll molecules. On illumination, the chlorophyll molecules absorb incoming photons, leading to conversion of light energy to kinetic energy in the form of excited electrons. This state of excitation is unstable, and the return to the original state requires one of three possible dissipation mechanisms: resonance transfer, meaning the transfer of energy to a nearby molecule and exciting its electron; oxidation with the resulting reduction of a second compound; or emission of fluorescent light (Kim and Gadd 2008). Within photosystems, the energy is transferred from the light-harvesting antennas to the reaction centres via resonance transfer.



**Figure 1.7: Simplified representation of light-dependent reactions of photosynthesis.** Black arrows represent linear electron flow, red arrows represent cyclic electron flow. Proton flow is shown in dashed arrows. PQ, plastoquinone; PC, plastocyanin; Fd, ferredoxin; FNR, ferredoxin-NADP<sup>+</sup> reductase; *hν*, light energy.

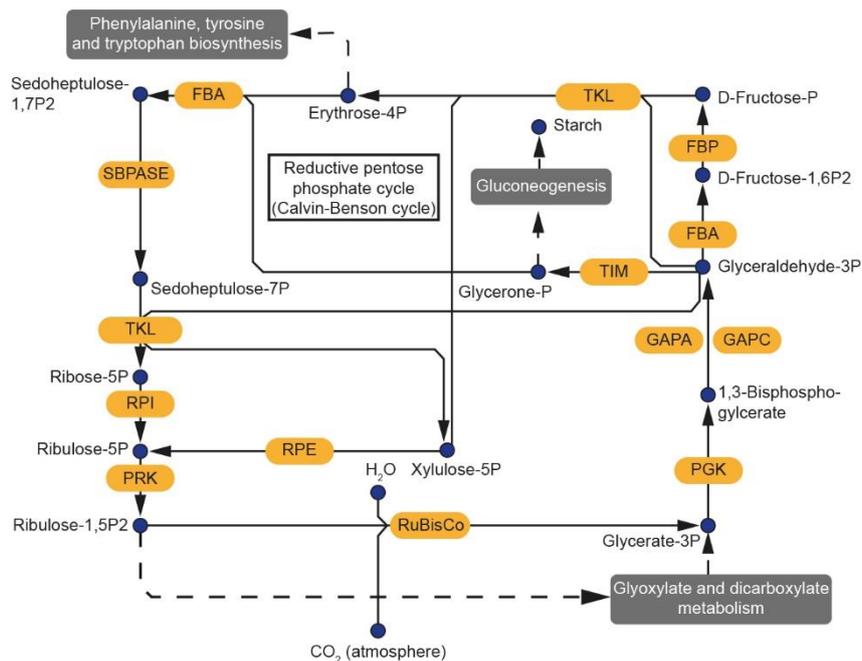
The reaction centres predominantly bind chlorophyll *a*, while the antenna complexes usually contain chlorophyll *b* in the outer parts and chlorophyll *a* in the inner parts. Additionally, the reaction centres contain  $\beta$ -carotene to quench excess energy, while the light-harvesting complexes (LHCs) contain xanthophylls which contribute to light harvesting and have structural functions (Cuttriss, Cazzonelli et al. 2011). At the reaction centre of PSII, the oxygen evolving complex with its  $\text{Mn}_4\text{CaO}_5$ -cluster is located and oxidation of water takes place in the thylakoid lumen, fuelled by the energy delivered via resonance transfer. This is a critical step in the light-dependent reactions of photosynthesis and yields electrons and protons required for the electron transport chain and building of the proton gradient, respectively. The reaction centre of PSII contains a special pair of chlorophyll *a* molecules called P680 and these become oxidized by the released electrons. They become reduced again by oxidizing electron donors of the electron transport chain. In this way, electrons are transferred via plastoquinone (PQ), the cytochrome *b6f* complex, plastocyanin (PC), and the other components of the electron transport chain to P700, the reaction centre of PSI, in a process called linear electron transport. Apart from transferring electrons from plastoquinone to plastocyanin, the cytochrome *b6f* complex functions as an electron-propelled proton pump involved in generating a proton gradient, resulting in acidification of the thylakoid membrane. P700 is excited by light and thus, the oxidized chlorophyll *a* molecules can transfer the electrons to membrane bound iron-sulfur-proteins and across the membrane to ferredoxin (Fd) in the thylakoid stroma. Eventually, ferredoxin supplies the electrons to reduce NADP together with protons ( $\text{H}^+$ ) to NADPH via ferredoxin-NADP<sup>+</sup> reductase (FNR). Alternatively, electrons are transported back from ferredoxin to P700 via plastoquinone and the cytochrome *b6f* complex in a process called cyclic electron transport, in contrast to the previously described linear electron transport. Cyclic electron transport is required to fuel the cytochrome *b6f* complex and thus, shuffle more protons from the thylakoid stroma to the lumen (Figure 1.7). The resulting electrochemical potential is called proton-motive force and is required to power ATP synthesis by the F-ATPase (Nelson and Ben-Shem 2004, Hasan, Yamashita et al. 2013). The main products of the light-dependent reaction are NADPH and ATP, and they are required for the carbon fixation in the carbon reactions (Herrmann 1999).

In angiosperms like *Arabidopsis thaliana*, the reaction centre of PSII consists of two core proteins, D1 and D2, encoded by the chloroplastic genes *PHOTOSYSTEM II REACTION CENTER PROTEIN A (AtPSBA)* and *PHOTOSYSTEM II REACTION CENTER PROTEIN D (AtPSBD)*, respectively (Supplemental Table S5.5). D1 and D2 are flanked by intrinsic light-harvesting proteins CP43 and CP47 encoded by the chloroplastic genes *PHOTOSYSTEM II REACTION CENTRE PROTEIN C (AtPSBC)* and *PHOTOSYSTEM II REACTION CENTRE PROTEIN B (AtPSBB)* (Barber 2002, Ferreira, Iverson et al. 2004). While CP43 and CP47 bind

14 and 16 chlorophyll *a* molecules, respectively, the majority of chlorophylls are bound by the antenna proteins of the light-harvesting complex II. These consist of the trimeric Lhcb1-3 proteins and bind 12-14 chlorophyll-*a* and -*b* molecules, as well as up to four carotenoid molecules each (Wang, Fujiyoshi et al. 1994, Liu, Yan et al. 2004). Lhcb4-7 are monomeric, minor light harvesting proteins and mediate the transfer of excitation from Lhcb1-3 to the reaction centre proteins. In contrast to the reaction centre proteins, the light harvesting proteins are encoded by nuclear genes. Similar to PSII, PSI consists of a reaction centre and a light-harvesting antenna complex. The reaction centre of PSI consists of the nuclear encoded subunits PsaA-PsaL, PsaN and PsaO. PsaA and PsaB form a heterodimer and act as primary electron acceptors binding the P700 chlorophyll pair. The other subunits are required for various other functions: PsaC binds two iron-sulfur clusters as terminal components of the electron transfer chain, PsaC-PsaE are required for the association with ferredoxin, PsaF associates plastocyanine and PsaF, PsaG, PsaJ and PsaK coordinate the LHCI (Nelson and Ben-Shem 2004). Interestingly, the reaction centre proteins of PSI are encoded by the nuclear- (*psaN*, *psaF*, *psaD*, *psaE*), as well as the plastid genome (*psaA*, *psaB*, *psaC*).

### 1.7.2 Carbon reactions of photosynthesis

Photoautotrophic organisms use the main products of the light-dependent reactions, ATP and NADPH, for the fixation of carbon in the so-called carbon reactions, formerly also known as “light-independent reactions”. In plants, these consist of the Calvin-Benson cycle, resulting in the formation of sugar in the form of glyceraldehyde 3-phosphate (GAP).



**Figure 1.8: Overview of the Calvin-Benson cycle.** Enzymes are shown in orange boxes, reactions as solid arrows and metabolites as blue dots. Pathways branching off the carbon fixation are indicated by dotted arrows and grey boxes. Pathway adapted from KEGG. See Supplemental Table S5.6 for abbreviations.

The Calvin-Benson cycle is a complex pathway consisting of 13 enzymatic reactions that can be divided in the three phases carboxylation, reduction, and regeneration (Figure 1.8; Supplemental Table S5.6). The carboxylation is catalysed by ribulose-1,5-bisphosphate carboxylase-oxygenase (RuBisCO) and yields 3-phosphoglycerate (3-PGA) through the addition of CO<sub>2</sub> to ribulose-1,5-bisphosphate (RuBP). Following the carboxylation, GAP, the final product of the Calvin-Benson cycle, is generated from 3-PGA in the reduction phase. The reduction phase consists of two enzymatic reactions catalysed by phosphoglycerate kinases (PGK) and a glyceraldehyde-3-phosphate dehydrogenases (GAPA and GAPC), for which the products of the light-dependent reaction, ATP and NADPH, are required. After the reduction

phase, the regeneration phase takes place. This phase comprises ten of the 13 enzymatic reactions of the Calvin-Benson cycle and results in the regeneration of ribulose-1,5-bisphosphate, so it can serve as acceptor molecule in the RuBisCO-mediated carboxylation (Raines 2003). Notably, all enzymes of the Calvin-Benson cycle are encoded by nuclear genes, except the large subunit of RuBisCO (*rbcL*), which is encoded by a chloroplastic gene.

## 1.8 Excess light energy causes photooxidative damage

Plants balance the light energy required for photosynthesis with the harmful properties of excess light that can result in the formation of singlet oxygen and other reactive oxygen species (ROS) (Pospíšil 2016). ROS are formed through the reactions of excited chlorophyll molecules with atmospheric oxygen and their accumulation can lead to damage of the photosynthetic apparatus, mainly of PSI and PSII (Eberhard, Finazzi et al. 2008, Li, Gao et al. 2009). Damage of the photosynthetic machinery through ROS leads to photoinhibition and the D1 protein, located at the reaction centre of PSII, is especially vulnerable to photooxidative damage. While repair mechanisms for PSII exist (Tikkanen, Mekala et al. 2014), such mechanisms seem to be lacking for the repair of PSI. Thus, photooxidative damage of PSI requires resynthesis of the entire supercomplex (Alboresi, Storti et al. 2019).

In the course of evolution, plants have developed a range of mechanisms to cope with exposure to light intensities that exceed their photosynthetic capacities. These include xanthophyll-mediated non-photochemical quenching (NPQ), which is the harmless dissipation of excess energy as heat, accumulation of antioxidants, and activation of antioxidant enzymes (Niyogi, Grossman et al. 1998, Bailey and Grossman 2008, Cazzonelli 2011, Cuttriss, Cazzonelli et al. 2011). Additionally, plants produce anthocyanins and other flavonoids. These red-to-blue non-photochemically active pigments are involved in the absorption of excess radiation, which would otherwise induce the formation of ROS (Steyn, Wand et al. 2002).

The *B-GATA* mutants of *M. polymorpha* and *A. thaliana* exhibited constitutive high-light stress phenotypes, thus I will shortly introduce the typical responses to high-light stress here.

### 1.8.1 EARLY LIGHT INDUCED PROTEINS respond to high-light stress

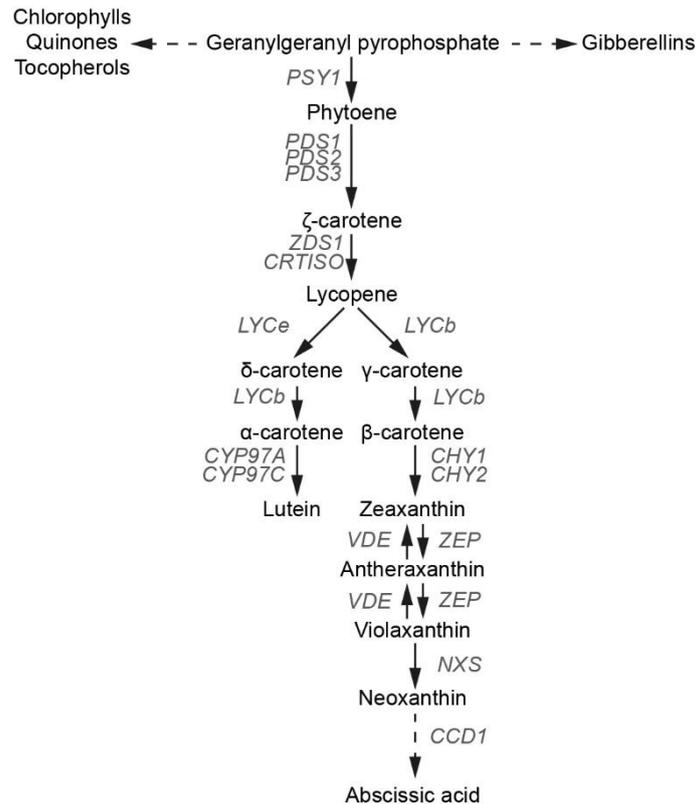
In order to reduce the formation of ROS and to prevent the resulting photooxidative damage, plants reduce the production of photosynthesis-related proteins like light-harvesting complexes and photosystem components, and simultaneously decrease the biosynthesis of chlorophylls and promote their degradation (Pöttter and Kloppstech 1993). At the same time, angiosperms increase the production of EARLY LIGHT INDUCED PROTEINs (ELIPs), a group of LHC-related chlorophyll AB-binding proteins. In *A. thaliana*, ELIPs accumulate transiently in plants exposed to high light intensities. In the *chaos* mutant, posttranslational targeting of light-harvesting complex-type proteins to the thylakoids is impaired and ELIPs cannot accumulate in response to high light (Hutin, Nussaume et al. 2003). The *chaos* mutant suffers severely from photooxidative damage resulting in leaf bleaching when exposed to high light intensities, whereas constitutive expression of *ELIPs* in the *chaos* background restored ELIP accumulation and photo-tolerance to the wild-type level. Notably, the photo-tolerance of the transgenic lines is directly proportional to *ELIP* expression levels, thus suggesting a photoprotective role of ELIPs (Hutin, Nussaume et al. 2003). In the *elip1 elip2* double mutant a photosensitive phenotype could not be observed, hinting at the existence of compensatory processes beyond *ELIP* gene expression (Rossini, Casazza et al. 2006). Similar to LHCS, ELIPs bind chlorophyll *a* and xanthophylls (Adamska, Roobol-Bóza et al. 1999) and ELIP-bound chlorophylls should be able to safely dissipate energy to the bound xanthophylls. However, this does not seem to be the main function of ELIPs, since NPQ is similar for wild-type, *chaos* mutants and the transgenic lines expressing *ELIPs* in the *chaos* background (Hutin, Nussaume et al. 2003). The same observation applies to the *elip1 elip2* double mutants, confirming that ELIPs do not modulate NPQ in high light (Rossini, Casazza et al. 2006).

Interestingly, overexpression of *ELIP2* in *A. thaliana* resulted in decreased chlorophyll accumulation and reduction in the number of photosystems in the thylakoid membranes (Tzvetkova-Chevolleau, Franck et al. 2007). Due to their chlorophyll binding capacities, ELIPs were thought to act as scavengers for free chlorophyll molecules (Hutin, Nussaume et al. 2003) or modulate chlorophyll biosynthesis in order to prevent accumulation of free chlorophyll, and

thus prevent photooxidative stress (Tzvetkova-Chevolleau, Franck et al. 2007). However, their function has not been fully elucidated to date.

### 1.8.2 Carotenoids are required for photoprotection in plants

Carotenoids are orange-yellow pigments that are present in all photosynthetically active organisms and absorb light between 400-500 nm (Cuttriss, Cazzonelli et al. 2011). Carotenoids act as accessory pigments in photosynthesis where they contribute to light harvesting in the blue region of the spectrum and help in transferring energy to chlorophylls. Their main role, however, is in photoprotection.



**Figure 1.9: The biosynthetic pathway of carotenoids in plants.** Intermediates are indicated by black text; enzymes are indicated by grey italic text; enzymatic reactions are shown as black arrows; dashed arrows indicate additional steps not shown here. The four main products are shown in coloured boxes.

Chemically, carotenoids are tetraterpenes and are synthesised via the isoprenoid pathway in chloroplasts. Their biosynthesis starts with the synthesis of phytoene from geranylgeranyl pyrophosphate (GGPP) catalysed by the PHYTOENE SYNTHASE PSY1 (Figure 1.9). Phytoene is converted to  $\zeta$ -carotene and then lycopene via the actions of phytoene desaturases PDS1-3, the  $\zeta$ -carotene-desaturase ZDS1, and the carotenoid isomerase CRTISO. After the synthesis of lycopene, the pathway divides into the formation of  $\alpha$ - and  $\beta$ -carotene and their respective derivatives.  $\alpha$ -carotene is synthesized from lycopene via  $\delta$ -carotene through the actions of the lycopene cyclases LYCe and LYCb and is in turn converted to lutein by the cytochromes CYP97A and CYP97C. For the formation of  $\beta$ -carotene, lycopene is first converted to  $\gamma$ -carotene and then to  $\beta$ -carotene through LYCb. Xanthophylls can be synthesized from  $\beta$ -carotene through the  $\beta$ -hydroxylases CHY1-2, giving rise to zeaxanthin. Under normal light conditions, the ZEAXANTHIN EPOXIDASE (ZEP) can convert zeaxanthin first to antheraxanthin and then to violaxanthin, while the reverse reaction catalysed by the VIOLAXANTHIN DE-EXPOXIDASE VDE is taking place under high-light conditions (Pfündel, Renganathan et al. 1994). These reactions are known as the xanthophyll cycle, and they are a means of dissipating excess light energy as heat. In high-light, the zeaxanthin pool is enriched and functions as quenching agent for excited chlorophylls, thus facilitating NPQ

(Jahns and Holzwarth 2012). In the absence of high light intensities, neoxanthin synthase NXS converts violaxanthin to neoxanthin, which can be cleaved by carotenoid cleavage dioxygenases (CCD) and 9-*cis*-epoxycarotenoid dioxygenases (NCED). The cleavage products undergo further modification and form apocarotenoids with various functions. One of these is abscisic acid, a plant hormone involved in stress responses (Cuttriss, Cazzonelli et al. 2011).

## 1.9 Aims of my project

In angiosperms, the family of B-GATAs is expanded and can be subdivided based on the existence of a HAN- or an LLM-domain. As described above, the members of both subfamilies have distinct functions. The two domains are characteristic for B-GATAs and have not been identified in any other proteins, suggesting that they might have specialized functions. So far, however, their biochemical functions remain elusive. In contrast to angiosperms, the family of B-GATAs in bryophytes is small and bryophyte B-GATAs contain both HAN- and LLM-domains, a constellation unique to bryophytes.

The aims of my project were to find the origin of HAN- and LLM-domains in the plant lineage and to investigate when the division of the two distinct B-GATA subfamilies occurred. Additionally, I wanted to analyse the functions of B-GATAs from the bryophytes *P. patens* and *M. polymorpha* in order to understand the evolutionary divergence or conservation of the B-GATA factors, as well as the HAN- and LLM-domains.

## 1.10 Publications created as part of the doctorate

The following publications were created during my doctorate:

Schwechheimer, C., Schröder, P. M., & Blaby-Haas, C. E. (2022). "Plant GATA factors: their biology, phylogeny, and phylogenomics." Annual Review of Plant Biology, 73, 123-148.

Schröder, P. M., Hsu, B.-Y., Gutsche, N., Winkler, J. B., Hedtke, B., Grimm, B., & Schwechheimer, C. (2023). "B-GATA factors are required to repress high-light stress responses in *Marchantia polymorpha* and *Arabidopsis thaliana*." Plant, Cell & Environment. (in press)<sup>1</sup>

---

<sup>1</sup> All figures of from this publication are licensed for reproduction in print and electronic form in this dissertation under license number 5560081279909 by John Wiley and Sons.

## 2 Results

### 2.1 GATA evolution in land plants

While GATAs have been described in a range of angiosperm species, there are hardly any reports on GATAs in early diverging plant lineages such as bryophytes, lycophytes and ferns (Chen, Shao et al. 2017, Gupta, Nutan et al. 2017, Wang, Yang et al. 2019, Zhang, Zou et al. 2019, An, Zhou et al. 2020, Duan, Zhang et al. 2020, Zhang, Huang et al. 2020, Manzoor, Sabir et al. 2021, Peng, Li et al. 2021, Zhu, Zhai et al. 2022).

In order to identify GATAs in these early diverging lineages, I searched the genomes or transcriptomes of the bryophytes *Marchantia polymorpha* and *Physcomitrium patens*, the lycophyte *Selaginella moellendorffii*, the fern *Salvinia cucullata*, and the streptophyte alga *Chara braunii* for proteins containing the GATA-type zinc finger domain as defined by the Pfam database entry PF00320 (Mistry, Chuguransky et al. 2021).

*A. thaliana* encodes 30 GATAs but I found fewer GATAs in the other species: I identified six GATAs in *M. polymorpha*, 15 in *P. patens*, eight in *S. moellendorffii*, 16 in *S. cucullata* and six in *C. braunii* (Table 2.1). Phylogenetic analysis of the GATAs using only the sequence of the zinc finger domain for the alignment recovered the classification introduced previously (Figure 2.1A; Reyes et al., 2004). The classes A-D were differently represented in these species. Class A was the largest in *A. thaliana* and *S. cucullata* with 13 and nine proteins, respectively. In *P. patens* the classes A-C were equally big with four members each. *S. moellendorffii* had three GATAs belonging to the classes A and B each, but, interestingly, no class D GATA. In the case of *C. braunii*, no class A GATAs could be found, but members of the classes B-D were present. One GATA from *C. braunii* clustered closely to AtGATA14, which was considered a member of class A, but a highly conserved motif close to the zinc finger suggests that it does not belong to this class but is phylogenetically related to it.

The six GATAs from *M. polymorpha* could be subdivided in the classes A-D and the classes A, B, and D contained a single *M. polymorpha* GATA each (Mp1g03950 in class A, Mp7g03490 in class B, and Mp5g09130 in class D), while class C contained three GATAs (Mp1g05310, MpUg00060 and MpVg00300). Two of these three GATAs from class C, MpUg00060 and MpVg00300, are encoded by genes located on the male and female sex chromosomes, respectively. Their genomic location implies roles in sex determination; however, their functions are still unknown.

**Table 2.1: Numbers of GATA factors in *A. thaliana*, *M. polymorpha*, *P. patens*, *Selaginella moellendorffii*, *Salvinia cucullata* and *Chara braunii*.**

	Class A	Class B	Class C	Class D	without class	total
<i>A. thaliana</i>	<b>13</b>	11	3	2	1	30
<i>M. polymorpha</i>	1	1	<b>3</b>	1	0	6
<i>P. patens</i>	<b>4</b>	<b>4</b>	<b>4</b>	3	0	15
<i>S. moellendorffii</i>	<b>3</b>	<b>3</b>	2	0	0	8
<i>S. cucullata</i>	<b>9</b>	5	1	1	0	16
<i>C. braunii</i>	0	<b>3</b>	1	1	1	6

Across species, GATAs belonging to the same class share certain characteristics, e.g., the approximate location of the zinc finger domain and existence of conserved domains. The domains were identified using InterProScan (Jones, Binns et al. 2014) or ScanProsite (De Castro, Sigrist et al. 2006) and domain identifiers are given as either Pfam-ID (PF), InterPro-ID (IPR) or PROSITE-ID (PS; Figure 2.1B).

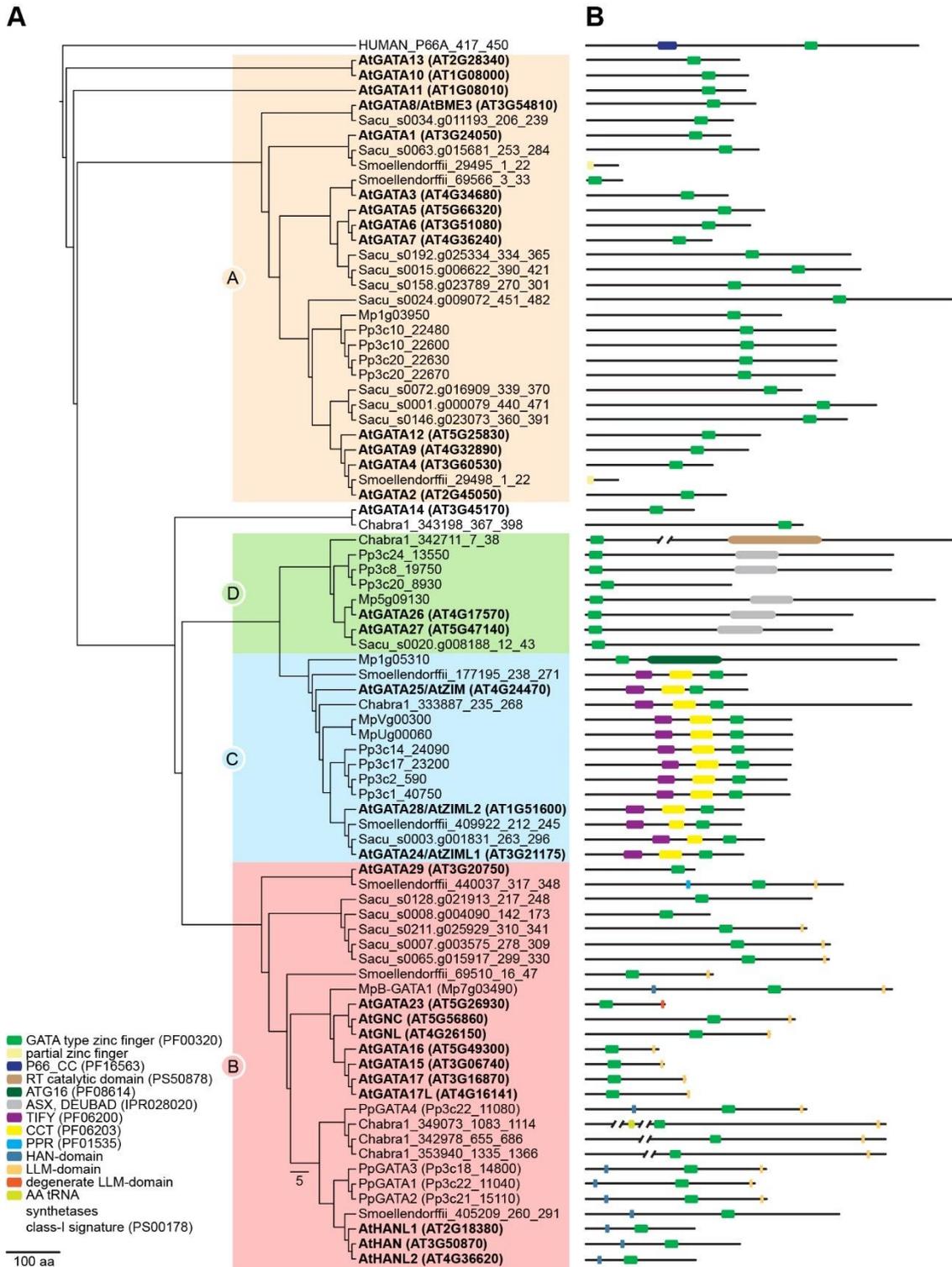
For most members of class A, the zinc finger domain is located in the C-terminal third of the protein, except for two GATAs from *S. moellendorffii*, that only contain N-terminally located partial zinc fingers. GATAs from class A do not contain domains other than the zinc finger. Class B GATAs contain a central zinc finger domain, as well as N-terminal HAN-domains or C-terminal LLM-domains. B-GATAs in *A. thaliana* and other angiosperms are frequently subdivided in HAN-domain B-GATAs and LLM-domain B-GATAs based on the existence of either of the two domains and members of the two subclasses have distinct biological functions.

Interestingly, B-GATAs from the bryophyte species *M. polymorpha* and *P. patens* contain both HAN- and LLM-domains on the same protein, a constellation unique to bryophytes. Except for one protein from *S. moellendorffii* (Smoellendorffii\_440037\_317-348), which contained a pentatricopeptide repeat (PPR; PF01535) domain, and one protein from *C. braunii* (Chabra1\_349073\_1083-1114), which contained an aminoacyl-transfer RNA synthetases class-I signature (PS00178), the other B-GATAs did not contain additional domains.

The members of class C contain TIFY (PF06200) and *CONSTANS*, *CO-like*, and *TOC1* (CCT; PF06203) domains, except Mp1g05310, which contains an AUTOPHAGY-RELATED PROTEIN 16 domain (ATG16; PF08614). The zinc finger of class C GATAs is located in the C-terminal third of the protein in close proximity to the CCT domain, except for Mp1g05310 where it is located close to the N-terminus in front of the ATG16 domain.

For all members of class D the zinc finger is located N-terminally and the majority of GATAs in this class contain an ASX DEUBiquitinase Adaptor (DEUBAD, ASHX; IPR028020) domain. One D-GATA from *P. patens*, Pp3c20\_8930, and one from *S. cucullata*, Sacu\_s0020.g008188\_12-43, do not contain domains other than the zinc finger and one D-GATA from *C. braunii*, Chabra1\_342711\_7-38, contains an RT catalytic domain (PS50878) instead of the DEUBAD, ASHX domain.

Overall, the GATA classes proposed previously for angiosperms (Reyes, Muro-Pastor et al. 2004) appear to be conserved across the plant lineage. Few exceptions are e.g., the missing D-GATAs in the case of *S. moellendorffii* or the missing A-GATAs in the case of *C. braunii*. The number of GATA transcription factors in the species analysed here varies from six to thirty and the differences are likely due to recent whole genome duplications. Most of our knowledge about the functions of GATA transcription factors comes from studies in *A. thaliana*, but due to their involvement in many aspects of development and physiology they are interesting candidates as players in the evolution of land plants and the water to land transition.



**Figure 2.1: GATAs from *A. thaliana* and early diverging plant species can be classified based on their zinc finger domain.**  
**A** Phylogeny of GATAs from *Arabidopsis thaliana* (At), *Marchantia polymorpha* (Mp), *Physcomitrium patens* (Pp), *Selaginella moellendorffii* (Smoellendorffii), *Salvinia cucullata* (Sacu) and *Chara braunii* (Chabra) based on the sequence of the GATA-type zinc finger domain as defined by PF00320. The proteins from all species can be classified according to Reyes et al., 2004 in the classes A – D, except for AtGATA14 and Chabra\_343198\_367\_398. GATAs from *A. thaliana* are shown in bold. Human GATAd2A was used as outgroup. Sequences were aligned using MUSCLE (Edgar 2004) and phylogeny construction was carried out with PhyML (Guindon, Dufayard et al. 2010) via ATGC (<http://www.atgc-montpellier.fr/phyml/>). The Q model was chosen using automatic model selection by SMS based on the Bayesian Information Criterion (Lefort, Longueville et al. 2017) and used for phylogeny construction. **B** Schematic representation of the proteins from **A** with their domains.

## 2.2 Novel functions for B-GATAs of *Arabidopsis thaliana*

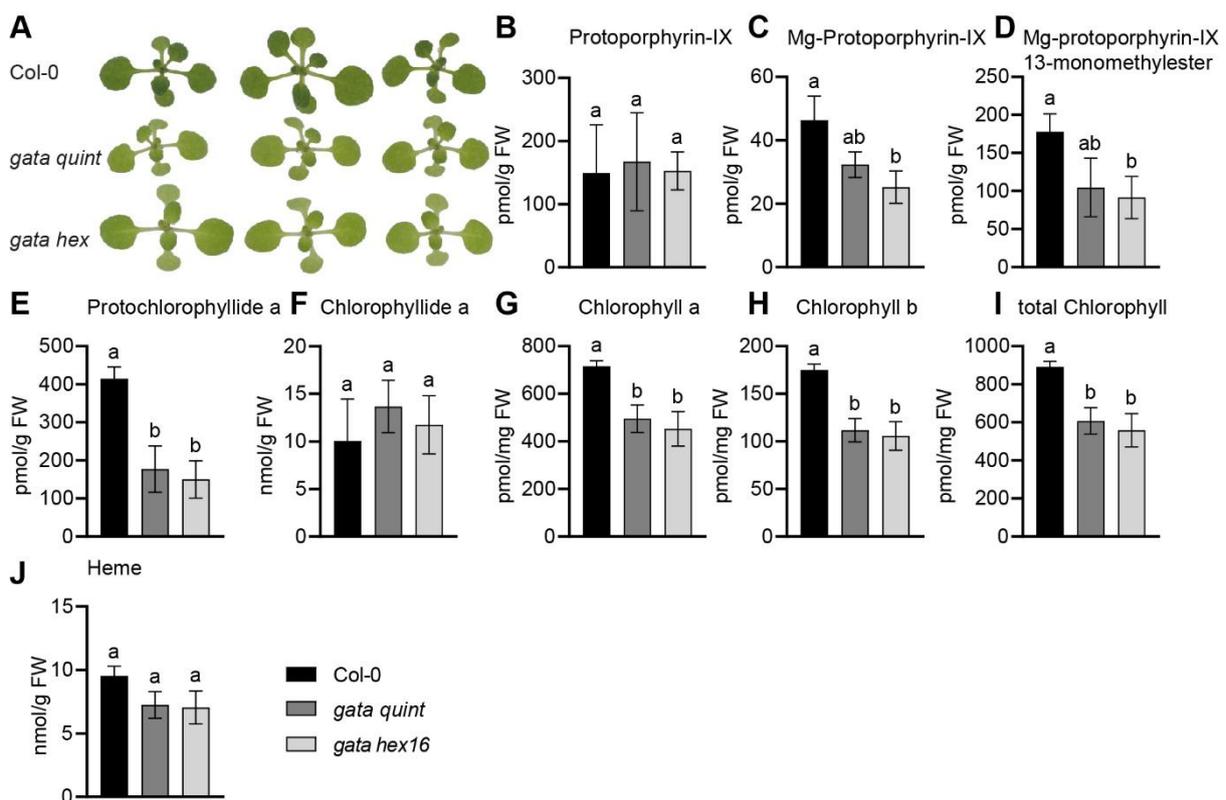
### 2.2.1 Generation of a *A. thaliana gata hexuple* mutant using CRISPR/Cas9

The six LLM-domain containing B-GATAs in *Arabidopsis thaliana* (GNC, GNL, GATA15, GATA16, GATA17 and GATA17L) have redundant roles in the regulation of multiple processes (Behringer, Bastakis et al. 2014, Behringer and Schwechheimer 2015, Ranftl, Bastakis et al. 2016). In order to fully understand their functions, it is desired to generate higher order mutants in which all LLM-domain B-GATAs are non-functional. At the start of my project, the *gata quint* mutant harbouring loss-of-function alleles of GNC, GNL, GATA17 and GATA17L and a weak allele of GATA15 was the highest order B-GATA mutant available. This was due to the incorrect annotation of T-DNA insertion lines used in the original effort to generate a *gata hexuple* (*gata hex*) mutant through crossings of the respective single mutants (Ranftl, Bastakis et al. 2016).

I have successfully generated a *gata hex* mutant in the previously available *gnc gnl gata17 gata17l* (*gata quad*) background (Ranftl, Bastakis et al. 2016). To this end, I employed CRISPR/Cas9-mediated genome editing targeting GATA15 and GATA16 with one sgRNA per gene (Wang, Xing et al. 2015). This approach yielded small InDel mutations leading to a frameshift in the coding sequence of the zinc finger domain of GATA15 and GATA16 (Supplemental Figure S5.2). Since the initial mutant screens of over 200 plants did not yield the desired *gata hex* mutant, I crossed two of the resulting homozygous *quintuple* mutants for *gata15* (*gnc, gnl, gata15, GATA16, gata17, gata17L*) or *gata16* (*gnc, gnl, GATA15, gata16, gata17, gata17L, gnc, gnl*), respectively. The crosses yielded several independent *gata hex* mutants in which all the already described phenotypes of previous *gata* mutants can be observed. The resulting *gata hex* mutants 16, 11 and 23 had a one-bp insertion for GATA15 after base pair 220 in the 2<sup>nd</sup> exon and a one-bp insertion for GATA16 after base pair 258 in the 2<sup>nd</sup> exon. Following this, I isolated the additional *gata hex* mutants 37 and 38, which had two-bp insertions for GATA15 after base pair 220 in the 2<sup>nd</sup> exon and a four-bp deletions for GATA16 after base pair 252 in the 2<sup>nd</sup> exon. In each case, the InDel mutations led to frame shifts causing premature stop codons (Supplemental Figure S5.2). The *gata hex* mutants are valuable resources for the further functional characterisation of LLM-domain containing B-GATAs and will allow me overcome the problem of functional redundancy that hindered previous studies.

## 2.2.2 Deficiency in chlorophyll accumulation is a phenotype of the *gata hex* mutant

It is known from previous research that mutants of LLM B-GATAs exhibit a pale green color, which can be explained by misregulation of the chlorophyll biosynthesis pathway, as well as chloroplast division (Naito, Kiba et al. 2007, Chiang, Zubo et al. 2012, Richter, Bastakis et al. 2013, Bastakis, Hedtke et al. 2018). In order to evaluate the greening defect of *gata hex*, chlorophyll biosynthesis intermediates were quantified in a collaboration with Dr. Boris Hedtke from the group of Prof. Dr. Bernhard Grimm (Department of Plant Physiology, Humboldt-Universität zu Berlin, Berlin, Germany). There was no visible difference in greening between the already established *gata quint* mutant and the new *gata hex* mutant, while both mutants exhibit a visible difference in comparison to the Columbia wildtype (Col-0; Figure 2.2A). High-performance liquid chromatography (HPLC) analysis of chlorophylls and chlorophyll biosynthesis intermediates from *gata quint* and *gata hex* revealed strong reductions for most measured compounds between the mutant lines and Col-0 (Figure 2.2B - J). Protoporphyrin-IX (Proto-IX), chlorophyllide a and heme contents were unchanged in both mutants compared to Col-0, while Mg-Protoporphyrin-IX (MgP) and Mg-Protoporphyrin-IX 13-monomethylester (MgPMME) contents were significantly reduced in *gata hex* but not in *gata quint*. Even though the differences between the two mutant lines are not significant, there seems to be a trend of further reduction in *gata hex* compared to *gata quint*, suggesting that *GATA15*, which is not mutated in *gata quint*, also plays a role in chlorophyll biosynthesis.

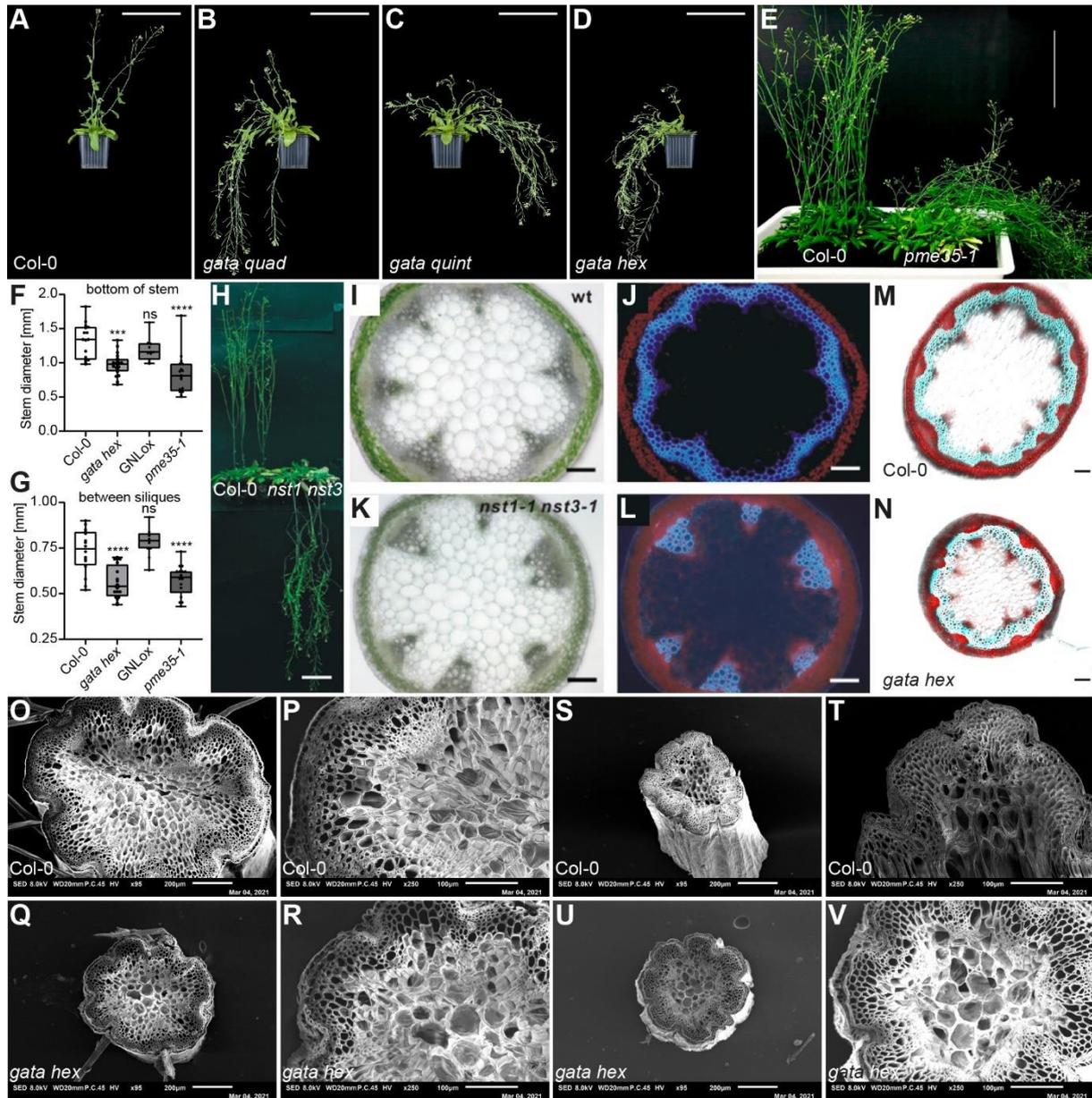


**Figure 2.2: The *gata hex* mutant is deficient in chlorophyll accumulation.** A Representative photograph of two-weeks old seedlings of *gata quint* and *gata hex16* show reduced greening in leaves and cotyledons compared to the Col-0 wild type. B - J HPLC-based measurements of chlorophylls and chlorophyll biosynthesis intermediates. Measurements were carried out by Dr. Boris Hedtke (Department of Plant Physiology, Humboldt-Universität zu Berlin, Berlin, Germany). Data sets with no statistical difference after two-way ANOVA and Tukey's HSD post-hoc test fall into one group and are labelled with identical letters. n = 5.

### 2.2.3 Decreasing stem stability is a novel phenotype of *A. thaliana* higher order *gata* mutants

Next to the slightly more pronounced greening defect (Figure 2.2A), *gata hex* shows strong instability and fragility of the entire plant, especially the stem. While the stem fragility can also be observed in *gata quad* and *gata quint* mutants, the strongest defects appear in *gata hex* (Figure 2.3A - D). A similar phenotype regarding stem stability is known from mutants of *PECTIN METHYLESTERASE35* (*PME35*; Figure 2.3E). In *pme35-1*, the demethylesterification of methylesterified homogalacturonans in the primary cell wall is suppressed, leading to fragile stems (Hongo, Sato et al. 2012). The diameter of the stem both close to the rosette of adult plants and in between the siliques was significantly reduced in *gata hex* and *pme35-1*, while stems of *GNLox* were not affected (Figure 2.3F - G).

Stem fragility can also be explained through defects in lignification of the secondary cell wall and double mutants of *NAC SECONDARY WALL THICKENING PROMOTING FACTOR1* (*NST1*) and *NST3*, the key regulators of the formation of secondary cell walls, are deficient in the secondary wall thickenings in interfascicular fibers and secondary xylem (Figure 2.3I - K) (Mitsuda, Iwase et al. 2007). Cross sections of Col-0 and *gata hex* illuminated with UV light revealed no difference in lignin autofluorescence (Figure 2.3M - N). Additionally, electron micrographs of horizontal stem sections of Col-0 and *gata hex* also revealed no differences in the composition of the different cell files in the stems both at the bottom and between the siliques (Figure 2.3O - V). This observation is reminiscent of the *pme35-1* mutant, in which the lignified cells in the secondary cell wall and the vasculature are unaffected (Hongo, Sato et al. 2012).

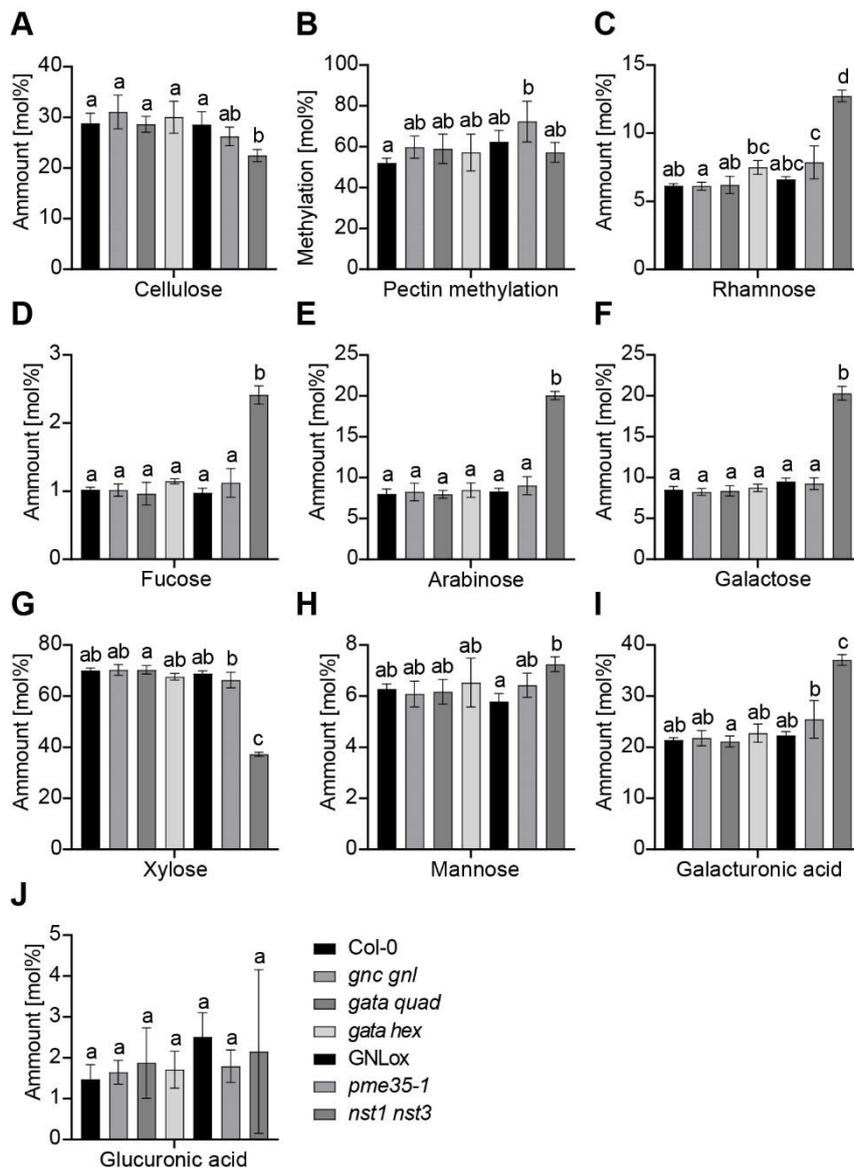


**Figure 2.3: *gata hex* mutants show decreased stem stability and stem diameter compared to Col-0.** A - D Representative photographs of 43-days old Col-0 (A), *gata quad* (B), *gata quint* (C) and *gata hex* (D) plants. Scale bar = 10 cm. E Photographs of 40-day-old Col-0 and the *pme35-1* plants. Scale bar = 10 cm. F and G Quantification of stem diameter in 43-days old plants of Col-0 (n = 16), *gata hex* (n = 20), GNLox (n = 11) and *pme35-1* (n = 17). Stem diameter was measured at the bottom of the stem (1 cm from the rosette) and close to the apical end of the stem between the siliques. H Photographs of the *nst1 nst3* double mutant compared to Col-0. Scale bar = 5 cm. I - L Cross sections of stems of Col-0 and *nst1 nst3* under UV illumination revealed lignification defects in the double mutant. Scale bar = 100  $\mu$ m. M and N Cross sections of stems of Col-0 and *gata hex* under UV illumination. Colours were amplified in post-production for better visibility. Scale bar = 100  $\mu$ m. O - V Electron micrographs of horizontal cross sections of stems from 43 days-old Col-0 and *gata hex* plants. O and P Section of the bottom part of the stem from Col-0 approximately 1cm from the rosette in x95 and x250 magnification, respectively. Q and R Section of the bottom part of the stem from *gata hex* approximately 1 cm from the rosette in x95 and x250 magnification, respectively. S and T Section of the apical part of the stem from Col-0 in between the siliques in x95 and x250 magnification, respectively. U and V Section of the apical part of the stem from *gata hex* in between the siliques in x95 and x250 magnification, respectively. Panel E was modified from Hongo, Sato et al. 2012<sup>2</sup>, panels H - L were adapted from Mitsuda, Iwase et al. 2007. One-way ANOVA comparing mutants and GNLox to Col-0: ns, not significant; \*P  $\leq$  0.05; \*\*P  $\leq$  0.01; \*\*\*P  $\leq$  0.001; \*\*\*\*P  $\leq$  0.0001.

In order to analyse the cell wall composition in *gata hex*, I collaborated with Dr. Timo Engelsdorf (Philipps-Universität Marburg, Marburg, Germany). The *pme35* mutant and the *nst1 nst3* double mutant, both established mutants with similar phenotypes regarding stem stability

<sup>2</sup> Reprinted from Hongo, S., Sato, K., Yokoyama, R., & Nishitani, K. (2012). "Demethylesterification of the primary wall by PECTIN METHYLESTERASE35 provides mechanical support to the *Arabidopsis* stem." *The Plant Cell*, 24(6), 2624-2634, by permission of Oxford University Press.

(Mitsuda, Iwase et al. 2007, Hongo, Sato et al. 2012), were used as controls. Stem fragments close to the rosette of 28 – 30 cm tall adult plants were used for the analysis. As expected, the degree of pectin methylation was significantly increased in *pme35-1*, as well as the amount of rhamnose (Figure 2.4B - C). The *nst1 nst3* double mutant displayed significant differences for all cell wall components, except for glucuronic acid, compared to Col-0 (Figure 2.4A - J). Interestingly, no statistical difference between the *gata* mutants (*gnc gnl*, *gata quad*, *gata quint* and *gata hex*) and Col-0 could be detected for any of the quantified components, suggesting that the stem instability cannot be caused by altered cell wall composition (Figure 2.4A - J).



**Figure 2.4: High performance anion exchange chromatography (HPAEC) with pulsed amperometric detection (PAD) based analysis of cell wall components from Col-0, GNLox and different mutants as specified. *pme35-1* and *nst1 nst3* were used as controls. Data from collaboration with Dr. Timo Engelsdorf (Molecular Plant Physiology, Philipps-Universität Marburg, Marburg, Germany). Data sets with no statistical difference after two-way ANOVA and Tukey's HSD post-hoc test fall into one group and are labelled with identical letters. n = 4.**

## 2.2.4 Mutants and overexpressors of *A. thaliana* B-GATAs show altered cytokinin contents

In order to test whether the phenotypes of *A. thaliana* and *M. polymorpha* GATA mutants and overexpressors were related to altered levels of endogenous cytokinins, I collaborated with Dr. Ondřej Novák (Palacky University, Olomouc, Czech Republic) for the quantification of endogenous cytokinins using high-performance liquid chromatography coupled with mass spectrometry (HPLC–MS; Supplemental Table S5.2). In *M. polymorpha*, *Mpb-gata1* displayed a strong increase in CK overall, as well as in total iP types, total tZ types and total cZ types compared to BoGa (Supplemental Table S5.3). Interestingly, *Mpb-gata1-2* did not exhibit the same increase in CK concentration. Total iP-type CKs were more abundant in both mutant alleles when compared to BoGa, but overall CK levels, total tZ-type, and total cZ-type CKs, *Mpb-gata1-2* did not show increased amounts. *MpB-GATA1ox3* also showed increased overall CK and total iP-type CK contents compared to BoGa, even though the increase was not as big as in *Mpb-gata1* in both cases. The levels of tZ- and cZ-type CKs were not increased in the overexpressor. Taken together, the CK levels in the *Mpb-gata1* mutants and *MpB-GATA1ox3* were increased compared to BoGa, while these lines exhibited opposing phenotypes regarding the number of gemma cups. Thus, the causal relationship between endogenous CK content and gemma cup formation as previously suggested did not seem to apply here (Aki, Mikami et al. 2019).

In *A. thaliana*, the analysis resolved a decrease in overall CK content in *gata hex* compared to Col-0. Total CK content was not altered in GNLox, but iP-type CKs and cZ-type CKs showed increased concentrations in GNLox, whereas tZ- and DHZ-types showed decreased CK concentrations. In *gata hex*, iP-, tZ-, and DHZ-type CK concentration was decreased, while cZ-type CK content was not altered. Often B-GATA mutants and overexpressors exhibit opposing phenotypes, but this observation only partly applied to the content of different types of cytokinins: iP-type CKs were decreased in *gata hex* and increased in GNLox, but tZ- and DHZ-types were decreased in both genotypes. Overall, the significant differences in endogenous CK content observed in *gata hex* and GNLox might explain some of the previously described phenotypes related to greening or senescence (Ranftl, Bastakis et al. 2016), but more experiments are needed to gain further insights on the impact of B-GATAs on endogenous CK levels and its downstream functions.

## 2.3 *M. polymorpha* MpB-GATA1 has developmental- and physiological functions

### 2.3.1 Increased gemma cup density and reduced plant size are phenotypes of *Mpb-gata1* mutants

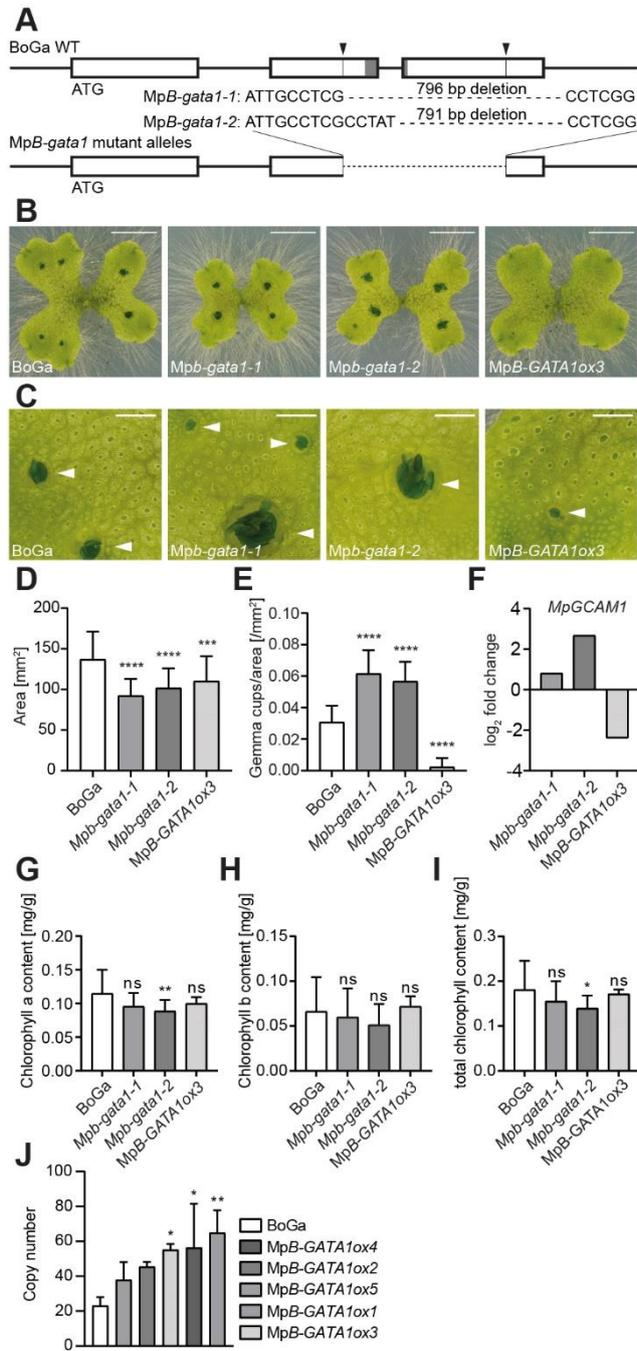
*Marchantia polymorpha* encodes a single B-GATA transcription factor MpB-GATA1 (Mp7g03490). As described above, *A. thaliana* B-GATAs can be subdivided in HAN- and LLM-domain containing B-GATAs based on the existence of the N-terminal HAN domain or the C-terminal LLM-domain, respectively. MpB-GATA1 shares characteristics of both subgroups in encoding the HAN-domain as well as the LLM-domain, a constellation unique to bryophytes (Figure 2.5).

	HAN-domain	zinc finger domain	LLM-domain
	S X - - - V D C T L S L G X	X R X C A X - C G T X K T P L W R N G P X G P K S L C N A C G I R F X K X	D F E X A A V L L M A L S C G
AtGNC	S N - - - S S S S S S I S S	I R V C S D - C N T T K T P L W R S G P R G P K S L C N A C G I R Q R K A	D E K E A A V L L M A L S Y G
AtGNL	Q A - - - S S N P S S L M S P	I R I C S D - C N T T K T P L W R S G P R G P K S L C N A C G I R Q R K K	D E K E A A I L L M A L S H G
AtGATA15	L T - - - S M D A I E H S S	K K S C A I - C G T S K T P L W R G G P A G P K S L C N A C G I R N R K K	E F E Q A A V L L M A L S Y A
AtGATA16	- - - R A E D M I E Q N N	K K T C A D - C G T S K T P L W R G G P V G P K S L C N A C G I R N R K K	E F E Q A A V L L M A L S Y G
AtGATA17	V D - - - N E N C S S S - - -	K R T C V D - C G T I R T P L W R G G P A G P K S L C N A C G I K S R K K	E F E R A A V L L M A L S C S
AtGATA17L	V D - - - N G N C S S S - - -	K K T C V D - C G T S R T P L W R G G P A G P K S L C N A C G I K S R K K	E F E R A A V L L M A L S C S
AtGATA23	- - - L L S C S S S Y V S	I R C C S E - C K T T K T P M W R G G P T G P K S L C N A C G I R H R K Q	E F E Q A A L C L L L S C S
AtHAN	S S - - - L M D C T L S L G T	A R R C A N - C D T T S T P L W R N G P R G P K S L C N A C G I R F K K E	D G A H G G M P F L S W R L N
AtHAN1	S S - - - S V D C T L S L G T	P R R C A S - C D T T S T P L W R N G P K G P K S L C N A C G I R F K K E	N V T A S F M S W N
AtHAN2	S S P Y A S V D C T L S L G T	A R R C A N - C D T T S T P L W R N G P R G P K S L C N A C G I R F K K E	D H D V T T D P F L S W R L N
AtGATA29	- - - Q G T N V V D G G E	K K C T N M N C N A L N T P M W R R G P L G P K S L C N A C G I K F R K E	K R N V M I V L D D
MpB-GATA1	S S E P G A V D C T L S L G T	P R M C A H - C G T S K T P L W R N G P D G P K S L C N A C G I R H K K L	N A E A A M C L M K M S S G
PpGATA1	S L - - - A M N C T L S L G S	V H V C A H - C G T S K T P L W R N G P G G P K S L C N A C G I R F K K A	D F E E G A V L L M A L S C G
PpGATA2	S L - - - A M D C T L S L G S	A R V C A H - C G T S K T P L W R N G P G G P K S L C N A C G I R F K K A	D F E E G A V L L M A L S C G
PpGATA3	S L - - - A V D C T L S L G S	A R V C A H - C G T S K T P L W R N G P G G P K S L C N A C G I R F K K A	D F E E G A V L L M A L S C G
PpGATA4	S L - - - A V D C T L S L G S	V R V C A H - C G T S K T P L W R N G P Q G P K S L C N A C G I R F K K A	D F E E G A E L L M A L S C G

**Figure 2.5: B-GATAs from *A. thaliana*, *M. polymorpha* and *P. patens*.** In *A. thaliana* and all other angiosperms analysed to date, only the HAN B-GATAs (HAN, HANL1, and HANL2) contain HAN domains (VDCTLSG), while only the LLM B-GATAs (GNC, GNL, GATA15-17, and GATA17L) contain LLM domains (AAXLLMXLSXG). The bryophytes *M. polymorpha* and *P. patens*, however, encode GATAs with both HAN- and LLM-domains on the same protein.

In order to investigate the functions of the single B-GATA ortholog in *M. polymorpha*, I generated *Mpb-gata1* mutants by CRISPR/Cas9-mediated gene targeting using two gRNAs. I isolated several mutants but continued working on only two of them which carried similar deletions. The two lines called *Mpb-gata1-1* and *Mpb-gata1-2* exhibited deletions of 796 bp and 791 bp, respectively (Figure 2.6A). The deletions affect the second and third exon including the region coding for the zinc finger domain. Without the DNA-binding zinc finger domain, it can be assumed that, even if the remaining transcript would be translated, the resulting protein would not function as transcriptional regulator. Additionally, the deletions cause a frame shift leading to premature stop codons (Figure S5.1). The mutants showed several phenotypes, among these was a significant reduction in thallus size compared to the BoGa wildtype (Figure 2.6B and D). Additionally, formation of gemma cups in the mutants preceded that of the wild type (Figure 2.6B, C). Despite the reduction in thallus size, the ratio of gemma cups per area was strongly increased in the mutant alleles (Figure 2.6E). Similar to the phenotype of *B-GATA* mutants in *A. thaliana*, I also observed reduced chlorophyll accumulation in the mutants compared to the BoGa wildtype. The difference in chlorophyll content is subtle compared to that of the *Arabidopsis* *LLM-domain B-GATA* mutants, and only statistically significant for *Mpb-gata1-2* (Figure 2.6G – I).

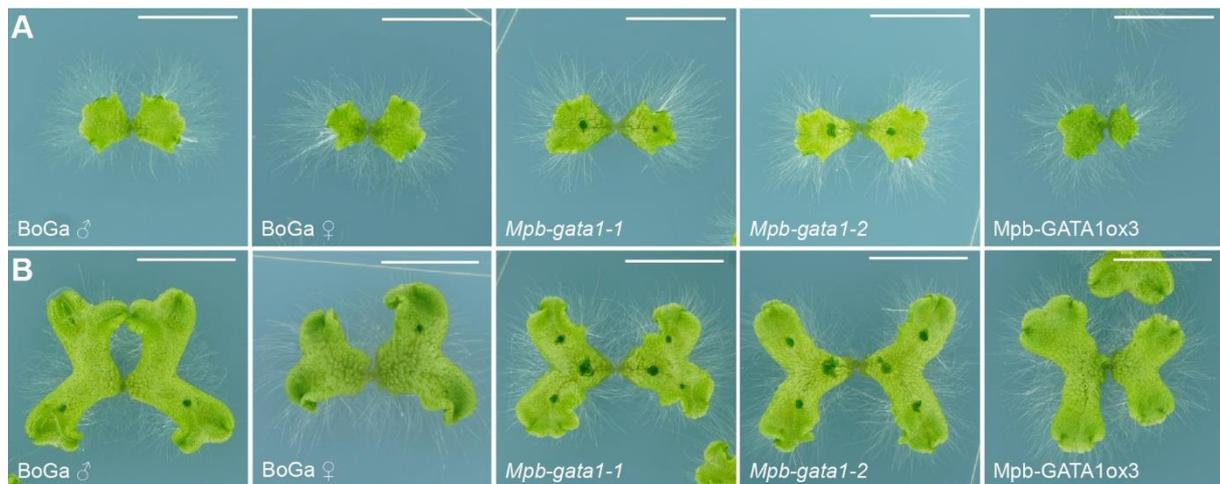
In *A. thaliana* HAN B-GATAs mainly have developmental functions in embryonic development and flower morphology (Zhao, Medrano et al. 2004, Nawy, Bayer et al. 2010, Kanei, Horiguchi et al. 2012, Zhang, Zhou et al. 2013), whereas LLM B-GATAs mainly exhibit physiological functions, e.g. related to greening, flowering time, or lateral shoot angle (Richter, Bastakis et al. 2013, Richter, Behringer et al. 2013, Ranftl, Bastakis et al. 2016, Bastakis, Hedtke et al. 2018, Schwechheimer, Schröder et al. 2022). The single B-GATA from *Marchantia polymorpha* seems to have developmental and physiological roles and might thus combine the functions of the distinct HAN- and LLM-domain containing B-GATAs as known from angiosperms as *A. thaliana*.



**Figure 2.6: The *Mpb-gata1* mutants and MpB-GATA1ox3 exhibit developmental phenotypes.** **A** Schematic representation of the *Mpb-gata1* mutants. The positions of the gRNAs are indicated by black arrow heads. Introns are displayed as lines, exons as boxes. The dark grey shading indicates the zinc finger-coding region in the second and third exon. **B** Representative photographs of three-weeks old BoGa, *Mpb-gata1-1*, *Mpb-gata1-2* and MpB-GATA1ox3 plants. Scale bar = 5 mm. **C** Close-up photographs of the wild type, mutant lines and overexpressor. White arrow heads indicate mature or developing gemma cups. Scale bar = 1 mm. **D** Area of BoGa, *Mpb-gata1* mutants and MpB-GATA1ox3 (n = 93). **E** Gemma cups per area of BoGa (n = 35), *Mpb-gata1-1* (n = 35), *Mpb-gata1-2* (n = 36) and MpB-GATA1ox3 (n = 34). **F** Expression of GCAM1 in the *Mpb-gata1* mutant lines and MpB-GATA1ox3. Data from RNA-seq experiments. **G – I** Contents of chlorophyll a (**G**), chlorophyll b (**H**) and total chlorophyll (**I**) in BoGa (n = 19), *Mpb-gata1-1* (n = 20), *Mpb-gata1-2* (n = 19) and MpB-GATA1ox3 (n = 4). **J** Expression of MpB-GATA1 under control of the EF1 $\alpha$  promoter in the MpB-GATA1ox3 overexpression lines. Expression level is shown as copy number and was calculated according to the method of Pfaffl (Pfaffl 2001). One-way ANOVA: ns, not significant; \*P  $\leq$  0.05; \*\*P  $\leq$  0.01; \*\*\*P  $\leq$  0.001; \*\*\*\*P  $\leq$  0.0001.

### 2.3.2 Reduced gemma cup formation is a phenotype of MpB-GATA1 overexpressors

In *A. thaliana*, LLM-domain B-GATA overexpressors often exhibit phenotypes opposing the ones of the mutant alleles. I generated overexpressors of MpB-GATA1 by expressing the gene under the promoter of *ELONGATION FACTOR 1α* (*EF1α*), routinely used for these analyses in *Marchantia* (Althoff, Kopischke et al. 2014). Even though expression of MpB-GATA1 in MpB-GATA1ox lines was relatively weak (three-fold compared to WT; Figure 2.6J), the gemma-cup density phenotype of the *Mpb-gata1* alleles seemed to be reverted in MpB-GATA1ox. While the number of gemma cups in *Mpb-gata1-1* and *Mpb-gata1-2* was increased compared to BoGa, the number of gemma cups, as well as the ratio of gemma cups per area, in MpB-GATA1ox was drastically reduced (Figure 2.6B, C, E). The higher number of gemma cups in the mutants may be explained by increased expression of *GEMMA CUP-ASSOCIATED MYB1* (*GCAM1*; Fig. 8F), an essential regulator of gemma cup development (Yasui, Tsukamoto et al. 2019). Interestingly, expression of *GCAM1* was strongly reduced in MpB-GATA1ox, hinting at involvement of MpB-GATA1 in gemma cup development upstream of *GCAM1* (Figure 2.6F). Additionally, the formation of gemma cups in the *Mpb-gata1* mutants could be observed after two weeks (Figure 2.7A), while gemma cups are first visible in BoGa after three-weeks (Figure 2.7B). In the overexpressor, gemma cups were still not visible after three weeks and only rarely found in even older plants (Figure 2.7). These data suggest that MpB-GATA1 is directly or indirectly involved in the regulation of gemma cup formation in *M. polymorpha*.



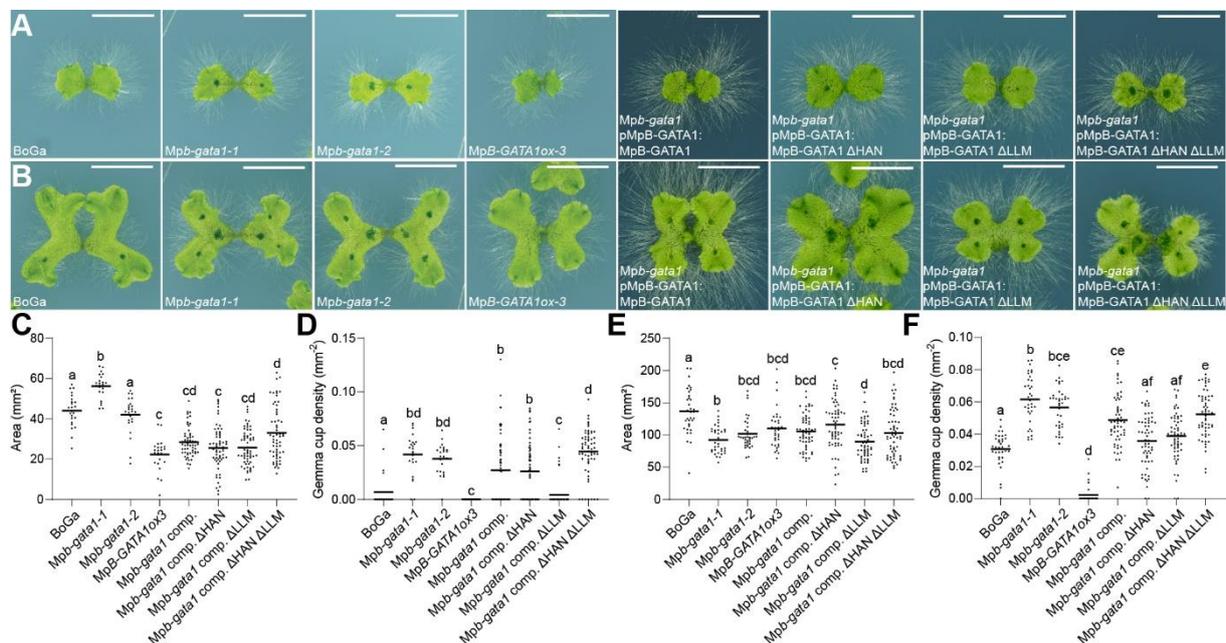
**Figure 2.7: Formation of gemma cups in the *Mpb-gata1* mutant alleles precedes the wildtype and the MpB-GATA1ox3 overexpressor.** **A** Representative photographs of two-weeks old plants. Gemma cups are already visible in both mutant alleles but not in the wild types or the overexpressor. **B** Representative photographs of three-weeks old plants. Gemma cups are visible in both mutant alleles and in BoGa but not in the overexpressor. Scale bar = 10mm.

## 2.4 Complementation lines of *Mpb-gata1* mutants lacking HAN- or LLM-domains did not give insights in domain functions

Since B-GATAs in bryophytes contain both HAN- and LLM-domains on the same protein, these organisms are good models to study the function of the two domains. In order to investigate the functions of HAN- and LLM-domain of the single B-GATA ortholog of *M. polymorpha*, I generated a set of domain-deletion complementation lines (*Mpb-gata1* pMpB-GATA1:MpB-GATA1 or *Mpb-gata1* comp., *Mpb-gata1* pMpB-GATA1:MpB-GATA1  $\Delta$ HAN or *Mpb-gata1* comp.  $\Delta$ HAN, *Mpb-gata1* pMpB-GATA1:MpB-GATA1  $\Delta$ LLM or *Mpb-gata1* comp.  $\Delta$ LLM, *Mpb-gata1* pMpB-GATA1:MpB-GATA1  $\Delta$ HAN  $\Delta$ LLM or *Mpb-gata1* comp.  $\Delta$ HAN  $\Delta$ LLM), using a genetic cross of the two mutants described above as background. Mutants of MpB-GATA1 exhibit both developmental- and physiological phenotypes, thus I hypothesized that the two domains might be responsible for the different phenotypes observed in the mutants.

After two weeks of growth, the complementation lines were all significantly smaller than the *Mpb-gata1* mutants (Figure 2.8A, C). Gemma cup density was slightly increased in the mutants, pMpB-GATA1:MpB-GATA1,  $\Delta$ HAN, and  $\Delta$ HAN  $\Delta$ LLM, but  $\Delta$ LLM showed a strong reduction in gemma cup density similar to MpB-GATA1ox3 (Figure 2.8A, D). After three weeks of growth, all complementation lines were significantly smaller than the wild type, suggesting that the complementation was only partial (Figure 2.8B, E). Analysis of gemma cup density revealed elevated levels for *Mpb-gata1* comp. and  $\Delta$ HAN  $\Delta$ LLM relative to BoGa, whereas gemma cup density in  $\Delta$ HAN and  $\Delta$ LLM was at the wild-type level (Figure 2.8B, F).

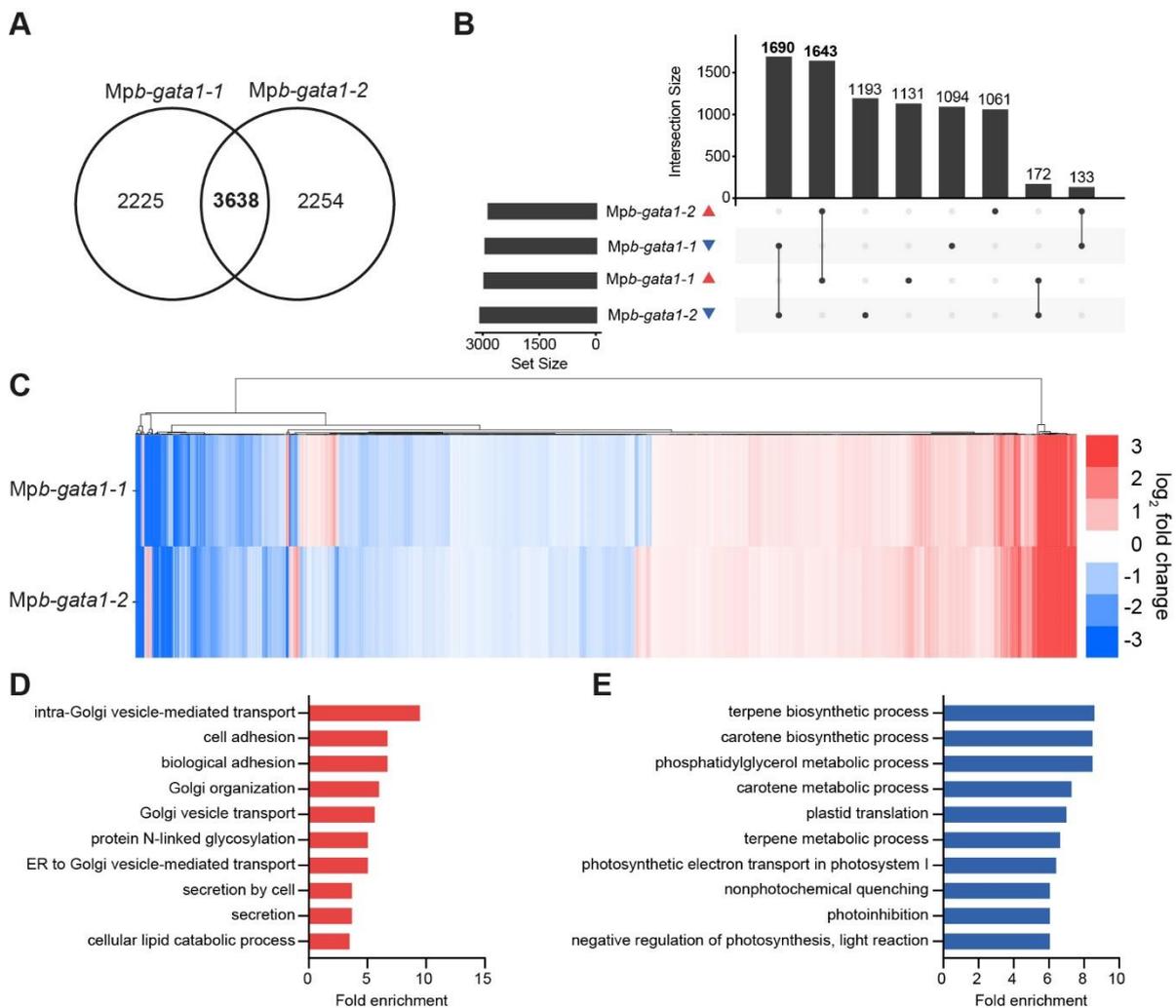
Taken together, these results are not consistent and do not give any information about the functions of the domains. I generated the complementation lines close to the end of my project with the help of Bang-Yu Hsu (Plant Systems Biology, Technical University of Munich, Freising, Germany), thus they were not included in the other experiments presented here. Due to time constraints, I did not continue to work on these lines. Further biochemical experiments are required to understand the roles of HAN- and LLM-domains.



**Figure 2.8: Domain deletion complementation lines did not give insights into functions of HAN- and LLM-domain.** A and B Representative photographs of two- and three weeks old plants of the indicated genotypes, respectively. Scale bar = 10mm. C and D Area and gemma cup density of two-weeks old plants, respectively. n = 24. E and F Area and gemma cup density of three-weeks old plants, respectively. n = 24. Data sets with no statistical difference after one-way ANOVA and Tukey's HSD posthoc test fall into one group and are labelled with identical letters.

## 2.4.1 Transcriptomic analysis of *MpB-gata1* suggests role in regulation of high-light stress

In order to understand gene expression downstream of *MpB-GATA1*, I performed an RNA-sequencing experiment with three-weeks old BoGa wild type and the mutant alleles *MpB-gata1-1* and *MpB-gata1-2*. Analysis of differentially expressed genes (DEGs) with filtering for FDR p-value  $\leq 0.05$  against the BoGa wild type yielded 5863 and 5892 DEGs for *Mpb-gata1-1* and *Mpb-gata1-2*, respectively. Out of these, 3638 DEGs or 62% of all DEGs for *Mpb-gata1-1* and 62% of all DEGs for *Mpb-gata1-2*, could be identified in both mutant alleles (Figure 2.9A). The majority of the shared 3638 DEGs exhibit the same direction of transcriptional regulation: 1643 or 45.2% DEGs were down regulated, while 1690 or 46.5% DEGs were upregulated in both *Mpb-gata1-1* and *Mpb-gata1-2*. Only 8.4% or 305 DEGs were regulated in opposite directions between the two mutant alleles (Figure 2.9B).



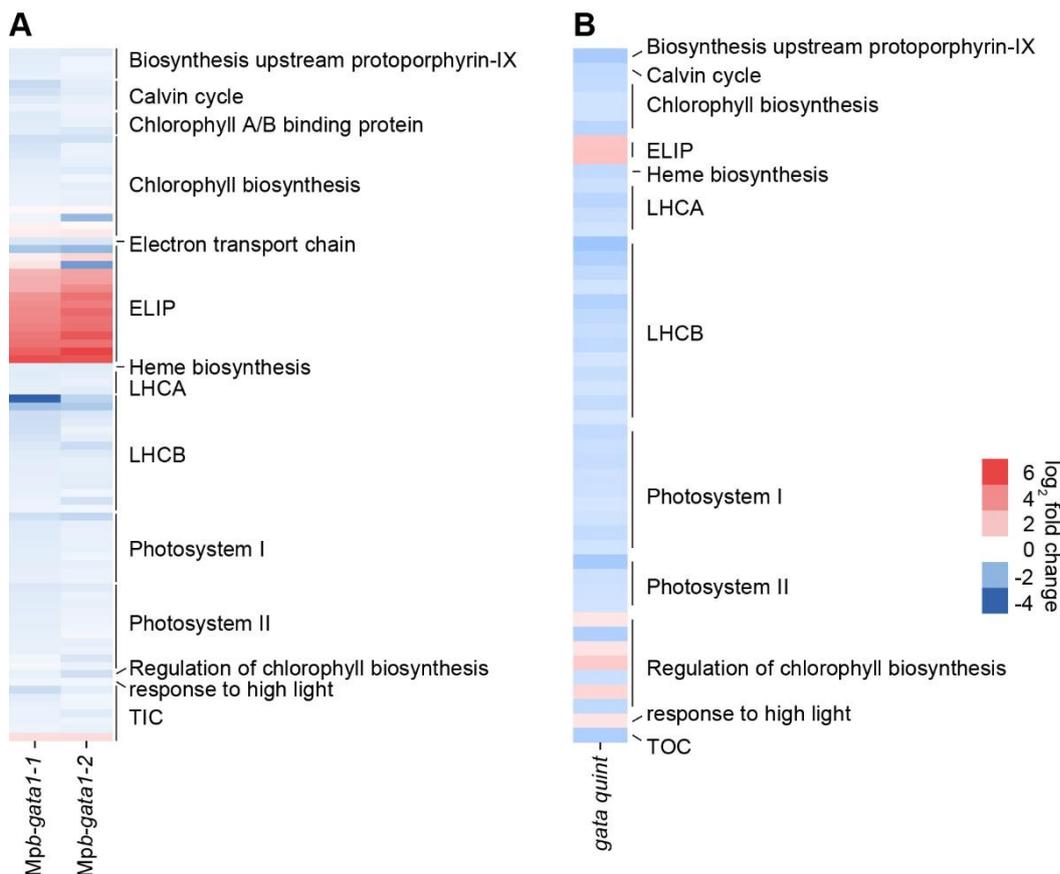
**Figure 2.9: Transcriptomic analysis of the *Marchantia polymorpha* *MpB-gata1* mutants.** **A** Venn diagram of DEGs from *Mpb-gata1-1* and *Mpb-gata1-2* after filtering for FDR p-value  $\leq 0.05$ . **B** UpSetR diagram showing the direction of the transcriptional regulation of all DEGs from *Mpb-gata1-1* and *Mpb-gata1-2*. **C** Heatmap of the 3638 genes shared by *Mpb-gata1-1* and *Mpb-gata1-2* and their regulation as log<sub>2</sub> fold change. Hierarchical clustering is based on Euclidian distance using complete linkage. **D** and **E** The ten most strongly enriched GO terms of the 1643 upregulated genes (**D**) and 1690 downregulated genes (**E**) shared by both mutant alleles.

Overall, the mutant alleles displayed overlapping but also distinct differential expression. The DEGs present in both alleles, however, showed a high degree of correspondence (Figure 2.9C).

GO term enrichment analysis of downregulated genes from *Mpb-gata1* mutants emphasized the mis-regulation of genes associated with photosynthesis (Figure 2.9E). Terms related to

photosynthetic electron transport, non-photochemical quenching, photoinhibition and negative regulation of photosynthesis were strongly enriched among the downregulated genes. Interestingly, terms related to carotene and terpene biosynthesis were among most strongly enriched GO terms here. GO enrichment of upregulated genes yielded terms related to Golgi vesicle transport and cell adhesion (Figure 2.9D). A striking observation of this transcriptomic experiment was the uniform downregulation of many chlorophyll-biosynthesis associated genes, as well as genes connected to photosynthesis (Figure 2.10A). Among these were the *LIGHT HARVESTING COMPLEX* genes, which encode the antenna proteins for photosystem I and II (*LHCA* and *LHCB*, respectively), as well as genes encoding the photosystem subunits themselves (*PSAD-2*, *PSAG*, *PSAF*, *PSAK* and *PSAO* for PSI and *PSBO2*, *PSBP-1*, *PSBR*, *PSBW* and *PSBX* for PSII). At the same time, 16 of 18 differentially expressed *EARLY LIGHT INDUCED PROTEIN (ELIP)* genes were strongly upregulated in *Mpb-gata1-1* and *Mpb-gata1-2*. *ELIP1* and *ELIP2*, the only two *A. thaliana* *ELIPs*, were similarly upregulated in the *gata quint* mutant (Figure 2.10B). In *A. thaliana*, *ELIPs* are known as high-light induced proteins with a proposed role in photoprotection (Hutin, Nussaume et al. 2003).

Interestingly, I made similar observations in the previously acquired microarray data for the *A. thaliana* LLM-GATA *quint* mutant (*gnc gnl gata15 gata17 gata17l*) (Ranfll, Bastakis et al. 2016). In *gata quint*, genes involved in chlorophyll biosynthesis and photosynthesis (PS compartments, LHCs, etc.) were downregulated, while *ELIPs* were strongly upregulated, suggesting that GATAs from *M. polymorpha* and *A. thaliana* are involved in high-light stress responses.

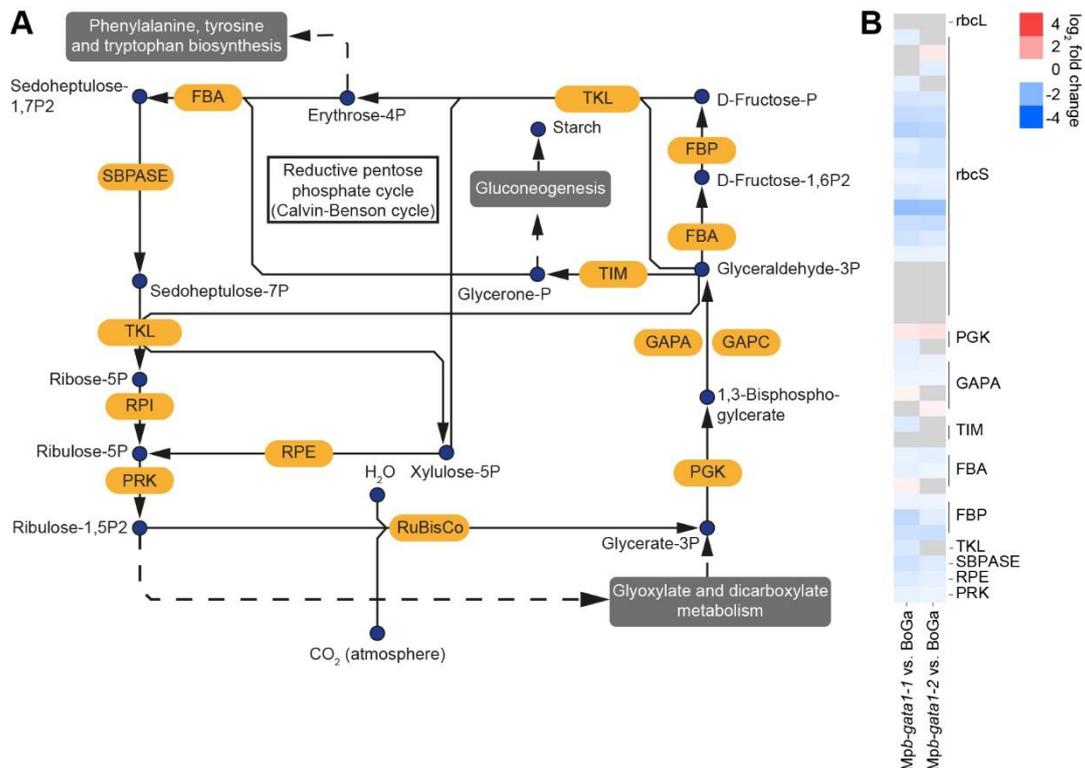


**Figure 2.10: Photosynthesis-related genes exhibit constitutive high-light regulation in the *Mpb-gata1* mutants of *M. polymorpha* and in the *A. thaliana gata quint* mutant.** **A** Differential expression of 88 photosynthesis-related genes in BoGa at  $t_{6HL}$  and *Mpb-gata1-1* and *Mpb-gata1-2* at  $t_{6LL}$  and  $t_{6HL}$ . **B** Differential expression of 48 photosynthesis-related genes in *gata quint*.



### 2.4.3 Differential regulation of the Calvin-Benson cycle in *Mpb-gata1* mutants

Carbon fixation is the process in which photoautotrophic organisms utilize atmospheric CO<sub>2</sub> to produce carbohydrates through photosynthesis. In the *Mpb-gata1* mutants, many genes encoding enzymes required for carbon fixation were differentially regulated (Figure 2.12). The majority of these DEGs were downregulated, except for the genes encoding one phosphoglycerate kinase ortholog (PGK). This regulation might be an indirect effect of the loss of *Mpb-gata1* caused by the downregulation of light reaction components (photosystems and antennae) and thus leading to reduced photosynthetic activity with reduced output of NADPH and ATP required for carbon fixation.

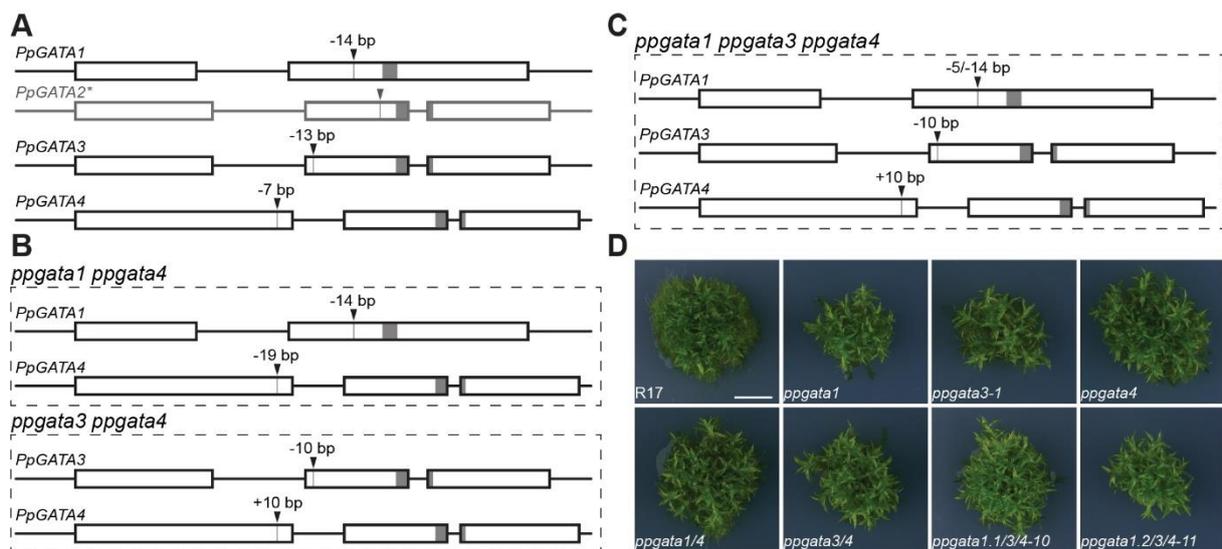


**Figure 2.12: Genes involved in carbon fixation are differentially regulated in *Mpb-gata1* mutants.** **A** Schematic representation of the carbon fixation in plants. Enzymes are shown in yellow boxes, reactions as solid arrows and metabolites as blue dots. Pathways branching off the carbon fixation are indicated by dotted arrows and grey boxes. See Supplemental Table S5.6 for abbreviations. **B** Heatmap of differentially expressed genes involved in carbon fixation from the transcriptomic experiment in *M. polymorpha*. Genes filtered by FDR p-value < 0.05.

## 2.5 *Physcomitrium patens* B-GATAs are involved in protonema formation

### 2.5.1 Deficiency in protonema formation is a phenotype of *P. patens gata* mutants

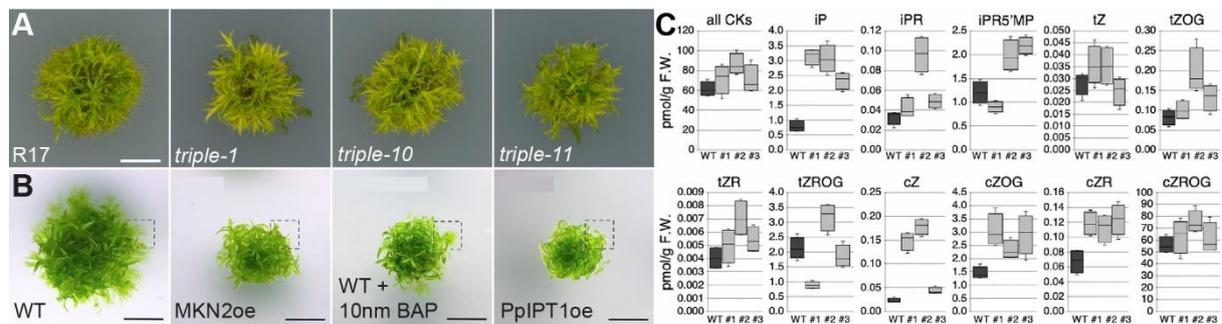
*P. patens* contains four B-GATAs, PpGATA1 - PpGATA4 (Figure 3.1-1A). Similar to the single B-GATA ortholog in *M. polymorpha*, the four B-GATAs from *P. patens* contain both HAN- and LLM-domains. In order to investigate the functions of these four B-GATA orthologs, I generated a suite of mutants through genome editing with CRISPR/Cas9. The aim of the strategy was the generation of single, double, triple, and quadruple mutants in the Reute wild type (R17; Hiss, Meyberg et al. 2017) by targeting all four genes at the same time with one gRNA per gene. After multiple independent transformations, I was able to successfully isolate the single mutants for *ppgata1* (14-bp deletion), *ppgata3* (13-bp deletion), and *ppgata4* (seven-bp deletion; Figure 2.13A and D). No mutant for PpGATA2 could be generated. CRISPR/Cas9 multiplex approaches yielded the double mutants *ppgata1/4* and *ppgata3/4* (Figure 2.13B and D), as well as the triple mutants *ppgata1.1/3/4-10* and *ppgata1.2/3/4-11*, harbouring different alleles for *ppgata1* (five-bp deletion in *ppgata1.1*, 14-bp deletion in *ppgata1.2*; Figure 2.13C, D; Supplemental Figure S5.3, Supplemental Figure S5.4). In each case, the mutations caused frame shifts in front of the zinc finger coding region, leading to premature stop codons (Supplemental Figure S5.3, Supplemental Figure S5.4).



**Figure 2.13: *P. patens gata* mutants.** **A - C** Schematic representation of the *P. patens gata* mutants generated through CRISPR/Cas9. **A** Single mutants for PpGATA1, PpGATA3 and PpGATA4, but not PpGATA2(\*) could be isolated. Introns are displayed as lines, exons as boxes. The dark grey shading indicates the zinc finger-coding region in the second- or second and third exon. The positions of the gRNAs are indicated by black arrow heads. **B** Available *ppgata* double mutants derived from multiplex CRISPR approach. **C** Available *ppgata* triple mutants derived from multiplex CRISPR approach. *ppgata1.1/3/4-10* and *ppgata1.2/3/4-11* harbour different alleles for *ppgata1*, a five-bp deletion for *ppgata1.1*, and a 14-bp deletion for *ppgata1.2* **D** Representative photographs of four-weeks old R17 wild type, *ppgata* single mutants (*ppgata1*, *ppgata3-1*, and *ppgata4*), *ppgata* double mutants (*ppgata1/4* and *ppgata3/4*), and *ppgata* triple mutants (*ppgata1.1/3/4-10* and *ppgata1.2/3/4-11*) grown on BCD medium under standard conditions. Scale bar = 5mm. Genotyping and photographs from Dario Zappone (Plant Systems Biology, Technical University of Munich, Freising, Germany).

## 2.5.2 Protonema defects of *ppgata* mutants might be related to cytokinin

When grown on regular BCD medium (Cove, Perroud et al. 2009), *ppgata* mutants exhibited limited protonema production compared to R17 (Figure 2.14A). This phenotype was reminiscent of plants overexpressing the moss *KNOX2* transcription factor, *PpMKN2*, or of plants overexpressing *PpIPT3* (Figure 2.14B; (Coudert, Novák et al. 2019). *PpMKN2* activates cytokinin biosynthesis via ISOPENTENYL TRANSFERASE, *PpIPT3*, and thus, overexpression of *PpMKN2* or *PpIPT3* lead to elevated levels of endogenous cytokinin (Figure 2.14C; (Coudert, Novák et al. 2019). The same observation was made for wild type plants treated with 10nm BAP (Coudert, Novák et al. 2019).



**Figure 2.14: *P. patens gata triple* mutants exhibit reduced protonema formation comparable to plants with high cytokinin content or treated with cytokinin. A** Representative images of three-weeks old plants grown on BCDAT medium under standard conditions. Protonema formation in the mutants is strongly reduced compared to R17 wild type. Scale bar = 5mm. **B** Protonema formation and overall plant spread was reduced in MKN2oe and PpIPT1oe transgenic lines and wildtype treated with 10 nM 6-benzylaminopurine (BAP) relative to WT and mock-treated controls. Scale bar = 5 mm. **C** Cytokinin quantification by LC-MS/MS showed a global increase in CK levels in three independent MKN2oe transgenic lines (#1, #2, and #3) relative to WT controls. MKN2oe #1 is the line shown in B. Boxplots represent data from four biological replicates, and CK levels are expressed in pmol/g of fresh weight (pmol/g F.W.). iP, isopentenyladenine; iPR, isopentenyladenosine; iPR5'MP, isopentenyladenosine-5'-monophosphate; tZ, trans-zeatin; tZR, trans-zeatin riboside; tZOG, trans-zeatin O-glucoside; tZROG, trans-zeatin riboside O-glucoside; cZ, cis-zeatin; cZR, cis-zeatin riboside; cZOG, cis-zeatin O-glucoside; cZROG, cis-zeatin riboside O-glucoside. Panels B and C were adapted from Coudert, Novák et al. 2019<sup>3</sup>.

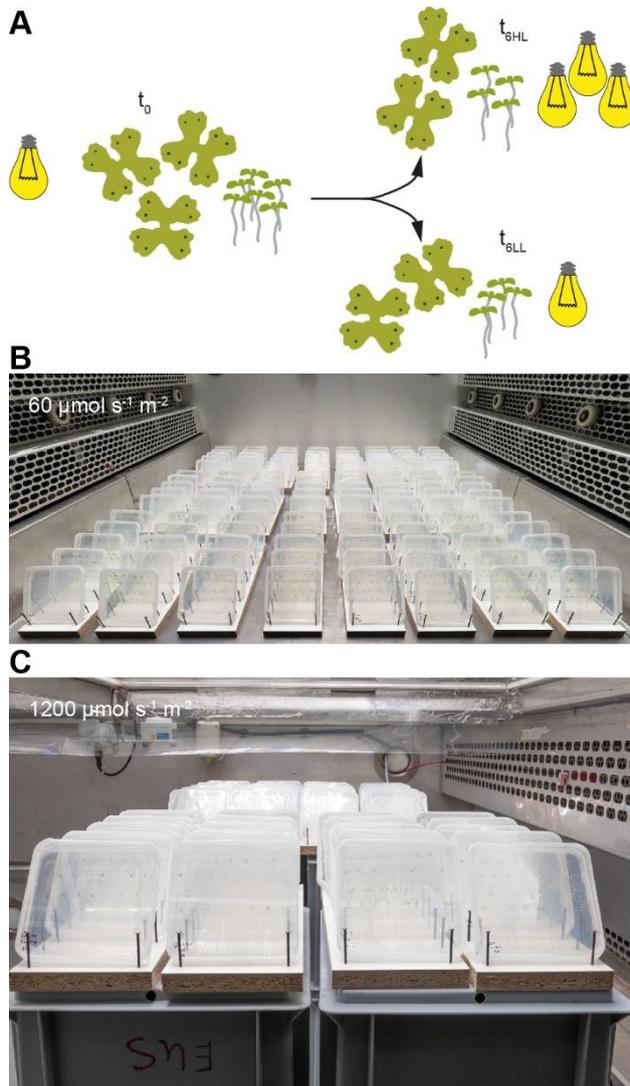
In order to test whether the defect in protonema formation observed in *ppgata* mutants was related to elevated levels of endogenous cytokinins, I collaborated with Dr. Ondřej Novák from Palacky University (Olomouc, Czech Republic) for the quantification of endogenous cytokinins using high-performance liquid chromatography coupled with mass spectrometry (HPLC–MS). A weak increase in overall CK content could be observed for *ppgata1.2/3/4-11* but not for *ppgata1.1/3/4-10*, while iP-type cytokinins were increased in *ppgata1.1/3/4-10* but not in *ppgata1.2/3/4-11*. At the same time, tZ- and cZ-type cytokinins increased significantly only in *ppgata1.2/3/4-11* relative to the wild type (Supplemental Table S6.1). Overall, the results of the CK 42quantification were not conclusive and the defects in protonema formation observed in the *ppgata* mutants were likely not due to differences in CK content. Due to time constraints, I did not continue to work on this project. The project was taken over by PhD candidate Dario Zappone (Plant Systems Biology, Technical University of Munich, Freising, Germany), who is characterizing the mutants and is working on the functional analysis of B-GATAs from *P. patens*.

<sup>3</sup> Reprinted from Coudert, Y., Novák, O., & Harrison, C. J. (2019). "A KNOX-cytokinin regulatory module predates the origin of indeterminate vascular plants." *Current biology*, 29(16), 2743-2750. Copyright 2023, with permission from Elsevier.

## 2.6 B-GATAs are involved in high-light stress response

### 2.6.1 *Mpb-gata1* mutants maintain high-light response

The transcriptomic profile of the *Mpb-gata1* mutants and *A. thaliana gata quint* mutant with downregulation of photosystem components and upregulation of *ELIP* genes suggests a constitutive high light stress response in B-GATA mutants of both species and thus a novel role for B-GATAs. In order to further investigate the high-light stress response on the molecular level, I performed an RNA-seq experiment using high-light exposed plants. In order to achieve suitable high-light intensities, I collaborated with Dr. Barbro Winkler from the Research Unit for Environmental Simulation at the Institute of Biochemical Plant Pathology at Helmholtz Munich (HMGU, Munich, Germany). For this experiment three-weeks-old wild type plants and the two *Mpb-gata1* alleles grown in low-light ( $60 \mu\text{mol s}^{-1} \text{m}^{-2}$ ) were exposed to either six hours of high-light ( $1200 \mu\text{mol s}^{-1} \text{m}^{-2}$ ;  $t_{6HL}$ ) or a corresponding low light ( $60 \mu\text{mol s}^{-1} \text{m}^{-2}$ ;  $t_{6LL}$ ) mock treatment in sun simulation chambers (Figure 2.15A - C).

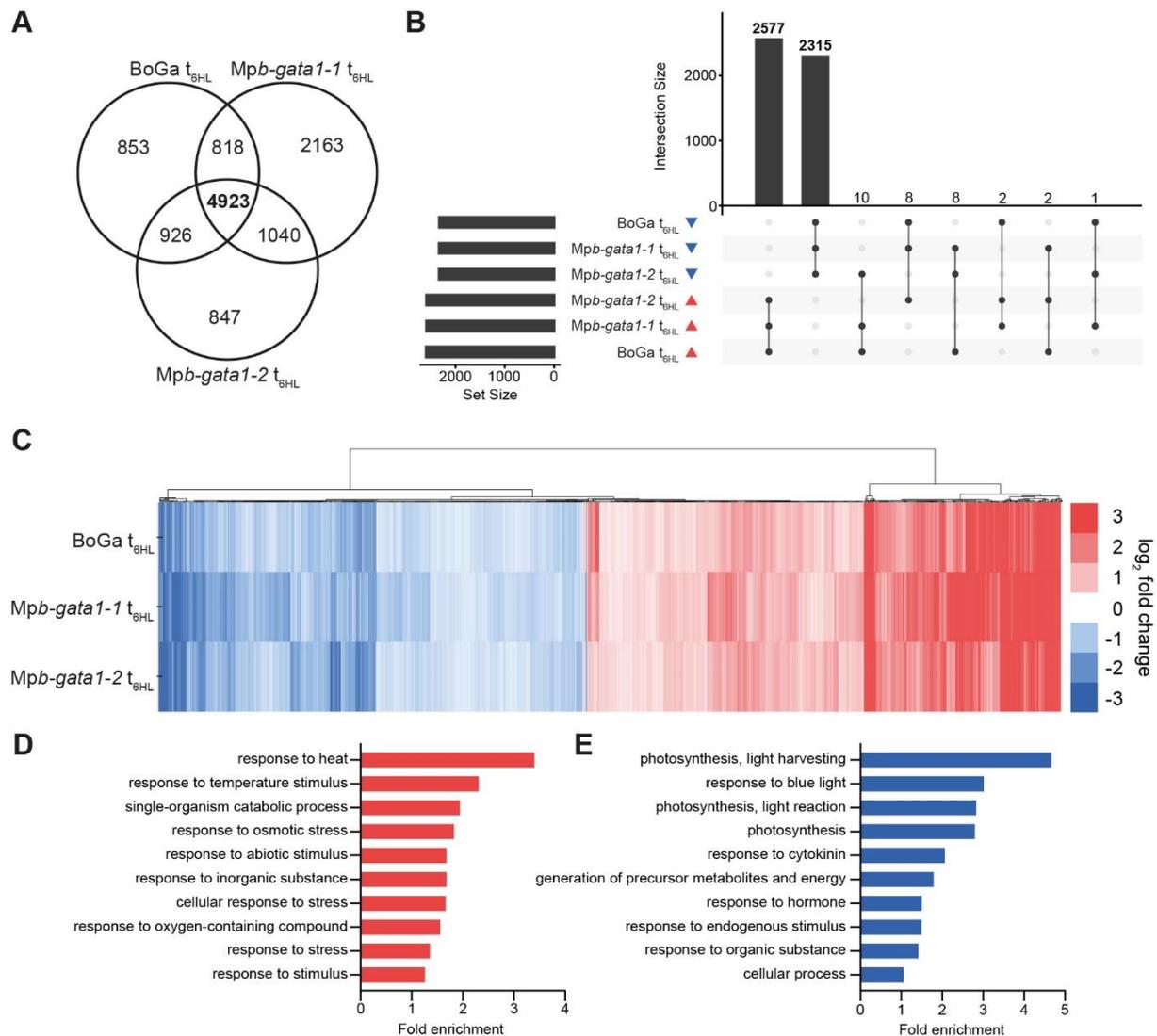


**Figure 2.15: Setup of the high-light experiment.** **A** Three-weeks-old *M. polymorpha* plants and one-week-old *A. thaliana* seedlings grown under low-light conditions at  $60 \mu\text{mol s}^{-1} \text{m}^{-2}$  ( $t_0$ ) were incubated for six hours under high-light at  $1200 \mu\text{mol s}^{-1} \text{m}^{-2}$  ( $t_{6HL}$ ) or six hours under low-light at  $60 \mu\text{mol s}^{-1} \text{m}^{-2}$  ( $t_{6LL}$ ). In the case of *M. polymorpha*, the *Mpb-gata1* mutant alleles and the BoGa wildtype was used. In the case of *A. thaliana*, the *gata hex* mutant, GNLox and Col-0 were used. **B** Photographs of the chamber used for the low-light treatment. Plates were placed in a  $60^\circ$  angle to allow for ventilation and to reduce condensation on the lids. **C** Photographs of the chamber used for the high-light treatment. The plates were elevated to be closer to the light source. Plates were placed in a  $60^\circ$  angle to allow for ventilation and to reduce condensation on the lids.

To establish how the transcriptome of *M. polymorpha* wild type plants reacts in response to high light, I first compared the transcriptome of the high-light exposed BoGa wildtype (BoGa  $t_{6HL}$ ) to mock treated BoGa (BoGa  $t_{6LL}$ ; Figure 2.16). Here, I identified 7520 genes that were differentially regulated in high-light conditions but not in the corresponding mock situation when filtering the data according to their FDR p-value  $\leq 0.05$  (Figure 2.16A). In order to improve the *gata* loss-of-function specificity of the read-out and to exclude second-site effects of the two mutant alleles, the overlapping DEGs of *Mpb-gata1-1* and *Mpb-gata1-2* were used for the following analyses.

In the *Mpb-gata1* alleles, 5195 shared DEGs were identified after six hours of high-light treatment compared to the corresponding mock treatment (Figure 2.16A). Comparison of the gene sets from BoGa  $t_{6HL}$  and *Mpb-gata1*  $t_{6HL}$  yielded overlapping but also substantially different transcriptomic responses to the high-light treatment. All genotypes responded to the treatment with striking changes in gene expression, even though the affected genes are only partially overlapping between the different genotypes. Of the 7520 DEGs from the wildtype (BoGa  $t_{6HL}$  vs BoGa  $t_{6LL}$ ) and 5963 DEGs from the mutant alleles

(mutants  $t_{6HL}$  vs mutants  $t_{6LL}$ ), 4923 genes are shared between all genotypes (Figure 2.16A).



**Figure 2.16: The *M. polymorpha* *Mpb-gata1* mutants maintain the same high-light response as observed in BoGa.** **A** Venn diagram of DEGs from BoGa, *Mpb-gata1-1* and *Mpb-gata1-2* at  $t_{6HL}$  after filtering for FDR  $p$ -value  $\leq 0.05$ . **B** UpSetR diagram showing the direction of the transcriptional regulation of the 4923 genes shared by BoGa, *Mpb-gata1-1* and *Mpb-gata1-2* after the high-light treatment. For simplicity, the genes in the overlaps of only two datasets and the genes unique to each dataset are not included. **C** Heatmap of the 4923 genes shared by BoGa, *Mpb-gata1-1* and *Mpb-gata1-2* at  $t_{6HL}$  and their regulation as  $\log_2$  fold change. Hierarchical clustering is based on Euclidian distance using complete linkage. **D** and **E** The ten most strongly enriched GO terms of the 2577 upregulated genes (**D**) and 2315 downregulated genes (**E**) shared by all genotypes.

This overlap is significantly greater than the numbers of DEGs unique to either BoGa (2597) or the mutant alleles (1040). Of these 4923 DEGs, 2577 (52.4%) were upregulated in all genotypes, while 2315 (47%) were downregulated in all genotypes. Only a small subset of 31 genes (0.6%) exhibited non-overlapping expression without any discernible pattern among the genotypes (Figure 2.16B - C). Among the most significantly enriched GO terms for the 2315 DEGs downregulated in the HL-treated wildtype and mutant alleles were terms related to photosynthesis (e.g., “photosynthesis, light harvesting”, “photosynthesis, light reaction” and “generation of precursor metabolites and energy”; Figure 2.16E).

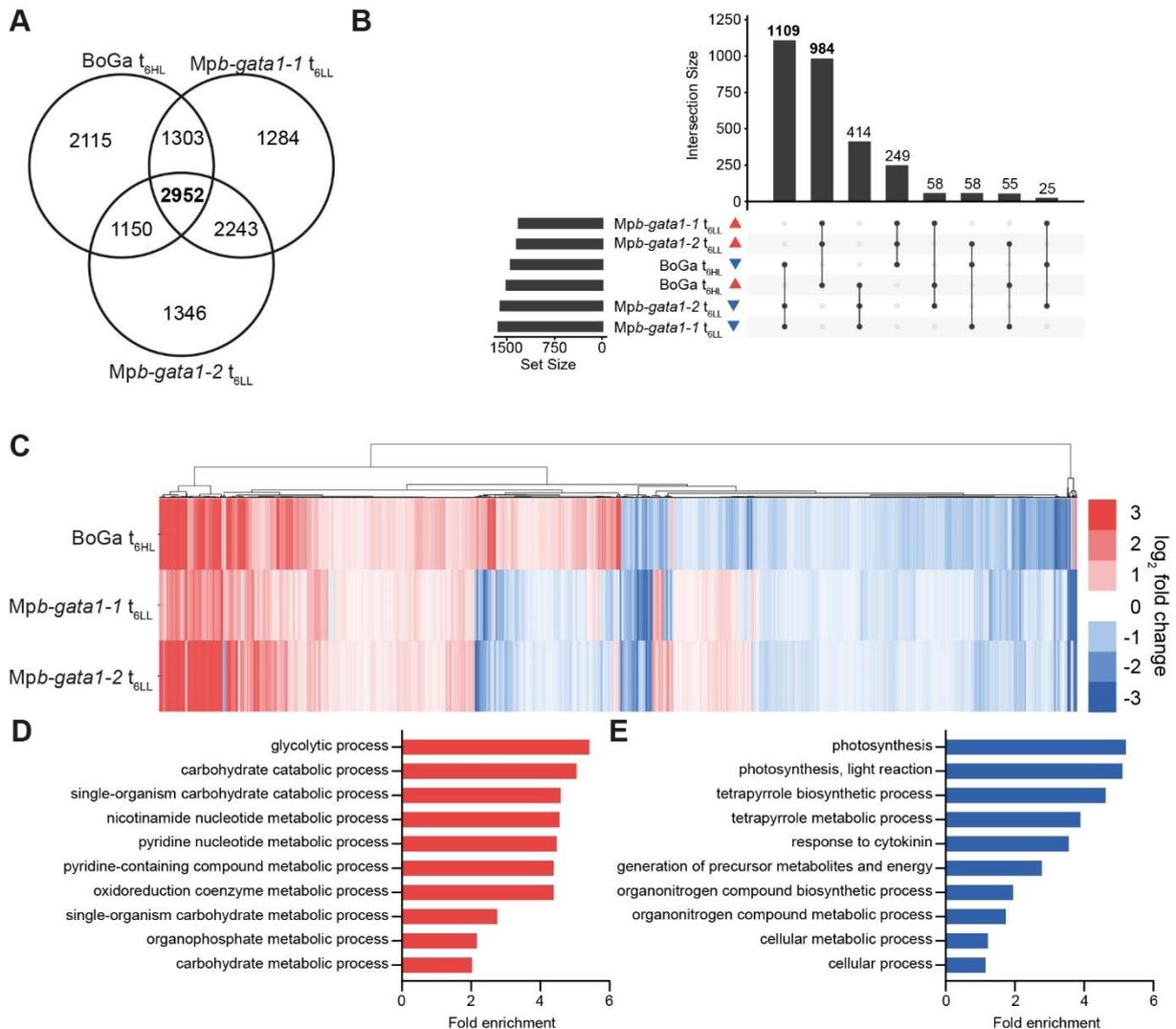
Interestingly, the most significantly enriched GO terms for the 2577 DEGs upregulated in the HL-treated wildtype and mutant alleles included terms related to temperature (e.g. “response to heat” and “response to temperature stimulus”) and general GO terms related to abiotic stresses (e.g. “response to osmotic stress”, “response to abiotic stimulus”, “cellular response to stress” and “response to stress”; Figure 2.16D).

As it is known from previous studies, many genes related to high light intensities are also regulated by heat (Huang et al., 2019). A maximum temperature of 24°C was measured directly

in the plates after six hours of high-light treatment. This temperature is unlikely to cause a heat stress response, thus the enrichment of the GO term “Response to heat” might also be linked to the high light intensity. Overall, the *Mpb-gata1* mutants maintained a high-light stress response, but this response was quantitatively impaired in comparison to the wild type.

## 2.6.2 *Marchantia Mpb-gata1* exhibit constitutive a high-light stress response

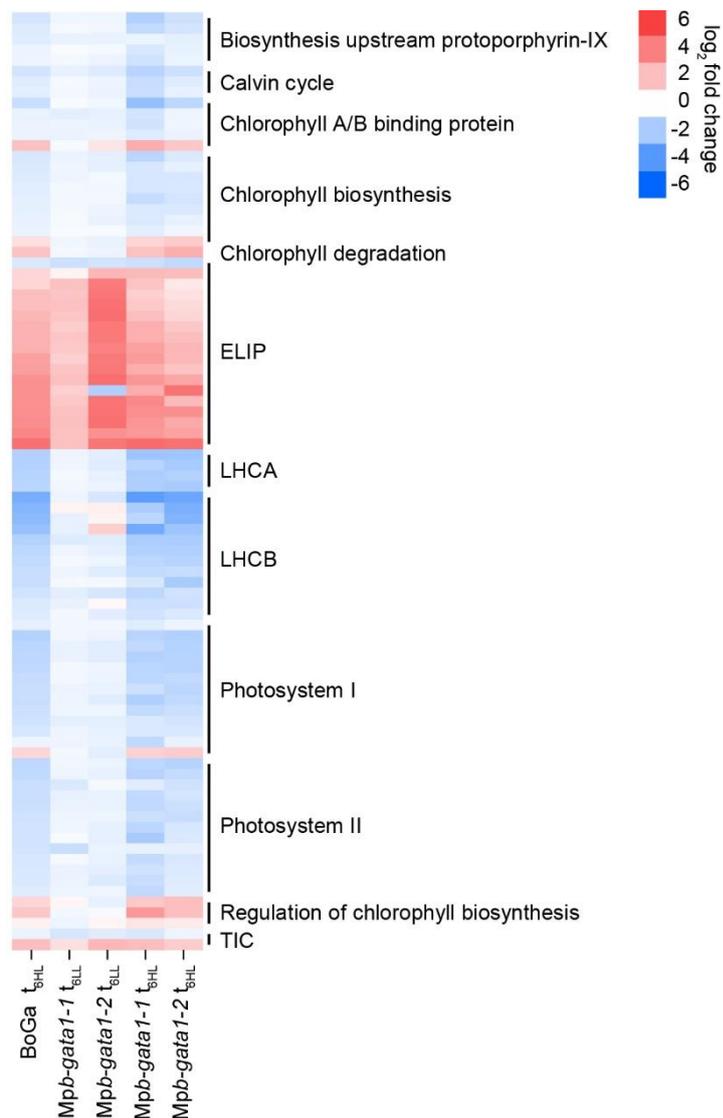
After confirming the ability of *Mpb-gata1* mutants to sense and react to high-light, I investigated a possible constitutive high-light stress response of the mutant alleles (Figure 2.17).



**Figure 2.17: The *Mpb-gata1* mutants exhibit a constitutive high-light stress transcriptome.** **A** Venn diagram of DEGs from BoGa at  $t_{6HL}$  and *Mpb-gata1-1* and *Mpb-gata1-2* at  $t_{6LL}$  after filtering for FDR p-value  $\leq 0.05$ . **B** UpSetR diagram showing the direction of the transcriptional regulation of the 2952 genes shared by BoGa at  $t_{6HL}$  and *Mpb-gata1-1* and *Mpb-gata1-2* at  $t_{6LL}$ . For simplicity, the genes in the overlaps of only two datasets and the genes unique to each dataset are not included. **C** Heatmap of the 2952 genes shared by BoGa at  $t_{6HL}$  and *Mpb-gata1-1* and *Mpb-gata1-2* at  $t_{6LL}$  and their regulation as  $\log_2$  fold change. Hierarchical clustering is based on Euclidian distance using complete linkage. **D** and **E** The ten most strongly enriched GO terms of the 984 upregulated genes (**D**) and 1109 downregulated genes (**E**) shared by all genotypes.

Of the 7520 high light stress-responsive DEGs identified in the wild type (BoGa  $t_{6HL}$  vs BoGa  $t_{6LL}$ ), 2952 were already differentially regulated in mock-treated *Mpb-gata1* mutants (*Mpb-gata1-1*/*Mpb-gata1-2*  $t_{6LL}$  vs BoGa  $t_{6LL}$ ; Figure 2.17A). Among these, 984 DEGs (33.3%) were upregulated in the high-light stressed BoGa wildtype and mock-treated mutant alleles, while 1109 (37.6%) were downregulated in the three genotypes. Overall, 71% of the genes differentially regulated in both, the mock-treated *Mpb-gata1* mutants ( $t_{6LL}$ ) and in the high-light treated BoGa wildtype ( $t_{6HL}$ ), exhibit the same direction of transcriptional regulation. Among the

859 DEGs (29.09%) that do not share the same direction of transcriptional regulation, no discernible pattern could be identified (Figure 2.17B – C).



**Figure 2.18: Photosynthesis associated genes were similarly regulated in high-light stressed and BoGa and unstressed *Mpb-gata1* mutants.** Heatmap of genes connected to photosynthesis showing their differential expression in the BoGa wild type at  $t_{6LL}$  and the mutants at  $t_{6LL}$  and  $t_{6HL}$ . Most genes exhibited the same direction of transcriptional regulation, but the response in the mutants at  $t_{6LL}$  is weaker compared to BoGa or the mutants at  $t_{6HL}$ .

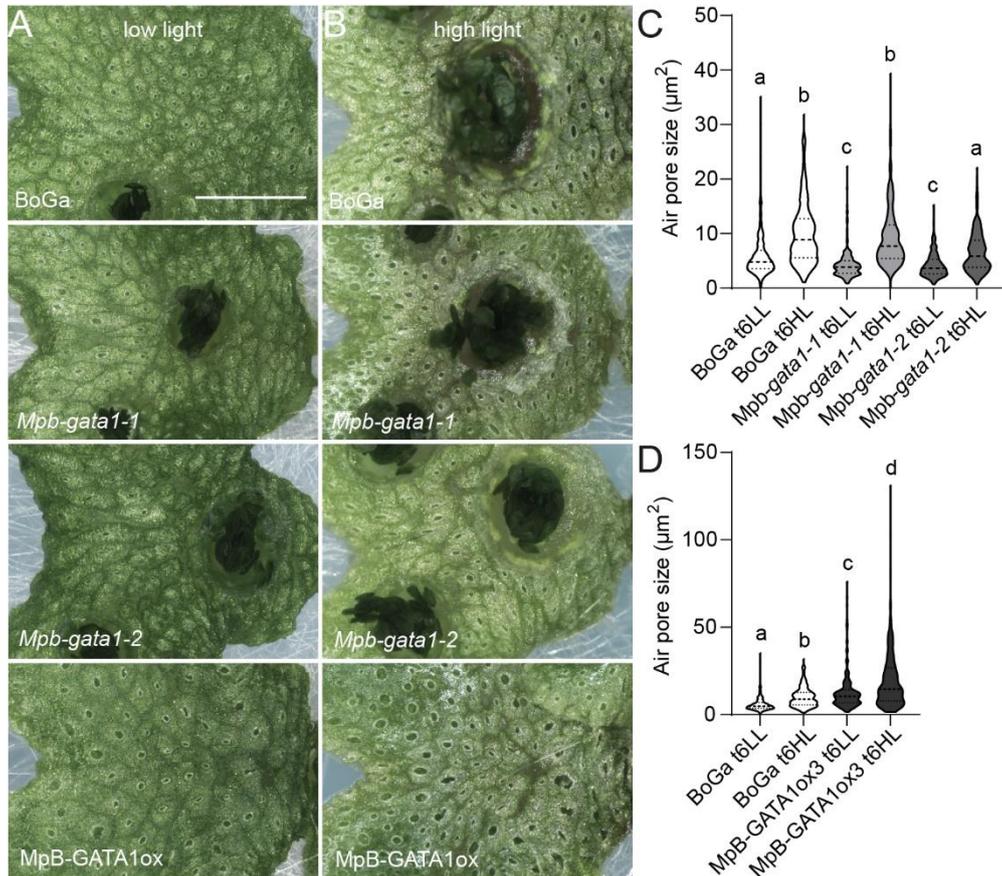
*gata1* mutant alleles exhibited a strikingly similar expression among high-light treated wildtype plants (BoGa  $t_{6HL}$ ), low-light treated mutant plants (*Mpb-gata1*  $t_{6LL}$ ) and high-light treated mutant plants (*Mpb-gata1*  $t_{6HL}$ ). In qualitative terms, the direction of transcriptional regulation was identical for almost all genes, while in quantitative terms, differential expression was always stronger in the high-light treated plants (Figure 2.18).

GO enrichment analysis of the downregulated genes shared by all three genotypes yielded many terms related to photosynthesis (e.g., “photosynthesis”, “photosynthesis, light reaction” and “generation of precursor metabolites and energy”), as well as terms related to chlorophyll biosynthesis (e.g., “tetrapyrrole biosynthetic process” and “tetrapyrrole metabolic process”; Figure 2.17E). Interestingly, GO terms related to cytokinin (“response to cytokinin”) and nitrogen metabolism (“organonitrogen compound biosynthetic process” and “organonitrogen compound metabolic process”) were also present in the ten most strongly enriched GO terms. GO enrichment analysis of the upregulated DEGs revealed a strong overrepresentation the carbohydrate metabolism. Five of the ten most strongly enriched GO terms were directly related to carbohydrate metabolism (e.g., “glycolytic process”, “carbohydrate catabolic process”, “single-organism carbohydrate catabolic process”, “single-organism carbohydrate metabolic process” and “carbohydrate metabolic process”; Figure 2.17D).

Interestingly, the photosynthesis-associated genes, that I had originally discovered in the initial transcriptomic analysis of the *Mpb-*

### 2.6.3 High-light stress induces increased air pore size in *M. polymorpha* genotypes

Air pores are structures at the surface of the thallus of *M. polymorpha* that allow for gas exchange analogous to the stomata in vascular plants. *Mpb-gata1* mutants exhibited decreased air pore size compared to BoGa, while air pores in *MpB-GATA1ox3* were significantly bigger compared to the wild type (Figure 2.19A, D).

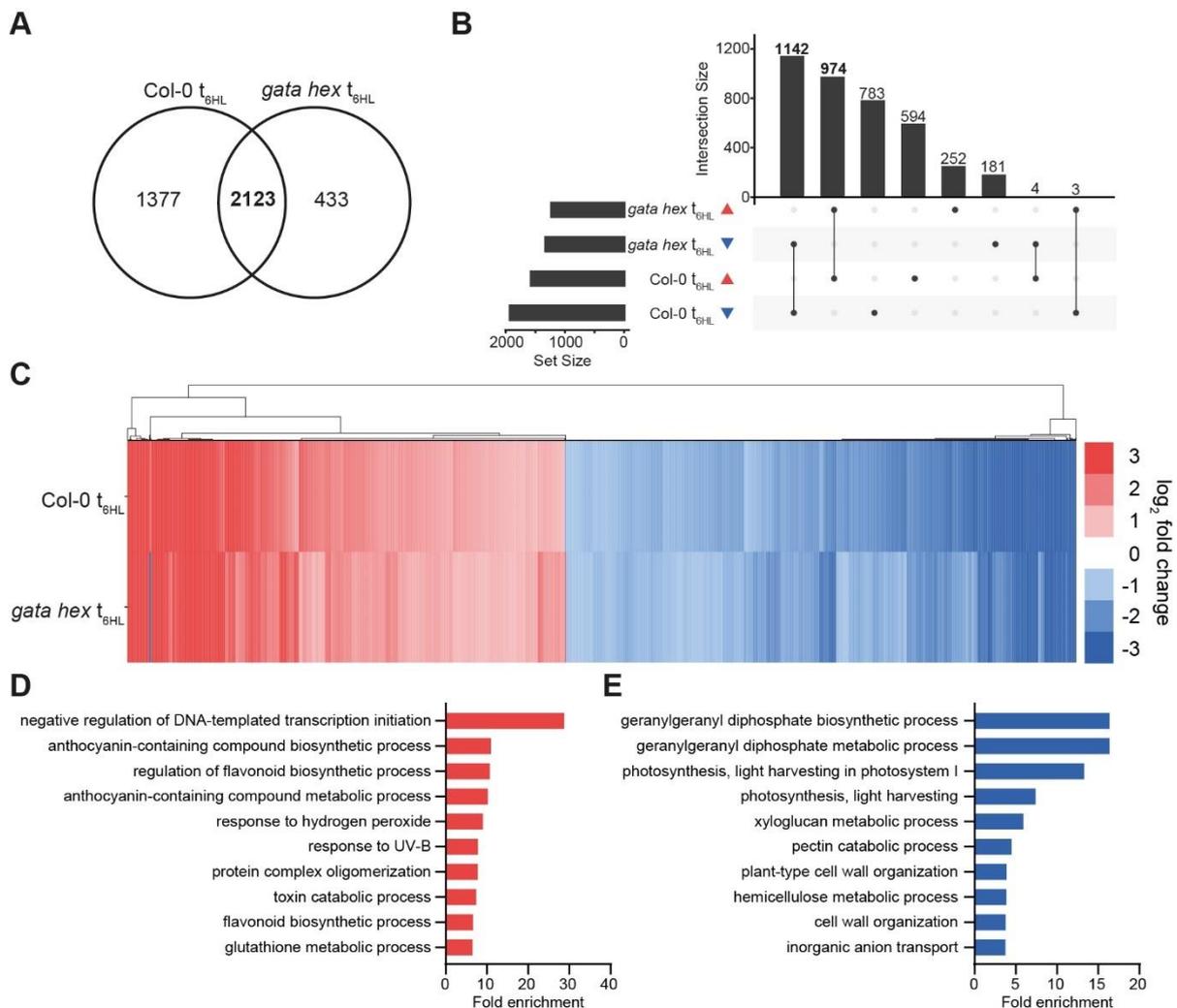


**Figure 2.19: Air pore size in *M. polymorpha* is modulated by high-light and MpB-GATA1.** **A** Representative photograph of three-week-old plants grown under regular growth conditions. **B** Representative photographs of three-week-old plants grown under regular growth conditions for one week before shift to moderate moderate high-light conditions ( $400-500 \mu\text{mol s}^{-1} \text{m}^{-2}$ ) for two weeks. **C** Air pore size measurements of BoGa and *Mpb-gata1* mutants reveal reduced air pore size in mutants compared to BoGa in low light and increased air pore size of all genotypes in high light. **D** Air pore size measurements of BoGa and *MpB-GATA1ox3* reveal increased air pore size in *MpB-GATA1ox3* compared to BoGa in low light and increased air pore size of all genotypes in high light. Scale bar = 2mm. Data sets with no statistical difference after one-way ANOVA and Tukey's HSD posthoc test fall into one group and are labelled with identical letters.

After exposure to moderate high-light conditions ( $400-500 \mu\text{mol s}^{-1} \text{m}^{-2}$ ), air pore size was increased in all genotypes, but most prominently in *MpB-GATA1ox3*, compared to plants grown in low-light (Figure 2.19A - D). These data suggest that high-light promotes air pore size and that MpB-GATA1 can quantitatively modulate this phenotype.

## 2.6.4 *Arabidopsis thaliana* LLM B-GATA *gata hex* maintains high-light stress response

To test whether the role of B-GATAs in high-light stress response is conserved in *A. thaliana*, I performed the same high-light experiment using one-week old seedlings of *A. thaliana* Col-0 and the *gata hex* mutant (Figure 2.20). The plants were grown in low light ( $60 \mu\text{mol s}^{-1} \text{m}^{-2}$ ) for one week and subsequently exposed to either six hours of high light ( $1200 \mu\text{mol s}^{-1} \text{m}^{-2}$ ) or a corresponding low light ( $60 \mu\text{mol s}^{-1} \text{m}^{-2}$ ) mock treatment. I identified 3500 DEGs in Col-0 and 2556 DEGs in *gata hex* after six hours of high-light treatment when filtering the data according to FDR p-value  $\leq 0.05$  and  $\log_2$  fold change  $\geq 1$  and  $\leq -1$ . Of these, 2123 DEGs were present in both genotypes (Figure 2.20A). This overlap is significantly greater than the number of DEGs unique to either Col-0  $t_{6\text{HL}}$  (1377) or *gata hex*  $t_{6\text{HL}}$  (433). Of the 2123 DEGs present in Col-0  $t_{6\text{HL}}$  and *gata hex*  $t_{6\text{HL}}$ , 974 (48.9%) were upregulated in both genotypes while 1142 (53.8%) were downregulated in both genotypes. Only a small subset of 7 genes (0.3%) exhibited non-overlapping expression without any discernible pattern between the genotypes (Figure 2.20B – C).

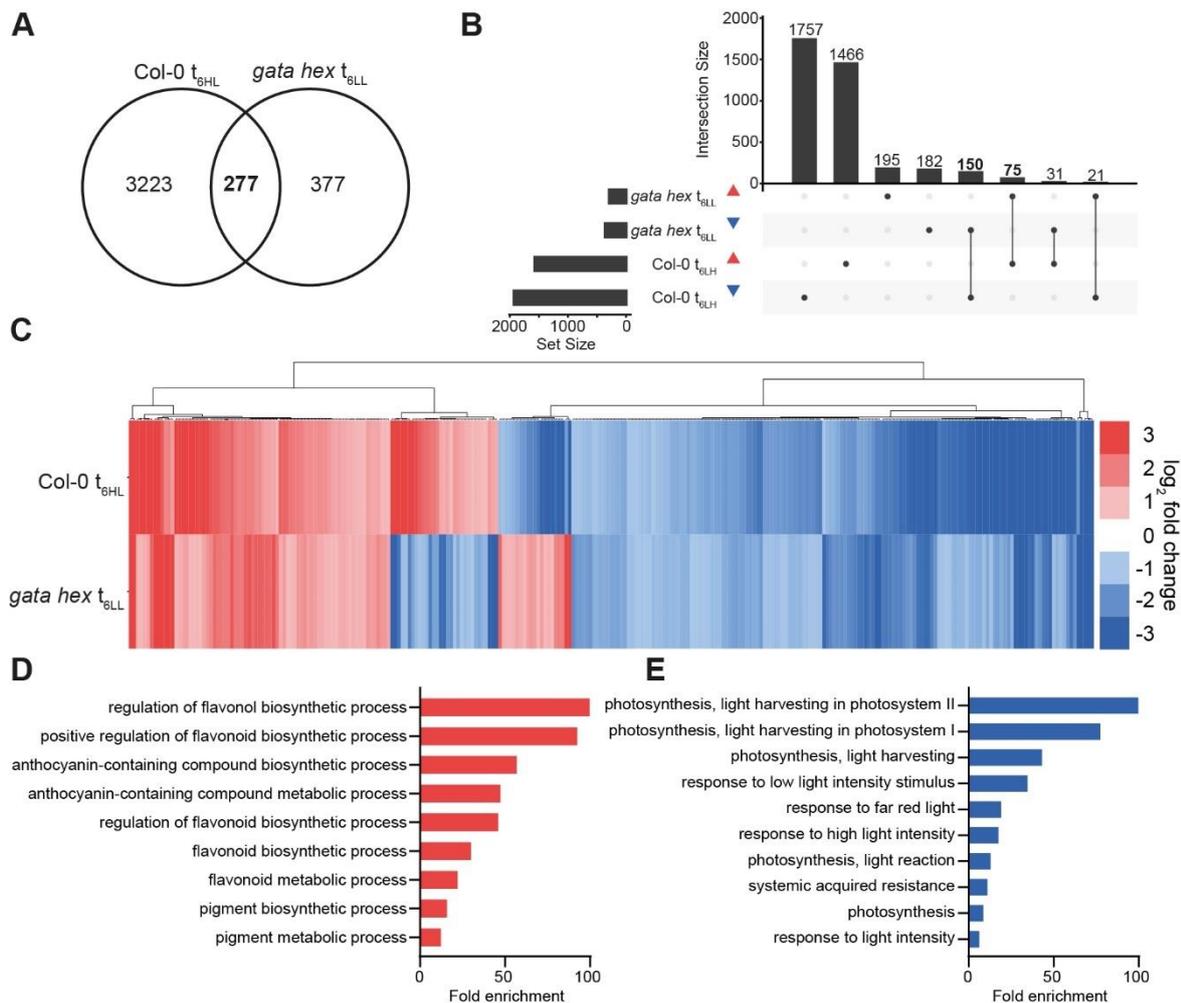


**Figure 2.20: The *A. thaliana gata hex* mutant maintains the same high-light response as observed in Col-0.** **A** Venn diagram of DEGs from Col-0 and *gata hex* at  $t_{6\text{HL}}$  after filtering for FDR p-value  $\leq 0.05$  and  $\log_2$  fold change  $\geq 1$  and  $\leq -1$ . **B** UpSetR diagram showing the direction of the transcriptional regulation of all genes from **A**. **C** Heatmap of the 2123 genes shared by Col-0 and *gata hex* at  $t_{6\text{HL}}$ . Hierarchical clustering is based on Euclidian distance using complete linkage. **D** and **E** The ten most strongly enriched GO terms of the 974 upregulated genes (**D**) and 1142 downregulated genes (**E**) shared by Col-0 and *gata hex* at  $t_{6\text{HL}}$ .

GO analysis of the genes upregulated in both Col-0 and *gata hex*  $t_{6\text{HL}}$  revealed a strong enrichment of terms related to flavonoids (e.g., “anthocyanin-containing compound biosynthetic

process”) and ROS (e.g., “response to hydrogen peroxide”; Figure 2.20D). As ROS are known to be produced under high-light stress and flavonoids are known to be important photoprotectors for excess energy dissipation and thus minimizers of photooxidative damage, the GO enrichment analysis confirms the effect of the high-light treatment. Among the down-regulated genes, GO terms related to photosynthesis (e.g., “photosynthesis, light harvesting in photosystem I” and “photosynthesis, light harvesting”) and cell wall composition (e.g., “xyloglucan metabolic process”, “pectin catabolic process”, “plant-type cell wall organization”, “hemicellulose metabolic process” and “cell wall organization”) were strongly enriched (Figure 2.20E). Similar to the case in *M. polymorpha*, *gata hex* maintained a high-light stress response, but this response was quantitatively reduced in comparison to the wild type.

### 2.6.5 The *A. thaliana* LLM B-GATA *gata hex* exhibits a constitutive high-light stress phenotype



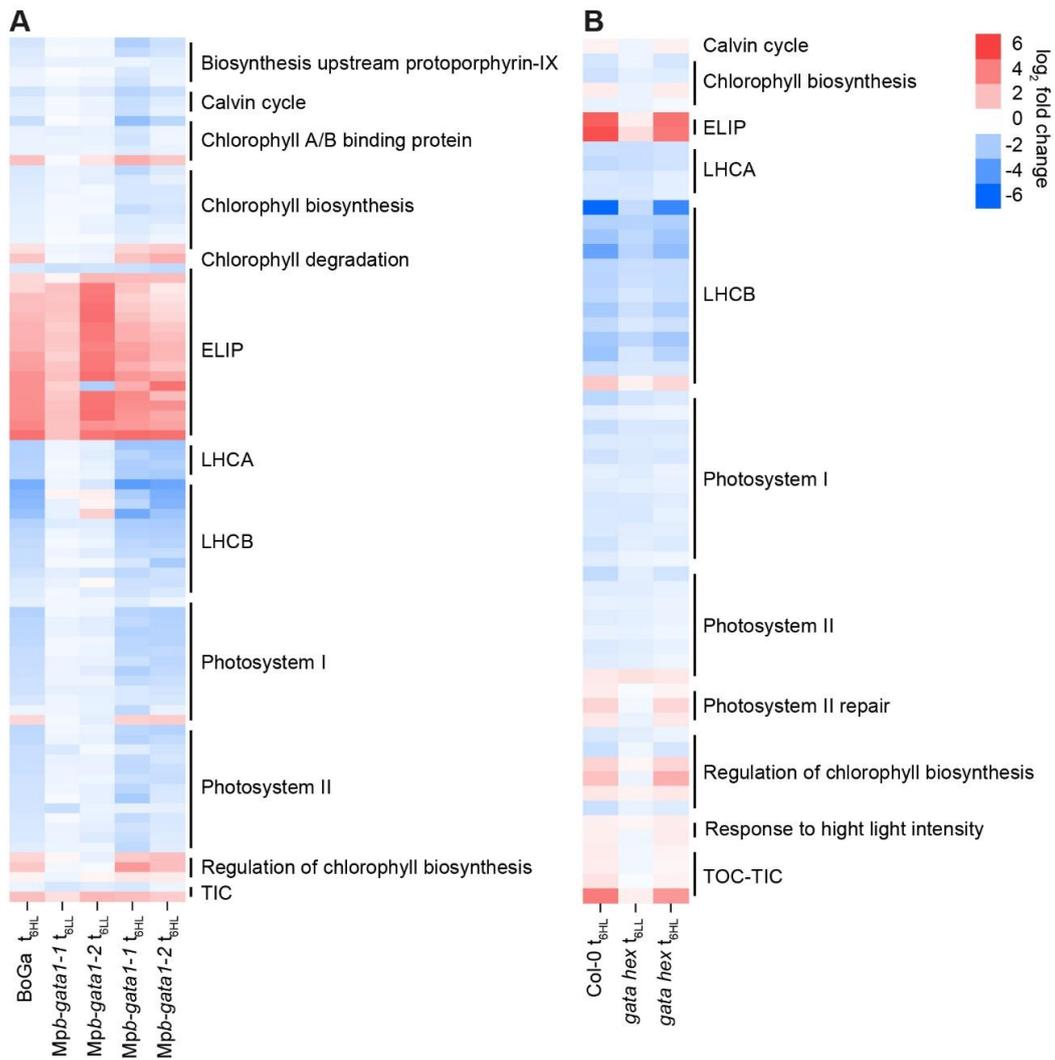
**Figure 2.21: The *A. thaliana gata hex* mutant exhibits a constitutive high-light transcriptome.** **A** Venn diagram of DEGs from Col-0 at  $t_{6HL}$  and *gata hex* at  $t_{6LL}$  after filtering for FDR p-value  $\leq 0.05$  and  $\log_2$  fold change  $\geq 1$  and  $\leq -1$ . **B** UpSetR diagram showing the direction of the transcriptional regulation of all genes from **A**. **C** Heatmap of the 277 genes genes shared by Col-0 at  $t_{6HL}$  and *gata hex* at  $t_{6LL}$  and their regulation as  $\log_2$  fold change. Hierarchical clustering is based on Euclidian distance using complete linkage. **D** and **E** The ten most strongly enriched GO terms of the 75 upregulated genes (**D**) and 150 downregulated genes (**E**) shared by Col-0 at  $t_{6HL}$  and *gata hex* at  $t_{6LL}$ .

After assessing the constitutive high-light response in *M. polymorpha Mpb-gata-1* mutant lines, I next tested whether the *A. thaliana gata hex* displayed a similar transcriptomic pattern. I identified 654 DEGs in *gata hex* after six hours of mock treatment compared to Col-0 after the same mock treatment (*gata hex*  $t_{6LL}$  vs Col-0  $t_{6LL}$ ; Figure 2.21A) when filtering the data according to FDR p-value  $\leq 0.05$  and  $\log_2$  fold change  $\geq 1$  and  $\leq -1$ . Of these 654 DEGs, 277 were also differentially regulated in Col-0 after six hours of high-light (Col-0  $t_{6HL}$  vs Col-0  $t_{6LL}$ ).

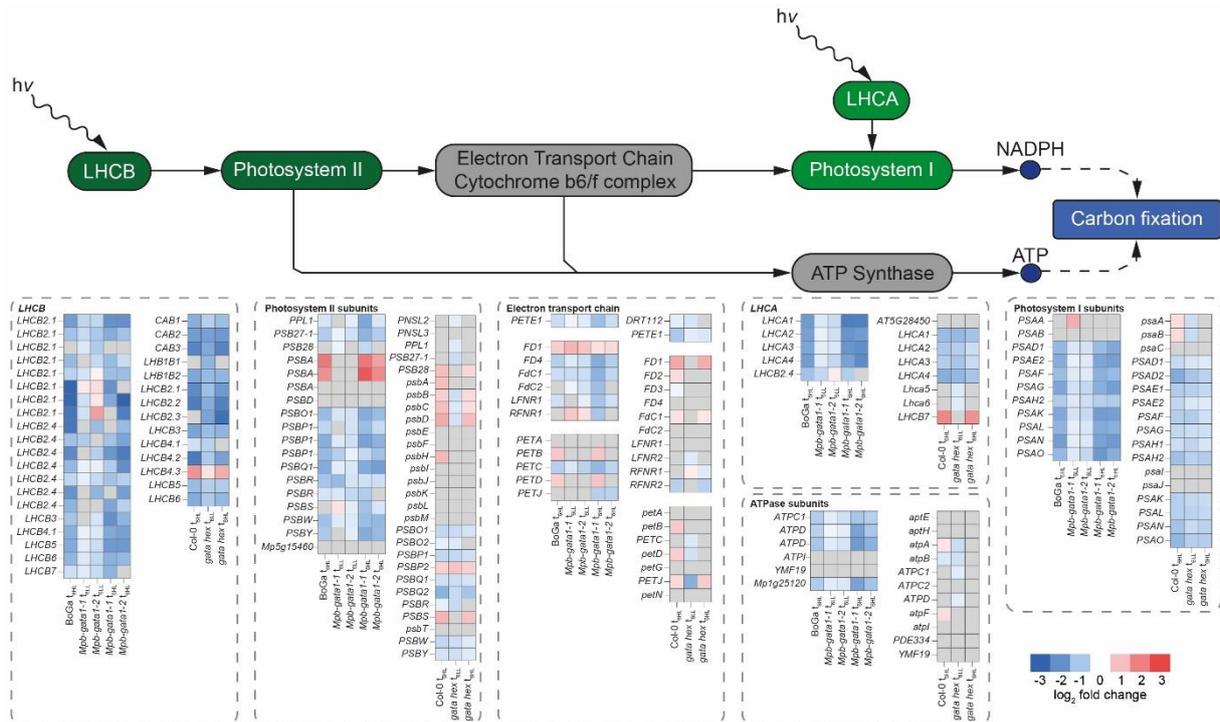
The big difference in the number of DEGs is likely due to the strong effect of the high-light stress treatment for Col-0  $t_{6HL}$ , which has a bigger impact on the transcriptome than the loss of the six LLM-domain containing B-GATAs in *gata hex*. This is not surprising, since light is the single most important environmental cue for plants and excess light is an important abiotic stress factor. Of the 277 DEGs present in BoGa  $t_{6HL}$  and *gata hex*  $t_{6LL}$ , 150 (54%) were upregulated in both genotypes, while 75 (27%) were upregulated in both genotypes. A subset of 52 genes (18%) displayed non overlapping transcriptional regulation. Overall, 81% or 225 of 277 DEGs present in both Col-0  $t_{6HL}$  and *gata hex*  $t_{6LL}$  display the same direction of transcriptional regulation (Figure 2.21B - C). GO enrichment analysis of the 75 upregulated genes yielded many terms related to biosynthesis of flavonoids and anthocyanins, important pigments for photoprotection during exposure to high light (Page, Sultana et al. 2012). Among the 150 downregulated genes, most enriched GO terms were connected to photosynthesis or light sensing (Figure 2.21E). Based on these findings, I concluded that *A. thaliana gata hex* also exhibited a constitutive high-light stress response.

### **2.6.6 The role of B-GATAs in high-light stress is conserved between *M. polymorpha* and *A. thaliana***

Direct comparison of the expression of genes involved in photosynthesis (e.g., PS subunits, LHCS, chlorophyll biosynthesis) revealed a similar pattern to the one already described for *M. polymorpha* (Figure 2.22A). For most genes, the direction of transcriptional regulation observed in Col-0  $t_{6HL}$  is identical to the one in *gata hex*  $t_{6LL}$  and *gata hex*  $t_{6HL}$ , suggesting that *gata hex* plants experience high-light stress in regular growth conditions (Figure 2.22B). Taken together, these results suggest the same constitutive high-light stress for *gata hex* as in the *M. polymorpha* *Mpb-gata1* mutant alleles (Figure 2.22).



**Figure 2.22: Photosynthesis-related genes exhibit constitutive high-light regulation in low-light grown *gata* mutants of *M. polymorpha* and *A. thaliana*.** **A** Differential expression of 88 photosynthesis-related genes in BoGa at  $t_{6HL}$  and Mpb-gata1-1 and Mpb-gata1-2 at  $t_{6LL}$  and  $t_{6HL}$ . **B** Differential expression of 59 photosynthesis-related genes in Col-0 at  $t_{6HL}$  and *gata hex* at  $t_{6LL}$  and  $t_{6HL}$ . The scale bar represents  $\log_2$  fold change.

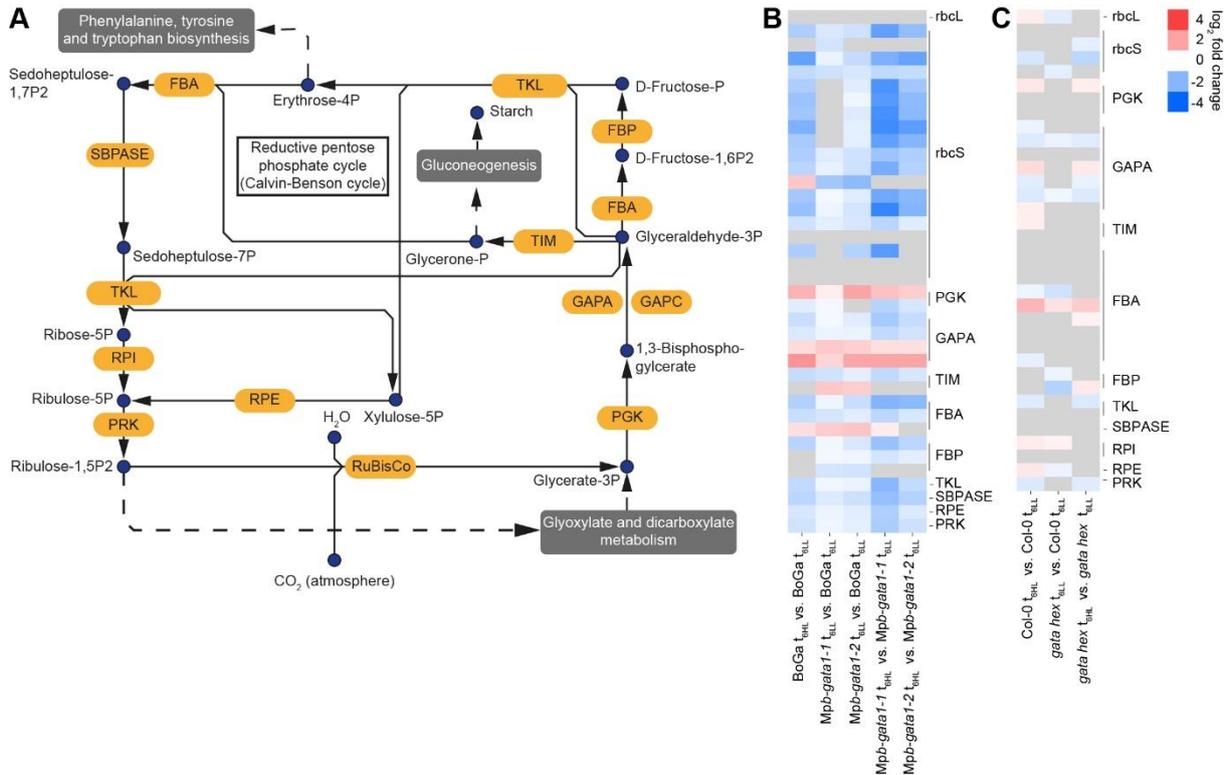


**Figure 2.23: Photosynthesis-associated genes were misregulated in *gata* mutants of *M. polymorpha* and *A. thaliana*.** Schematic representation of photosynthesis with heatmaps displaying the differential expression of genes in the *M. polymorpha* and *A. thaliana* genotypes used in the high-light experiment. Grey boxes indicate no differential expression. See Supplemental Table S5.5 for abbreviations. hv = light energy.

I next analysed the expression of all other genes related to photosynthesis as annotated in the KEGG database for *A. thaliana* in the different *gata* mutants and light conditions, including genes involved in the electron transport chain or genes encoding ATPase subunits. *M. polymorpha* photosynthesis-associated genes were identified via their Arabidopsis orthologs. Strikingly, the majority of all genes was differentially regulated in the *gata* mutants of both species, and the direction of the transcriptional regulation was in most cases conserved for mutants at  $t_{6LL}$  when compared with the wild types and mutants at  $t_{6HL}$  (Figure 2.23). Based on these results, I concluded that regulation of photosynthesis-associated genes is a conserved function of *B-GATAs*.

## 2.6.7 Genes involved in carbon fixation are similarly regulated in *Mpb-gata1* mutants and in response to high-light

After identifying the striking similarity in expression of photosynthesis-associated genes in low-light grown *gata* mutants from *M. polymorpha* and *A. thaliana* to their respective wildtypes, I also analysed gene expression in carbon fixation downstream of the light-dependent reactions of photosynthesis (Figure 2.24A).

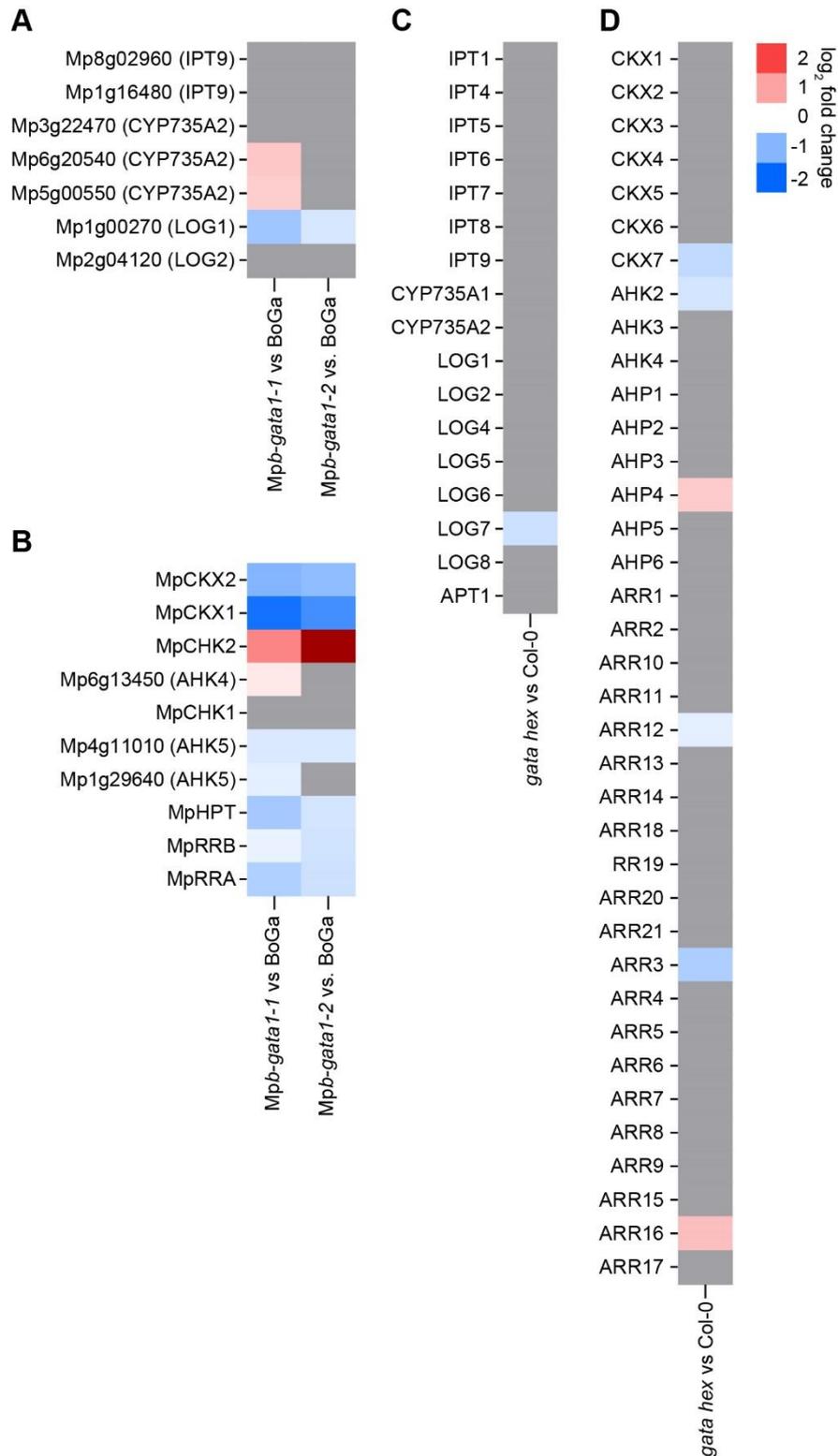


**Figure 2.24: Genes involved in carbon fixation are differentially regulated in high-light and *gata* mutants.** **A** Schematic representation of the carbon fixation in plants. Enzymes are shown in yellow boxes, reactions as solid arrows and metabolites as blue dots. Pathways branching off the carbon fixation are indicated by dotted arrows and grey boxes. **B** Heatmap of differentially expressed genes involved in carbon fixation from the high-light experiment in *M. polymorpha*. **C** Heatmap of differentially expressed genes involved in carbon fixation from the high-light experiment in *A. thaliana*. See Supplemental Table S5.6 for abbreviations. Genes filtered by FDR p-value < 0.05.

Interestingly, most genes in the *Mpb-gata1* mutants at t<sub>6LL</sub> were similarly regulated compared to BoGa at t<sub>6HL</sub> or the mutants at t<sub>6HL</sub> (Figure 2.24B), while this was only the case for few genes in *A. thaliana* (Figure 2.24C). One notable example being the genes encoding RuBisCo small chain subunits (*rbcS*). The family comprises 19 members in *M. polymorpha* and only four members in *A. thaliana*, and most of them were similarly regulated in BoGa at t<sub>6HL</sub> and the *Mpb-gata1* mutants at t<sub>6LL</sub> and t<sub>6HL</sub> (Figure 2.24B). In *A. thaliana*, there was no discernible pattern of regulation across Col-0 and *gata hex* across the different light conditions (Figure 2.24C). Notably, in both species only the RuBisCo small subunits (*rbcS*) were differentially regulated, while the large subunits (*rbcL*; one gene in each *M. polymorpha* and *A. thaliana*) were only affected in *A. thaliana*. The overlap of differential regulation of genes involved in carbon fixation in the *Mpb-gata1* mutants without exposure to high-light and BoGa after high-light treatment is striking, and future analyses are required to investigate the role of *MpB-GATA1* in carbon fixation.

### 2.6.8 Transcriptomic analysis yields differences in cytokinin signalling of *gata* mutants in *M. polymorpha* and *A. thaliana*

Many phenotypes of *A. thaliana b-gata* mutants, e.g. greening and senescence, can be explained through cytokinin signalling (Ranftl, Bastakis et al. 2016). Recently, the role of CK in gemma cup formation of *M. polymorpha* was partly elucidated (Aki, Mikami et al. 2019, Aki, Nishihama et al. 2019). In order to test whether the CK signalling is affected in the *gata* mutants of *M. polymorpha* and *A. thaliana*, I analysed the expression of the genes involved in CK biosynthesis and signalling. In *M. polymorpha*, the genes involved in CK biosynthesis were not consistently regulated among the mutant lines, only the supposed ortholog of *LOG2* was downregulated in both alleles (Figure 2.25A). The expression of the cytokinin oxidase genes *MpCKX1* and *MpCKX2* was strongly reduced in the mutants and the cytokinin receptor *MpCHK2* was upregulated in both mutants (Figure 2.25B). The gene encoding the histidine-containing phosphor-transfer protein HPT, as well as the genes encoding the response regulators *RRA* and *RRB* were downregulated in the mutants (Figure 2.25A). In *A. thaliana*, the genes involved in CK biosynthesis were not differentially regulated in *gata hex*, except for *LOG7*, which was slightly downregulated (Figure 2.25C). Only six of 37 CK signalling genes were differentially regulated in *A. thaliana gata hex*. Additionally, one type-A response regulator (*ARR12*) and one type-B response regulator (*ARR3*) were downregulated in *gata hex* (Figure 2.25D). In conclusion, the differential expression of genes involved in CK biosynthesis and signalling revealed distinct patterns in *M. polymorpha* and *A. thaliana*. Based on the transcriptomic data, it can be speculated that *MpB-GATA1* might be involved in CK signalling.



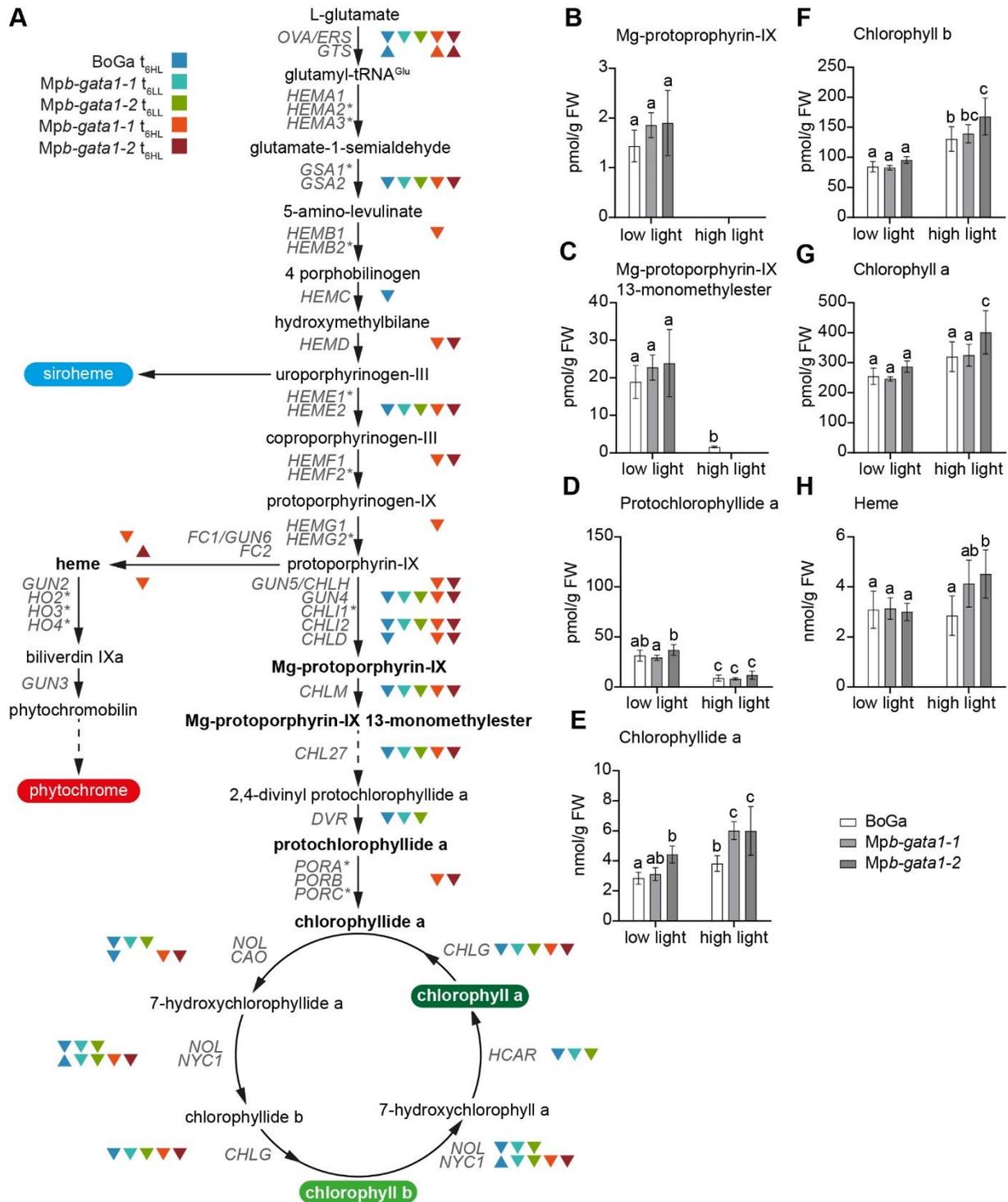
**Figure 2.25: Differential expression of genes involved in cytokinin signalling in *M. polymorpha* and *A. thaliana* gata mutants.** **A** Differential expression of genes involved in cytokinin biosynthesis in *M. polymorpha*. The genes were identified by Aki et al. (2019) or by me based on their orthologs in *A. thaliana* shown in brackets. **B** Differential expression of genes involved in cytokinin signalling in *M. polymorpha*. The genes were identified by Aki et al. (2019) or by me based on their orthologs in *A. thaliana* shown in brackets. **C** Differential expression of genes involved in cytokinin biosynthesis in *A. thaliana*. **D** Differential expression of genes involved in cytokinin signalling in *A. thaliana*. **G** Differential expression of genes involved in cytokinin signalling. IPT, isopentenyl transferase; LOG, cytokinin nucleoside 5'-monophosphate phosphoribohydrolase LONELY GUY; APT1, adenine phosphoribosyl transferase; CKX, cytokinin oxidase; CHK, CHASE domain-containing histidine kinase receptor, AHK in *A. thaliana*; AHP, authentic histidine phosphotransferase, HPT in *M. polymorpha*; RRA, type-A response regulator; RRB, type-B response regulator; ARR, Arabidopsis Response Regulator; P, phosphate.

### 2.6.9 The tetrapyrrole biosynthesis pathway is differentially regulated in high-light stressed BoGa and not-stressed *Mpb-gata1* mutants

Since the analysis of DEGs in *M. polymorpha* BoGa  $t_{6HL}$  and *Mpb-gata1-1/Mpb-gata1-2*  $t_{6LL}$  suggested involvement of MpB-GATA1 in the biosynthesis of chlorophylls, I analysed the response of the tetrapyrrole pathway in plants exposed to high-light in more detail (Figure 2.26A). In BoGa  $t_{6HL}$ , 17 of 27 genes directly involved in the tetrapyrrole biosynthesis pathway were differentially regulated. The genes encoding glutamyl/glutaminyl-tRNA synthetase (GTS, Mp2g10560) and NON-YELLOW COLORING 1 (NYC1, Mp2g20360) were upregulated, while all remaining DEGs were downregulated. In the case of the *Mpb-gata1* mutant alleles at  $t_{6LL}$ , 13 of 27 genes were differentially regulated and all DEGs were downregulated compared to BoGa  $t_{6LL}$ . Most of the differentially expressed genes for BoGa  $t_{6HL}$  were also differentially expressed in *Mpb-gata1*  $t_{6LL}$ , except for *GTS* (Mp2g10560), *HEMC* (Mp2g07740), *CHLD* (Mp6g04260), and *CAO* (Mp2g09760), which were only differentially expressed in BoGa  $t_{6HL}$  but not in the mutant alleles at  $t_{6LL}$ . After six hours of high-light treatment, most genes were downregulated in the *Mpb-gata1* mutant alleles, except for the genes encoding GTS, FERROCHELATASE 2 (FC2, Mp7g00290), GENOMES UNCOUPLED 3 (GUN3, Mp8g10280) and the 7-HYDROXYMETHYL CHLOROPHYLL A REDUCTASE (HCAR, Mp7g02790) which were upregulated in at least one dataset. Based on these data, I concluded that downregulation of genes involved in chlorophyll biosynthesis is a high-light response in *M. polymorpha* and that this response is constitutively exhibited by *Mpb-gata1* mutants.

### 2.6.10 Quantification of chlorophyll intermediates reveals no differences in *Mpb-gata1* mutants

In order to test whether the changes in gene expression affected the concentrations of chlorophylls and their biosynthesis intermediates, high-performance liquid chromatography (HPLC) was performed with three-weeks-old BoGa and *Mpb-gata1* mutant plants from the high-light experiment at  $t_{6LL}$  and  $t_{6HL}$  by Dr. Boris Hedtke from the group of Prof. Dr. Bernhard Grimm (HU Berlin). Despite the striking expression changes of genes from the chlorophyll biosynthesis pathway, there were no measurable differences for Mg-protoporphyrin-IX, Mp-protoporphyrin-IX 13-monomethylester (MME), chlorophyll *b* and chlorophyll *a* in the *Mpb-gata1* mutants under regular growth conditions ( $t_{6LL}$ ). Only protochlorophyllide *a* and chlorophyllide *a* exhibited slightly elevated levels in *Mpb-gata1-2* (Figure 2.26B – H). After six hours of high-light treatment, the concentration of Mg-protoporphyrin-IX was drastically decreased in all genotypes to the point where they could not be quantified. The same applied for Mp-protoporphyrin-IX 13-monomethylester in the case of the *Mpb-gata1* mutant alleles. Interestingly, the *Mpb-gata1* mutant alleles exhibited increased content of chlorophyllide *a* (*Mpb-gata1-1* and *Mpb-gata1-2*), chlorophyll *b* and chlorophyll *a* (only *Mpb-gata1-2*) after six hours of high-light treatment. Additionally, *Mpb-gata1-2* exhibited an increase in the amount of heme at  $t_{6HL}$  (Figure 2.26B - H). Overall, these data suggest that exposure to high-light led to a depletion of the chlorophyll biosynthesis pathway upstream of chlorophyllide *a* in BoGa and the *Mpb-gata1* mutants. Additionally, the changes in gene expression observed in the *Mpb-gata1* mutants did not translate to the amounts of chlorophylls and chlorophyll biosynthesis intermediates.



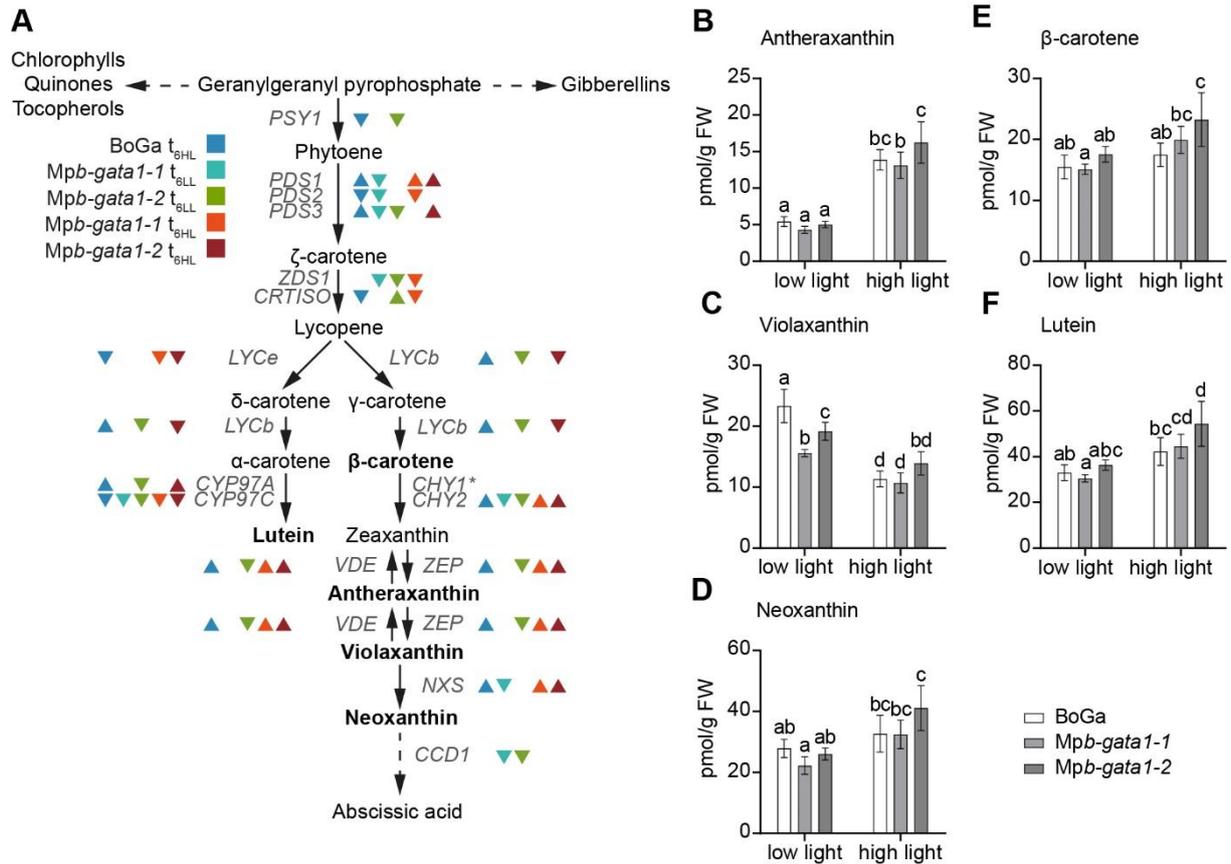
**Figure 2.26: The tetrapyrrole biosynthesis pathway is differentially regulated in *Mpb-gata1* mutants but hardly affected on the level of biosynthesis intermediates.** **A** Schematic representation of the tetrapyrrole biosynthesis pathway with arrows representing enzymatic reactions. Biosynthesis intermediates are indicated by black regular letters, enzymes by grey italic letters. Biosynthesis intermediates in bold face were quantified by HPLC (**B - H**). Enzymes marked with asterisks (\*) are not present in *M. polymorpha*. Coloured triangles indicate differential expression of genes from the respective datasets filtered for FDR  $p$ -value  $\leq 0.05$  with upward facing triangles indicating upregulation and downward facing triangles indicating downregulation. Mutant data at  $t_{6LL}$  were always compared to BoGa at  $t_{6LL}$ , while mutant data at  $t_{6HL}$  was compared to the respective genotype at  $t_{6LL}$ . BoGa  $t_{6HL}$  was compared to BoGa  $t_{6LL}$ . **B - H** Barplots showing HPLC-based quantification of biosynthesis intermediates from six replicates of three-weeks old plants used in the high-light experiment. Error bars show standard deviation. Data sets with no statistical difference after two-way ANOVA and Tukey's HSD posthoc test fall into one group and are labelled with identical letters.

### 2.6.11 Differential expression of the carotenoid biosynthesis pathway

Carotenoids are important pigments for photoprotection in plants. Lutein and xanthophylls like zeaxanthin, violaxanthin and neoxanthin are known to be involved in nonphotochemical quenching (Choudhury and Behera 2001), a means of dissipating excess energy as heat. I analysed the genes involved in carotenoid biosynthesis and found 13 of 15 genes differentially regulated in BoGa after six hours of high-light stress (Figure 2.27A). Five of the 13 genes were downregulated, while eight were upregulated. Noticeably, genes encoding the xanthophyll cycle enzymes VIOLAXANTHIN DE-EPOXIDASE (VDE, Mp7g15960) and ZEAXANTHIN EPOXIDASE (ZEP, Mp2g00670) were both upregulated in BoGa  $t_{6HL}$ . VDE and ZEP are involved in the light-dependent conversion of epoxidized xanthophylls to de-epoxidized ones in high light conditions and their reversion to epoxidized xanthophylls in low light by (Latowski, Kuczynska et al. 2011), respectively, and are thus important components of the high-light response in plants. In the case of the *Mpb-gata1* mutant alleles at  $t_{6LL}$ , eight and 11 genes were differentially regulated for *Mpb-gata1-1* and *Mpb-gata1-2*, respectively, and the genes of the xanthophyll cycle were only downregulated in *Mpb-gata1-2*. After the high-light treatment, seven genes were differentially expressed in both mutant alleles and three additional genes were only differentially regulated in *Mpb-gata1-1* and not in *Mpb-gata1-2*. In all cases, the DEGs from the *Mpb-gata1* alleles at  $t_{6HL}$  had the same direction of transcriptional regulation as the wildtype after the high-light treatment (Figure 2.27A). Taken together, exposure to high-light intensities resulted in the expected upregulation of genes involved in the xanthophyll cycle regardless of genotype. *MpB-GATA1* seems to be involved in the regulation of carotenoid biosynthesis, but *Mpb-gata1* mutants did not exhibit a constitutive high-light regulation of the pathway.

### 2.6.12 Quantification of carotenoids and xanthophylls

Similar to the quantification of chlorophyll intermediates, carotenoid biosynthesis intermediates were analysed by Dr. Boris Hedtke from the group of Prof. Dr. Bernhard Grimm (HU Berlin). HPLC-based quantification of lutein,  $\beta$ -carotene, antheraxanthin, violaxanthin and neoxanthin revealed reduced contents only of violaxanthin in *Mpb-gata1-1* and *Mpb-gata1-2* at  $t_{6LL}$  in comparison to BoGa (Figure 2.27B - F). After the high-light treatment, the violaxanthin content was significantly reduced in all genotypes, while the antheraxanthin content drastically increased in all genotypes compared to  $t_{6LL}$ . For violaxanthin, antheraxanthin, and neoxanthin, there were no significant differences between the mutants and BoGa at  $t_{6HL}$ . Lutein,  $\beta$ -carotene and neoxanthin contents increased in the *Mpb-gata1* mutant alleles after the high-light treatment, but only significantly in *Mpb-gata1-2*. The differences observed between BoGa and the *Mpb-gata1* mutants were subtle, indicating that *MpB-GATA1* only plays a minor role in the regulation of carotenoid biosynthesis.



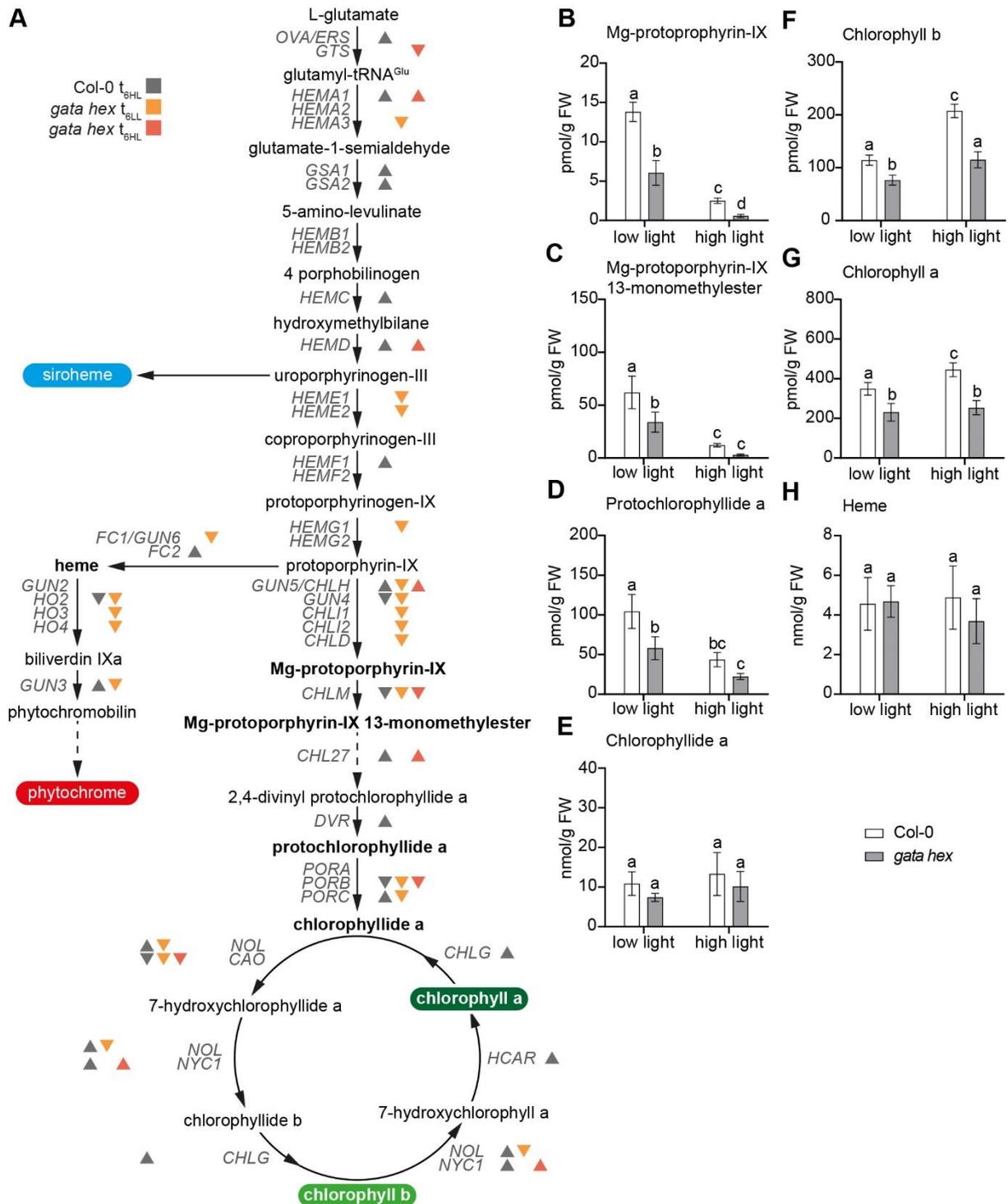
**Figure 2.27: The carotenoid biosynthesis pathway is differentially regulated in *Mpb-gata1* mutants but hardly affected on the level of biosynthesis intermediates.** **A** Schematic representation of the carotenoid biosynthesis pathway with arrows representing enzymatic reactions. Biosynthesis intermediates are indicated by black regular letters, enzymes by grey italic letters. Biosynthesis intermediates in bold face were quantified by HPLC (**B - F**). Enzymes marked with asterisks (\*) are not present in *M. polymorpha*. Coloured triangles indicate differential expression of genes from the respective datasets filtered for FDR p-value  $\leq 0.05$  with upward facing triangles indicating upregulation and downward facing triangles indication downregulation. Mutant data at  $t_{6LL}$  were always compared to BoGa at  $t_{6LL}$ , while mutant data at  $t_{6HL}$  was compared to the respective genotype at  $t_{6LL}$ . BoGa  $t_{6HL}$  was compared to BoGa  $t_{6LL}$ . **B - E** Barplots showing HPLC-based quantification of biosynthesis intermediates from six replicates of three-weeks old plants used in the high-light experiment. Error bars show standard deviation. Data sets with no statistical difference after two-way ANOVA and Tukey's HSD posthoc test fall into one group and are labelled with identical letters.

### **2.6.13 The tetrapyrrole biosynthesis pathway is differentially regulated in *A. thaliana gata hex* and in response to high-light**

In order to test whether the drastic changes in the chlorophyll biosynthesis pathway were conserved among *M. polymorpha* and *A. thaliana*, I also analysed the genes involved in the tetrapyrrole pathway in *A. thaliana* (Figure 2.28A). In Col-0  $t_{6HL}$ , 21 of 40 genes were differentially regulated. Noticeably, sixteen of these genes were upregulated and only five were downregulated. This transcriptomic profile strongly opposes the one from BoGa at  $t_{6HL}$ , in which 15 of 17 DEGs were downregulated (Figure 2.26A). The *gata hex* mutant exhibited 19 differentially regulated genes after six hours of low light treatment, all of which were downregulated. Interestingly, the majority of these genes (15 of 19) were downstream of protoporphyrin-IX, both in the biosynthesis pathway of phytochromes and chlorophylls. After the high-light treatment, nine genes were differentially regulated in *gata hex* (Figure 2.28A). Similar to the case in *M. polymorpha* BoGa and the *Mpb-gata1* alleles, the direction of transcriptional regulation observed in *gata hex* at  $t_{6HL}$  and Col-0 at  $t_{6HL}$  was the same for all DEGs (Figure 2.26A). The changes in gene expression confirm that LLM B-GATAs are involved in transcriptional regulation of tetrapyrrole biosynthesis genes. However, the differential expression of genes in *gata hex* at  $t_{6LL}$  did not resemble the expression pattern of genes from Col-0 at  $t_{6HL}$ , suggesting that loss of all six LLM B-GATAs did not cause constitutive high-light regulation of the tetrapyrrole biosynthesis pathway.

### **2.6.14 Quantification of chlorophyll intermediates confirms role of *A. thaliana* B-GATAs in greening**

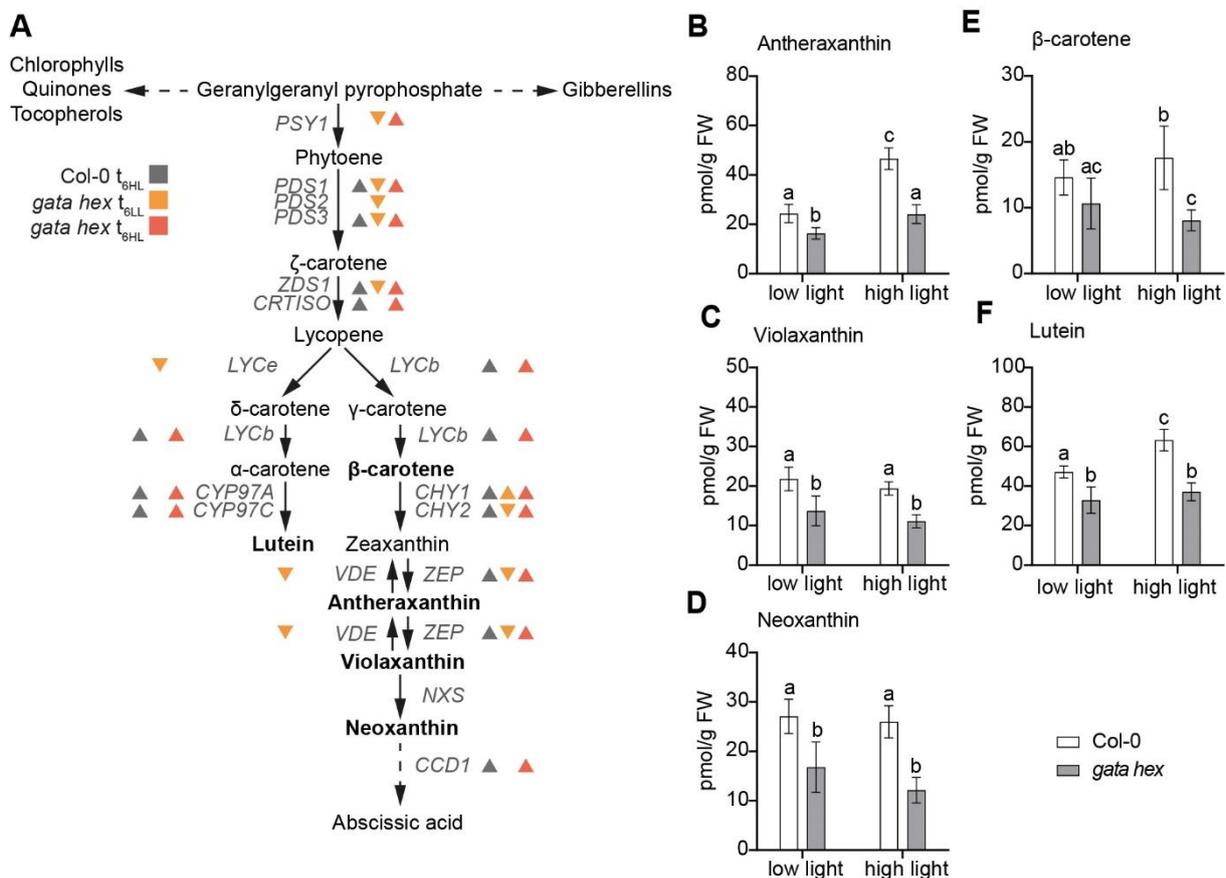
The role of LLM-domain containing B-GATAs in greening and, more specifically, chlorophyll biosynthesis, has already been described in previous publications from our laboratory (Behringer, Bastakis et al. 2014, Ranftl, Bastakis et al. 2016, Bastakis, Hedtke et al. 2018). In accordance with our knowledge of established *gata* mutants and the gene expression data, *gata hex* displayed significantly lower concentrations of Mp-protoporphyrin-IX, MME, protochlorophyllide *a*, chlorophyllide *a*, as well as chlorophyll *b* and chlorophyll *a* at  $t_{6LL}$  compared to Col-0 at  $t_{6LL}$  in HPLC analysis of intermediates in the tetrapyrrole pathway. Heme contents were unchanged in *gata hex* at  $t_{6LL}$ . After six hours of high-light, the concentrations of Mp-protoporphyrin-IX, MME and protochlorophyllide *a* were drastically reduced in Col-0 and *gata hex*. Chlorophyllide *a* content remained almost unchanged, while chlorophyll *b* was significantly increased in all genotypes. The levels of chlorophyll *a* remained constant in *gata hex* but increased in Col-0 at  $t_{6HL}$  compared to  $t_{6LL}$  (Figure 2.28B - H). Overall, *gata hex* displayed the expected reduction of chlorophyll content and content of chlorophyll biosynthesis intermediates. High-light led to increased chlorophyll *a* and *b* content in Col-0, but only chlorophyll *b* increased in *gata hex*, suggesting that *gata hex* is deficient in the conversion of chlorophyll *b* to chlorophyll *a*.



**Figure 2.28: The tetrapyrrole biosynthesis pathway is affected in *A. thaliana gata hex* on the level of transcriptional regulation and biosynthesis intermediates.** **A** Schematic representation of the tetrapyrrole biosynthesis pathway with arrows representing enzymatic reactions. Biosynthesis intermediates are indicated by black regular letters, enzymes by grey italic letters. Biosynthesis intermediates in bold face were quantified by HPLC (**B - F**). Coloured triangles indicate differential expression of genes from the respective datasets filtered for FDR p-value  $\leq 0.05$  with upward facing triangles indicating upregulation and downward facing triangles indicating downregulation. Mutant data at  $t_{6LL}$  were always compared to Col-0 at  $t_{6LL}$ , while mutant data at  $t_{6HL}$  was compared to the respective genotype at  $t_{6LL}$ . Col-0  $t_{6HL}$  was compared to Col-0  $t_{6LL}$ . **B - E** Barplots showing HPLC-based quantification of biosynthesis intermediates from six replicates of seven-days old plants used in the high-light experiment. Error bars show standard deviation. Data sets with no statistical difference after two-way ANOVA and Tukey's HSD posthoc test fall into one group and are labelled with identical letters.

## 2.6.15 Differential expression of the carotenoid biosynthesis pathway in *A. thaliana gata hex*

In order to test whether LLM B-GATAs from *A. thaliana* are involved in the regulation of carotenoid biosynthesis, I analysed the expression of genes involved in the biosynthesis pathway. After six hours of high-light treatment, 12 of 16 genes of the carotenoid biosynthesis pathway were differentially regulated in Col-0 (Figure 2.29A). All these genes were upregulated. In *gata hex* at  $t_{6LL}$ , nine genes were downregulated, and one gene was upregulated. Noticeably, the genes of the xanthophyll cycle, *AtZEP* and *AtVDE*, were both downregulated in *gata hex* at  $t_{6LL}$ , while *AtZEP* was upregulated and *AtVDE* was not differentially regulated in Col-0  $t_{6HL}$ . At  $t_{6HL}$ , the direction of transcriptional regulation of all genes was the same in Col-0 and *gata hex* (Figure 2.29A). Based on these results, it seems that LLM B-GATAs are involved in the regulation of genes of the carotenoid biosynthesis pathway. Similar to the case described above for the tetrapyrrole biosynthesis pathway however, loss of all six LLM B-GATAs did not lead to a constitutive high-light regulation of the pathway or impair the high-light response observed in Col-0.



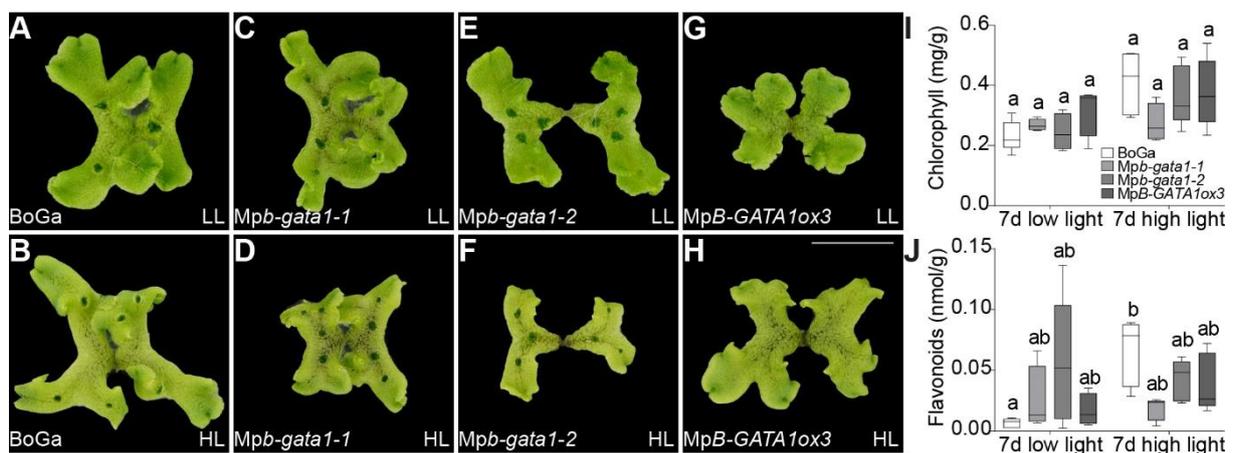
**Figure 2.29: The carotenoid biosynthesis pathway is affected in *A. thaliana gata hex* on the level of transcriptional regulation and biosynthesis intermediates.** **A** Schematic representation of the carotenoid biosynthesis pathway with arrows representing enzymatic reactions. Biosynthesis intermediates are indicated by black regular letters, enzymes by grey italic letters. Biosynthesis intermediates in bold face were quantified by HPLC (**B - F**). Coloured triangles indicate differential expression of genes from the respective datasets filtered for FDR p-value  $\leq 0.05$  with upward facing triangles indicating upregulation and downward facing triangles indicating downregulation. Mutant data at  $t_{6LL}$  were always compared to Col-0 at  $t_{6LL}$ , while mutant data at  $t_{6HL}$  was compared to the respective genotype at  $t_{6LL}$ . Col-0  $t_{6HL}$  was compared to Col-0  $t_{6LL}$ . **B - F** Barplots showing HPLC-based quantification of biosynthesis intermediates from six replicates of seven-days old plants used in the high-light experiment. Error bars show standard deviation. Data sets with no statistical difference after two-way ANOVA and Tukey's HSD post-hoc test fall into one group and are labelled with identical letters.

## 2.6.16 Quantification of carotenoids suggests impairment of the xanthophyll cycle in *gata hex*

After showing that LLM B-GATAs are involved in the transcriptional regulation of carotenoid biosynthesis, I analysed whether this regulation translated to the metabolite level. In low light conditions, *gata hex* contained lower amounts of violaxanthin, neoxanthin and lutein, while amounts of antheraxanthin and  $\beta$ -carotene remained unchanged compared to Col-0 (Figure 2.29B - F). After exposure to high light the amount of violaxanthin was unchanged in Col-0 and *gata hex* compared to the respective genotypes in low light. The amount of antheraxanthin, however, was strongly increased in Col-0 at  $t_{6HL}$ , but not in *gata hex*. This suggests an impairment of the xanthophyll cycle in *gata hex*, despite the fact that the xanthophyll cycle genes *ZEP* and *VDE* were upregulated in *gata hex* at  $t_{6HL}$ . Similar observations were made for the amount of lutein, which was strongly increased after exposure to high-light in Col-0 but remained unchanged in *gata hex*. Unfortunately, not all intermediates of the carotenoid pathway could be analysed here. Further analyses are needed to gain information about the amount of zeaxanthin and thus help in understanding the impairment of the xanthophyll cycle in of LLM-domain B-GATA mutants.

## 2.6.17 Long term high-light stress did not reveal GATA effect in *M. polymorpha*

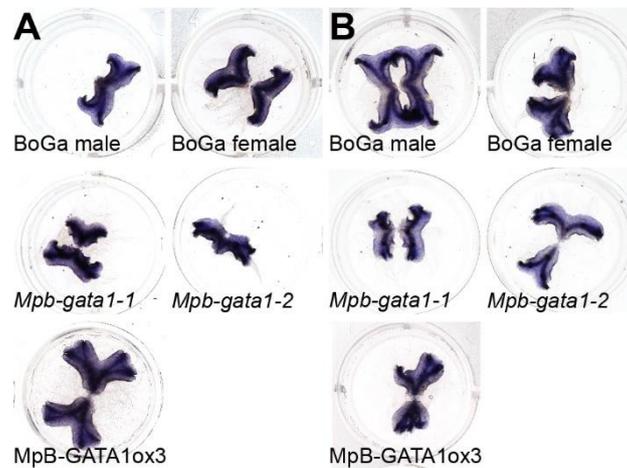
After assessing the constitutive high-light response of the *Mpb-gata1* mutants, I exposed plants to a prolonged growing period under moderate high-light ( $400 - 500 \mu\text{mol s}^{-1} \text{m}^{-2}$ ) in order to test whether *MpB-GATA1* plays a role in the long-term high-light-stress response. To this end, I analysed chlorophyll-, flavonoid-, and ROS-accumulation. Exposure to constant moderate high-light for one week lead to inward curling of the thallus in all genotypes exhibited, which is likely to be a light-avoidance strategy. The effects were equally pronounced in the wild type and the mutants, but visibly less so in the overexpression line (Figure 2.30A – H). Chlorophyll content did not change significantly in any of the genotypes when comparing plants exposed to low- or high light (Figure 2.30I). There were no discernible differences in flavonoid content between the genotypes in low light and flavonoid content only significantly increased in BoGa but not in the mutants or the overexpressor when subjected to high-light (Figure 2.30J).



**Figure 2.30: *M. polymorpha* exhibits high-light avoidance unaffected by loss- or overexpression of *MpB-GATA1*.** A, C, E, and G Representative photographs of three-weeks-old BoGa, *Mpb-gata1* mutants and *MpB-GATA1ox3* exposed to constant low-light for seven days. B, D, F, and H Representative photographs of three-weeks old BoGa, *Mpb-gata1* mutants and *MpB-GATA1ox3* exposed to constant moderate high-light ( $400 - 500 \mu\text{mol s}^{-1} \text{m}^{-2}$ ) for seven days exhibit inward curling of the thallus, a putative light-avoidance mechanism. The midrib of the *Mpb-gata1* mutants and *MpB-GATA1ox3* is visibly darker compared to BoGa, which might suggest increased flavonoid content. Scale bar = 1cm. I Chlorophyll quantification of plants of the specified genotypes after one week in low- or high light. J Flavonoid quantification of plants of the specified genotypes after one week in low- or high light. Data sets with no statistical difference after two-way ANOVA and Tukey's HSD post-hoc test fall into one group and are labelled with identical letters. n = 5.

In order to test whether *M. polymorpha* plants exposed to moderate high-light exhibit formation of ROS, I subjected the plants to either zero (Figure 2.31A) or six hours (Figure 2.31B) of moderate high-light and performed NBT staining (Grellet Bournonville and Díaz-Ricci 2011) for

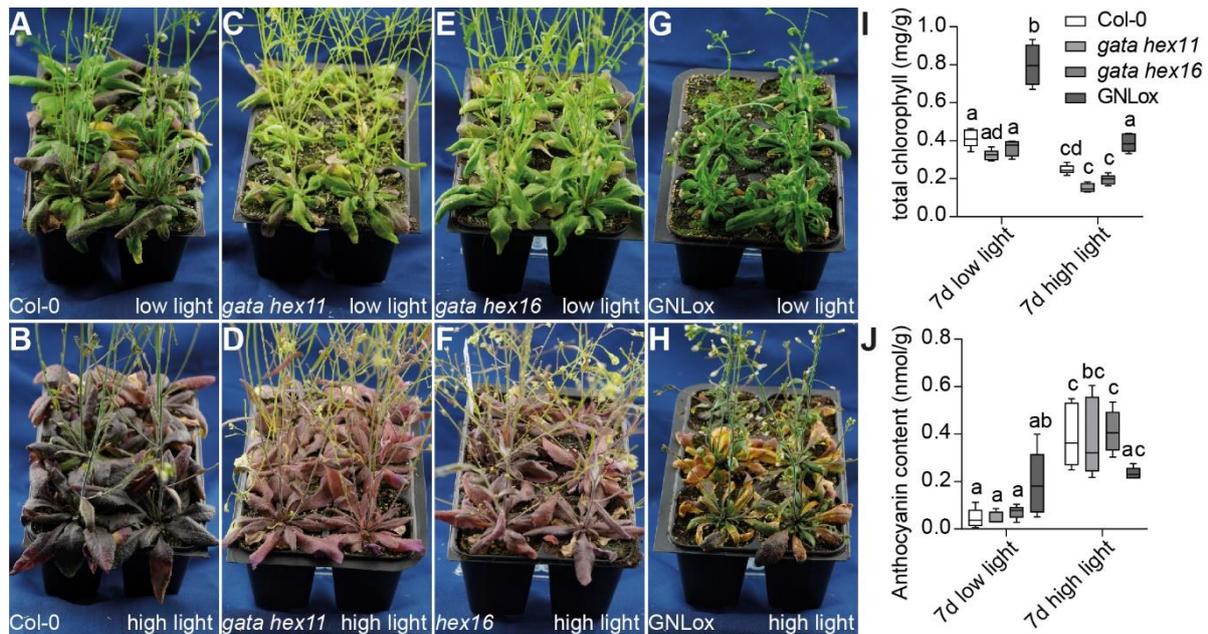
the detection of superoxide. Even plants not subjected to high-light were fully stained (Figure 2.31A - B), suggesting constitutively high levels of superoxide in *M. polymorpha*. Due to this over-staining, I did not test further time points.



**Figure 2.31: NBT staining of *M. polymorpha* BoGa, *Mpb-gata1* mutants and MpB-GATA1ox3 suggests constitutively high superoxide content without exposure to moderate high-light. A** NBT staining of BoGa, *Mpb-gata1* mutants and MpB-GATA1ox3 without exposure to high-light resulted in fully stained plants. **B** NBT staining of BoGa, *Mpb-gata1* mutants and MpB-GATA1ox3 after six hours of high-light resulted in fully stained plants. n = 6.

### 2.6.18 *A. thaliana* GNLox is deficient in anthocyanin biosynthesis

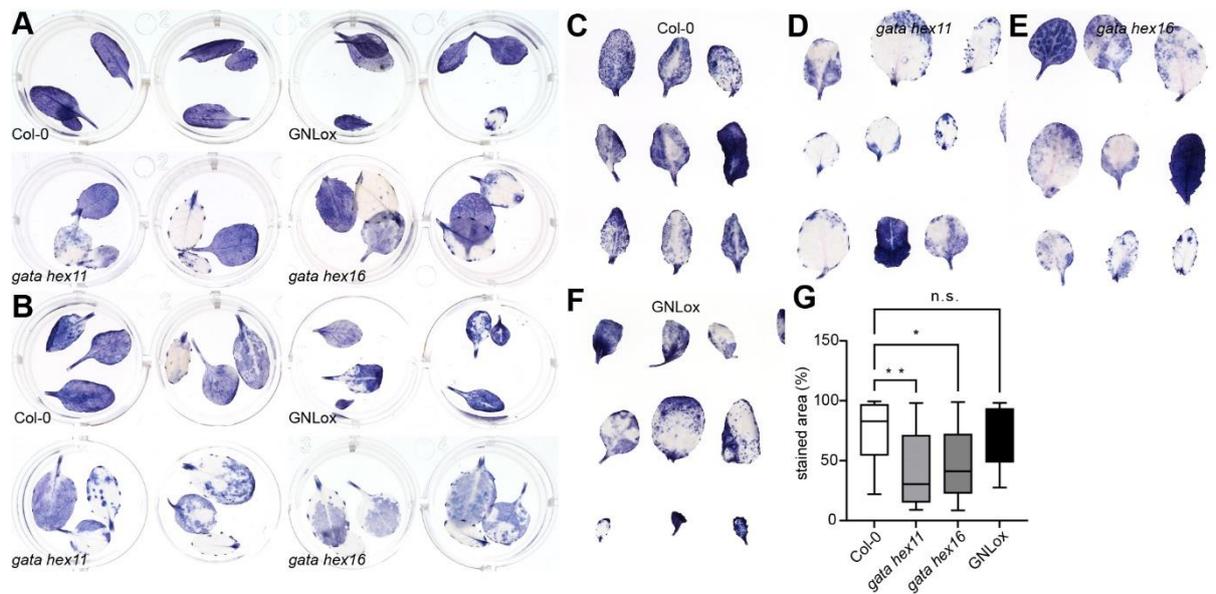
In order to test the involvement of the LLM B-GATAs in the stress response to prolonged high-light, I exposed *A. thaliana* plants grown for 35 days under standard conditions to moderate high light ( $400 - 500 \mu\text{mol s}^{-1} \text{m}^{-2}$ ) or a corresponding low light treatment for one week. In low light, the *gata hex* mutants were visibly paler compared to Col-0, while GNLox was visibly darker (Figure 2.32A – G), but chlorophyll quantification did not yield significant differences in chlorophyll content between Col-0 and the mutants, whereas chlorophyll content was strongly increased in GNLox (Figure 2.32I). After the high-light treatment, chlorophyll content of all genotypes was significantly reduced compared to low-light treated plants. Also here, there were no significant differences for chlorophyll content between Col-0 and the mutants, while GNLox still showed increased chlorophyll content (Figure 2.32I). After exposure to high-light, the *gata hex* mutants exhibited a visibly lighter colour, but overall normal growth, compared to Col-0, while GNLox showed severe yellowing and curling of leaves after the same treatment (Figure 2.32B - H). In low light, anthocyanin content was unchanged for all genotypes and increased significantly for Col-0 and the *gata hex* mutants but not GNLox after exposure to high light (Figure 2.32J). These results suggest that prolonged exposure to high-light, in contrast to the six-hour treatments used in the transcriptomic experiments, results in reduced chlorophyll content, independently of LLM B-GATAs. Additionally, I concluded that anthocyanin biosynthesis *per se* is unaffected by the loss of LLM B-GATAs, but that the high-light-mediated anthocyanin production might be negatively regulated by B-GATAs, or at least GNL.



**Figure 2.32: *A. thaliana* plants are differentially affected by high light intensities.** **A, C, E, and G** Representative photographs of thirty-five days-old *A. thaliana* Col-0, *gata hex* and GNLox exposed to constant low light for one week. **B, D, F, and H** Representative photographs of thirty-five days-old *A. thaliana* Col-0, *gata hex* and GNLox exposed to constant moderate high light for one week. **I** Quantification of chlorophyll content of plants exposed to seven days of low- or high light. Entire rosettes were used for the analysis, stems and roots were removed. **J** Quantification of anthocyanin content of plants exposed to seven days of low- or high light. Entire rosettes were used for the analysis, stems and roots were removed. Data sets with no statistical difference after two-way ANOVA and Tukey's HSD post-hoc test fall into one group and are labelled with identical letters.  $n = 5$ .

### 2.6.19 *A. thaliana gata hex* accumulates less ROS compared to Col-0 after exposure to high light

Exposure to light intensities that exceed the photosynthetic capacities of plants can lead to the formation of ROS through reactions of excited chlorophyll molecules with atmospheric oxygen (Eberhard, Finazzi et al. 2008). To test whether LLM B-GATAs play a role in high-light induced ROS formation, I exposed three-weeks-old *A. thaliana* plants to constant moderate high-light ( $400-500\mu\text{mol s}^{-1} \text{m}^{-2}$ ) for zero hours, six hours, and 30 hours (Figure 2.33A – F). After zero hours of high-light, the *gata hex* alleles exhibited less NBT staining compared to Col-0 and GNLox (Figure 2.33A). The result was very similar for six hours and 30 hours (Figure 2.33B - F). Quantification of the stained area after 30 hours of high-light in respect to leaf size yielded significantly less staining for the mutants, while GNLox did not show any differences to Col-0 (Figure 2.33C - G). This suggests a lower amount of superoxide in the *gata hex* mutants compared to Col-0, possibly because their constitutive high-light responses act as a priming mechanism.



**Figure 2.33: *gata hex* produces less ROS after exposure to moderate high-light compared to Col-0 and GNLox.** A and B NBT staining of adult plants of *Arabidopsis thaliana* Col-0, *gata hex* and GNLox for zero hours and six hours, respectively. C – F Leaves of Col-0, *gata hex* and GNLox after exposure for moderate high-light for 30 hours. G Quantification of the NBT-stained area in relation to leaf size showed significantly less staining for *gata hex* compared to Col-0. Asterisks indicate statistically significant difference in mutant and overexpression lines versus wild types in a Student's t-test. \*,  $0.05 > p > 0.01$ ; \*\*,  $0.01 > p > 0.001$ . n = 12.

### 3 Discussion

The aims of my project were to investigate the origin of B-GATAs with HAN- and LLM-domain in the plant lineage and to find out when the division of the two distinct B-GATA subfamilies occurred. In comparison to the previously characterized angiosperm B-GATAs, these stand out because they contain both HAN- and LLM-domains, a constellation only found in bryophytes. Additionally, I wanted to analyse the functions of B-GATAs from the bryophytes *P. patens* and *M. polymorpha* in order to understand the evolutionary divergence or conservation of the B-GATA factors, as well as the HAN- and LLM-domains. In order to gain insights into the functions of these transcription factors, I combined a range of different techniques and approaches. I generated mutants in *P. patens*, *M. polymorpha*, and *A. thaliana* making use of CRISPR/Cas9-mediated genome editing and found novel roles for B-GATAs in all three species. I performed several transcriptomic experiments and uncovered a potentially conserved role of GATAs in the repression of high-light stress in both *M. polymorpha* and *A. thaliana*.

#### 3.1 Bryophytes contain B-GATAs with both HAN- and LLM-domains

My phylogenetic analysis of GATA zinc-finger sequences from the bryophytes *M. polymorpha* and *P. patens*, the lycophyte *S. moellendorffii*, the fern *S. cucullata*, and the streptophyte alga *C. braunii* revealed the conservation of the four GATA classes A-D originally described for *A. thaliana* and *O. sativa* across the entire plant lineage (Figure 2.1A; Reyes, Muro-Pastor et al. 2004). In most cases, the characteristic additional protein domains of the distinct GATA classes and the relative location of the zinc finger domain was conserved in the early diverging plant species used for the construction of the phylogeny (Figure 2.1B). This analysis was the initial phylogenomic description of GATA transcription factors from the plant lineage and, based on my observations, we were invited to present and discuss these findings in a review article in collaboration with Crysten E. Blaby-Haas (Brookhaven National Laboratory, Upton, New York, USA; Schwechheimer, Schröder et al. 2022).

I was able to show that the number of GATAs was highly variable among the species included, with *A. thaliana* encoding the most GATAs (30) and *C. braunii* and *M. polymorpha* encoding the smallest number of GATAs (six; Figure 2.1). The large numbers of GATAs in *A. thaliana*, *P. patens*, and *S. cucullata* are likely due to recent whole genome duplication events (Rensing, Ick et al. 2007, Bomblies and Madlung 2014, Li, Brouwer et al. 2018). Interestingly, I could not identify any D-GATAs from *S. moellendorffii*. Class D contains the fewest members in all species analysed here, and since nothing is known about their biological or biochemical functions to date, it is difficult to speculate why this class is absent in *S. moellendorffii*. *C. braunii* does not encode any A-GATAs, but one of its GATAs was resolved in a clade together with AtGATA14, which was previously considered an A-GATA. In *A. thaliana* and other embryophytes, class A is largest of the GATA classes and its members are involved in de-etiolation, root apical meristem development, formation of xylem vessels, and primary dormancy (Liu, Koizuka et al. 2005, Luo, Lin et al. 2010, Endo, Yamaguchi et al. 2015, Jiang, Yung et al. 2018) but these processes do not take place in streptophyte algae like *C. braunii*. This finding suggests that A-GATAs evolved after the water-to-land transition, possibly in order to perform new developmental functions required for life on land.

In *A. thaliana*, B-GATAs are the second largest class with 11 members and the proteins usually contain either a HAN- or an LLM-domain. I found the same arrangement in the genome of *S. moellendorffii*, which encodes two LLM-domains containing B-GATAs and one GATA with a HAN-domain. The genome of the fern *S. cucullata* contained three B-GATAs with an LLM-domain, one without any known secondary domains, and one with only a partial HAN-domain (Sacu\_s0008.g004090\_142-173), while *C. braunii* encoded three LLM-domain containing B-GATAs but none with a HAN-domain. The presence of the LLM-domain and absence of the

HAN-domain in *C. braunii* B-GATAs suggests that the LLM-domain may have arisen first in green algae, and that the HAN-domain evolved in the ancestor of land plants.

The bryophytes *M. polymorpha* and *P. patens* encode B-GATAs with both HAN- and LLM-domains on the same protein. As of now, this constellation has not been reported before in other plant lineages and is thus unique to bryophytes. While the LLM-domain was already present in algae like *C. braunii*, the HAN-domain can only be found in the genomes of land plants. It seems likely that the HAN-domain evolved in the course of terrestrialization and was present in the last common ancestor of land plants. It can be speculated that this ancestral plant encoded a B-GATA with both domains on the same protein and that this arrangement was maintained in bryophytes but not in embryophytes. Since bryophyte B-GATAs contain both HAN- and LLM-domains, it can be suggested that they are functionally interdependent and might physically interact. In the case of B-GATAs from vascular plants, this could be fulfilled through heterodimerization of HAN- and LLM-domains from the respective B-GATA subfamilies. However, there is no indication for a shared biological role so far and this hypothesis thus remains speculative. Alternatively, and probably more likely, the increasing complexity of the body plan and life cycle of embryophytes, compared to bryophytes, might have required the division of the HAN- and LLM-domains into two separate subfamilies in order to allow for subfunctionalization.

### **3.2 B-GATAs in *P. patens* regulate protonema formation independently of cytokinin**

CRISPR/Cas9-mediated gene targeting of *B-GATA* orthologs in *P. patens* resulted in the generation of single, double, and triple mutants, however no mutant for *PpGATA2* could be identified even after multiple rounds of transformation using two different gRNAs. So far, it is difficult to judge whether the inability to generate a *PpGATA2* mutant is due to inefficient gRNAs or whether the loss of *PpGATA2* is potentially lethal.

*Ppgata134* triple mutants exhibited defects in protonema formation in comparison to the wild type and this phenotype was reminiscent of plants with elevated levels of endogenous CK or CK-treated plants (Coudert, Novák et al. 2019). Quantification of endogenous CK levels in the *Ppgata134* triple mutants yielded heterogenous and ambiguous results. The results of the cytokinin quantification were not consistent for the two analysed *ppgata* mutants, thus it is currently not possible to confirm the involvement of cytokinin in the phenotypes observed in the *ppgata* mutants.

Previously, *gata quint* was the highest order *LLM-domain B-GATA* mutant available in *A. thaliana*. In order to fully understand B-GATA function in *A. thaliana*, I generated a *gata hexuple* mutant through CRISPR/Cas9-mediated gene targeting of *AtGATA15* and *AtGATA16* in the *gata quad* background and subsequent crossing of the resulting quintuple mutants. Phenotypically, *gata hex* resembled the previously available *gata quint* and both mutants exhibit the pale green phenotype caused by reduced chlorophyll accumulation (Ranftl, Bastakis et al. 2016, Bastakis, Hedtke et al. 2018). Analysis of intermediates of the tetrapyrrole biosynthesis pathway yielded similar results for both mutants, however the amounts of MgP and MgPMME were only significantly reduced in *gata hex* but not in *gata quint* in comparison to Col-0. It is known from previous studies, that LLM-domain B-GATAs control greening redundantly (Ranftl, Bastakis et al. 2016), and my results suggest that *AtGATA15*, which is not mutated in *gata quint*, is also involved in tetrapyrrole biosynthesis downstream of MgP. The additional loss of *GATA15* might be causal for the further reduction of MgP and MgPMME observed in *gata hex*.

Phenotypic analysis of *gata hex* further revealed striking instability of the stems resulting in plants falling over and not being able to maintain the typical vertical growth (Figure 2.3). Observation of *gata quad* and *gata quint* revealed similar stem fragility, however the phenotype

was most strongly pronounced in *gata hex*, suggesting that B-GATA have redundant, yet unknown, functions in stem stability. The *nst1 nst3* double is mutant deficient in lignification of the secondary cell wall and exhibits a similar fragility of the stem (Mitsuda, Iwase et al. 2007), however microscopic analyses of stem cross sections revealed no obvious lignification defects in *gata hex*. In the *pme35* mutant, demethylesterification of homogalacturonans in the primary cell wall is suppressed, leading to stem fragility (Hongo, Sato et al. 2012). Analysis of the stem diameter in adult plants revealed a strong reduction for both *gata hex* and *pme35* at the bottom of the stem and between the siliques, while no differences could be observed for GNLox. Quantification of cell wall components revealed a significant increase of pectin methylation in *pme35* and strong differences for several sugars in *nst1 nst3*, however no differences for any cell wall component could be observed in *gnc gnl*, *gata quad*, *gata quint*, and *gata hex*. Thus, the stem fragility phenotype of *gata* mutants in *A. thaliana* cannot be explained through pectin methylation, defects in secondary cell wall formation or cell wall composition.

### 3.3 The single B-GATA ortholog of *M. polymorpha* has physiological and developmental functions

Mutants of MpB-GATA1 exhibited a number of morphological and physiological defects. For example, in both *A. thaliana* and *M. polymorpha*, LLM-domain B-GATA mutants exhibited related phenotypes in respect to reduced plant size (Zubo, Blakley et al. 2018). Additionally, I observed increased size of air pores in plants overexpressing MpB-GATA1 visible by eye, and measurements of air pore size not only confirmed this observation but also identified decreased air pore size in the *Mpb-gata1* mutant alleles grown under regular light conditions (Figure 2.19). When plants were grown under moderate high light intensities for extended time periods, the size of air pores increased in BoGa, *Mpb-gata1* mutants and most strongly in MpB-GATA1ox3.

Little is known about air pores in liverworts, but they likely assist in gas exchange, transpiration and respiration (Green and Snelgar 1982). Despite their proposed involvement in these essential processes, air pores do not seem to be required for growth under laboratory conditions, since mutants of the E3 ubiquitin ligase NOPPERABO1 show complete absence of air pores as well as air chambers but are not impaired in thallus growth compared to the wild type (Ishizaki, Mizutani et al. 2013). Air pores may however be required for plant growth in challenging conditions, such as high-light environments. Therefore, a dual function of MpB-GATA1 can be proposed in the regulation of abiotic stress responses and developmental processes. Interestingly, in *A. thaliana*, the formation of stomata, which regulate gas exchange in vascular plants, is regulated by LLM-domain B-GATAs. This suggests that B-GATAs may regulate the formation of analogous structures in both *M. polymorpha* and *A. thaliana*.

### 3.4 Gemma cup formation is regulated by MpB-GATA1

*Mpb-gata1* mutants showed a significant increase in gemma cup density and gemma cup formation in the mutants preceded the wild type by one week, while MpB-GATA1 overexpressors showed a strong reduction in the formation of gemma cups and this phenotype was maintained after longer growing periods. Gemma cup formation in *M. polymorpha* is regulated by CK, and overexpressors of the cytokinin degrading oxidase CKX2 exhibit reduced CK content and fewer gemma cups (Aki, Mikami et al. 2019). The results of the CK quantification carried out in cooperation with Dr. Ondřej Novák (Palacky University, Olomouc, Czech Republic) yielded increased CK levels for the *Mpb-gata1* mutants and MpB-GATA1ox3 in comparison to BoGa, but mutants and overexpressor displayed opposing phenotypes regarding gemma cup abundance.

The genes involved in CK biosynthesis were not consistently regulated between the mutant alleles, but the overexpressor exhibited increased expression of an *IPT* ortholog and two *CYP735A2* orthologs, which correlates with the increase in CK levels. In *Mpb-gata1-1*, two

orthologs of *CYP735A2* showed slight upregulation and these plants also exhibited a strong increase in CK contents. In *A. thaliana gata hex*, only the expression of *LOG7* was slightly reduced and the plants exhibited a reduction in total CKs. In GNLox, two *IPTs*, *CYP375A2* and six out of eight *LOGs* were differentially regulated, however the total CK content was not altered in these plants. These data suggest that, in the present case, the expression of CK biosynthesis genes alone cannot explain the differences in CK contents measured in the different genotypes.

### **3.5 The functions of HAN- and LLM-domain remain unknown in *M. polymorpha MpB-GATA1***

As discussed above, the single B-GATA of *M. polymorpha*, as well as the four B-GATA orthologs of *P. patens*, contains both HAN- and LLM-domains in contrast to B-GATAs from angiosperms that contain only one of the two domains. Both HAN- and LLM-domain are characteristic and specific for plant B-GATAs but their biochemical functions are unknown. Previous analyses from my host laboratory found that the LLM-domain is required for the complementation of the prominent greening phenotype of the *gnc gnl* double mutant (Behringer, Bastakis et al. 2014), but nothing is known about the HAN-domain to date.

Physical interaction of the HAN- and LLM-domain was proposed after a yeast-two-hybrid screen (Zhang, Zhou et al. 2013), but this hypothesis was later dismissed when the original finding could not be replicated using the identical material. Interestingly, *A. thaliana* mutants of HAN B-GATAs seem to mainly exhibit developmental phenotypes regarding embryonic development and flower morphology (Zhao, Medrano et al. 2004, Nawy, Bayer et al. 2010, Kanei, Horiguchi et al. 2012, Zhang, Zhou et al. 2013), whereas mutants of LLM B-GATAs mainly exhibit physiological phenotypes e.g. related to greening, flowering time or lateral shoot angle (Richter, Bastakis et al. 2013, Richter, Behringer et al. 2013, Ranftl, Bastakis et al. 2016, Bastakis, Hedtke et al. 2018, Schwechheimer, Schröder et al. 2022). Since B-GATAs in bryophytes contain both HAN- and LLM-domains on the same protein, these organisms are good models to study the function of the two domains.

I generated a set of complementation lines for the *Mpb-gata1* mutants with transgenes lacking the HAN- or LLM-domain, in order to infer the involvement of the two domains on the phenotypes I described for the mutants. I acquired these complementation lines close to the end of my project, thus they were not included in any other experiments.

Analysis of plant size revealed that all domain-deletion complementation lines, regardless of the presence or absence of HAN- or LLM-domain, were significantly smaller than the wild type after two or three weeks of growth, thus only partially rescuing the phenotype of the mutants (Figure 2.8A, B, C, E). After two weeks of growth, gemma cup density of *Mpb-gata1* pMpB-GATA1:MpB-GATA1  $\Delta$ LLM was at the level of MpB-GATA1ox3, while all other complementation lines showed gemma cup densities similar to the mutants (Figure 2.8A, D). The complementation lines lacking the HAN- or LLM-domain exhibited a gemma cup density similar to the wild type after three weeks of growth, whereas the lines lacking both or none of the domains showed an increase in gemma cup density comparable to the *Mpb-gata1* mutants (Figure 2.8B, F).

Overall, these results are inconclusive, and the phenotypes are not sufficient to deduce the functions of HAN- and LLM-domains. Expression level of the respective transgenes might influence the gemma cup density in the rescue lines since it is not clear whether the HAN- and LLM-domains are involved in the regulation of gemma cup formation and overexpressors of *MpB-GATA1* exhibit a strong gemma cup phenotype. In order to circumvent this problem, a better approach for future experiments might be the CRISPR/Cas9-mediated deletion of the respective domains in the wild type background.

### **3.6 Gene expression is not sufficient to explain altered CK levels in GATA mutants and overexpressors of *M. polymorpha* and *A. thaliana***

In *M. polymorpha*, the genes encoding the cytokinin oxidases *CKX1* and *CKX2* were downregulated in the mutant alleles, while in the overexpressor, *CKX1* was upregulated and *CKX2* was downregulated. In *A. thaliana gata hex*, only *CKX7* was slightly downregulated, while *CKX2-4* were downregulated and *CKX7* was upregulated in GNLox. The CK levels in *gata hex* were strongly decreased compared to Col-0, while GNLox did not exhibit changes in CK accumulation despite the strong downregulation of three *CKX* genes and in *M. polymorpha*, both *Mpb-gata1* and *MpB-GATA1ox3* exhibited increases in CK levels. Cytokinin oxidases catalyse the degradation of CKs, and plants of different species expressing low levels of *CKX* genes have been shown to exhibit CK-overproducing phenotypes (Ashikari, Sakakibara et al. 2005, Bartrina, Otto et al. 2011, Yeh, Chen et al. 2015). In the data presented here however, the expression of *CKX* genes and genes involved in CK biosynthesis cannot be directly correlated with CK content, suggesting that CK content in the B-GATA mutants and overexpressors might be controlled on a different level, possibly through the availability of precursor molecules upstream of *IPTs*.

In *M. polymorpha*, overexpressors of the type-A response regulator *MpRRA* were shown to have significantly fewer gemma cups compared to the wild type, while *Mprra* mutants produce more gemma cups. Mutants of *MpRRB* do not produce any gemma cups (Flores-Sandoval, Dierschke et al. 2016). The expression of *MpRRA* and *MpRRB* was significantly decreased in the *Mpb-gata1* mutants, and these lines produced more gemma cups compared to BoGa. The expression of *MpRRA* and *MpRRB* was unchanged in *MpB-GATA1ox3* however and can thus not explain the strong decrease in gemma cup formation in this genetic background. These data suggest that *MpB-GATA1* acts downstream or independently of CK in the regulation of gemma cup formation.

### **3.7 *gata* mutants exhibit high-light stress response in low-light conditions**

In a transcriptomic experiment with the *Mpb-gata1* mutants, many GO terms connected to photosynthesis, terpenes and carotenoids were enriched among the downregulated genes. Both, carotenoids and terpenes play important roles in light harvesting and quenching of excess energy in cases of high irradiance and are thus also connected to photosynthesis. Instigated by this observation, I analysed the expression of genes connected to photosynthesis. The majority of the 88 photosynthesis associated genes I identified among the differentially expressed genes were downregulated, while 16 of the 18 differentially expressed *ELIP* genes were strongly upregulated in both *Mpb-gata1* mutants. The role of ELIPs is not fully understood to date, but it has been hypothesized that they act as structure proxys for light harvesting complexes during chloroplast biogenesis. While their structure and pigment binding capacities are similar to that of LHCs, ELIPs differ from LHCs in their transient expression upon exposure to high light intensities (Hutin, Nussaume et al. 2003). In this context, they are thought to act as scavengers binding free chlorophyll molecules. By doing so, they could prevent photooxidative damage and facilitate energy dissipation that protects the photosynthetic machinery from photoinhibition in high-light situations (Montané and Kloppstech 2000). *A. thaliana* encodes two ELIPs, ELIP1 and ELIP2. In *M. polymorpha* however, the family expanded drastically with 23 members and the majority of these genes was strongly upregulated in the *Mpb-gata1* mutants. ELIPs have not been studied in *Marchantia* to date, but it can be speculated that a typically shade-dwelling plant like *Marchantia* would benefit from a larger number of proteins involved in photoprotection in order to tackle periods with increased light intensities.

On the phenotypic level, it is noteworthy that the phenotypes of *gata* mutants e.g., reduced chlorophyll accumulation and reduced plant size, are also caused by overexpression of *ELIP2* in *A. thaliana* (Tzvetkova-Chevolleau, Franck et al. 2007). Additionally, plants overexpressing *ELIP2* showed a strong reduction in the abundance of both photosystems and the corresponding light harvesting complexes on the protein level. Interestingly, genes encoding the subunits of both photosystems, as well as their antenna proteins of the light harvesting complexes, were among the downregulated genes in the *Mpb-gata1* mutants. The transcriptomic data of the *A. thaliana gata quint* mutant gathered earlier (Ranftl, Bastakis et al. 2016) exhibited similar expression of genes encoding photosystems, light harvesting complexes and ELIPs. This expression pattern is usually found in plants exposed to high light intensities and is thought to act as a protective mechanism avoiding ROS formation through over-excitation of chlorophyll molecules (Rossel, Wilson et al. 2002, Kleine, Kindgren et al. 2007, Jung, Crisp et al. 2013). Taking together the strong upregulation of *ELIPs*, downregulation of photosystem compartments and antenna proteins with the results of the GO enrichment analysis, the transcriptomic profiles of the *Mpb-gata1* mutants and the *A. thaliana gata quint* mutant, I hypothesized that *gata* mutants of *M. polymorpha* and *A. thaliana* display high-light stress on the transcriptomic level in regular, low-light conditions.

### **3.8 B-GATA factors are required to repress high-light stress responses in *M. polymorpha* and *A. thaliana***

In order to test whether *gata* mutants in *M. polymorpha* and *A. thaliana* display a constitutive high-light stress response, I performed additional RNA-seq experiments. At the time, there were no publicly available transcriptomic datasets for high-light stress in *M. polymorpha*, so I started my analysis by establishing the transcriptome of BoGa after exposure to high light.

I found a substantial, but not complete overlap of differentially regulated genes in BoGa and the *Mpb-gata1* mutants after exposure to six hours of high light. Of the genes present in BoGa and the mutants, 52% and 47% were up- and downregulated in all genotypes, respectively. Only a small subset of genes did not share the same direction of transcriptional regulation, suggesting that the *Mpb-gata1* mutants maintained the ability to sense high light. GO enrichment analysis of the upregulated genes yielded terms connected to heat and abiotic stress. Even though there is a number of genes that are differentially regulated in response to both heat and high-light, the overall transcriptomic responses to heat and high-light can be separated (Huang, Zhao et al. 2019). The experimental setup allowed for stable temperatures at the chamber level however, temperature increases within the plates could not be avoided due to technical constraints. During the experiment, the temperature at the chamber level was stable at 19°C, while the temperature in plates rose to 24-25°C (data not shown). The temperatures used in heat stress experiments generally exceed these values significantly, with temperatures ranging from 35°C to 45°C in most publications (Li, Zhou et al. 2010, Pecinka, Dinh et al. 2010, Evrard, Kumar et al. 2013, Liu, Zhang et al. 2013, Pan, Sun et al. 2018, Huang, Zhao et al. 2019). Thus, even though the temperature within the plates in the high-light chamber was increased compared to the chamber used for the corresponding low light (mock) treatment at 22°C, it is unlikely that this slight increase in temperature caused a heat stress response, and the observed GO term enrichment might instead be an artifact of gene annotation that does not distinguish between heat and HL. The GO enrichment analysis of the downregulated genes however yielded many terms connected to photosynthesis and is reminiscent of the well characterized response to middle-term high-light in *A. thaliana* (Huang, Zhao et al. 2019).

I next tested whether the *Mpb-gata1* mutants exhibit a constitutive high-light stress transcriptome in low-light conditions by comparing the DEGs of mutant plants grown in low light conditions to high-light stressed wild type plants. Of the 7520 DEGs from BoGa after the high-light treatment, 2952 were also differentially regulated in the *Mpb-gata1* mutants without

exposure to HL. Over 70% of these 2952 DEGs shared the direction of transcriptional regulation in BoGa and the mutants and GO enrichment of these genes yielded terms connected to carbohydrate metabolism among the upregulated genes and terms connected to photosynthesis and tetrapyrrole biosynthesis among the downregulated genes. The substantial gene expression overlaps for the low-light grown mutants and the high-light stressed wild type suggests that MpB-GATA1 acts as repressor of the transcriptomic high-light stress response in BoGa. It is however likely that other, yet unknown regulators of high-light stress exist in *M. polymorpha*, since the high-light response observed in BoGa is only partially present in the mutants.

I analysed the expression of genes connected to photosynthesis, originally identified through my previous transcriptomic experiment, and found a striking overlap for BoGa  $t_{6HL}$ , *Mpb-gata1* mutants  $t_{6LL}$  and *Mpb-gata1* mutants  $t_{6HL}$ . The majority of the genes exhibited the same direction of transcriptional regulation, but the response in the mutants at  $t_{6LL}$  was weaker compared to BoGa or the mutants at  $t_{6HL}$ . Two of the three *psbA* orthologs in *Marchantia* (*Mp1g09910*, *Mp5g10950* and *Mp8g01990*; the expression of *Mp8g01990* was not altered) were upregulated in the wild type after high-light treatment but not differentially regulated in the mutant alleles after mock treatment, indicating that the plants did not actually suffer from high-light stress, as the *psbA*-encoded D1 protein usually is the first PSII component to be damaged by high-light induced ROS formation and would thus require re-synthesis (Tikkanen, Mekala et al. 2014).

In order to test whether the role of B-GATAs in the suppression of high-light response was conserved in *A. thaliana*, I performed an additional RNA-seq experiment with the *gata hex* mutant. Similar to *M. polymorpha*, the *A. thaliana gata hex* mutant was capable of sensing HL, as Col-0 and *gata hex* display a significant overlap of DEGs after the high-light treatment. GO enrichment of the upregulated genes yielded terms connected to flavonoids and anthocyanins, while many terms connected to photosynthesis were enriched among the downregulated genes. Interestingly, the most strongly enriched GO terms were related to the biosynthesis of geranylgeranyl diphosphates, compounds that are the precursors of carotenoids, chlorophylls and plastoquinones (Ruiz-Sola, Coman et al. 2016). Similar to *M. polymorpha*, the expression of *psbA* was not altered in the *hexuple* mutant after six hours of mock treatment, suggesting that low light intensities did not cause PSII damage in the mutant. Interestingly, the expression of *psbA* was also not changed in *gata hex* after the high-light treatment.

Comparison of Col-0  $t_{6HL}$  and *gata hex t\_{6LL} yielded 277 DEGs present in both datasets and the majority of these (81%) shared the same direction of transcriptional regulation. Terms related to flavonoid and anthocyanin biosynthesis were strongly enriched among the upregulated genes, while the downregulated genes exhibited mainly overrepresentation of terms connected to photosynthesis. The differential expression of photosynthesis-associated genes was strikingly parallel in *M. polymorpha* and *A. thaliana*. In both species, the *gata* mutants grown under low light conditions exhibited the same transcriptomic pattern as the respective wild type or mutant plants exposed to HL, but the response in low-light grown plants was generally weaker compared to the high-light stressed plants. In all cases, genes involved in chlorophyll biosynthesis, photosystem reaction centre genes and light harvesting complex genes were uniformly downregulated, while *ELIPs* were upregulated. I additionally analysed the entire set of genes involved in the light-dependent reaction of photosynthesis in order to elucidate which proportion of them were affected in the *gata* mutants at  $t_{6LL}$ . Interestingly, almost all genes encoding the light harvesting complexes of both photosystems were downregulated in both *M. polymorpha* and *A. thaliana* and this response was identical for mutants or wild types exposed to high-light. Most of the photosystem I subunits were equally affected in both species, while genes encoding photosystem II subunits were mainly affected in *M. polymorpha*, and many of them did not exhibit differential regulation in *A. thaliana*. The same applied to the genes encoding proteins of the electron transport chain and ATPase subunits.*

Thus, in both species, the loss of the *LLM-domain B-GATAs* results in a constitutive high-light transcriptome, suggesting that B-GATAs act as repressors of the high-light response in the wild type and that this might be an ancestral role of B-GATAs. In an evolutionary context, high light intensities were a novel abiotic stress for the algal ancestors of extant embryophytes (Becker and Marin 2009). In contrast to their algal ancestors, the earliest land plants were surrounded by air, which filters the light differently compared to water. On land, light can fluctuate drastically and can easily reach harmful intensities capable of damaging the photosynthetic apparatus. In order to adjust to these new challenges, plants had to develop strategies to deal with high-light stress in the course of terrestrialization. Among these strategies are the reduced production of photosynthesis-related proteins like light-harvesting complexes and photosystem components and decreased biosynthesis of chlorophylls in order to protect the photosynthetic machinery. The fact that multiple genes are only differentially regulated in the *M. polymorpha* *Mpb-gata1* mutants but not in *A. thaliana* *hex* suggests that *A. thaliana* encodes other regulators of the high-light response not present in *M. polymorpha*. Currently, it is difficult to evaluate whether the observed gene expression patterns are interrelated or a direct consequence of the loss of B-GATAs due to the absence of other analyses.

### **3.9 B-GATAs regulate tetrapyrrole biosynthesis in *A. thaliana* and *M. polymorpha***

#### **3.9.1 The role of B-GATAs in regulation of chlorophyll biosynthesis is conserved**

Previous research from my host laboratory clearly elucidated the role of LLM-domain containing B-GATAs from *A. thaliana* in the regulation of chlorophyll biosynthesis on the transcriptional level (Bastakis, Hedtke et al. 2018). In my first transcriptomic experiment, analysis of the genes involved in the tetrapyrrole biosynthesis pathway revealed a striking similarity between the *Mpb-gata1* mutants and the previously published transcriptomic data from the *A. thaliana* *gnc gnl* and *gata quad* mutants (Figure 2.11; Richter, Behringer et al. 2010; Ranftl, Bastakis et al. 2016). In all three datasets, the majority of genes was downregulated. The high degree of similarity between the *gata* mutants from *M. polymorpha* and *A. thaliana*, two species separated by more than 450 million years, suggest that regulation of chlorophyll biosynthesis presents an ancestral function of B-GATAs in land plants.

As shown in my phylogenetic analyses and our recently published review article (Schwechheimer, Schröder et al. 2022), algae also encode LLM-domain B-GATAs. As of now, there are no studies on the functions of these algal LLM-domain GATAs, but it can be speculated that chlorophyll biosynthesis in green algae might also be regulated by GATAs. In *A. thaliana*, the chlorophyll- and heme biosynthesis pathways were equally affected by the loss of B-GATAs, but only one gene involved in heme biosynthesis, the ortholog of *AtGUN2*, was differentially regulated in mutants of *Mpb-gata1*. Interestingly, heme biosynthesis in animals is also regulated by GATAs and HsGATA-1 is the master regulator of heme biosynthesis and erythropoiesis in humans (Fujiwara, Browne et al. 1996, Bresnick, Katsumura et al. 2012). Without analyses of B-GATAs from other non-vascular and vascular plants, it is currently difficult to judge whether regulation of the heme biosynthesis pathway also was an ancestral role of plant B-GATAs which was lost in *M. polymorpha*, or whether this function was independently acquired by GATAs in animals and angiosperms.

#### **3.9.2 MpB-GATA1 represses the high-light response of the tetrapyrrole biosynthesis pathway in *M. polymorpha***

High-light treatment of BoGa resulted in a very similar expression pattern of the genes in the tetrapyrrole biosynthesis pathway compared to the *Mpb-gata1* mutants grown in low-light conditions (Figure 2.26A). In both cases, most of the genes were downregulated, especially

downstream of protoporphyrin-IX, while the heme branch of the pathway was virtually unaffected. Only the ortholog of *FERROCHELATASE 2 (FC2)* was upregulated in *Mpb-gata1-2* at  $t_{6HL}$  and heme content was increased in this line compared to BoGa, but downregulation of *FC1* in *Mpb-gata1-1* resulted in similar heme levels (Figure 2.26A, H). Exposure to high-light did not change the direction of expression in the *Mpb-gata1* mutants, however the intensity of the regulation increased in comparison to  $t_{6LL}$  (data not shown), suggesting that there are other, redundant mechanisms suppressing the high-light response in the wild type under low-light conditions.

Interestingly, the clear transcriptomic pattern observed in the mutants at  $t_{6LL}$  did not translate to the metabolite level (Figure 2.26A - H). Despite the uniform downregulation of most genes of the pathway, there was no measurable difference in metabolite levels between BoGa and the *Mpb-gata1* mutants at  $t_{6LL}$ . Exposure to high light had a strong effect, resulting in a significant reduction of MgP, MME, and protochlorophyllide *a* regardless of genotype (Figure 2.26A - D). The amounts of chlorophyllide *a* and chlorophyll *a* increased significantly in the *Mpb-gata1* mutants at  $t_{6HL}$  but stayed constant in BoGa, while chlorophyll *b* levels increased in BoGa and the mutants (Figure 2.26E - G). These data suggest that high-light leads to a depletion of tetrapyrrole biosynthesis intermediates upstream of chlorophyllide *a*. In the absence of quantification data for the remaining biosynthesis intermediates however and without a better understanding of heme biosynthesis in *M. polymorpha*, it is currently not possible to elucidate the effect of high-light on the tetrapyrrole biosynthesis pathway in more detail.

### **3.9.3 *A. thaliana gata hex* did not copy the high-light response of the tetrapyrrole biosynthesis pathway**

In contrast to *M. polymorpha*, where exposure to high-light resulted in downregulation of the genes involved in the tetrapyrrole biosynthesis pathway, many of these genes were upregulated in *A. thaliana* Col-0 after the same high-light treatment (Figure 2.28A). A possible explanation for this observation is that *A. thaliana* and *M. polymorpha* usually occupy different natural environments. Liverworts like *M. polymorpha* are usually found in shaded areas with high humidity (Mache and Loiseaux 1973), in contrast to *A. thaliana*, which can be found throughout temperate regions of the world with light intensities easily exceeding  $1000\mu\text{mol m}^{-2} \text{s}^{-1}$  (Huang, Zhao et al. 2019). Thus, plants like *A. thaliana* are naturally more tolerant towards high light intensities, which might account for the transcriptomic response of the tetrapyrrole biosynthesis pathway observed here.

The *gata hex* mutant displayed a similar expression profile at  $t_{6LL}$  to the *Mpb-gata1* mutants in *M. polymorpha*. However, it is notable that in *A. thaliana* the majority of the differentially regulated genes was downstream of protoporphyrin-IX in the heme- and the chlorophyll branch of the pathway (Figure 2.26A and Figure 2.28A). As already mentioned above, the overlap of gene expression between the *gata* mutants in *A. thaliana* and *M. polymorpha* suggests that regulation of the tetrapyrrole biosynthesis pathway is a conserved, and possibly ancestral, function of B-GATAs.

At  $t_{6HL}$ , *gata hex* and Col-0 had highly similar expression profiles but overall, more genes were affected in Col-0 compared to *gata hex*, suggesting that LLM-domain B-GATAs in *A. thaliana* are required for the high-light response of at least some of the genes of the tetrapyrrole biosynthesis pathway. All chlorophyll biosynthesis intermediates except for chlorophyllide *a* were significantly reduced in *gata hex* compared to Col-0 at  $t_{6LL}$  and this observation falls in line with the previous analyses on mutants of LLM-domain B-GATAs that have been shown to be deficient in chlorophyll biosynthesis (Figure 2.28A – H; Behringer, Bastakis et al. 2014, Ranftl, Bastakis et al. 2016, Bastakis, Hedtke et al. 2018). Interestingly, heme levels were not reduced in *gata hex* despite the downregulation of *FERROCHELATASE 1*. After exposure to high-light, MgP, MME, and protochlorophyllide *a* were drastically reduced in Col-0 and *gata*

*hex* compared to the respective genotype at  $t_{6LL}$ , whereas chlorophyllide *a* and heme levels stayed constant (Figure 2.28B – D, H). Chlorophyll *b* levels increased in *gata gata hex* and Col-0 at  $t_{6HL}$  compared to the respective genotype at  $t_{6LL}$ , whereas chlorophyll *a* levels only increased in Col-0 and remained constant in *gata hex* (Figure 2.28F, G). Similar to *M. polymorpha*, the exposure to high light seems to have led to the depletion of the chlorophyll biosynthesis pathway upstream of chlorophyllide *a*.

### **3.10 B-GATAs are not involved in the high-light response of the carotenoid biosynthesis pathway**

#### **3.10.1 The xanthophyll cycle functions independently of MpB-GATA1**

Carotenoids are a group of pigments that are widely distributed in plants and play important roles in various physiological processes, including photoprotection (Choudhury and Behera 2001). In *M. polymorpha* BoGa, exposure to high light resulted in upregulation of most genes involved in carotenoid biosynthesis, most prominently downstream of lycopene (Figure 2.27). Interestingly, after the high-light treatment, antheraxanthin levels increased strongly in BoGa (Figure 2.27B). At the same time, the reduction in violaxanthin levels suggests that the light-dependent de-epoxidation of violaxanthin to zeaxanthin via antheraxanthin (xanthophyll cycle) was taking place (Figure 2.27B - C). However, since the levels of zeaxanthin were not quantified, this remains speculative.

Most genes were downregulated at  $t_{6LL}$  in the *Mpb-gata1* mutants, suggesting that MpB-GATA1 regulates the expression of the carotenoid biosynthesis pathway but does not repress the usual high-light response of these genes observed in BoGa. At  $t_{6HL}$ , the genes were regulated similar to BoGa, except for *LYCb*, which was upregulated in BoGa but downregulated in *Mpb-gata1-2* (Figure 2.27A). Analysis of intermediates of the carotenoid pathway resulted in reduced levels of violaxanthin in the mutants compared to BoGa at  $t_{6LL}$ , which might be explained through the downregulation of most genes of the pathway (Figure 2.27A, C). At  $t_{6HL}$ , the mutants displayed the same increase in antheraxanthin accompanying the reduction of violaxanthin, suggesting that the xanthophyll cycle functions independently of MpB-GATA1 (Figure 2.27B - C).

In contrast to BoGa,  $\beta$ -carotene and lutein levels were significantly increased in the mutants at  $t_{6HL}$  compared to  $t_{6LL}$ , despite the similar expression of the genes involved (Figure 2.27E - F). Lutein is an important and unique photoprotective agent in plants functioning in the deactivation of excited triplet chlorophyll states, whereas zeaxanthin as product of the light-induced reactions of xanthophyll cycle is required for the deactivation of excited singlet chlorophyll as part of the NPQ processes (Jahns and Holzwarth 2012). Presently, it is not clear why lutein contents were increased in the mutants at  $t_{6HL}$  despite the downregulation of most upstream genes (Figure 2.27A). The difference to BoGa at  $t_{6HL}$  is small and only statistically significant for *Mpb-gata1-2*, thus the elevated lutein levels probably do not result in enhanced photoprotection. In order to understand the involvement of MpB-GATA1 in the biosynthesis of carotenoids in more detail, a comprehensive analysis of all biosynthesis intermediates is required. Presently, the data suggest MpB-GATA1 is not required for the high-light response of genes involved in carotenoid biosynthesis and the xanthophyll cycle.

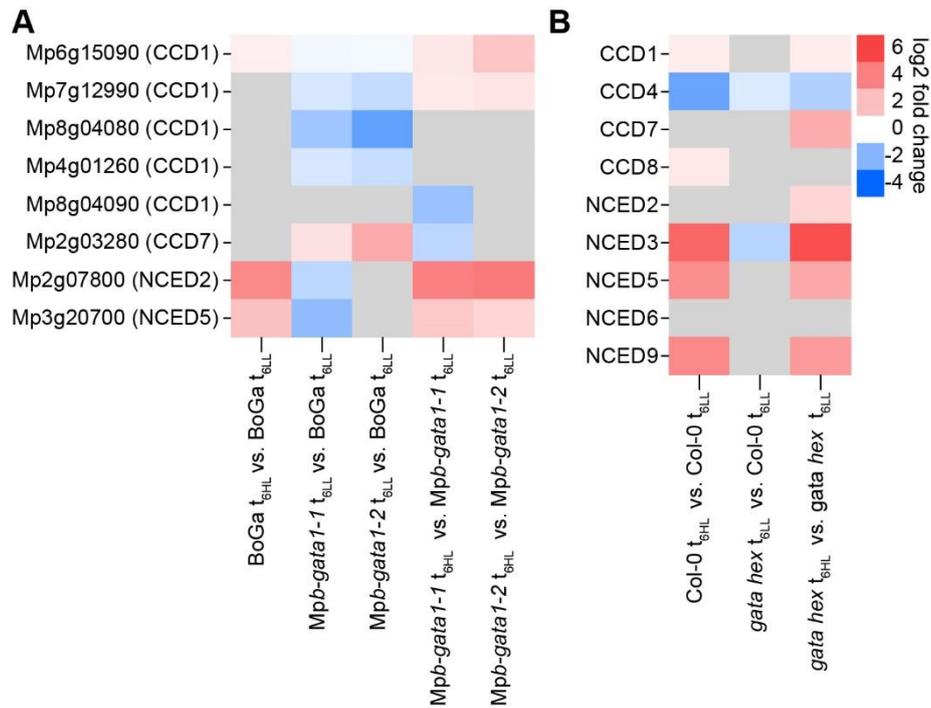
### 3.10.2 Regulation of the carotenoid biosynthesis pathway is a conserved function of LLM B-GATAs

Similar to *M. polymorpha*, most genes of the carotenoid biosynthesis pathway were upregulated in Col-0 after exposure to high-light, emphasizing the importance of carotenoids for photoprotection in plants. Antheraxanthin content increased in Col-0 after the high-light treatment, while violaxanthin and neoxanthin remained unchanged, suggesting that also in *A. thaliana* the reverse-reactions of the xanthophyll cycle were taking place (Figure 2.29B – D). Most genes of the pathway were downregulated in *gata hex* at  $t_{6LL}$ , similar to the *Mpb-gata1* mutants in *M. polymorpha* (Figure 2.27A and Figure 2.29A). In contrast to *M. polymorpha*, however this transcriptomic regulation was clearly reflected on the metabolite level. Antheraxanthin, violaxanthin, neoxanthin, and lutein contents were significantly reduced in *gata hex* compared to Col-0 (Figure 2.29B – D, F). At  $t_{6HL}$ , most genes of the pathway were upregulated in *gata hex* and the gene expression profile of *gata hex* perfectly overlapped with Col-0, suggesting that the high-light response was not impaired in *gata hex* (Figure 2.29A).

Overall, my data suggest that a conserved function of LLM B-GATAs is the promotion of carotenoid biosynthesis on the transcriptomic level, as the majority of the genes was downregulated in mutants of both species under regular light conditions. Since exposure to high-light and loss of LLM B-GATAs led to opposing gene expression profiles of the carotenoid biosynthesis pathway in both *M. polymorpha* and *A. thaliana*, I concluded that LLM B-GATAs do not repress the usual high-light response of the carotenoid biosynthesis pathway on the transcriptomic level. Loss of B-GATAs in *M. polymorpha* and *A. thaliana* did not impair the high-light response of the pathway on the transcriptomic level, hinting at the light-related regulatory pathways impacting carotenoid biosynthesis (Llorente, Martinez-Garcia et al. 2017).

Quantification of carotenoids revealed a slight increase of antheraxanthin levels in *gata hex* at  $t_{6HL}$  compared to  $t_{6LL}$ , but violaxanthin, neoxanthin,  $\beta$ -carotene, and lutein levels remained constant (Figure 2.29C, D - F). Interestingly, the amounts of all measured metabolites in *gata hex* at  $t_{6HL}$  were significantly reduced compared to Col-0 at  $t_{6HL}$ , despite the identical transcriptomic regulation of the genes required for their biosynthesis (Figure 2.29A). These data suggest that either carotenoid catabolism was increased in *gata hex*, or that availability of upstream precursor molecules was limited in *gata hex*.

In *A. thaliana*, carotenoids are degraded by carotenoid cleavage dioxygenases (CCD) and 9-cis-exopoxycarotenoid dioxygenases (NCED) (Cuttriss, Cazzonelli et al. 2011). Analysis of *CCD* and *NCED* orthologs in *M. polymorpha* revealed downregulation of most *CCD* genes in the *Mpb-gata1* mutants at  $t_{6LL}$  (Figure 3.1A). The ortholog of *CCD7* was slightly upregulated in both mutants and *NCED2* and *NCED5* were only downregulated in *Mpb-gata1-1* (Figure 3.1A). In *A. thaliana*, only *CCD4* and *NCED3* were slightly downregulated in *gata hex* at  $t_{6LL}$  (Figure 3.1B). In both species, the differential regulation of genes encoding carotenoid cleavage enzymes was not sufficient to explain the differences observed on the metabolite level. After exposure to high-light, the *GATA* mutants of both species behaved similar to their respective wild types. However, in *M. polymorpha*, *CCD7* and one ortholog of *CCD1* (*Mp2g03280*) were only differentially regulated in *Mpb-gata1-1* and another ortholog of *CCD1* (*Mp7g12990*) was upregulated in both mutants but not in BoGa (Figure 3.1A). In *A. thaliana gata hex*, *CCD7* and *NCED2* were upregulated at  $t_{6HL}$  but not in Col-0 (Figure 3.1B). *CCD7* is involved in strigolactone biosynthesis (Vogel, Walter et al. 2010), whereas *NCEDs* are associated with ABA (Auldridge, McCarty et al. 2006, Walter, Floss et al. 2010), however without quantification of these compounds in the *gata* mutants of *M. polymorpha* and *A. thaliana*, it is not possible to judge the effect of the differential expression described here. Since the carotenoid deficiency of *A. thaliana gata hex* at  $t_{6HL}$  could not be explained through regulation of carotenoid catabolism genes, it is likely that precursor molecules upstream of the metabolites measured here are limited in *gata hex*.



**Figure 3.1: Differential expression of genes involved in carotenoid catabolism in *M. polymorpha* and *A. thaliana gata* mutants.** **A** Heatmaps showing differential expression of genes involved in carotenoid catabolism in *M. polymorpha*. The genes were identified based on their orthologs in *A. thaliana* shown in brackets. **B** Heatmaps showing differential expression of genes involved in carotenoid catabolism in *A. thaliana*. CCD, carotenoid cleavage dioxygenase; NCED, 9-cis-epoxycarotenoid dioxygenase.

### 3.11 Loss of MpB-GATA1 affects carbon fixation

When GNC was first described, it was reported to be involved in the carbon metabolism based on microarray expression data. In this context, GNC was proposed to act upstream of sugar transporters, galactosidases, cellulose synthases, chitinases, glucanases, and glycosyl-transferases (Bi, Zhang et al. 2005). Studies in *A. thaliana*, rice, and poplar reported effects on starch biosynthesis indirectly through the number of chloroplasts, or directly in correlation with GATA transcript abundance (Chiang, Zubo et al. 2012, Hudson, Guevara et al. 2013, An, Han et al. 2014). Here, I present a novel role of the single B-GATA from *M. polymorpha* in the regulation of carbon fixation in the Calvin-Benson cycle. The majority of genes encoding enzymes of the Calvin-Benson cycle was downregulated in mutants of MpB-GATA1 (Figure 2.12B), most prominently so in case of the RuBisCO small subunits. With the present data however, it is not clearly discernible whether these transcriptional changes are a direct consequence of the loss of MpB-GATA1, or an indirect effect of the downregulation of genes involved in the light-dependent reaction of photosynthesis discussed above. Reduced photosynthetic activity would result in reduced production of ATP and NADPH required for carbon fixation, thus it can be speculated that plants would adapt to this complication by adjusting the expression of genes involved in the downstream processes e.g., the Calvin-Benson cycle. Further experiments are required to elucidate whether MpB-GATA1 regulates the Calvin-Benson cycle actively or passively.

### 3.12 The Calvin-Benson cycle is high-light regulated in *M. polymorpha* but not in *A. thaliana*

In the first transcriptomic experiment, I observed differential regulation of genes involved in carbon fixation (Figure 2.12) and in order to find out whether this regulation is also part of the constitutive high-light response in B-GATA mutants, I analysed expression of the respective genes in *M. polymorpha* and *A. thaliana* after exposure to high-light (Figure 2.24). Interestingly, in *M. polymorpha* BoGa most of the genes of the Calvin-Benson cycle were affected by high-light, and the mock-treated *Mpb-gata1* mutants exhibited a striking similarity to high-light exposed plants (Figure 2.24B). In *A. thaliana* however, only few genes were affected by high-light or in the mock-treated *gata hex* and the similarity of gene expression in high-light and low-light grown plants was negligible in comparison to *M. polymorpha* (Figure 2.24C). Some of the genes of the Calvin-Benson cycle are known to be indirectly light regulated via the ferredoxin/thioredoxin system (Michelet, Zaffagnini et al. 2013), however in *A. thaliana* most of the genes did not change their expression in response to high-light (Figure 2.24C; Huang, Zhao et al. 2019). There are no studies on the regulation of the Calvin-Benson cycle genes in *M. polymorpha*, but my data suggests that the majority of the genes is differentially regulated by high-light. Since the genes are similarly regulated in BoGa at  $t_{6HL}$  and *Mpb-gata1* mutants at  $t_{6LL}$ , it seems that this high-light response is directly or indirectly suppressed by MpB-GATA1 in the wild type.

### 3.13 GATAs influence flavonoid accumulation differently in *M. polymorpha* and *A. thaliana*.

Long term exposure to moderate high-light did not influence chlorophyll content in *M. polymorpha* BoGa, *Mpb-gata1* mutants or MpB-GATA1ox3, but led to drastic reduction of chlorophyll levels in *A. thaliana* Col-0, *gata hex*, and GNLox (Figure 2.30 and Figure 2.32). At the same time, flavonoid content increased only in BoGa at  $t_{6HL}$  compared to BoGa at  $t_{6LL}$ , but not in the *Mpb-gata1* mutants or MpB-GATA1ox3 (Figure 2.30J). In plants, flavonoids are produced as protective agents against harmful UV light exposure, and their biosynthesis in response to light is regulated via the UV RESISTANCE LOCUS 8 (UVR8) (Rizzini, Favory et al. 2011, Ferreyra, Serra et al. 2021). Since flavonoid levels were not significantly altered in any of these lines under regular light conditions, it seems likely that mutants and

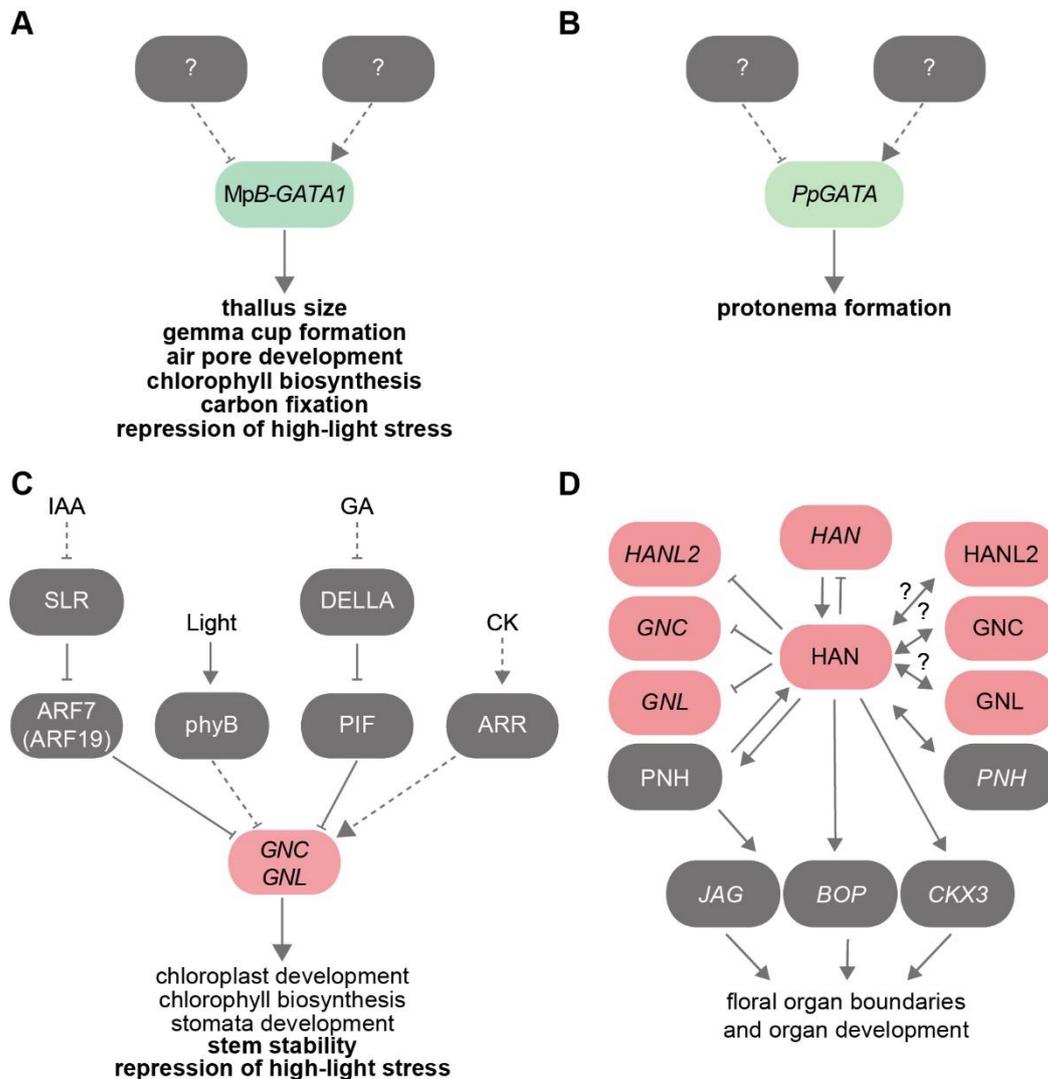
overexpressors of *MpB-GATA1* are not deficient in flavonoid biosynthesis *per se* but might be impaired in the high-light mediated flavonoid accumulation, possibly in the UVR8 mediated pathway. In *A. thaliana*, GNLox exhibited slightly, but not significantly, elevated anthocyanin contents in low-light, while anthocyanin content in *gata hex* was not altered (Figure 2.32J). After exposure to moderate high-light for one week, anthocyanin content of Col-0 and *gata hex* increased significantly but remained constant in GNLox (Figure 2.32J). Rosette leaves of Col-0 and *gata hex* plants appeared visibly purple after exposure to the high-light treatment due to the accumulation of anthocyanins, while GNLox rosette leaves appeared brown-yellowish, similar to plants experiencing drought (Figure 2.32B, D, F, H). Presently, these phenotypes cannot be explained, however the data suggest that B-GATAs might be involved in flavonoid biosynthesis or at least in the high-light induced accumulation of flavonoids.

### **3.14 Loss of B-GATAs might prime *A. thaliana* for high-light stress**

Exposure to light intensities that exceed the photosynthetic capacities of plants can lead to the formation of ROS. In order to test whether B-GATAs influence the superoxide accumulation in *A. thaliana* and *M. polymorpha*, I stained the leaves or thalli of plants after exposure to moderate high-light with NBT (Figure 2.31 and Figure 2.33). In *M. polymorpha*, the plants were fully stained without exposure to high-light and no difference was observed when comparing to plants exposed to six hours of moderate high-light, suggesting that *M. polymorpha* naturally contains high levels of superoxide (Figure 2.31). NBT staining of *A. thaliana* rosette leaves revealed significantly smaller stained areas in *gata hex* and no difference for GNLox when compared to Col-0 after six hours (data not shown) and most prominently after thirty 30 hours of exposure to moderate high-light (Figure 2.33A – G). The differences in NBT staining suggest that *gata hex* accumulated less superoxide in comparison to Col-0 and GNLox. Currently, it is not clear whether *gata hex* produced less superoxide or whether the superoxide-scavenging capacity was increased compared to Col-0. A possible explanation for this observation might be that the loss of B-GATAs in *gata hex* and the accompanying transcriptomic changes discussed above might act as primer for high-light stress. The typical transcriptomic pattern observable in plants exposed to high-light, downregulation of genes encoding the components of the photosynthetic apparatus and the upregulation of *ELIP* genes, is already present in *gata hex* without exposure to high-light and might be advantageous in the context of high-light stress in that it might lead to reduced ROS formation through the lower availability of chlorophyll molecules and increased ROS scavenging capacities through upregulation of *ELIP* genes. At the moment, this hypothesis is mostly speculative but future experiments may shed light on the potential priming effect of B-GATA mutants in the context of high-light.

### 3.15 Summary and outlook

The GATA transcription factors are a class of highly interesting candidates for comparative evolutionary biology due to their involvement in a plethora of pathways and biological processes. In the course of my project, I uncovered their conservation within the plant lineage and identified novel functions for B-GATAs from *M. polymorpha*, *P. patens*, and *A. thaliana*.



**Figure 3.2: Overview of B-GATA functions from *M. polymorpha*, *P. patens*, and *A. thaliana*.** **A** Biological functions of MpB-GATA1 from *M. polymorpha*. **B** Biological function of the PpGATAs from *P. patens*. **C** Selected biological functions of LLM B-GATAs from *A. thaliana*. **D** Selected biological functions of HAN B-GATAs from *A. thaliana*. Bold text indicates novel functions uncovered through my project. Genes that are transcriptionally regulated are shown in italics. Arrows indicate activation, blunt arrows indicate inhibition. Dashed lines represent indirect activation or repression. Abbreviations: ARF, AUXIN RESPONSE FACTOR; ARR, ARABIDOPSIS RESPONSE REGULATOR; BOP, BLADE-ON-PETIOLE; CK, cytokinin; CKX3, CYTOKININ OXIDASE3; GA, gibberellin; IAA, auxin (indole-3-acetic acid); JAG, JAGGED; phyB; phytochrome B; PIF, PHYTOCHROME-INTERACTING FACTOR; PNH, PINHEAD; SLR, SOLITARY ROOT. Panels C and D adapted from Schwechheimer, Schröder et al. 2022.

I provided an analysis of GATA transcription factors in the green lineage, which suggests that the HAN-domain was acquired by plants in the course of terrestrialization. I found conserved roles of MpB-GATA1 from *M. polymorpha* compared to *A. thaliana* B-GATAs in greening and downstream transcriptional regulation, as well as a potential role of MpB-GATA1 in gemma cup formation and thallus growth (Figure 3.2A). Analysis of B-GATA mutants from *P. patens* suggested involvement of B-GATAs in protonema formation (Figure 3.2B), but I did not continue to investigate this phenotype due to time constraints. Additionally, I uncovered new

potential functions of *A. thaliana* LLM B-GATAs in stem stability, using the *gata hex* mutants I generated in the course of my project (Figure 3.2C). The transcriptional profiles of *gata* mutants in both *A. thaliana* and *M. polymorpha* exhibited a striking similarity to plants experiencing high-light stress, thus I proposed a novel role of B-GATAs in the repression of the high-light response in regular light conditions, a previously unknown function of GATA transcription factors (Figure 3.2A, C). I tried to analyze the functions of HAN- and LLM-domain through complementation of the *Mpb-gata1* mutants with transgenes lacking one or the other domain, but this approach did not yield conclusive results. To date, knowledge is lacking about the functions of HAN- and LLM-domains or whether they are independent or function together, at least in certain tissues or at certain time points in plant development. More experiments are required to shed light on the functions of HAN- and LLM-domains.

## 4 Materials and Methods

### 4.1 Biological material

All experiments were carried out with *Marchantia polymorpha* L. ecotype BoGa (Buschmann, Holtmannspötter et al. 2016). *Physcomitrium patens* (Hedw.) B.S.G. ecotype Reute was used for all experiments (Hiss, Meyberg et al. 2017). *Arabidopsis thaliana* (L.) Heynh. ecotype Columbia (Col-0) was used in all experiments. The following mutants were used in this study: *gnc* (SALK\_001778) *gnl* (SALK\_003995) (Bi, Zhang et al. 2005, Richter, Behringer et al. 2010), quadruple mutant (*quad*) *gnc* (SALK\_001778) *gnl* (SALK\_003995) *gata17* (SALK\_049041) *gata17l* (SALK\_026798) (Ranftl, Bastakis et al. 2016), GNLox (Richter, Behringer et al. 2010).

### 4.2 Plant culture

For asexual propagation and phenotyping, *M. polymorpha* gemmae were plated on ½ Gamborg's B5 medium containing 1.4% plant agar (Duchefa, Harlem, The Netherlands). For routine subculture, plants were grown under 60  $\mu\text{mol m}^{-2} \text{s}^{-1}$  white light (L58W/830, Lumiluxwarm white, 5200 lm, Osram, Munich, Germany) at 22°C under long-day conditions with a 16-hour light/8-hour dark period in MLR-351 plant growth chambers (Sanyo Denki, Eschborn, Germany). Experiments with moderate high light were carried out in custom walk-in plant growth chambers (Vötsch Industrietechnik, Balingen-Frommern, Germany) equipped with True Daylight dual white-light LED panels with warm white (2700K) and cold white (6500K) LEDs (Poly Klima, Freising, Germany) at the maximum light intensity of 400 - 500  $\mu\text{mol m}^{-2} \text{s}^{-1}$  PPFD and ambient temperature of 20 - 22°C.

*P. patens* was plated on BCD medium containing 0.8% plant agar (Duchefa, Harlem, The Netherlands) supplemented with 5 mM ammonium tartrate for routine subculture (Thelander, Nilsson et al. 2007). Plants were grown under 60  $\mu\text{mol m}^{-2} \text{s}^{-1}$  white light (L58W/830, Lumiluxwarm white, 5200 lm, Osram, Munich, Germany) at 22°C under long-day conditions with a 16-hour light/8-hour dark period in MLR-351 plant growth chambers (Sanyo Denki, Eschborn, Germany).

For *A. thaliana* growth, seeds were stratified for three days in the dark at 4°C, before being plated on ½ strength Murashige & Skoog medium containing 0.8% plant agar (Duchefa, Harlem, The Netherlands). Unless otherwise stated, plants were grown under 100  $\mu\text{mol m}^{-2} \text{s}^{-1}$  white light (L58W/830, Lumiluxwarm white, 5200 lm, Osram, Munich, Germany) at 22°C in long day conditions with a 16-hour light/8-hour dark cycle.

### 4.3 Phylogeny construction

GATAs from *A. thaliana*, *M. polymorpha*, *P. patens*, *S. moellendorffii*, *S. cucullata*, and *C. braunii* were identified based on the existence of the GATA-type zinc finger domains as defined by PF00320 using HMMER (hmmer.org) and BLAST searches. Zinc-finger sequences were aligned using MUSCLE (Edgar 2004) and phylogeny construction was carried out with PhyML (Guindon, Dufayard et al. 2010) via ATGC (<http://www.atgc-montpellier.fr/phyml/>). The Q model was chosen using automatic model selection by SMS based on the Bayesian Information Criterion (Lefort, Longueville et al. 2017) and used for phylogeny construction.

### 4.4 Thallus size quantification

Thallus size of *M. polymorpha* plants was quantified using the SmartGrain high-throughput phenotyping software (Tanabata, Shibaya et al. 2012). Gemma cups were manually counted and the number of gemma cups was divided by thallus size for the gemma cups per area quantification.

## 4.5 Air pore size quantification

Air pores size of *M. polymorpha* plants was quantified by manually tracing individual air pores in ImageJ and calculating their size based on the scale provided with the photographs (Schneider, Rasband et al. 2012).

## 4.6 Molecular cloning and CRISPR/Cas9-mediated genome editing

*M. polymorpha* B-GATA mutants *Mpb-gata1-1* and *Mpb-gata1-2* were generated in the BoGa ecotype by CRISPR/Cas9 mediated gene targeting of Mp7g03490 using two gRNAs. The gRNAs were synthesized as oligos with overhangs for *Bsal* and cloned into pMpGE\_En03 as previously described (Sugano, Shirakawa et al. 2014). LR reaction of the resulting Entry clones with pMpGE010 and pMpGE011, respectively, was carried out according to the manual (LR Clonase, ThermoFisher Scientific, Waltham, USA). The expression clones were transformed into *Agrobacterium tumefaciens* strain GV2260 and *M. polymorpha* sporeling transformation was performed with *Agrobacterium* cultures of both expression clones as previously described (Althoff, Kopschke et al. 2014). Selection of the transformed sporelings was carried out using both 0.5  $\mu$ M chlorsulfuron and 10 mg/L hygromycin and 100 mg/L Cefotaxime. DNA of resistant plants was extracted and genotyping by PCR was performed. PCR products were analyzed by gel electrophoresis and Sanger sequencing.

For CRISPR/Cas9 mediated genome editing in *P. patens*, each gene was targeted with one gRNA. A 500 bp fragment containing attB sites, the snRNA U6 promoter (Collonnier, Epert et al. 2017), and the gRNA was chemically synthesized (GeneArt, ThermoFisher Scientific, Waltham, USA). After synthesis, the fragment was cloned into pDONR207 (ThermoFisher Scientific, Waltham, MA). *P. patens* protoplasts were extracted and transformed as previously described (Ashton, Grimsley et al. 1979, Schaefer and Zryd 1997, Lopez-Obando, Hoffmann et al. 2016). After transformation, plants were regenerated on BCDAT medium supplemented with 0.33 M mannitol and overlaid with cellophane disks for one week. Selection of regenerated protoplasts was carried out by transferring the cellophane disks onto new plates containing 50 mg/l G418 (Duchefa, Harlem, The Netherlands). DNA of resistant plants was extracted and genotyping by PCR was performed. PCR products were analyzed by gel electrophoresis and Sanger sequencing.

The *A. thaliana gata hexuple (hex)* mutants were generated by CRISPR/Cas9 mediated gene disruption of *GATA15* and *GATA16* in the *gata quad* mutant. One guide-RNA (gRNA) per gene (for *GATA15* and for *GATA16*) was used, resulting in frame shifts before (*GATA15*) or in the zinc finger-coding-region (*GATA16*) leading to missense mutations and premature stop codons. The gRNAs were cloned into pHEE401E as previously described (Wang, Xing et al. 2015) and introduced into the quadruple mutant by floral dipping. Seeds of the T1 generation were plated on half-strength MS medium containing 25 mg/l hygromycin. The resistant seedlings were transferred to soil and gDNA was extracted for high resolution melt analysis (HRM) (Gundry, Vandersteen et al. 2003), PCR and subsequent sanger sequencing. The melt curves were analyzed in Precision Melt Analysis Software (Biorad, Hercules, USA) by comparing the melt curves of potential mutants to the ones of Col-0 samples. PCR was carried out for the putative mutant plants identified via HRM. The PCR products were purified and sequenced. Estimation of indels in the heterozygous T1 plants was conducted using Tracking of Indels by Decomposition (TIDE) software (Brinkman, Chen et al. 2014) and resulted in several homozygous mutants for either *GATA15* or *GATA16*. Out of 234 screened plants, nine quintuple mutants *gnc gnl gata15 GATA16 gata17 gata17l* and two quintuple mutants *gnc gnl GATA15 gata16 gata17 gata17l* but no hexuple mutants could be identified. Thus, the quintuple mutants, *gnc gnl gata15 GATA16 gata17 gata17l* and *gnc gnl GATA15 gata16 gata17 gata17l*, were crossed. The F1 generation was genotyped as described above and yielded several independent *gata hex* mutant lines *gnc gnl gata15 gata16 gata17 gata17l*.

To obtain Mp*B-GATA1ox3*, the coding sequence of Mp*B-GATA1* (*Mp7g03490*) was amplified from cDNA for Gateway-cloning and subcloned via pDONR221 (ThermoFisher Scientific, Waltham, MA) into pMpGWB224 (Ishizaki, Nishihama et al. 2016).

#### 4.7 Quantitative real-time PCR

RNA was extracted using the NucleoSpin RNA kit (Machery-Nagel, Düren, Germany). cDNA was generated through reverse transcription using oligo(dT) primers and RevertAid reverse transcriptase (ThermoFisher Scientific, Waltham, USA). qRT-PCR was carried out using Takyon No ROX SYBR MasterMix (Eurogentec, Seraing, Belgium) in a CFX96 or CFX384 Real-Time System Cyclers (Bio-Rad) with the following protocol: step 1, 95°C for 3:00 min; step 2, 95°C for 0:10 min; step 3, 55°C for 0:30 min followed by 39 repeats of steps 2 and 3. Expression was normalized to *MpACT7* (*Mp6g11010*) and copy number was calculated as previously described (Pfaffl 2001).

#### 4.8 RNA sequencing

For transcriptomic analysis comparing the two *Mpb-gata1* alleles to the BoGa wildtype, gemmae of BoGa, *Mpb-gata1-1* and *Mpb-gata1-2* were grown on half-strength Gamborg's B5 medium for three weeks under long day conditions ( $60 \mu\text{mol m}^{-2} \text{s}^{-1}$ ) at 22°C. Three biological replicates were prepared for each genotype, each replicate consisting of three entire plants. Plants were frozen in liquid nitrogen and RNA was extracted using the NucleoSpin RNA kit (Machery-Nagel, Düren, Germany). The quality and quantity of the extracted RNA was checked on a Bioanalyzer 2100 microfluidics system (Agilent, Santa Clara, USA) using the RNA Nano chip. Library construction was performed using the TruSeq stranded mRNA Kit (Illumina, San Diego, USA) with TruSeq RNA Adapter Box Set B (Illumina, San Diego, USA) and the protocol 1000000040498v00 (Elution-2-Frag-Prime step: 4 minutes at 94°C; 15 cycles enrichment). Paired-end sequencing (2x75 bp, single-index) was carried out on a HiSeq2500 (Illumina, San Diego, USA) at the Chair of Animal Breeding (TUM, Freising, Germany) with HiSeq Rapid SBS Kit v2 (2x 50 cycles), HiSeq PE Rapid Cluster Kit v2 (Illumina, San Diego, USA) and onboard clustering. The reads were mapped onto the Marchantia genome V5.1 (<http://marchantia.info>) using the CLC Genomics Workbench (Qiagen, Hilden, Germany). For the discovery of differentially expressed genes (DEGs), genes with a FDR p-value < 0.05 were considered. No fold change filter was applied.

For the high-light intensity RNA-seq experiment, *M. polymorpha* lines BoGa, *Mpb-gata1-1*, and *Mpb-gata1-2* were grown on half-strength Gamborg's B5 medium under standard conditions described above for two weeks. Following this initial growth period, plants were shifted to the high-light sun simulation chambers (Research Unit for Environmental Simulation, Biochemical Plant Pathology, HMGU, Munich, Germany) for an additional growth- and acclimatization period of one week under constant light with  $60\text{-}70 \mu\text{mol m}^{-2} \text{s}^{-1}$  (low-light; LL) at 20-22°C. Seeds of the *A. thaliana* lines Col-0, *gata hex* and GNLox were plated on half-strength MS medium and stratified for 3 days in the dark at 4°C. Plants were grown in the high-light sun simulation chambers for one week under constant light with  $60\text{-}70 \mu\text{mol m}^{-2} \text{s}^{-1}$  (low-light; LL) at 20-22°C. The light treatments consisted of either six hours of high-light (HL) with  $1200 \mu\text{mol m}^{-2} \text{s}^{-1}$  at 20-22°C ( $t_{6\text{HL}}$ ) or a six-hour mock treatment under constant light with  $60\text{-}70 \mu\text{mol m}^{-2} \text{s}^{-1}$  at 20-22°C ( $t_{6\text{LL}}$ ). One third of the plants were harvested and immediately frozen in liquid nitrogen each before the shift to high-light ( $t_0$ ), after six hours of high-light ( $t_{6\text{HL}}$ ) or six hours of mock ( $t_{6\text{LL}}$ ). Three biological replicates were prepared for each genotype, each replicate consisting of three entire plants for *M. polymorpha* or 30 seedlings for *A. thaliana*. RNA was extracted using the NucleoSpin RNA kit (Machery-Nagel, Düren, Germany). The quality and quantity of the extracted RNA was checked on a Bioanalyzer 2100 microfluidics system (Agilent, Santa Clara, USA) using the RNA Nano chip. For the *M. polymorpha* samples, library construction was performed using the TruSeq stranded mRNA Kit (Illumina, San Diego, USA)

with TruSeq RNA Adapter Box Set A and B (Illumina, San Diego, USA) and the protocol 1000000040498v00 (Elution-2-Frag-Prime step: 4 minutes at 94°C; 15 cycles enrichment). Paired-end sequencing (2x100 bp, single index) was carried out on a NovaSeq 6000 at the NGS Core Facility Helmholtz Zentrum München (Munich, Germany) with NovaSeq 6000 S1 Reagent Kit v1.5 (200 cycles). For the *A. thaliana* samples, library construction was performed using the Illumina stranded mRNA Kit (Illumina, San Diego, USA) with IDT for Illumina RNA UD Indexes Set A and the protocol 1000000124518v01 with 11 cycles of enrichment. Single-end sequencing (1x100 bp, dual-index) was performed on NovaSeq 6000 at IMG/M Laboratories (Martinsried, Germany) with NovaSeq 6000 S1 Reagent Kit v1.5 (100 cycles). Reads were mapped onto the *Marchantia* genome V6.1 (<http://marchantia.info>) or the *A. thaliana* TAIR10 genome (<http://arabidopsis.org>), respectively, using the CLC Genomics Workbench (Qiagen, Hilden, Germany). For the discovery of differentially expressed genes (DEGs), genes with a FDR p-value  $\leq 0.05$  were considered. In case of the *A. thaliana* data, DEGs were filtered by  $\log_2$  fold change  $\geq 1$  and  $\leq -1$ . No fold change filter was applied for the *M. polymorpha* data. Heat maps were generated using the Cluster 3.0 and Treeview tools (De Hoon, Imoto et al. 2004, Saldanha 2004).

#### 4.8.1 Identification of orthologs

Orthofinder2 (Emms and Kelly 2018) was used to identify *A. thaliana* orthologs of *M. polymorpha* genes. The program was run using default settings. *M. polymorpha* genome version v6.1 and *A. thaliana* genome version TAIR10 were used as input. *M. polymorpha* genes that did not yield an Orthofinder2 result, were annotated by running a DIAMOND search against the TAIR10 genome (Buchfink, Xie et al. 2015). The program was run using default settings. The *A. thaliana* Best-Hit gene was considered an ortholog to the corresponding *M. polymorpha* input gene.

#### 4.8.2 GO enrichment analysis

Gene ontology (GO) enrichment analysis for *M. polymorpha* was performed by first converting gene IDs to genome version v3.1 nomenclature. Enrichment analysis was performed at PlantRegMap (<http://plantregmap.gao-lab.org/>) (Tian, Yang et al. 2020). GO terms with p-value  $\leq 0.05$  were considered statistically significant. The *A. thaliana* datasets were analyzed using PANTHER (<http://pantherdb.org/>) (Mi, Ebert et al. 2021) with default settings.

#### 4.8.3 Accession Numbers

Data from the RNA-seq experiments have been deposited in NCBI-GEO (<https://www.ncbi.nlm.nih.gov/geo/>) under accession series number GSE208563.

### 4.9 Chlorophyll quantification

Chlorophylls were extracted as previously described with slight modifications (Hu, Tanaka et al. 2013). Shortly, 30-100mg of *M. polymorpha* frozen thallus tissue was homogenized using steel beads in a TissueLyser II (Qiagen, Hilden, Germany). Ground material was incubated in 80% acetone for 30 min in the dark and centrifugated for 15 min at 16.000g. The supernatant was mixed in a 1:1 ration with methanol and incubated for 30 min in the dark. The mixture was again centrifugated for 15 min at 16.000g and the absorbance of the supernatant was measured at 470 nm, 645 nm and 663 nm in a Tristar 5 plate reader (Berthold, Bad Wildbad, Germany). Chlorophyll and carotenoid contents were calculated as previously described with modifications as follows (Lichtenthaler and Buschmann 2001):

$$\text{Chl}_a \text{ (mg/mg FW)} = [12.25A_{663\text{nm}} - 2.79A_{645\text{nm}}] * V/W$$

$$\text{Chl}_b \text{ (mg/mg FW)} = [21.5A_{645\text{nm}} - 5.10A_{663\text{nm}}] * V/W$$

$$\text{Chl}_{a+b} \text{ (mg/mg FW)} = [7.15A_{663\text{nm}} + 18.71A_{645\text{nm}}] * V/W$$

$$C_{x+c} \text{ (mg/mg FW)} = \frac{[[1000A_{470\text{nm}} - 1.63\text{Chl}_a - 85.0\text{Chl}_b]/198]*V}{W}$$

## 4.10 Detection of superoxide with NBT staining

NBT staining was carried out as previously described (Juszczak and Baier 2014). Shortly, leaves of thalli were placed in 10 mM sodium azide immediately after harvesting and incubated under vacuum for 10 min. The sodium azide solution was discarded, and leaves or thalli were incubated in NBT under vacuum for 12 h. Afterwards, leaves or thalli were de-stained and photographed. Area quantification of stained leaves was carried out using ImageJ (Schneider, Rasband et al. 2012).

## 4.11 gRNAs used for CRISPR/Cas9-mediated genome editing

### 4.11.1 *M. polymorpha* gRNAs

Table 4.1: gRNAs used for CRISPR/Cas9 mediated genome editing in *M. polymorpha*.

Name	Target	Sequence
MpB-GATA1 gRNA 1	Mp7g03490	5'-CTGTCCATAAACGAGCGGCG-3'
MpB-GATA1 gRNA 2	Mp7g03490	5'-GCTATTGCCTCGCCTATCAG-3'

### 4.11.2 *P. patens* gRNAs

Table 4.2: gRNAs used for CRISPR/Cas9 mediated genome editing in *P. patens*.

Name	Target	Sequence
PpGATA1 gRNA	Pp3c22_11040	5'-TTTTTTTGGAGCTCGTCGACC-3'
PpGATA3 gRNA	Pp3c18_14800	5'-ATCAGTGATTCGGCCTCGTT-3'
PpGATA4 gRNA	Pp3c22_11080	5'-CTCGCACGTGTGGCGCGTGA-3'

### 4.11.3 *A. thaliana* gRNAs

Table 4.3: gRNAs used for CRISPR/Cas9 mediated genome editing in *A. thaliana*.

Name	Target	Sequence
GATA15 gRNA	AT3G06740	5'-AAGAAGAGTTGTGCCATTTG-3'
GATA16 gRNA	AT5G49300	5'-GTATGAACCTTTGGACCAAC-3'

## 4.12 Primers used in this study

### 4.12.1 *Mpb-gata1* genotyping primers

Table 4.4: Primers for genotyping of *Mpb-gata1* mutants.

Name	Target	Sequence
MpB-GATA1 fw	Mp7g03490	5'-CGAAATTGAGACAAGCGAGG-3'
MpB-GATA1 rv	Mp7g03490	5'-GACGCTAGTGGCTCGATCTC-3'

### 4.12.2 *P. patens* genotyping primers

Table 4.5: Primers for genotyping of PpGATA mutants.

Name	Target	Sequence
PpGATA1 fw	Pp3c22_11040	5'-ATTCAGTGATGGCGGCAGAAAG-3'
PpGATA1 rv	Pp3c22_11040	5'-CATGCCTGTCCGGCCTCTTG-3'
PpGATA3 fw	Pp3c18_14800	5'-TCTGTTTGCACAATGTCTCTCG-3'
PpGATA3 rv	Pp3c18_14800	5'-ATTTGGAGGCATTTCGCTTCG-3'
PpGATA4 fw	Pp3c22_11080	5'-ATGGACGAGTCAATCTGCGG-3'
PpGATA4 rv	Pp3c22_11080	5'-TCTTGCATCAACCGACTGGA-3'

### 4.12.3 *A. thaliana* SALK primers

Table 4.6: Primers for genotyping of *A. thaliana* SALK lines.

Name	Target	Sequence
GNC SALK fw	AT5G56860	5'-TCTCTGAATCTCCGGCAACC-3'
GNC SALK rv	AT5G56860	5'-CTGATCATCTCCATCTTTCTC-3'
GNL SALK fw	AT4G26150	5'-ATGCTAGATCATCGAAATAGATATTG-3'
GNL SALK rv	AT4G26150	5'-TATCTGATGGTGGTTCATCATCAAG-3'
GATA17 SALK fw	AT3G16870	5'-TCAAACTTTGTGGTTGAGGC-3'
GATA17 SALK rv	AT3G16870	5'-TCGCGATCCATTGATTAGATC-3'
GATA17L SALK fw	AT4G16141	5'-CAGCTGGACTTTATTTTGGAGC-3'
GATA17L SALK rv	AT4G16141	5'-TTGTGTGCTTAGCTCTGTTGC-3'
AtLBp1.3	SALK TDNA	5'-ATTTTGCCGATTTCCGGAAC-3'

### 4.12.4 *A. thaliana gata hex* genotyping primers

Table 4.7: Primers for genotyping of the *A. thaliana gata hex* mutant.

Name	Target	Sequence
GATA15 HRM fw	AT3G06740	5'-ATCGAAGAACACAGCAGCAGTA-3'
GATA15 HRM rv	AT3G06740	5'-GGAAGAGACTAACCTTGGGACC-3'
GATA16 HRM fw	AT5G49300	5'-TGCTGATTGTGGAACCAGT-3'
GATA16 HRM rv	AT5G49300	5'-TCAAATGATCAACCTCGATTT-3'
GATA15 PCR fw	AT3G06740	5'-ACGCGCTACTACTTCTGTGT-3'
GATA15 PCR rv	AT3G06740	5'-AAGAGACTAACCTTGGGACCG-3'
GATA16 PCR fw	AT5G49300	5'-TGTTGATTCAGAAACCATGAA-3'
GATA16 PCR rv	AT5G49300	5'-CGACTACACAGAACCAAACCA-3'

## 4.13 qRT-PCR primers

Table 4.8: Primers used for qualitative real-time PCR.

Name	Target	Sequence
MpACT7 qRT fw	<i>Mp6g11010</i>	5'-AGGCATCTGGTATCCACGAG-3'
MpACT7 qRT rv	<i>Mp6g11010</i>	5'-ACATGGTCGTTCCCTCCAGAC-3'
MpB-GATA1 qRT fw	<i>Mp7g03490</i>	5'-TCGTTTCCCTCAAAGGTCCAC-3'
MpB-GATA1 qRT rv	<i>Mp7g03490</i>	5'-CATTGCACAGAGACTTCGGG-3'

## 4.14 Cloning of MpB-GATA1

Table 4.9: Primers for molecular cloning of MpB-GATA1 with the Gateway system.

Name	Target	Sequence
MpB-GATA1 GW fw	<i>Mp7g03490</i>	5'-GGGGACAAGTTTGTACAAAAA- AGCAGGCTTAATGTTGCAAGCCCCATGCTATC-3'
MpB-GATA1 GW rv	<i>Mp7g03490</i>	5'-GGGGACCACTTTGTACAAG- AAAGCTGGGTACATGAGCCTTGGGGCGA-3'

## 5 Supplemental materials

### 5.1 Supplemental Data

**Supplemental Table S5.1: Quantification of cytokinins from five replicates of *P. patens* R17, *gata triple 10* and *gata triple 12*.** Cytokinin levels in 1 g of one-month old gametophores extracted (pmol/g fresh weight; Mean  $\pm$  SD). Asterisks indicate statistically significant difference in mutant lines versus wild types in a paired Student's t-test (\*, \*\*, and \*\*\* correspond to P-values of 0.05 > p > 0.01, 0.01 > p > 0.001, and p < 0.001, respectively). u.q., under quantification limit.

	R17		<i>ppgata1.1/3/4-10</i>			<i>ppgata1.2/3/4-11</i>		
<b>CK types</b>			pmol/g fresh weight					
total cytokinins	440.46	$\pm$ 125.39	428.73	$\pm$ 74.94	ns	<b>647.79</b>	$\pm$ <b>122.22</b>	*
CK Bases	5.78	$\pm$ 1.11	5.34	$\pm$ 0.37	ns	7.18	$\pm$ 1.34	ns
CK Ribosides	3.20	$\pm$ 0.59	4.37	$\pm$ 1.08	ns	2.45	$\pm$ 0.33	ns
CK Nucleotides	16.27	$\pm$ 4.17	23.60	$\pm$ 4.57	*	16.00	$\pm$ 1.97	ns
CK O-glucosides	415.21	$\pm$ 122.37	395.43	$\pm$ 72.21	ns	622.16	$\pm$ 123.85	*
CK N-glucosides	u.q.		u.q.			u.q.		
<b>iP types</b>								
total iP types	20.05	$\pm$ 5.21	<b>28.81</b>	$\pm$ <b>5.13</b>	*	21.96	$\pm$ 2.66	ns
iP	3.64	$\pm$ 0.88	4.16	$\pm$ 0.35	ns	5.69	$\pm$ 1.14	*
iPR	1.73	$\pm$ 0.47	2.98	$\pm$ 0.71	*	1.61	$\pm$ 0.36	ns
iPRMP	14.67	$\pm$ 4.32	21.67	$\pm$ 4.54	ns	14.66	$\pm$ 2.24	ns
iP7G	u.q.		u.q.			u.q.		
iP9G	u.q.		u.q.			u.q.		
<b>tZ types</b>								
total tZ types	6.67	$\pm$ 1.21	7.57	$\pm$ 0.84	ns	<b>13.11</b>	$\pm$ <b>2.91</b>	**
tZ	0.02	$\pm$ 0.01	0.03	$\pm$ 0.00	*	0.03	$\pm$ 0.01	ns
tZR	0.04	$\pm$ 0.01	0.05	$\pm$ 0.01	ns	0.05	$\pm$ 0.01	ns
tZRMP	u.q.		u.q.			u.q.		
tZOG	0.41	$\pm$ 0.09	0.70	$\pm$ 0.16	*	1.00	$\pm$ 0.24	**
tZROG	6.20	$\pm$ 1.13	6.78	$\pm$ 0.80	ns	12.04	$\pm$ 2.66	**
tZ7G	u.q.		u.q.			u.q.		
tZ9G	u.q.		u.q.			u.q.		
<b>cZ types</b>								
total cZ types	413.74	$\pm$ 121.17	392.34	$\pm$ 71.55	ns	612.72	$\pm$ 122.17	*
cZ	2.12	$\pm$ 0.31	1.14	$\pm$ 0.14	***	1.47	$\pm$ 0.22	**
cZR	1.43	$\pm$ 0.29	1.34	$\pm$ 0.39	ns	0.80	$\pm$ 0.15	**
cZRMP	1.59	$\pm$ 0.40	1.92	$\pm$ 0.19	ns	1.33	$\pm$ 0.36	ns
cZOG	7.20	$\pm$ 1.47	10.37	$\pm$ 2.24	*	9.79	$\pm$ 2.12	ns
cZROG	401.40	$\pm$ 121.16	377.57	$\pm$ 69.41	ns	599.33	$\pm$ 120.06	*
cZ7G	u.q.		u.q.			u.q.		
cZ9G	u.q.		u.q.			u.q.		

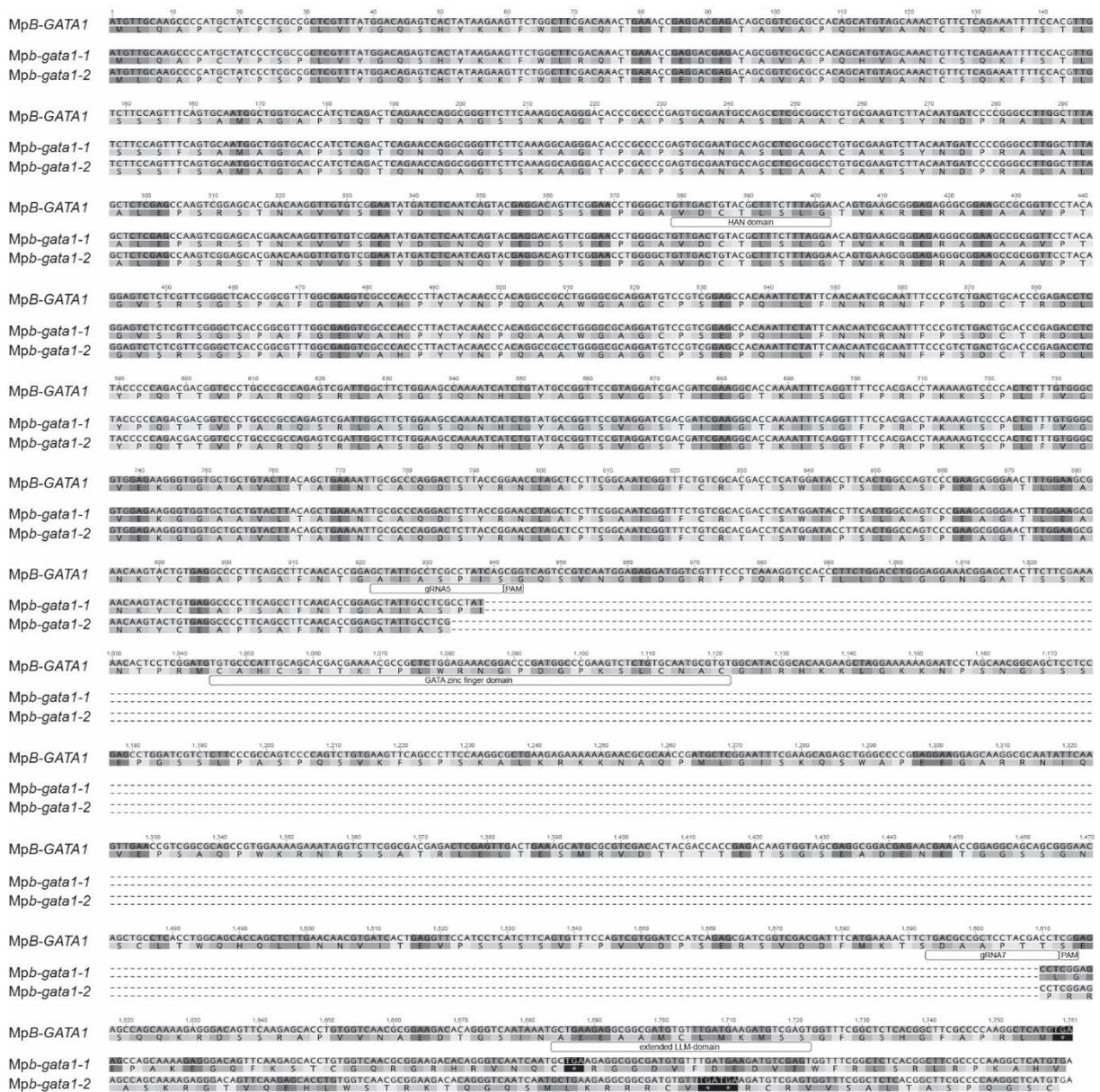
**Supplemental Table S5.2: Quantification of cytokinins from *A. thaliana* Col-0, *gata hex* and *GNLox*.** Cytokinin levels in 1g of eight-days old seedlings as pmol/g fresh weight; Mean  $\pm$  SD. Asterisks indicate statistically significant difference in mutant lines versus wild types in a paired Student's t-test (\*, \*\*, and \*\*\* correspond to p-values of 0.05 > p > 0.01, 0.01 > p > 0.001, and p < 0.001, respectively). u.q., under quantification limit.

	<b>Col-0 (n = 10)</b>		<b><i>gata hex</i> (n = 10)</b>			<b><i>GNLox</i> (n = 7)</b>		
<b>CK types</b>	pmol/g fresh weight							
total cytokinins	110.43 $\pm$ 12.66		<b>77.97 <math>\pm</math> 8.64</b> *			123.77 $\pm$ 12.07	ns	
CK bases	1.12 $\pm$ 0.17		0.74 $\pm$ 0.15	***		0.90 $\pm$ 0.21	*	
CK ribosides	8.62 $\pm$ 1.25		4.38 $\pm$ 1.03	***		9.71 $\pm$ 1.54	ns	
CK nucleotides	39.74 $\pm$ 6.49		17.14 $\pm$ 3.37	***		47.03 $\pm$ 9.36	ns	
CK O-glucosides	12.79 $\pm$ 2.06		12.15 $\pm$ 1.59	ns		15.63 $\pm$ 1.77	*	
CK N-glucosides	48.15 $\pm$ 5.37		43.56 $\pm$ 3.82	ns		50.50 $\pm$ 7.42	ns	
<b>iP types</b>								
total iP types	35.36 $\pm$ 5.21		<b>21.88 <math>\pm</math> 2.57</b> ***			<b>48.70 <math>\pm</math> 6.58</b> ***		
iP	0.43 $\pm$ 0.13		0.28 $\pm$ 0.07	**		0.36 $\pm$ 0.12	ns	
iPR	3.69 $\pm$ 0.87		1.52 $\pm$ 0.50	***		4.64 $\pm$ 1.16	ns	
iPRMP	22.71 $\pm$ 4.17		9.13 $\pm$ 1.90	***		29.53 $\pm$ 7.22	*	
iP7G	6.63 $\pm$ 1.00		8.92 $\pm$ 1.47	**		12.03 $\pm$ 3.14	***	
iP9G	1.89 $\pm$ 0.36		2.03 $\pm$ 0.34	ns		2.13 $\pm$ 0.39	ns	
<b>tZ types</b>								
total tZ types	48.66 $\pm$ 6.77		<b>33.82 <math>\pm</math> 5.24</b> ***			<b>41.44 <math>\pm</math> 5.47</b> *		
tZ	0.36 $\pm$ 0.09		0.18 $\pm$ 0.04	***		0.26 $\pm$ 0.03	*	
tZR	2.96 $\pm$ 0.49		1.19 $\pm$ 0.32	***		2.73 $\pm$ 0.40	ns	
tZRMP	9.20 $\pm$ 2.55		3.24 $\pm$ 0.77	***		8.21 $\pm$ 1.91	ns	
tZOG	6.78 $\pm$ 0.94		7.18 $\pm$ 1.08	ns		5.86 $\pm$ 0.69	ns	
tZROG	2.19 $\pm$ 0.49		1.58 $\pm$ 0.21	**		2.35 $\pm$ 0.46	ns	
tZ7G	14.39 $\pm$ 1.36		12.06 $\pm$ 2.01	**		11.36 $\pm$ 1.43	***	
tZ9G	12.78 $\pm$ 1.93		8.41 $\pm$ 1.17	***		10.68 $\pm$ 1.73	*	
<b>cZ types</b>								
total cZ types	22.52 $\pm$ 3.03		19.92 $\pm$ 2.49	ns		<b>31.71 <math>\pm</math> 3.59</b> ***		
cZ	0.33 $\pm$ 0.08		0.28 $\pm$ 0.09	ns		0.28 $\pm$ 0.09	ns	
cZR	1.89 $\pm$ 0.48		1.61 $\pm$ 0.43	ns		2.23 $\pm$ 0.55	ns	
cZRMP	7.83 $\pm$ 0.92		4.78 $\pm$ 1.21	***		9.28 $\pm$ 1.14	*	
cZOG	1.40 $\pm$ 0.36		1.21 $\pm$ 0.19	ns		3.53 $\pm$ 0.81	***	
cZROG	2.17 $\pm$ 0.60		2.05 $\pm$ 0.49	ns		3.75 $\pm$ 0.76	***	
cZ7G	8.52 $\pm$ 1.86		9.63 $\pm$ 1.13	ns		12.31 $\pm$ 2.36	**	
cZ9G	0.38 $\pm$ 0.07		0.38 $\pm$ 0.09	ns		0.32 $\pm$ 0.10	ns	
<b>DHZ types</b>								
total DHZ types	3.90 $\pm$ 0.67		<b>2.34 <math>\pm</math> 0.59</b> ***			<b>1.93 <math>\pm</math> 0.22</b> ***		
DHZ	u.q.		u.q.			u.q.		
DHZR	0.08 $\pm$ 0.02		0.05 $\pm$ 0.02	**		0.11 $\pm$ 0.03	*	
DHZRMP	u.q.		u.q.			u.q.		
DHZOG	0.26 $\pm$ 0.06		0.15 $\pm$ 0.05	***		0.15 $\pm$ 0.04	**	
DHZROG	u.q.		u.q.			u.q.		
DHZ7G	3.33 $\pm$ 0.59		1.97 $\pm$ 0.52	***		1.49 $\pm$ 0.20	***	
DhZ9G	0.23 $\pm$ 0.04		0.16 $\pm$ 0.04	**		0.18 $\pm$ 0.03	*	

**Supplemental Table S5.3: Quantification of cytokinins from five replicates of *M. polymorpha* BoGa, *Mpb-gata1-1*, *Mpb-gata1-2* and *MpB-GATA1ox3*.** Cytokinin levels in 1 g of tissue surrounding the apical notches (1 – 3 mm) of three-weeks old plants as pmol/g fresh weight; Mean  $\pm$  SD. Asterisks indicate statistically significant difference in mutant lines versus wild types in a paired Student's t-test (\*, \*\*, and \*\*\* correspond to P-values of 0.05 > p > 0.01, 0.01 > p > 0.001, and p < 0.001, respectively). u.q., under quantification limit.

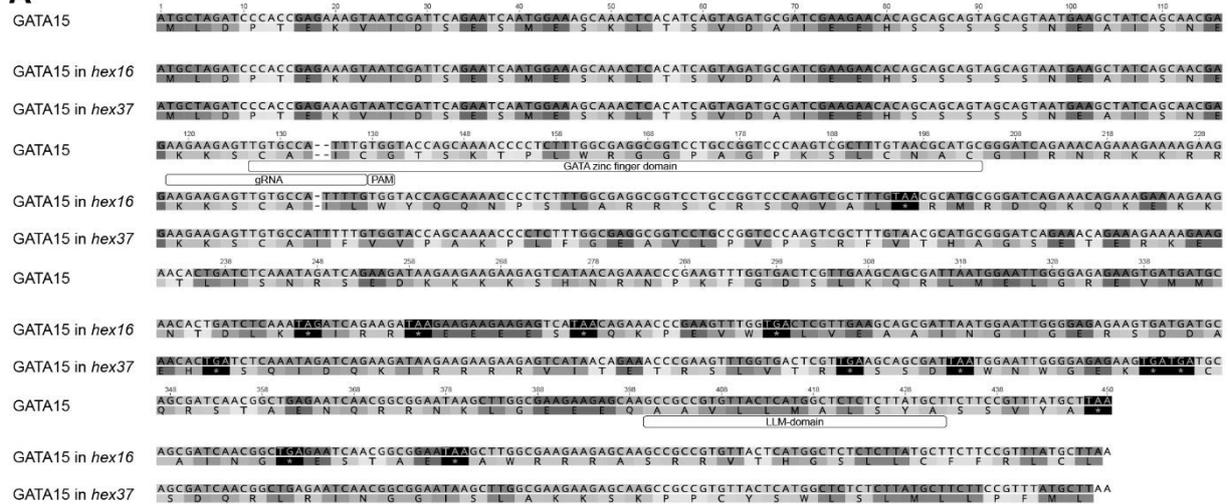
	BoGa			<i>Mpb-gata1-1</i>			<i>Mpb-gata1-2</i>			<i>MpB-GATA1ox3</i>					
<b>CK types</b>	pmol/g fresh weight														
total cytokinins	40.36	$\pm$	5.74	122.68	$\pm$	16.33	***	48.47	$\pm$	6.45	ns	62.04	$\pm$	13.49	*
CK bases	1.49	$\pm$	0.38	5.65	$\pm$	0.89	***	1.62	$\pm$	0.32	ns	2.74	$\pm$	0.41	**
CK ribosides	0.98	$\pm$	0.17	3.55	$\pm$	0.49	***	1.53	$\pm$	0.24	**	1.43	$\pm$	0.33	*
CK nucleotides	36.42	$\pm$	5.17	110.97	$\pm$	14.89	***	44.29	$\pm$	6.12	ns	56.41	$\pm$	12.73	*
CK O-glucosides	1.47	$\pm$	0.24	2.50	$\pm$	0.29	***	1.05	$\pm$	0.14	*	1.45	$\pm$	0.18	ns
CK N-glucosides	u.q.			u.q.				u.q.				u.q.			
<b>iP types</b>															
total iP types	10.91	$\pm$	1.05	45.91	$\pm$	8.10	***	20.63	$\pm$	3.91	**	24.28	$\pm$	5.19	***
iP	0.85	$\pm$	0.24	2.61	$\pm$	0.42	***	0.97	$\pm$	0.16	ns	1.34	$\pm$	0.21	*
iPR	0.45	$\pm$	0.10	1.95	$\pm$	0.32	***	0.86	$\pm$	0.14	*	0.87	$\pm$	0.23	**
iPRMP	9.61	$\pm$	0.99	41.35	$\pm$	7.60	***	18.80	$\pm$	3.88	**	22.08	$\pm$	4.83	***
iP7G	u.q.			u.q.				u.q.				u.q.			
iP9G	u.q.			u.q.				u.q.				u.q.			
<b>tZ types</b>															
total tZ types	0.31	$\pm$	0.07	0.53	$\pm$	0.05	***	0.25	$\pm$	0.03	ns	0.25	$\pm$	0.04	ns
tZ	0.02	$\pm$	0.01	0.03	$\pm$	0.01	ns	0.02	$\pm$	0.00	ns	0.03	$\pm$	0.01	ns
tZR	0.10	$\pm$	0.02	0.12	$\pm$	0.03	ns	0.08	$\pm$	0.02	ns	0.07	$\pm$	0.02	*
tZRMP	u.q.			u.q.				u.q.				u.q.			
tZOG	0.19	$\pm$	0.04	0.37	$\pm$	0.03	***	0.16	$\pm$	0.02	ns	0.15	$\pm$	0.02	ns
tZROG	u.q.			u.q.				u.q.				u.q.			
tZ7G	u.q.			u.q.				u.q.				u.q.			
tZ9G	u.q.			u.q.				u.q.				u.q.			
<b>cZ types</b>															
total cZ types	29.15	$\pm$	4.92	76.25	$\pm$	10.94	***	27.60	$\pm$	3.45	ns	37.51	$\pm$	10.07	ns
cZ	0.62	$\pm$	0.16	3.02	$\pm$	0.50	***	0.63	$\pm$	0.15	ns	1.38	$\pm$	0.24	***
cZR	0.44	$\pm$	0.09	1.48	$\pm$	0.37	***	0.59	$\pm$	0.12	ns	0.49	$\pm$	0.10	ns
cZRMP	26.81	$\pm$	4.54	69.62	$\pm$	10.06	***	25.49	$\pm$	3.23	ns	34.33	$\pm$	9.72	ns
cZOG	0.80	$\pm$	0.16	1.57	$\pm$	0.19	***	0.74	$\pm$	0.07	ns	0.94	$\pm$	0.09	ns
cZROG	0.49	$\pm$	0.11	0.56	$\pm$	0.10	ns	0.25	$\pm$	0.02	*	0.36	$\pm$	0.10	ns
cZ7G	u.q.			u.q.				u.q.				u.q.			
cZ9G	u.q.			u.q.				u.q.				u.q.			

## 5.2 Supplemental Figures

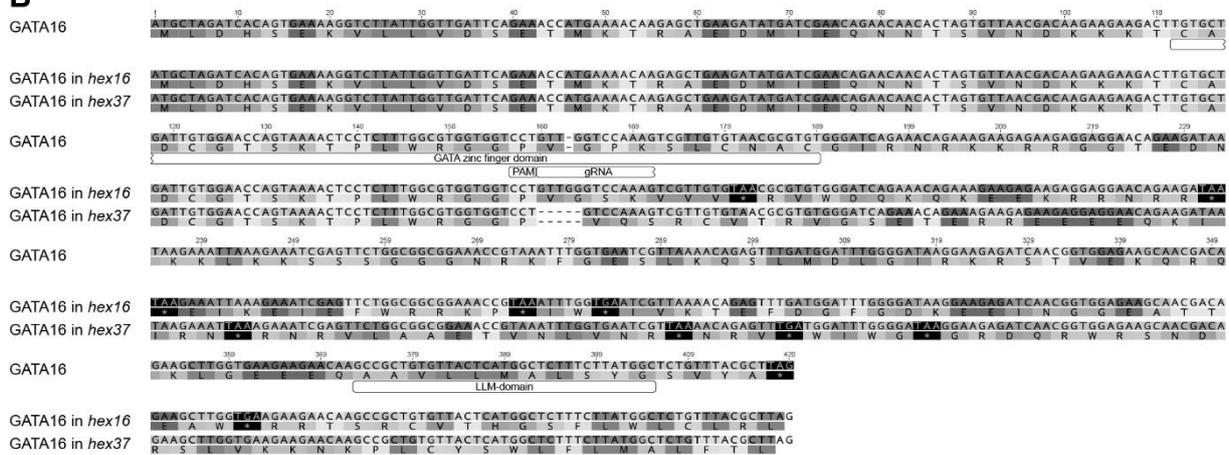


**Figure S5.1: Alignment of the *M. polymorpha* *MpB-GATA1* with the *Mpb-gata1-1* and *Mpb-gata1-2* mutant alleles. Shown is the coding sequence and its translation to the protein sequence using the single letter amino acid code. Codons and their corresponding amino acids are shown in the same shade of grey. Domains and CRISPR/Cas9 guide RNAs (gRNAs) are shown in boxes. Asterisks (\*) indicate stop codons.**

**A**

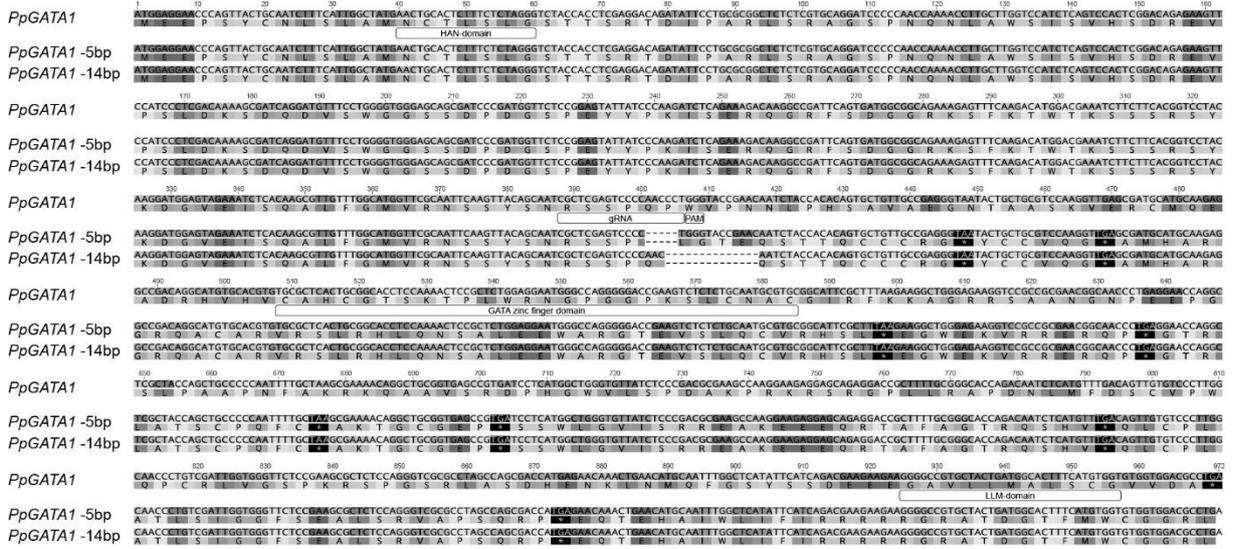


**B**

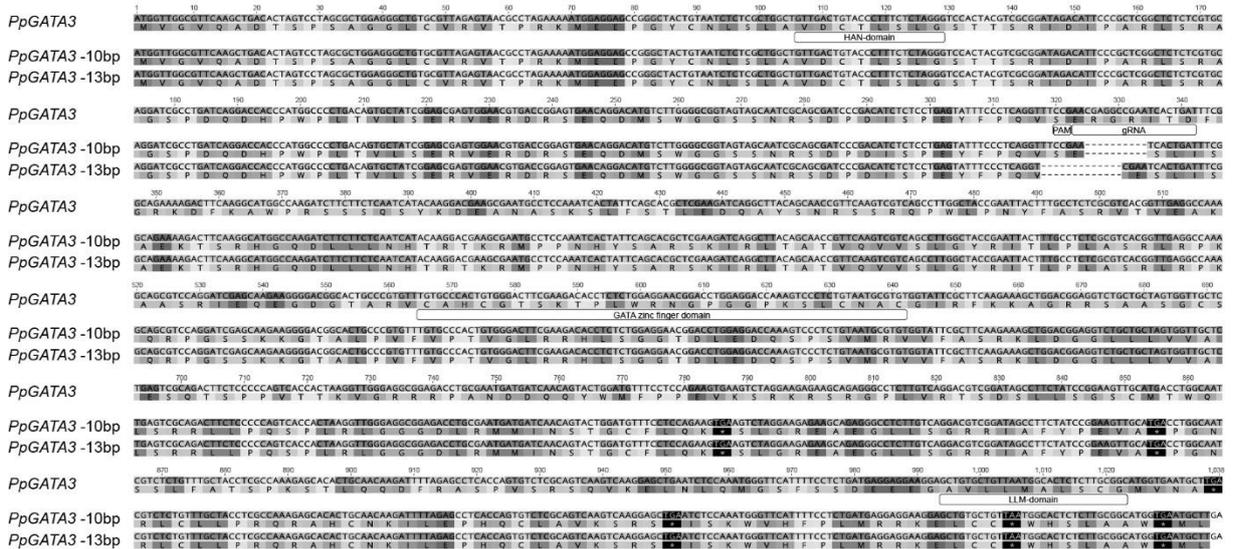


**Figure S5.2: Alignment of the *A. thaliana* wildtype alleles of *GATA15* and *GATA16* with the mutant alleles for both genes in *gata hex*. Shown is the coding sequence and its translation to the protein sequence using the single letter amino acid code for *GATA15* (A) and *GATA16* (B). *A gata hex16* has a one-bp insertion in *GATA15*, *gata hex37* has a two-bp insertion. *B gata hex16* has a one-bp insertion in *GATA16*, *gata hex37* has a four-bp deletion. Codons and their corresponding amino acids are shown in the same shade of grey. Domains and CRISPR/Cas9 guide RNAs (gRNAs) are shown in boxes. The gRNA used for *GATA16* is partially located in an intron and is thus shown as truncated. Asterisks (\*) indicate stop codons.**

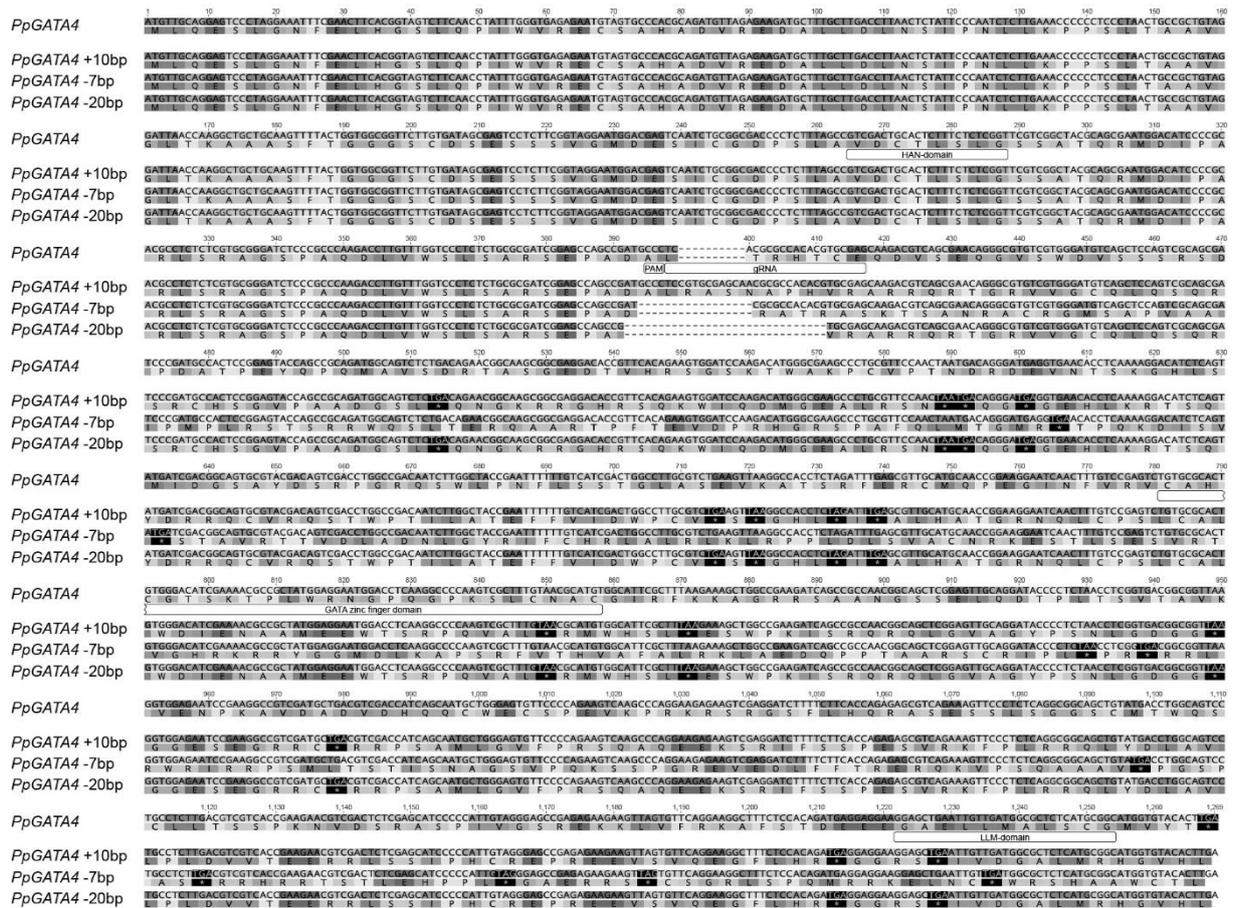
**A**



**B**



**Figure S5.3: Alignment of the *P. patens* R17 wildtype alleles of *PpGATA1* and *PpGATA3* with their mutant alleles in the different *ppgata* mutants. Shown is the coding sequence and its translation to the protein sequence using the single letter amino acid code for *PpGATA1* (A) and *PpGATA3* (B). A For *Ppgata1*, there are two alleles with either five-bp or 14-bp deletions. B For *Ppgata3*, there are two alleles with either 10-bp or 13-bp deletions. Codons and their corresponding amino acids are shown in the same shade of grey. Domains and CRISPR/Cas9 guide RNAs (gRNAs) are shown in boxes. Asterisks (\*) indicate stop codons.**



**Figure S5.4: Alignment of the *P. patens* R17 wildtype allele of *PpGATA4* with the respective mutant alleles in the different *ppgata* mutants.** Shown is the coding sequence and its translation to the protein sequence using the single letter amino acid code for *PpGATA4*. For *PpGata4*, there are three alleles with either a 10-bp insertion or seven-bp and 20-bp deletions. Codons and their corresponding amino acids are shown in the same shadow if grey. Domains and CRISPR/Cas9 guide RNAs (gRNAs) are shown in boxes. Asterisks (\*) indicate stop codons.

## 5.3 Supplemental Tables

**Supplemental Table S5.4: GATA transcription factors from *A. thaliana*.** All 30 GATA transcription factors from *A. thaliana* with their name, TAIR gene ID and classification.

Name	TAIR ID	Class
AtGATA1	AT3G24050	A
AtGATA2	AT2G45050	A
AtGATA3	AT4G34680	A
AtGATA4	AT3G60530	A
AtGATA5	AT5G66320	A
AtGATA6	AT3G51080	A
AtGATA7	AT4G36240	A
AtGATA8/AtBME3	AT3G54810	A
AtGATA9	AT4G32890	A
AtGATA10	AT1G08000	A
AtGATA11	AT1G08010	A
AtGATA12	AT5G25830	A
AtGATA13	AT2G28340	A
AtGATA14	AT3G45170	A
AtGATA15	AT3G06740	B
AtGATA16	AT5G49300	B
AtGATA17	AT3G16870	B
AtGATA17L	AT4G16141	B
AtGATA23	AT5G26930	B
AtGATA29	AT3G20750	B
AtGNC	AT5G56860	B
AtGNL	AT4G26150	B
AtHAN	AT3G50870	B
AtHANL1	AT2G18380	B
AtHANL2	AT4G36620	B
AtGATA24/AtZIML1	AT3G21175	C
AtGATA25/AtZIM	AT4G24470	C
AtGATA28/AtZIML2	AT1G51600	C
AtGATA26	AT4G17570	D
AtGATA27	AT5G47140	D

**Supplemental Table S5.5: *A. thaliana* proteins of the light-dependent reactions of photosynthesis.**

Name	TAIR10 ID	Description
PSAA	ATCG00350	photosystem I P700 chlorophyll a apoprotein A1
PSAB	ATCG00340	photosystem I P700 chlorophyll a apoprotein A2
PSAC	ATCG01060	photosystem I subunit VII
PSAD1	AT4G02770	photosystem I subunit D-1
PSAD2	AT1G03130	photosystem I subunit D-2
PSAE1	AT4G28750	Photosystem I reaction centre subunit IV / PsaE protein
PSAE2	AT2G20260	photosystem I subunit E-2
PSAF	AT1G31330	photosystem I subunit F
PSAG	AT1G55670	photosystem I subunit G
PSAH1	AT3G16140	photosystem I subunit H-1
PSAH2	AT1G52230	photosystem I subunit H2
PSAI	ATCG00510	photosystem I subunit VIII
PSAJ	ATCG00630	photosystem I subunit IX
PSAK	AT1G30380	photosystem I subunit K
PSAL	AT4G12800	photosystem I subunit I
PSAN	AT5G64040	photosystem I reaction center subunit PSI-N
PSAO	AT1G08380	photosystem I subunit O
PSB27-1	AT1G03600	photosystem II family protein
PSB28	AT4G28660	photosystem II reaction center PSB28 protein
PSBA	ATCG00020	photosystem II protein D1
PSBB	ATCG00680	photosystem II 47 kDa protein
PSBC	ATCG00280	photosystem II 44 kDa protein
PSBD	ATCG00270	photosystem II protein D2
PSBE	ATCG00580	photosystem II protein V
PSBF	ATCG00570	photosystem II protein VI
PSBH	ATCG00710	photosystem II protein H
PSBI	ATCG00080	photosystem II protein I
PSBJ	ATCG00550	photosystem II protein J
PSBK	ATCG00070	photosystem II protein K
PSBL	ATCG00560	photosystem II protein L
PSBM	ATCG00220	photosystem II protein M
PSBO2	AT3G50820	photosystem II subunit O-2
PSBP1	AT1G06680	photosystem II subunit P-1
PSBP2	AT2G30790	photosystem II subunit P-2
PSBQ2	AT4G05180	photosystem II subunit Q-2
PSBQ1	AT4G21280	photosystem II subunit QA
PSBR	AT1G79040	photosystem II subunit R
PSBT	ATCG00690	photosystem II protein T

PSBW	AT2G30570	photosystem II reaction centre W
PSBY	AT1G67740	photosystem II BY
PSBZ	ATCG00300	photosystem II protein Z
ATPA	ATCG00120	ATP synthase CF1 alpha subunit
ATPB	ATCG00480	ATP synthase CF1 beta subunit
ATPC1	AT4G04640	ATPase, F1 complex, gamma subunit protein
ATPC2	AT1G15700	ATPase, F1 complex, gamma subunit protein
ATPD	AT4G09650	F-type H <sup>+</sup> -transporting ATPase subunit delta
ATPE	ATCG00470	ATP synthase CF1 epsilon subunit
ATPF	ATCG00130	ATP synthase CF0 B subunit
ATPH	ATCG00140	ATP synthase CF0 C subunit
ATPI	ATCG00150	ATP synthase CF0 A subunit
PDE334	AT4G32260	ATPase, F0 complex, subunit B/B', bacterial/chloroplast
YMF19	AT2G07707	Plant mitochondrial ATPase, F0 complex, subunit 8 protein
FD1	AT1G10960	ferredoxin 1
FD3	AT2G27510	ferredoxin 3
FD4	AT5G10000	ferredoxin 4
FdC2	AT4G14890	2Fe-2S ferredoxin-like superfamily protein
FdC1	AT1G32550	2Fe-2S ferredoxin-like superfamily protein
FD2	AT1G60950	2Fe-2S ferredoxin-like superfamily protein
LFNR1	AT5G66190	ferredoxin-NADP <sup>+</sup> -oxidoreductase 1
LFNR2	AT1G20020	ferredoxin-NADP <sup>+</sup> -oxidoreductase 2
PETJ	AT5G45040	Cytochrome c
PETA	ATCG00540	cytochrome f
PETB	ATCG00720	cytochrome b6
PETD	ATCG00730	cytochrome b6/f complex subunit IV
PETG	ATCG00600	cytochrome b6/f complex subunit V
PETN	ATCG00210	cytochrome b6/f complex subunit VIII
DRT112	AT1G20340	Cupredoxin superfamily protein
PSBS	AT1G44575	Chlorophyll A-B binding family protein
PETC	AT4G03280	photosynthetic electron transfer C
PETE1	AT1G76100	plastocyanin 1
PNSL2	AT1G14150	PsbQ-like 2
PNSL3	AT3G01440	PsbQ-like 1
PPL1	AT3G55330	PsbP-like protein 1
PSBO1	AT5G66570	PS II oxygen-evolving complex 1
RFNR1	AT4G05390	root FNR 1
RFNR2	AT1G30510	root FNR 2

**Supplemental Table S5.6: *A. thaliana* proteins of the Calvin-Benson cycle.**

Name	TAIR10	Description	Phase
rbcL	ATCG00490	Ribulose biphosphate carboxylase large chain	Carboxylation
RBCS1A	AT1G67090	Ribulose biphosphate carboxylase small chain	Carboxylation
RBCS1B	AT5G38430	Ribulose biphosphate carboxylase small chain 1B	Carboxylation
RBCS2B	AT5G38420	Ribulose biphosphate carboxylase small chain 2B	Carboxylation
RBCS3B	AT5G38410	Ribulose biphosphate carboxylase small chain	Carboxylation
PGK	AT1G79550	Phosphoglycerate kinase	Reduction
PGK1	AT3G12780	Phosphoglycerate kinase	Reduction
PGKX	AT1G56190	Phosphoglycerate kinase	Reduction
GAPA	AT3G26650	Glyceraldehyde-3-phosphate dehydrogenase GAPA1	Reduction
GAPA2	AT1G12900	glyceraldehyde 3-phosphate dehydrogenase A subunit 2	Reduction
GAPB	AT1G42970	Glyceraldehyde-3-phosphate dehydrogenase GAPB	Reduction
GAPC1	AT3G04120	Glyceraldehyde-3-phosphate dehydrogenase GAPC1, cytosolic	Reduction
GAPC2	AT1G13440	Glyceraldehyde-3-phosphate dehydrogenase GAPC2, cytosolic	Reduction
GAPCP1	AT1G79530	Glyceraldehyde-3-phosphate dehydrogenase GAPCP1	Reduction
GAPCP2	AT1G16300	Glyceraldehyde-3-phosphate dehydrogenase GAPCP2	Reduction
TIM	AT2G21170	TIM	Regeneration
TPI	AT3G55440	Triosephosphate isomerase, cytosolic	Regeneration
FBA1	AT2G21330	Fructose-bisphosphate aldolase 1, chloroplastic	Regeneration
FBA2	AT4G38970	Fructose-bisphosphate aldolase	Regeneration
FBA4	AT5G03690	Fructose-bisphosphate aldolase 4, cytosolic	Regeneration
FBA5	AT4G26530	Fructose-bisphosphate aldolase	Regeneration
FBA6	AT2G36460	Fructose-bisphosphate aldolase 6, cytosolic	Regeneration
FBA7	AT4G26520	Aldolase superfamily protein	Regeneration
FBA8	AT3G52930	Fructose-bisphosphate aldolase	Regeneration
FBAX	AT1G63290	Ribulose-phosphate 3-epimerase	Regeneration
FBAX	AT3G01850	Ribulose-phosphate 3-epimerase	Regeneration
FBP	AT1G43670	Fructose-1,6-bisphosphatase, cytosolic	Regeneration
FBPX	AT5G64380	AT5G64380/MSJ1_22	Regeneration
TKL2	AT2G45290	Transketolase-2, chloroplastic	Regeneration
TKL1	AT3G60750	Transketolase-1, chloroplastic	Regeneration
SBPASE	AT3G55800	Sedoheptulose-1,7-bisphosphatase, chloroplastic	Regeneration
RPI2	AT2G01290	Probable ribose-5-phosphate isomerase 2	Regeneration
RSW10	AT1G71100	Probable ribose-5-phosphate isomerase 1	Regeneration
RPE	AT5G61410	RPE	Regeneration
PRK	AT1G32060	Phosphoribulokinase	Regeneration

## 6 Literature

- Adamska, I., M. Roobol-Bóza, M. Lindahl and B. Andersson (1999). "Isolation of pigment-binding early light-inducible proteins from pea." European Journal of Biochemistry **260**(2): 453-460.
- Aki, S. S., T. Mikami, S. Naramoto, R. Nishihama, K. Ishizaki, M. Kojima, Y. Takebayashi, H. Sakakibara, J. Kyojuka, T. Kohchi and M. Umeda (2019). "Cytokinin Signaling Is Essential for Organ Formation in *Marchantia polymorpha*." Plant Cell Physiol **60**(8): 1842-1854.
- Aki, S. S., R. Nishihama, T. Kohchi and M. Umeda (2019). "Cytokinin signaling coordinates development of diverse organs in *Marchantia polymorpha*." Plant Signal Behav **14**(11): 1668232.
- Alboresi, A., M. Storti and T. Morosinotto (2019). "Balancing protection and efficiency in the regulation of photosynthetic electron transport across plant evolution." New Phytologist **221**(1): 105-109.
- Alonso, J. M., A. N. Stepanova, T. J. Lisse, C. J. Kim, H. Chen, P. Shinn, D. K. Stevenson, J. Zimmerman, P. Barajas and R. Cheuk (2003). "Genome-wide insertional mutagenesis of *Arabidopsis thaliana*." Science **301**(5633): 653-657.
- Althoff, F., S. Kopischke, O. Zobell, K. Ide, K. Ishizaki, T. Kohchi and S. Zachgo (2014). "Comparison of the MpEF1alpha and CaMV35 promoters for application in *Marchantia polymorpha* overexpression studies." Transgenic Res **23**(2): 235-244.
- An, Y., X. Han, S. Tang, X. Xia and W. Yin (2014). "Poplar GATA transcription factor PdGNC is capable of regulating chloroplast ultrastructure, photosynthesis, and vegetative growth in *Arabidopsis* under varying nitrogen levels." Plant Cell, Tissue and Organ Culture (PCTOC) **119**(2): 313-327.
- An, Y., Y. Zhou, X. Han, C. Shen, S. Wang, C. Liu, W. Yin and X. Xia (2020). "The GATA transcription factor GNC plays an important role in photosynthesis and growth in poplar." J Exp Bot **71**(6): 1969-1984.
- Argueso, C. T., F. J. Ferreira and J. J. Kieber (2009). "Environmental perception avenues: the interaction of cytokinin and environmental response pathways." Plant, cell & environment **32**(9): 1147-1160.
- Ashikari, M., H. Sakakibara, S. Lin, T. Yamamoto, T. Takashi, A. Nishimura, E. R. Angeles, Q. Qian, H. Kitano and M. Matsuoka (2005). "Cytokinin oxidase regulates rice grain production." Science **309**(5735): 741-745.
- Ashton, N., N. Grimsley and D. Cove (1979). "Analysis of gametophytic development in the moss, *Physcomitrella patens*, using auxin and cytokinin resistant mutants." Planta **144**: 427-435.
- Auldridge, M. E., D. R. McCarty and H. J. Klee (2006). "Plant carotenoid cleavage oxygenases and their apocarotenoid products." Current opinion in plant biology **9**(3): 315-321.
- Bailey, S. and A. Grossman (2008). "Photoprotection in cyanobacteria: regulation of light harvesting." Photochemistry and Photobiology **84**(6): 1410-1420.
- Barber, J. (2002). "Photosystem II: a multisubunit membrane protein that oxidises water." Current opinion in structural biology **12**(4): 523-530.
- Bartrina, I., E. Otto, M. Strnad, T. Werner and T. Schmülling (2011). "Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in *Arabidopsis thaliana*." The Plant Cell **23**(1): 69-80.
- Bastakis, E., B. Hedtke, C. Klermund, B. Grimm and C. Schwechheimer (2018). "LLM-Domain B-GATA Transcription Factors Play Multifaceted Roles in Controlling Greening in *Arabidopsis*." Plant Cell **30**(3): 582-599.
- Becker, B. and B. Marin (2009). "Streptophyte algae and the origin of embryophytes." Annals of Botany **103**(7): 999-1004.
- Behringer, C., E. Bastakis, Q. L. Ranftl, K. F. Mayer and C. Schwechheimer (2014). "Functional diversification within the family of B-GATA transcription factors through the leucine-leucine-methionine domain." Plant Physiol **166**(1): 293-305.

- Behringer, C. and C. Schwechheimer (2015). "B-GATA transcription factors - insights into their structure, regulation, and role in plant development." Front Plant Sci **6**: 90.
- Bi, Y. M., Y. Zhang, T. Signorelli, R. Zhao, T. Zhu and S. Rothstein (2005). "Genetic analysis of Arabidopsis GATA transcription factor gene family reveals a nitrate-inducible member important for chlorophyll synthesis and glucose sensitivity." Plant J **44**(4): 680-692.
- Bishopp, A., S. El-Showk, D. Weijers, B. Scheres, J. Friml, E. Benková, A. P. Mähönen and Y. Helariutta (2011). "A mutually inhibitory interaction between auxin and cytokinin specifies vascular pattern in roots." Current biology **21**(11): 917-926.
- Blunt, W. and S. Raphael (1994). Illustrated herbal, Frances Lincoln.
- Bomblyes, K. and A. Madlung (2014). "Polyploidy in the Arabidopsis genus." Chromosome Research **22**: 117-134.
- Bowman, J. L. (2015). "A Brief History of Marchantia from Greece to Genomics." Plant and Cell Physiology **57**(2): 210-229.
- Bowman, J. L. (2022). "The liverwort *Marchantia polymorpha*, a model for all ages." Curr. Top. Dev. Biol **147**: 1-32.
- Bowman, J. L., T. Araki and T. Kohchi (2016). "Marchantia: Past, Present and Future." Plant Cell Physiol **57**(2): 205-209.
- Bowman, J. L., M. Arteaga-Vazquez, F. Berger, L. N. Briginshaw, P. Carella, A. Aguilar-Cruz, K. M. Davies, T. Dierschke, L. Dolan, A. E. Dorantes-Acosta, T. J. Fisher, E. Flores-Sandoval, K. Futagami, K. Ishizaki, R. Jibrán, T. Kanazawa, H. Kato, T. Kohchi, J. Levins, S. S. Lin, H. Nakagami, R. Nishihama, F. Romani, S. Schornack, Y. Tanizawa, M. Tsuzuki, T. Ueda, Y. Watanabe, K. T. Yamato and S. Zachgo (2022). "The renaissance and enlightenment of Marchantia as a model system." Plant Cell **34**(10): 3512-3542.
- Bresnick, E. H., K. R. Katsumura, H. Y. Lee, K. D. Johnson and A. S. Perkins (2012). "Master regulatory GATA transcription factors: mechanistic principles and emerging links to hematologic malignancies." Nucleic Acids Res **40**(13): 5819-5831.
- Brinkman, E. K., T. Chen, M. Amendola and B. Van Steensel (2014). "Easy quantitative assessment of genome editing by sequence trace decomposition." Nucleic acids research **42**(22): e168-e168.
- Buchfink, B., C. Xie and D. H. Huson (2015). "Fast and sensitive protein alignment using DIAMOND." Nature Methods **12**(1): 59-60.
- Buschmann, H., M. Holtmannspötter, A. Borchers, M. T. O'Donoghue and S. Zachgo (2016). "Microtubule dynamics of the centrosome-like polar organizers from the basal land plant *Marchantia polymorpha*." New Phytologist **209**(3): 999-1013.
- Casson, S. A. and A. M. Hetherington (2010). "Environmental regulation of stomatal development." Current opinion in plant biology **13**(1): 90-95.
- Cazzonelli, C. I. (2011). "Carotenoids in nature: insights from plants and beyond." Functional Plant Biology **38**(11): 833-847.
- Chang, L., E. Ramireddy and T. Schmölling (2015). "Cytokinin as a positional cue regulating lateral root spacing in Arabidopsis." Journal of Experimental Botany **66**(15): 4759-4768.
- Chen, F., A. Ludwiczuk, G. Wei, X. Chen, B. Crandall-Stotler and J. L. Bowman (2018). "Terpenoid Secondary Metabolites in Bryophytes: Chemical Diversity, Biosynthesis and Biological Functions." Critical Reviews in Plant Sciences **37**(2-3): 210-231.
- Chen, H., H. Shao, K. Li, D. Zhang, S. Fan, Y. Li and M. Han (2017). "Genome-wide identification, evolution, and expression analysis of GATA transcription factors in apple (*Malus domestica* Borkh.)." Gene **627**: 460-472.

- Chiang, Y. H., Y. O. Zubo, W. Tapken, H. J. Kim, A. M. Lavanway, L. Howard, M. Pilon, J. J. Kieber and G. E. Schaller (2012). "Functional characterization of the GATA transcription factors GNC and CGA1 reveals their key role in chloroplast development, growth, and division in Arabidopsis." Plant Physiol **160**(1): 332-348.
- Chickarmane, V. S., S. P. Gordon, P. T. Tarr, M. G. Heisler and E. M. Meyerowitz (2012). "Cytokinin signaling as a positional cue for patterning the apical–basal axis of the growing Arabidopsis shoot meristem." Proceedings of the National Academy of Sciences **109**(10): 4002-4007.
- Choudhury, N. K. and R. K. Behera (2001). "Photoinhibition of Photosynthesis: Role of Carotenoids in Photoprotection of Chloroplast Constituents." Photosynthetica **39**(4): 481-488.
- Collonnier, C., A. Epert, K. Mara, F. Maclot, A. Guyon-Debast, F. Charlot, C. White, D. G. Schaefer and F. Nogu  (2017). "CRISPR-Cas9-mediated efficient directed mutagenesis and RAD 51-dependent and RAD 51-independent gene targeting in the moss *Physcomitrella patens*." Plant Biotechnology Journal **15**(1): 122-131.
- Coudert, Y., O. Nov k and C. J. Harrison (2019). "A KNOX-Cytokinin Regulatory Module Predates the Origin of Indeterminate Vascular Plants." Curr Biol **29**(16): 2743-2750 e2745.
- Cove, D. (2005). "The moss *Physcomitrella patens*." Annu Rev Genet **39**: 339-358.
- Cove, D. J., P. F. Perroud, A. J. Charron, S. F. McDaniel, A. Khandelwal and R. S. Quatrano (2009). "Culturing the moss *Physcomitrella patens*." Cold Spring Harb Protoc **2009**(2): pdb prot5136.
- Cuttriss, A., C. Cazzonelli, E. Wurtzel and B. Pogson (2011). "Carotenoids." Advances in Botanical Research **58**: 1-36.
- De Castro, E., C. J. Sigrist, A. Gattiker, V. Bulliard, P. S. Langendijk-Genevaux, E. Gasteiger, A. Bairoch and N. Hulo (2006). "ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins." Nucleic acids research **34**(suppl\_2): W362-W365.
- De Hoon, M. J., S. Imoto, J. Nolan and S. Miyano (2004). "Open source clustering software." Bioinformatics **20**(9): 1453-1454.
- De Rybel, B., M. Adibi, A. S. Breda, J. R. Wendrich, M. E. Smit, O. Nov k, N. Yamaguchi, S. Yoshida, G. Van Isterdael and J. Palovaara (2014). "Integration of growth and patterning during vascular tissue formation in Arabidopsis." Science **345**(6197): 1255215.
- de Sousa, F., P. G. Foster, P. C. Donoghue, H. Schneider and C. J. Cox (2019). "Nuclear protein phylogenies support the monophyly of the three bryophyte groups (Bryophyta Schimp.)." New Phytologist **222**(1): 565-575.
- de Vries, J. and J. M. Archibald (2018). "Plant evolution: landmarks on the path to terrestrial life." New Phytol **217**(4): 1428-1434.
- Ding, L., S. Yan, L. Jiang, W. Zhao, K. Ning, J. Zhao, X. Liu, J. Zhang, Q. Wang and X. Zhang (2015). "HANABA TARANU (HAN) bridges meristem and organ primordia boundaries through PINHEAD, JAGGED, BLADE-ON-PETIOLE2 and CYTOKININ OXIDASE 3 during flower development in Arabidopsis." PLoS genetics **11**(9): e1005479.
- Duan, Z., Y. Zhang, J. Tu, J. Shen, B. Yi, T. Fu, C. Dai and C. Ma (2020). "The *Brassica napus* GATA transcription factor BnA5.ZML1 is a stigma compatibility factor." Journal of Integrative Plant Biology **62**(8): 1112-1131.
- Duckett, J. G., R. Ligrone, K. S. Renzaglia and S. Pressel (2014). "Pegged and smooth rhizoids in complex thalloid liverworts (Marchantiopsida): structure, function and evolution." Botanical Journal of the Linnean Society **174**(1): 68-92.
- Eberhard, S., G. Finazzi and F.-A. Wollman (2008). "The dynamics of photosynthesis." Annual review of genetics **42**: 463-515.

- Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Research **32**(5): 1792-1797.
- Emms, D. and S. Kelly (2018). "OrthoFinder2: fast and accurate phylogenomic orthology analysis from gene sequences." BioRxiv **466201**.
- Endo, H., M. Yamaguchi, T. Tamura, Y. Nakano, N. Nishikubo, A. Yoneda, K. Kato, M. Kubo, S. Kajita and Y. Katayama (2015). "Multiple classes of transcription factors regulate the expression of VASCULAR-RELATED NAC-DOMAIN7, a master switch of xylem vessel differentiation." Plant and Cell Physiology **56**(2): 242-254.
- Evrard, A., M. Kumar, D. Lecourieux, J. Lucks, P. von Koskull-Döring and H. Hirt (2013). "Regulation of the heat stress response in Arabidopsis by MPK6-targeted phosphorylation of the heat stress factor HsfA2." PeerJ **1**: e59.
- Ferreira, K. N., T. M. Iverson, K. Maghlaoui, J. Barber and S. Iwata (2004). "Architecture of the photosynthetic oxygen-evolving center." Science **303**(5665): 1831-1838.
- Ferreyra, M. L. F., P. Serra and P. Casati (2021). "Recent advances on the roles of flavonoids as plant protective molecules after UV and high light exposure." Physiologia Plantarum **173**(3): 736-749.
- Flores-Sandoval, E., T. Dierschke, T. J. Fisher and J. L. Bowman (2016). "Efficient and Inducible Use of Artificial MicroRNAs in *Marchantia polymorpha*." Plant Cell Physiol **57**(2): 281-290.
- Fujiwara, Y., C. P. Browne, K. Cunniff, S. C. Goff and S. H. Orkin (1996). "Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1." Proc Natl Acad Sci U S A **93**.
- Fürst-Jansen, J. M. R., S. de Vries and J. de Vries (2020). "Evo-physio: on stress responses and the earliest land plants." Journal of Experimental Botany **71**(11): 3254-3269.
- Gajdošová, S., L. Spíchal, M. Kamínek, K. Hoyerová, O. Novák, P. I. Dobrev, P. Galuszka, P. Klíma, A. Gaudinová and E. Žižková (2011). "Distribution, biological activities, metabolism, and the conceivable function of cis-zeatin-type cytokinins in plants." Journal of Experimental Botany **62**(8): 2827-2840.
- Galuszka, P., H. Popelková, T. Werner, J. Frébortová, H. Pospíšilová, V. Mik, I. Köllmer, T. Schmölling and I. Frébort (2007). "Biochemical characterization of cytokinin oxidases/dehydrogenases from *Arabidopsis thaliana* expressed in *Nicotiana tabacum* L." Journal of Plant Growth Regulation **26**: 255-267.
- Goodstein, D. M., S. Shu, R. Howson, R. Neupane, R. D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam and D. S. Rokhsar (2012). "Phytozome: a comparative platform for green plant genomics." Nucleic Acids Res **40**(Database issue): D1178-1186.
- Green, T. and W. Snelgar (1982). "A comparison of photosynthesis in two thalloid liverworts." Oecologia **54**: 275-280.
- Grellet Bournonville, C. F. and J. C. Díaz-Ricci (2011). "Quantitative determination of superoxide in plant leaves using a modified NBT staining method." Phytochemical Analysis **22**(3): 268-271.
- Guindon, S., J. Dufayard, V. Lefort, M. Anisimova, W. Hordijk and O. Gascuel (2010). "PhyML 3.0: new algorithms, methods and utilities." Systematic Biology **59**(3): 307-321.
- Gundry, C. N., J. G. Vandersteen, G. H. Reed, R. J. Pryor, J. Chen and C. T. Wittwer (2003). "Amplicon melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes." Clinical chemistry **49**(3): 396-406.
- Gupta, P., K. K. Nutan, S. L. Singla-Pareek and A. Pareek (2017). "Abiotic Stresses Cause Differential Regulation of Alternative Splice Forms of GATA Transcription Factor in Rice." Front Plant Sci **8**: 1944.
- Han, X., X. Chang, Z. Zhang, H. Chen, H. He, B. Zhong and X. W. Deng (2019). "Origin and Evolution of Core Components Responsible for Monitoring Light Environment Changes during Plant Terrestrialization." Molecular Plant **12**(6): 847-862.

- Harris, B. J., C. J. Harrison, A. M. Hetherington and T. A. Williams (2020). "Phylogenomic evidence for the monophyly of bryophytes and the reductive evolution of stomata." Current Biology **30**(11): 2001-2012. e2002.
- Hasan, S. S., E. Yamashita, D. Baniulis and W. A. Cramer (2013). "Quinone-dependent proton transfer pathways in the photosynthetic cytochrome b6/f complex." Proc Natl Acad Sci U S A **110**(11): 4297-4302.
- Herrmann, R. G. (1999). "Biogenesis and evolution of photosynthetic (thylakoid) membranes." Bioscience reports **19**: 355-365.
- Hiss, M., R. Meyberg, J. Westermann, F. B. Haas, L. Schneider, M. Schallenberg-Rudinger, K. K. Ullrich and S. A. Rensing (2017). "Sexual reproduction, sporophyte development and molecular variation in the model moss *Physcomitrella patens*: introducing the ecotype Reute." Plant J **90**(3): 606-620.
- Hongo, S., K. Sato, R. Yokoyama and K. Nishitani (2012). "Demethylesterification of the primary wall by PECTIN METHYLESTERASE35 provides mechanical support to the Arabidopsis stem." Plant Cell **24**(6): 2624-2634.
- Hothorn, M., T. Dabi and J. Chory (2011). "Structural basis for cytokinin recognition by *Arabidopsis thaliana* histidine kinase 4." Nature chemical biology **7**(11): 766-768.
- Hu, X., A. Tanaka and R. Tanaka (2013). "Simple extraction methods that prevent the artifactual conversion of chlorophyll to chlorophyllide during pigment isolation from leaf samples." Plant Methods **9**(1): 19.
- Huang, J., X. Zhao and J. Chory (2019). "The Arabidopsis Transcriptome Responds Specifically and Dynamically to High Light Stress." Cell Rep **29**(12): 4186-4199 e4183.
- Hudson, D., D. R. Guevara, A. J. Hand, Z. Xu, L. Hao, X. Chen, T. Zhu, Y. M. Bi and S. J. Rothstein (2013). "Rice cytokinin GATA transcription Factor1 regulates chloroplast development and plant architecture." Plant Physiol **162**(1): 132-144.
- Hutin, C., L. Nussaume, N. Moise, I. Moya, K. Kloppstech and M. Havaux (2003). "Early light-induced proteins protect Arabidopsis from photooxidative stress." PNAS **100**(8): 4921-4926.
- Hwang, I. and J. Sheen (2001). "Two-component circuitry in Arabidopsis cytokinin signal transduction." Nature **413**(6854): 383-389.
- Ioio, R. D., C. Galinha, A. G. Fletcher, S. P. Grigg, A. Molnar, V. Willemsen, B. Scheres, S. Sabatini, D. Baulcombe and P. K. Maini (2012). "A PHABULOSA/cytokinin feedback loop controls root growth in Arabidopsis." Current Biology **22**(18): 1699-1704.
- Ishizaki, K., M. Mizutani, M. Shimamura, A. Masuda, R. Nishihama and T. Kohchi (2013). "Essential Role of the E3 Ubiquitin Ligase NOPPERABO1 in Schizogenous Intercellular Space Formation in the Liverwort *Marchantia polymorpha*." The Plant Cell **25**(10): 4075-4084.
- Ishizaki, K., R. Nishihama, K. T. Yamato and T. Kohchi (2016). "Molecular Genetic Tools and Techniques for *Marchantia polymorpha* Research." Plant Cell Physiol **57**(2): 262-270.
- Jahns, P. and A. R. Holzwarth (2012). "The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II." Biochimica et Biophysica Acta (BBA) - Bioenergetics **1817**(1): 182-193.
- Jiang, K., V. Yung, T. Chiba and L. J. Feldman (2018). "Longitudinal patterning in roots: A GATA2–auxin interaction underlies and maintains the root transition domain." Planta **247**: 831-843.
- Jones, P., D. Binns, H.-Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell and G. Nuka (2014). "InterProScan 5: genome-scale protein function classification." Bioinformatics **30**(9): 1236-1240.
- Jones, R., H. Ougham, H. Thomas and S. Waaland (2012). Molecular life of plants, Wiley-Blackwell.

- Jung, H. S., P. A. Crisp, G. M. Estavillo, B. Cole, F. Hong, T. C. Mockler, B. J. Pogson and J. Chory (2013). "Subset of heat-shock transcription factors required for the early response of Arabidopsis to excess light." Proc Natl Acad Sci U S A **110**(35): 14474-14479.
- Juszczak, I. and M. Baier (2014). "Quantification of superoxide and hydrogen peroxide in leaves." Plant Cold Acclimation: Methods and Protocols: 217-224.
- Kamerling, Z. (1897). Zur Biologie und Physiologie der Marchantiaceen, Druck von V. Höfling.
- Kamisugi, Y., K. Schlink, S. A. Rensing, G. Schween, M. von Stackelberg, A. C. Cuming, R. Reski and D. J. Cove (2006). "The mechanism of gene targeting in *Physcomitrella patens*: homologous recombination, concatenation and multiple integration." Nucleic Acids Research **34**(21): 6205-6214.
- Kanei, M., G. Horiguchi and H. Tsukaya (2012). "Stable establishment of cotyledon identity during embryogenesis in Arabidopsis by ANGUSTIFOLIA3 and HANABA TARANU." Development **139**(13): 2436-2446.
- Kasahara, H., K. Takei, N. Ueda, S. Hishiyama, T. Yamaya, Y. Kamiya, S. Yamaguchi and H. Sakakibara (2004). "Distinct isoprenoid origins of cis- and trans-zeatin biosyntheses in Arabidopsis." Journal of Biological Chemistry **279**(14): 14049-14054.
- Kato, H., Y. Yasui and K. Ishizaki (2020). "Gemma cup and gemma development in *Marchantia polymorpha*." New Phytol **228**(2): 459-465.
- Keller, T., J. Abbott, T. Moritz and P. Doerner (2006). "Arabidopsis REGULATOR OF AXILLARY MERISTEMS1 controls a leaf axil stem cell niche and modulates vegetative development." The Plant Cell **18**(3): 598-611.
- Kieber, J. J. and G. E. Schaller (2014). "Cytokinins." Arabidopsis Book **12**: e0168.
- Kieber, J. J. and G. E. Schaller (2018). "Cytokinin signaling in plant development." Development **145**(4): dev149344.
- Kim, B. H. and G. M. Gadd (2008). Photosynthesis. Bacterial Physiology and Metabolism. Cambridge, Cambridge University Press: 386-407.
- Kleine, T., P. Kindgren, C. Benedict, L. Hendrickson and A. Strand (2007). "Genome-wide gene expression analysis reveals a critical role for CRYPTOCHROME1 in the response of Arabidopsis to high irradiance." Plant Physiol **144**(3): 1391-1406.
- Klermund, C., Q. L. Ranftl, J. Diener, E. Bastakis, R. Richter and C. Schwechheimer (2016). "LLM-Domain B-GATA Transcription Factors Promote Stomatal Development Downstream of Light Signaling Pathways in *Arabidopsis thaliana* Hypocotyls." Plant Cell **28**(3): 646-660.
- Kny, L. (1890). Bau und Entwicklung von *Marchantia Polymorpha* L., editor unknown.
- Koizumi, K., R. Yokoyama and K. Nishitani (2009). "Mechanical load induces upregulation of transcripts for a set of genes implicated in secondary wall formation in the supporting tissue of *Arabidopsis thaliana*." Journal of Plant Research **122**: 651-659.
- Kowalska, M., P. Galuszka, J. Frébortová, M. Šebela, T. Béres, T. Hluska, M. Šmehilová, K. D. Bilyeu and I. Frébort (2010). "Vacuolar and cytosolic cytokinin dehydrogenases of *Arabidopsis thaliana*: heterologous expression, purification and properties." Phytochemistry **71**(17-18): 1970-1978.
- Kurakawa, T., N. Ueda, M. Maekawa, K. Kobayashi, M. Kojima, Y. Nagato, H. Sakakibara and J. Kyojuka (2007). "Direct control of shoot meristem activity by a cytokinin-activating enzyme." Nature **445**(7128): 652-655.
- Kuroha, T., H. Tokunaga, M. Kojima, N. Ueda, T. Ishida, S. Nagawa, H. Fukuda, K. Sugimoto and H. Sakakibara (2009). "Functional analyses of LONELY GUY cytokinin-activating enzymes reveal the importance of the direct activation pathway in Arabidopsis." The Plant Cell **21**(10): 3152-3169.
- Laibach, F. (1943). "*Arabidopsis thaliana* (L.) Heynh. als Objekt für genetische und entwicklungsphysiologische Untersuchungen." Bot Arch **44**: 439-455.

- Latowski, D., P. Kuczynska and K. Strzalka (2011). "Xanthophyll cycle - a mechanism protecting plants against oxidative stress." Redox Report **16**(2): 78-90.
- Lefort, V., J.-E. Longueville and O. Gascuel (2017). "SMS: Smart Model Selection in PhyML." Molecular Biology and Evolution **34**(9): 2422-2424.
- Li, F.-W., P. Brouwer, L. Carretero-Paulet, S. Cheng, J. de Vries, P.-M. Delaux, A. Eily, N. Koppers, L.-Y. Kuo, Z. Li, M. Simenc, I. Small, E. Wafula, S. Angarita, M. S. Barker, A. Bräutigam, C. dePamphilis, S. Gould, P. S. Hosmani, Y.-M. Huang, B. Huettel, Y. Kato, X. Liu, S. Maere, R. McDowell, L. A. Mueller, K. G. J. Nierop, S. A. Rensing, T. Robison, C. J. Rothfels, E. M. Sigel, Y. Song, P. R. Timilsena, Y. Van de Peer, H. Wang, P. K. I. Wilhelmsson, P. G. Wolf, X. Xu, J. P. Der, H. Schluempmann, G. K. S. Wong and K. M. Pryer (2018). "Fern genomes elucidate land plant evolution and cyanobacterial symbioses." Nature Plants **4**(7): 460-472.
- Li, F.-W., T. Nishiyama, M. Waller, E. Frangedakis, J. Keller, Z. Li, N. Fernandez-Pozo, M. S. Barker, T. Bennett and M. A. Blázquez (2020). "Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts." Nature Plants **6**(3): 259-272.
- Li, F. W. and A. Harkess (2018). "A guide to sequence your favorite plant genomes." Appl Plant Sci **6**(3): e1030.
- Li, G., H. Gao, B. Zhao, S. Dong, J. Zhang, J. Yang, J. Wang and P. Liu (2009). "Effects of drought stress on activity of photosystems in leaves of maize at grain filling stage." Acta Agronomica Sinica **35**(10): 1916-1922.
- Li, S., X. Zhou, L. Chen, W. Huang and D. Yu (2010). "Functional characterization of *Arabidopsis thaliana* WRKY39 in heat stress." Molecules & Cells (Springer Science & Business Media BV) **29**(5).
- Lichtenthaler, H. K., C. Buschmann (2001). "Chlorophylls and carotenoids: Measurement and characterization by UV-VIS spectroscopy." Current protocols in food analytical chemistry. **1**(1): F4. 3.1-F4. 3.8
- Liu, P. P., N. Koizuka, R. C. Martin and H. Nonogaki (2005). "The BME3 (Blue Micropylar End 3) GATA zinc finger transcription factor is a positive regulator of Arabidopsis seed germination." The Plant Journal **44**(6): 960-971.
- Liu, Y., C. Zhang, J. Chen, L. Guo, X. Li, W. Li, Z. Yu, J. Deng, P. Zhang and K. Zhang (2013). "Arabidopsis heat shock factor HsfA1a directly senses heat stress, pH changes, and hydrogen peroxide via the engagement of redox state." Plant Physiology and Biochemistry **64**: 92-98.
- Liu, Z., H. Yan, K. Wang, T. Kuang, J. Zhang, L. Gui, X. An and W. Chang (2004). "Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution." Nature **428**(6980): 287-292.
- Llorente, B., J. F. Martinez-Garcia, C. Stange and M. Rodriguez-Concepcion (2017). "Illuminating colors: regulation of carotenoid biosynthesis and accumulation by light." Current Opinion in Plant Biology **37**: 49-55.
- Lomin, S. N., D. M. Krivosheev, M. Y. Steklov, D. V. Arkhipov, D. I. Osolodkin, T. Schmülling and G. A. Romanov (2015). "Plant membrane assays with cytokinin receptors underpin the unique role of free cytokinin bases as biologically active ligands." Journal of Experimental Botany **66**(7): 1851-1863.
- Lopez-Obando, M., B. Hoffmann, C. Gery, A. Guyon-Debast, E. Teoule, C. Rameau, S. Bonhomme and F. Nogue (2016). "Simple and Efficient Targeting of Multiple Genes Through CRISPR-Cas9 in *Physcomitrella patens*." G3 (Bethesda) **6**(11): 3647-3653.
- Luo, X.-M., W.-H. Lin, S. Zhu, J.-Y. Zhu, Y. Sun, X.-Y. Fan, M. Cheng, Y. Hao, E. Oh and M. Tian (2010). "Integration of light-and brassinosteroid-signaling pathways by a GATA transcription factor in Arabidopsis." Developmental cell **19**(6): 872-883.
- Mache, R. and S. Loiseaux (1973). "Light saturation of growth and photosynthesis of the shade plant *Marchantia polymorpha*." Journal of cell science **12**(2): 391-401.

- Mähönen, A. P., A. Bishopp, M. Higuchi, K. M. Nieminen, K. Kinoshita, K. Törmäkangas, Y. Ikeda, A. Oka, T. Kakimoto and Y. Helariutta (2006). "Cytokinin signaling and its inhibitor AHP6 regulate cell fate during vascular development." Science **311**(5757): 94-98.
- Manzoor, M. A., I. A. Sabir, I. H. Shah, H. Wang, Z. Yu, F. Rasool, M. Z. Mazhar, S. Younas, M. Abdullah and Y. Cai (2021). "Comprehensive Comparative Analysis of the GATA Transcription Factors in Four Rosaceae Species and Phytohormonal Response in Chinese Pear (*Pyrus bretschneideri*) Fruit." International Journal of Molecular Sciences **22**(22): 12492.
- Mara, C. D. I., Vivian F. (2008). "Two GATA Transcription Factors Are Downstream Effectors of Floral Homeotic Gene Action in Arabidopsis." Plant Physiology **147**(2): 707-718.
- McConaha, M. (1941). "Ventral structures effecting capillarity in the Marchantiales." American Journal of Botany: 301-306.
- Meyerowitz, E. M. (1987). "*Arabidopsis thaliana*." Annual Review of Genetics **21**(1): 93-111.
- Mi, H., D. Ebert, A. Muruganujan, C. Mills, L. P. Albu, T. Mushayamaha and P. D. Thomas (2021). "PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API." Nucleic Acids Res **49**(D1): D394-D403.
- Michelet, L., M. Zaffagnini, S. Morisse, F. Sparla, M. E. Pérez-Pérez, F. Francia, A. Danon, C. H. Marchand, S. Fermani and P. Trost (2013). "Redox regulation of the Calvin–Benson cycle: something old, something new." Frontiers in Plant Science **4**: 470.
- Micheli, F. (2001). "Pectin methylesterases: cell wall enzymes with important roles in plant physiology." Trends in Plant Science **6**(9): 414-419.
- Mistry, J., S. Chuguransky, L. Williams, M. Qureshi, G. A. Salazar, E. L. L. Sonnhammer, S. C. E. Tosatto, L. Paladin, S. Raj, L. J. Richardson, R. D. Finn and A. Bateman (2021). "Pfam: The protein families database in 2021." Nucleic Acids Res **49**(D1): D412-D419.
- Mitsuda, N., A. Iwase, H. Yamamoto, M. Yoshida, M. Seki, K. Shinozaki and M. Ohme-Takagi (2007). "NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis." Plant Cell **19**(1): 270-280.
- Montané, M.-H. and K. Kloppstech (2000). "The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPs): was the harvesting of light their primary function?" Gene **258**(1-2): 1-8.
- Moubayidin, L., S. Perilli, R. D. Ioio, R. Di Mambro, P. Costantino and S. Sabatini (2010). "The rate of cell differentiation controls the Arabidopsis root meristem growth phase." Current Biology **20**(12): 1138-1143.
- Naito, T., T. Kiba, N. Koizumi, T. Yamashino and T. Mizuno (2007). "Characterization of a unique GATA family gene that responds to both light and cytokinin in *Arabidopsis thaliana*." Biosci Biotechnol Biochem **71**(6): 1557-1560.
- Nawy, T., M. Bayer, J. Mravec, J. Friml, K. D. Birnbaum and W. Lukowitz (2010). "The GATA factor HANABA TARANU is required to position the proembryo boundary in the early Arabidopsis embryo." Developmental cell **19**(1): 103-113.
- Nelson, N. and A. Ben-Shem (2004). "The complex architecture of oxygenic photosynthesis." Nature Reviews Molecular Cell Biology **5**(12): 971-982.
- Neutelings, G. (2011). "Lignin variability in plant cell walls: Contribution of new models." Plant Science **181**(4): 379-386.
- Niyogi, K. K., A. R. Grossman and O. Björkman (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.
- Oliver, M. J., Z. Tuba and B. D. Mishler (2000). "The evolution of vegetative desiccation tolerance in land plants." Plant Ecology **151**(1): 85-100.

- Paajanen, P., G. Kettleborough, E. Lopez-Girona, M. Giolai, D. Heavens, D. Baker, A. Lister, F. Cugliandolo, G. Wilde, I. Hein, I. Macaulay, G. J. Bryan and M. D. Clark (2019). "A critical comparison of technologies for a plant genome sequencing project." Gigascience **8**(3).
- Pächt, O. (1950). "Early Italian nature studies and the early calendar landscape." Journal of the Warburg and Courtauld Institutes **13**(1-2): 13-47.
- Page, M., N. Sultana, K. Paszkiewicz, H. Florance and N. Smirnov (2012). "The influence of ascorbate on anthocyanin accumulation during high light acclimation in *Arabidopsis thaliana*: further evidence for redox control of anthocyanin synthesis." Plant, Cell & Environment **35**(2): 388-404.
- Pan, T., X. Sun, Y. Liu, H. Li, G. Deng, H. Lin and S. Wang (2018). "Heat stress alters genome-wide profiles of circular RNAs in Arabidopsis." Plant Molecular Biology **96**: 217-229.
- Pecinka, A., H. Q. Dinh, T. Baubec, M. Rosa, N. Lettner and O. M. Scheid (2010). "Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in Arabidopsis." The Plant Cell **22**(9): 3118-3129.
- Pelloux, J., C. Rusterucci and E. J. Mellerowicz (2007). "New insights into pectin methylesterase structure and function." Trends in Plant Science **12**(6): 267-277.
- Peng, W., W. Li, N. Song, Z. Tang, J. Liu, Y. Wang, S. Pan, L. Dai and B. Wang (2021). "Genome-Wide Characterization, Evolution, and Expression Profile Analysis of GATA Transcription Factors in *Brachypodium distachyon*." International Journal of Molecular Sciences **22**(4): 2026.
- Persson, B., B. Esberg, O. Olafsson and G. Björk (1994). "Synthesis and function of isopentenyl adenosine derivatives in tRNA." Biochimie **76**(12): 1152-1160.
- Pfaffl, M. W. (2001). "A new mathematical model for relative quantification in real-time RT-PCR." Nucleic Acids Research **29**(9): 2002 - 2007.
- Pfündel, E. E., M. Renganathan, A. M. Gilmore, H. Y. Yamamoto and R. A. Dilley (1994). "Intrathylakoid pH in isolated pea chloroplasts as probed by violaxanthin deepoxidation." Plant Physiology: 1647-1658.
- Pospíšil, P. (2016). "Production of reactive oxygen species by photosystem II as a response to light and temperature stress." Frontiers in plant science **7**: 1950.
- Pötter, E. and K. Kloppstech (1993). "Effects of light stress on the expression of early light-inducible proteins in barley." European Journal of Biochemistry **214**(3): 779-786.
- Puttick, M. N., J. L. Morris, T. A. Williams, C. J. Cox, D. Edwards, P. Kenrick, S. Pressel, C. H. Wellman, H. Schneider and D. Pisani (2018). "The interrelationships of land plants and the nature of the ancestral embryophyte." Current Biology **28**(5): 733-745. e732.
- Raines, C. A. (2003). "The Calvin cycle revisited." Photosynthesis Research **75**(1): 1-10.
- Ranftl, Q. L., E. Bastakis, C. Klermund and C. Schwechheimer (2016). "LLM-Domain Containing B-GATA Factors Control Different Aspects of Cytokinin-Regulated Development in *Arabidopsis thaliana*." Plant Physiol **170**(4): 2295-2311.
- Raven, J. A. and D. Edwards (2013). Photosynthesis in early land plants: adapting to the terrestrial environment. In: Hanson, D., Rice, S. (eds) Photosynthesis in Bryophytes and Early Land Plants. Advances in Photosynthesis and Respiration, vol 37. Springer, Dordrecht: 29-58.
- Rensing, S. A., B. Goffinet, R. Meyberg, S.-Z. Wu and M. Bezanilla (2020). "The Moss *Physcomitrium (Physcomitrella) patens*: A Model Organism for Non-Seed Plants[OPEN]." The Plant Cell **32**(5): 1361-1376.
- Rensing, S. A., J. Ick, J. A. Fawcett, D. Lang, A. Zimmer, Y. Van de Peer and R. Reski (2007). "An ancient genome duplication contributed to the abundance of metabolic genes in the moss *Physcomitrella patens*." BMC evolutionary biology **7**(1): 1-10.

- Rensing, S. A., D. Lang, A. D. Zimmer, A. Terry, A. Salamov, H. Shapiro, T. Nishiyama, P. F. Perroud, E. A. Lindquist, Y. Kamisugi, T. Tanahashi, K. Sakakibara, T. Fujita, K. Oishi, T. Shin-I, Y. Kuroki, A. Toyoda, Y. Suzuki, S. Hashimoto, K. Yamaguchi, S. Sugano, Y. Kohara, A. Fujiyama, A. Anterola, S. Aoki, N. Ashton, W. B. Barbazuk, E. Barker, J. L. Bennetzen, R. Blankenship, S. H. Cho, S. K. Dutcher, M. Estelle, J. A. Fawcett, H. Gundlach, K. Hanada, A. Heyl, K. A. Hicks, J. Hughes, M. Lohr, K. Mayer, A. Melkozernov, T. Murata, D. R. Nelson, B. Pils, M. Prigge, B. Reiss, T. Renner, S. Rombauts, P. J. Rushton, A. Sanderfoot, G. Schween, S. H. Shiu, K. Stueber, F. L. Theodoulou, H. Tu, Y. Van de Peer, P. J. Verrier, E. Waters, A. Wood, L. X. Yang, D. Cove, A. C. Cuming, M. Hasebe, S. Lucas, B. D. Mishler, R. Reski, I. V. Grigoriev, R. S. Quatrano and J. L. Boore (2008). "The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants." *Science* **319**(5859): 64-69.
- Reski, R. (1998). "Development, genetics and molecular biology of mosses." *Botanica Acta* **111**(1): 1-15.
- Reyes, J. C., M. I. Muro-Pastor and F. J. Florencio (2004). "The GATA family of transcription factors in *Arabidopsis* and rice." *Plant Physiol* **134**(4): 1718-1732.
- Rich, M. K. and P. M. Delaux (2020). "Plant Evolution: When *Arabidopsis* Is More Ancestral Than *Marchantia*." *Curr Biol* **30**(11): R642-R644.
- Richter, R., E. Bastakis and C. Schwechheimer (2013). "Cross-repressive interactions between SOC1 and the GATAs GNC and GNL/CGA1 in the control of greening, cold tolerance, and flowering time in *Arabidopsis*." *Plant Physiol* **162**(4): 1992-2004.
- Richter, R., C. Behringer, I. K. Muller and C. Schwechheimer (2010). "The GATA-type transcription factors GNC and GNL/CGA1 repress gibberellin signaling downstream from DELLA proteins and PHYTOCHROME-INTERACTING FACTORS." *Genes Dev* **24**(18): 2093-2104.
- Richter, R., C. Behringer, M. Zourelidou and C. Schwechheimer (2013). "Convergence of auxin and gibberellin signaling on the regulation of the GATA transcription factors GNC and GNL in *Arabidopsis thaliana*." *Proc Natl Acad Sci U S A* **110**(32): 13192-13197.
- Rizzini, L., J.-J. Favory, C. Cloix, D. Faggionato, A. O'Hara, E. Kaiserli, R. Baumeister, E. Schäfer, F. Nagy and G. I. Jenkins (2011). "Perception of UV-B by the *Arabidopsis* UVR8 protein." *Science* **332**(6025): 103-106.
- Rossel, J. B., I. W. Wilson and B. J. Pogson (2002). "Global changes in gene expression in response to high light in *Arabidopsis*." *Plant Physiol* **130**(3): 1109-1120.
- Rossini, S., A. P. Casazza, E. C. Engelmann, M. Havaux, R. C. Jennings and C. Soave (2006). "Suppression of both ELIP1 and ELIP2 in *Arabidopsis* does not affect tolerance to photoinhibition and photooxidative stress." *Plant Physiology* **141**(4): 1264-1273.
- Ruiz-Sola, M. A., D. Coman, G. Beck, M. V. Barja, M. Colinas, A. Graf, R. Welsch, P. Rutimann, P. Buhmann, L. Bigler, W. Gruissem, M. Rodriguez-Concepcion and E. Vranova (2016). "Arabidopsis GERANYLGERANYL DIPHOSPHATE SYNTHASE 11 is a hub isozyme required for the production of most photosynthesis-related isoprenoids." *New Phytol* **209**(1): 252-264.
- Sakai, H., T. Honma, T. Aoyama, S. Sato, T. Kato, S. Tabata and A. Oka (2001). "ARR1, a transcription factor for genes immediately responsive to cytokinins." *Science* **294**(5546): 1519-1521.
- Sakakibara, H. (2006). "Cytokinins: activity, biosynthesis, and translocation." *Annu. Rev. Plant Biol.* **57**: 431-449.
- Saldanha, A. J. (2004). "Java Treeview—extensible visualization of microarray data." *Bioinformatics* **20**(17): 3246-3248.
- Schaefer, D. G. and J. P. Zrýd (1997). "Efficient gene targeting in the moss *Physcomitrella patens*." *The Plant Journal* **11**(6): 1195-1206.
- Scheller, H. V. and P. Ulvskov (2010). "Hemicelluloses." *Annu Rev Plant Biol* **61**: 263-289.
- Schiffner, V. F. (1909). *Studien über die rhizoïden der Marchantiales.*

- Schneider, C. A., W. S. Rasband and K. W. Eliceiri (2012). "NIH Image to ImageJ: 25 years of image analysis." Nature methods **9**(7): 671-675.
- Schröder, P. M., Hsu, B.-Y., Gutsche, N., Winkler, J. B., Hedtke, B., Grimm, B., & Schwechheimer, C. (2023). "B-GATA factors are required to repress high-light stress responses in *Marchantia polymorpha* and *Arabidopsis thaliana*." Plant, Cell & Environment. (in press)
- Schwechheimer, C., P. M. Schröder and C. E. Blaby-Haas (2022). "Plant GATA Factors: Their Biology, Phylogeny, and Phylogenomics." Annu Rev Plant Biol **73**: 123-148.
- Schween, G., G. Gorr, A. Hohe and R. Reski (2003). "Unique tissue-specific cell cycle in *Physcomitrella*." Plant Biology **5**(01): 50-58.
- Shimamura, M. (2015). "*Marchantia polymorpha* : Taxonomy, Phylogeny and Morphology of a Model System." Plant and Cell Physiology **57**(2): 230-256.
- Spíchal, L., N. Y. Rakova, M. Riefler, T. Mizuno, G. A. Romanov, M. Strnad and T. Schmülling (2004). "Two cytokinin receptors of *Arabidopsis thaliana*, CRE1/AHK4 and AHK3, differ in their ligand specificity in a bacterial assay." Plant and Cell Physiology **45**(9): 1299-1305.
- Steyn, W., S. Wand, D. Holcroft and G. Jacobs (2002). "Anthocyanins in vegetative tissues: a proposed unified function in photoprotection." New Phytologist **155**(3): 349-361.
- Sugano, S. S., M. Shirakawa, J. Takagi, Y. Matsuda, T. Shimada, I. Hara-Nishimura and T. Kohchi (2014). "CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* L." Plant Cell Physiol **55**(3): 475-481.
- Takei, K., T. Yamaya and H. Sakakibara (2004). "*Arabidopsis* CYP735A1 and CYP735A2 encode cytokinin hydroxylases that catalyze the biosynthesis of trans-zeatin." Journal of Biological Chemistry **279**(40): 41866-41872.
- Tanabata, T., T. Shibaya, K. Hori, K. Ebana and M. Yano (2012). "SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis." Plant Physiology **160**(4): 1871-1880.
- Tanaka, A. and R. Tanaka (2006). "Chlorophyll metabolism." Curr Opin Plant Biol **9**(3): 248-255.
- Taniguchi, M., N. Sasaki, T. Tsuge, T. Aoyama and A. Oka (2007). "ARR1 directly activates cytokinin response genes that encode proteins with diverse regulatory functions." Plant Cell Physiol **48**(2): 263-277.
- Taylor-Teeples, M., L. Lin, M. de Lucas, G. Turco, T. W. Toal, A. Gaudinier, N. F. Young, G. M. Trabucco, M. T. Veling, R. Lamothe, P. P. Handakumbura, G. Xiong, C. Wang, J. Corwin, A. Tsoukalas, L. Zhang, D. Ware, M. Pauly, D. J. Kliebenstein, K. Dehesh, I. Tagkopoulos, G. Breton, J. L. Pruneda-Paz, S. E. Ahnert, S. A. Kay, S. P. Hazen and S. M. Brady (2015). "An *Arabidopsis* gene regulatory network for secondary cell wall synthesis." Nature **517**(7536): 571-575.
- The *Arabidopsis* Genome Initiative. (2000). "Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*." Nature **408**(6814): 796-815.
- Thelander, M., A. Nilsson, T. Olsson, M. Johansson, P.-A. Girod, D. G. Schaefer, J.-P. Zrýd and H. Ronne (2007). "The moss genes PpSKI1 and PpSKI2 encode nuclear SnRK1 interacting proteins with homologues in vascular plants." Plant Molecular Biology **64**: 559-573.
- Tian, F., D. C. Yang, Y. Q. Meng, J. Jin and G. Gao (2020). "PlantRegMap: charting functional regulatory maps in plants." Nucleic Acids Res **48**(D1): D1104-D1113.
- Tikkanen, M., N. R. Mekala and E.-M. Aro (2014). "Photosystem II photoinhibition-repair cycle protects Photosystem I from irreversible damage." Biochimica et Biophysica Acta (BBA)-Bioenergetics **1837**(1): 210-215.
- Titova, N. N. (1935). Sovietskaya Botanik.

- To, J. P., J. Deruere, B. B. Maxwell, V. F. Morris, C. E. Hutchison, F. J. Ferreira, G. E. Schaller and J. J. Kieber (2007). "Cytokinin regulates type-A Arabidopsis Response Regulator activity and protein stability via two-component phosphorelay." Plant Cell **19**(12): 3901-3914.
- To, J. P., G. Haberer, F. J. Ferreira, J. Deruere, M. G. Mason, G. E. Schaller, J. M. Alonso, J. R. Ecker and J. J. Kieber (2004). "Type-A Arabidopsis response regulators are partially redundant negative regulators of cytokinin signaling." Plant Cell **16**(3): 658-671.
- To, J. P. C. and J. J. Kieber (2008). "Cytokinin signaling: two-components and more." Trends in Plant Science **13**(2): 85-92.
- Tripathy, B. C. and G. K. Pattanayak (2012). "Chlorophyll biosynthesis in higher plants." Photosynthesis: plastid biology, energy conversion and carbon assimilation: 63-94.
- Tzvetkova-Chevolleau, T., F. Franck, A. E. Alawady, L. Dall'Osto, F. Carriere, R. Bassi, B. Grimm, L. Nussaume and M. Havaux (2007). "The light stress-induced protein ELIP2 is a regulator of chlorophyll synthesis in *Arabidopsis thaliana*." Plant J **50**(5): 795-809.
- Van Bel, M., T. Diels, E. Vancaester, L. Kreft, A. Botzki, Y. Van de Peer, F. Coppens and K. Vandepoele (2018). "PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics." Nucleic Acids Res **46**(D1): D1190-D1196.
- Vogel, J. T., M. H. Walter, P. Giavalisco, A. Lytovchenko, W. Kohlen, T. Charnikhova, A. J. Simkin, C. Goulet, D. Strack and H. J. Bouwmeester (2010). "SICCD7 controls strigolactone biosynthesis, shoot branching and mycorrhiza-induced apocarotenoid formation in tomato." The Plant Journal **61**(2): 300-311.
- Walter, M. H., D. S. Floss and D. Strack (2010). "Apocarotenoids: hormones, mycorrhizal metabolites and aroma volatiles." Planta **232**: 1-17.
- Wang, D., Y. Fujiyoshi and W. Kuhlbrandt (1994). An atomic model of plant LIGHT-HARVESTING COMPLEX determined by electron crystallography. Biophysical Journal, Biophysical Society 9650 Rockville Pike, Bethesda, MD 20814-3998.
- Wang, T., Y. Yang, S. Lou, W. Wei, Z. Zhao, Y. Ren, C. Lin and L. Ma (2019). "Genome-Wide Characterization and Gene Expression Analyses of GATA Transcription Factors in Moso Bamboo (*Phyllostachys edulis*)." International Journal of Molecular Sciences **21**(1): 14.
- Wang, Z. P., H. L. Xing, L. Dong, H. Y. Zhang, C. Y. Han, X. C. Wang and Q. J. Chen (2015). "Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation." Genome Biol **16**: 144.
- Waters, M. T., P. Wang, M. Korkaric, R. G. Capper, N. J. Saunders and J. A. Langdale (2009). "GLK Transcription Factors Coordinate Expression of the Photosynthetic Apparatus in Arabidopsis." The Plant Cell **21**(4): 1109-1128.
- Werner, T., I. Köllmer, I. Bartrina, K. Holst and T. Schmölling (2006). "New insights into the biology of cytokinin degradation." Plant Biology **8**(03): 371-381.
- Werner, T., V. Motyka, M. Strnad and T. Schmölling (2001). "Regulation of plant growth by cytokinin." Proceedings of the National Academy of Sciences **98**(18): 10487-10492.
- Werner, T., E. Nehnevajova, I. Köllmer, O. Novák, M. Strnad, U. Krämer and T. Schmölling (2010). "Root-specific reduction of cytokinin causes enhanced root growth, drought tolerance, and leaf mineral enrichment in Arabidopsis and tobacco." The Plant Cell **22**(12): 3905-3920.
- Wickett, N. J., S. Mirarab, N. Nguyen, T. Warnow, E. Carpenter, N. Matasci, S. Ayyampalayam, M. S. Barker, J. G. Burleigh, M. A. Gitzendanner, B. R. Ruhfel, E. Wafula, J. P. Der, S. W. Graham, S. Mathews, M. Melkonian, D. E. Soltis, P. S. Soltis, N. W. Miles, C. J. Rothfels, L. Pokorny, A. J. Shaw, L. DeGironimo, D. W. Stevenson, B. Surek, J. C. Villarreal, B. Roure, H. Philippe, C. W. dePamphilis, T. Chen, M. K. Deyholos, R. S. Baucom, T. M. Kutchan, M. M. Augustin, J. Wang, Y. Zhang, Z. Tian, Z.

- Yan, X. Wu, X. Sun, G. K. Wong and J. Leebens-Mack (2014). "Phylotranscriptomic analysis of the origin and early diversification of land plants." Proc Natl Acad Sci U S A **111**(45): E4859-4868.
- Witte, C.-P. and M. Herde (2020). "Nucleotide metabolism in plants." Plant Physiology **182**(1): 63-78.
- Wolf, S., G. Mouille and J. Pelloux (2009). "Homogalacturonan methyl-esterification and plant development." Molecular plant **2**(5): 851-860.
- Woodward, A. W. and B. Bartel (2018). "Biology in Bloom: A Primer on the *Arabidopsis thaliana* Model System." Genetics **208**(4): 1337-1349.
- Yamaoka, S., K. Inoue and T. Araki (2021). "Regulation of gametangia and gametangiophore initiation in the liverwort *Marchantia polymorpha*." Plant Reprod **34**(4): 297-306.
- Yasui, Y., S. Tsukamoto, T. Sugaya, R. Nishihama, Q. Wang, H. Kato, K. T. Yamato, H. Fukaki, T. Mimura, H. Kubo, K. Theres, T. Kohchi and K. Ishizaki (2019). "GEMMA CUP-ASSOCIATED MYB1, an Ortholog of Axillary Meristem Regulators, Is Essential in Vegetative Reproduction in *Marchantia polymorpha*." Curr Biol **29**(23): 3987-3995 e3985.
- Yeh, S. Y., H. W. Chen, C. Y. Ng, C. Y. Lin, T. H. Tseng, W. H. Li and M. S. Ku (2015). "Down-Regulation of Cytokinin Oxidase 2 Expression Increases Tiller Number and Improves Rice Yield." Rice (N Y) **8**(1): 36.
- Yokoyama, R. and K. Nishitani (2006). "Identification and characterization of *Arabidopsis thaliana* genes involved in xylem secondary cell walls." Journal of Plant Research **119**: 189-194.
- Zalabák, D., P. Galuszka, K. Mrízová, K. Podlešáková, R. Gu and J. Frébortová (2014). "Biochemical characterization of the maize cytokinin dehydrogenase family and cytokinin profiling in developing maize plantlets in relation to the expression of cytokinin dehydrogenase genes." Plant Physiology and Biochemistry **74**: 283-293.
- Zhang, C., Y. Huang, Z. Xiao, H. Yang, Q. Hao, S. Yuan, H. Chen, L. Chen, S. Chen, X. Zhou and W. Huang (2020). "A GATA Transcription Factor from Soybean (*Glycine max*) Regulates Chlorophyll Biosynthesis and Suppresses Growth in the Transgenic *Arabidopsis thaliana*." Plants **9**(8): 1036.
- Zhang, X., Y. Zhou, L. Ding, Z. Wu, R. Liu and E. M. Meyerowitz (2013). "Transcription repressor HANABA TARANU controls flower development by integrating the actions of multiple hormones, floral organ specification genes, and GATA3 family genes in Arabidopsis." Plant Cell **25**(1): 83-101.
- Zhang, Z., X. Zou, Z. Huang, S. Fan, G. Qun, A. Liu, J. Gong, J. Li, W. Gong, Y. Shi, L. Fan, Z. Zhang, R. Liu, X. Jiang, K. Lei, H. Shang, A. Xu and Y. Yuan (2019). "Genome-wide identification and analysis of the evolution and expression patterns of the GATA transcription factors in three species of Gossypium genus." Gene **680**: 72-83.
- Zhao, Y., L. Medrano, K. Ohashi, J. C. Fletcher, H. Yu, H. Sakai and E. M. Meyerowitz (2004). "HANABA TARANU is a GATA transcription factor that regulates shoot apical meristem and flower development in Arabidopsis." Plant Cell **16**(10): 2586-2600.
- Zhu, H., H. Zhai, S. He, H. Zhang, S. Gao and Q. Liu (2022). "A novel sweetpotato GATA transcription factor, IbGATA24, interacting with IbCOP9-5a positively regulates drought and salt tolerance." Environmental and Experimental Botany **194**: 104735.
- Zubo, Y. O., I. C. Blakley, J. M. Franco-Zorrilla, M. V. Yamburenko, R. Solano, J. J. Kieber, A. E. Loraine and G. E. Schaller (2018). "Coordination of Chloroplast Development through the Action of the GNC and GLK Transcription Factor Families." Plant Physiol **178**(1): 130-147.