Received: 7 February 2023 | Accepted: 15 May 2023

DOI: 10.1111/pce.14629

ORIGINAL ARTICLE

Plant, Cell & Environment PC&E WILEY

B-GATA factors are required to repress high-light stress responses in *Marchantia polymorpha* and *Arabidopsis thaliana*

Peter Schröder[1] | Bang-Yu Hsu[1] | Nora Gutsche[2] | Jana Barbro Winkler[3] | Boris Hedtke[4] | Bernhard Grimm[4] | Claus Schwechheimer[1]

[1]Plant Systems Biology, School of Life Sciences, Technical University of Munich, Freising, Germany

[2]Department of Botany, Osnabrück University, Osnabrück, Germany

[3]Research Unit Environmental Simulation, Institute of Biochemical Plant Pathology, Helmholtz Zentrum München, Neuherberg, Germany

[4]Department of Plant Physiology, Humboldt University Berlin, Berlin, Germany

Correspondence
Claus Schwechheimer, Plant Systems Biology, School of Life Sciences, Technical University of Munich, Emil-Ramann-Strasse 8, 85354 Freising, Germany.
Email: claus.schwechheimer@tum.de

Funding information
Deutsche Forschungsgemeinschaft

Abstract

GATAs are evolutionarily conserved zinc-finger transcription factors from eukaryotes. In plants, GATAs can be subdivided into four classes, A–D, based on their DNA-binding domain, and into further subclasses based on additional protein motifs. B-GATAs with a so-called leucine-leucine-methionine (LLM)-domain can already be found in algae. In angiosperms, the B-GATA family is expanded and can be subdivided in to LLM- or HAN-domain B-GATAs. Both, the LLM- and the HAN-domain are conserved domains of unknown biochemical function. Interestingly, the B-GATA family in the liverwort *Marchantia polymorpha* and the moss *Physcomitrium patens* is restricted to one and four family members, respectively. And, in contrast to vascular plants, the bryophyte B-GATAs contain a HAN- as well as an LLM-domain. Here, we characterise mutants of the single B-GATA from *Marchantia polymorpha*. We reveal that this mutant has defects in thallus growth and in gemma formation. Transcriptomic studies uncover that the B-GATA mutant displays a constitutive high-light (HL) stress response, a phenotype that we then also confirm in mutants of *Arabidopsis thaliana* LLM-domain B-GATAs, suggesting that the B-GATAs have a protective role towards HL stress.

KEYWORDS

Arabidopsis, *Arabidopsis thaliana*, chlorophyll, EARLY LIGHT-INDUCED PROTEIN, ELIP, GATA, high-light stress, Marchantia, *Marchantia polymorpha*, photosystem

1 | INTRODUCTION

During evolution on terrestrial habitats, the algal ancestors of extant land plants had to adjust to a broad range of new environmental stimuli and abiotic stresses. Besides the evolution of water-conducting tissues and strategies for desiccation tolerance, land plants had to adapt to the exposure to high sun light intensities (Trojak & Skowron, 2017). While sun light provides photosynthetic energy and important cues for plant growth, high-intensity sun light damages the photosynthetic machinery and has detrimental effects on plant survival (Demmig-Adams & Adams, 1992). This is caused by singlet oxygen, a form of reactive oxygen species (ROS) that is produced in high-light (HL) by energy transfer from triplet chlorophyll to molecular oxygen (Eberhard et al., 2008; Li et al., 2009; Pospisil, 2016). While repair mechanisms for photosystem II exist, photodamaged photosystem I requires resynthesis of the entire supercomplex to regain functionality (Alboresi et al., 2010; Tikkanen et al., 2014). To protect themselves against high-intensity sun light,

land plants developed several strategies. These include the synthesis of anthocyanins and flavonoids, important photoprotectors (Landry et al., 1995). Another important response to high light intensities is non-photochemical quenching, a mechanism for the dissipation of excess light energy of singlet chlorophyll in the form of heat to prevent ROS-related damage (Alboresi et al., 2010). Additionally, land plants produce photosynthesis-related complexes, like light-harvesting complexes and photosynthesis components, while reducing chlorophyll biosynthesis and promoting chlorophyll degradation to protect the photosynthetic machinery (Huang et al., 2019). EARLY LIGHT-INDUCED PROTEINS (ELIPs) are a group of light harvesting complex-related chlorophyll AB-binding proteins (Harari-Steinberg et al., 2001; Hutin et al., 2003). ELIPs are thought to act as scavengers for free chlorophyll molecules to allow the free energy transfer from excited chlorophyll to xanthophylls (Hutin et al., 2003; Tzvetkova-Chevolleau et al., 2007).

GATA factors are evolutionarily conserved zinc-finger transcription factors with essential biological functions in all eukaryotes examined to date (Schwechheimer et al., 2022). Among others, GATAs have been implicated in the development of the hematopoietic system in mammals or in nitrogen responses in fungi (Lentjes et al., 2016; Scazzocchio, 2000). When compared to the GATA families of animals and fungi, the GATA family in angiosperms is expanded but relatively little is still known about the biological roles of most of these angiosperm GATAs and their roles during plant evolution and terrestrialization (Reyes et al., 2004; Schwechheimer et al., 2022).

Plant GATA factors are subdivided into four major classes, A-, B-, C- and D-GATAs, and further subclasses can be defined based on the presence of additional domains (Reyes et al., 2004; Schwechheimer et al., 2022). We have previously described *Arabidopsis thaliana* B-class GATAs with an leucine-leucine-methionine (LLM) domain, a domain of unknown biochemical function found at the very C-terminus of B-GATAs (Bastakis et al., 2018; Klermund et al., 2016; Ranftl et al., 2016; Richter et al., 2010). GNC (GATA, NITRATE-INDUCIBLE, CARBON METABOLISM INVOLVED) and GNL/CGA1 (GNC-LIKE/CYTOKININ-RESPONSIVE GATA FACTOR1; from here GNL) are two well studied functionally redundant B-GATA family members (Bastakis et al., 2018; Bi et al., 2005; Klermund et al., 2016; Naito et al., 2007; Ranftl et al., 2016; Richter et al., 2010). The *gnc gnl* double mutant displays already comparatively strong defects, but these are enhanced when additional family members are mutated in higher order mutants of the LLM-domain B-GATAs (Bastakis et al., 2018; Klermund et al., 2016; Ranftl et al., 2016; Richter et al., 2010). The current findings suggest that the six *Arabidopsis thaliana* LLM-domain B-GATAs have very likely identical biochemical and redundant biological functions in the regulation of chlorophyll biosynthesis, stomata formation, phyllotaxis regulation, senescence and lateral branch angle control (Klermund et al., 2016; Ranftl et al., 2016).

HAN-domain B-GATAs represent the second subclass of B-GATAs in angiosperms (Nawy et al., 2010; Schwechheimer et al., 2022; Zhao et al., 2004). The conserved HAN-domain has no known biochemical function and was first described in the floral development regulator HANABA TARANU (Zhao et al., 2004). *HAN* gene mutants were identified in genetic analyses in different plant species. Arabidopsis *han* mutants have defects in maintaining boundaries between floral organs (Zhao et al., 2004). The loss of all three Arabidopsis HAN B-GATAs results in strong embryogenesis phenotypes associated with defects in auxin transport and distribution (Nawy et al., 2010). In rice, barley and maize, mutants of orthologous *HAN* genes develop bract leaves, whose outgrowth is suppressed in the respective reference cultivars (Whipple et al., 2010). The HAN- as well as the LLM-domains are specific for B-GATAs and for plants (Schwechheimer et al., 2022).

Phylogenomic analysis revealed that *Marchantia polymorpha* and *Physcomitrium patens* contain B-GATAs with HAN- as well as LLM-domains, which is a constellation unique to bryophytes (Schwechheimer et al., 2022). Here, we examine the molecular function of Mp*B-GATA1*, the single B-GATA orthologue from Marchantia (Schwechheimer et al., 2022). Through mutant analysis, we uncover roles of Mp*B-GATA1* in gemma cup formation and in HL stress responses. We then also find that proper HL stress responses in *Arabidopsis thaliana* are dependent on LLM-domain B-GATAs, suggesting that the protection from HL stress represents an evolutionarily conserved role of B-GATAs.

2 | MATERIALS AND METHODS

2.1 | Biological material

Marchantia polymorpha experiments were carried out with the ecotype BoGa (Buschmann et al., 2016). *Arabidopsis thaliana* ecotype Columbia (Col-0) was used in all experiments, as well as the following previously published mutants: *gnc* (SALK_001778) *gnl* (SALK_003995) (Richter et al., 2010), *gnc* (SALK_001778) *gnl* (SALK_003995) *gata17* (SALK_049041) *gata17l* (SALK_026798) (*gata quadruple*, *gata quad*) (Ranftl et al., 2016), GNLox (Richter et al., 2010).

2.2 | Phylogeny construction

GATA protein sequences from *Arabidopsis thaliana*, *Marchantia polymorpha*, *Physcomitrium patens*, *Selaginella moellendorffii*, *Salvinia cucullata* and *Chara braunii* were identified based on the presence of their GATA-type zinc-finger domains as defined by PF00320 using HMMER (hmmer.org) and BLAST searches (Altschul et al., 1990). Human GATAd2A was used as an outgroup. Zinc-finger sequences were aligned using MUSCLE (Edgar, 2004) and phylogeny construction was carried out with PhyML (Guindon 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 and used for phylogeny construction (Lefort et al., 2017). Branch support was calculated using the aLRT implementation in PhyML (Anisimova & Gascuel, 2006).

2.3 | Plant culture

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

For *Arabidopsis thaliana* growth, seeds were stratified for 3 days in the dark at 4°C, before being plated on ½ Murashige & Skoog medium containing 0.8% plant agar (Duchefa). Unless otherwise stated, plants were grown under 100 µmol m^{-2} s^{-1} white light (L58W/830, Lumiluxwarm white, 5200 lm, Osram) at 22°C in long day conditions with a 16-h light/8-h dark cycle.

2.4 | Thallus size quantification

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

2.5 | Air pore size quantification

Air pores size of *Marchantia 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 et al., 2012).

2.6 | Molecular cloning and CRISPR/Cas9-mediated mutagenesis

Mpb-gata1-1 and *Mpb-gata1-2* were obtained in the BoGa ecotype by CRISPR/Cas9-mutagenesis using two guide RNAs (gRNAs) targeting Mp*B-GATA1* (Mp7g03490). The transformation constructs were obtained by aligning complementary oligonucleotides, generating *BsaI* overhangs, and cloned into the Gateway entry vector pMpGE_En03 and subcloned into the Gateway™ (Thermo Fisher Scientific) system-compatible pMpGE010 and pMpGE011 (Table S1) (Sugano et al., 2014). For *Marchantia polymorpha* sporeling transformation, the pMpGE010 and pMpGE011 expression clones were transformed into *Agrobacterium tumefaciens* strain GV2260 (Althoff et al., 2014). Both Agrobacterium strains were used in parallel in a sporeling transformation (Althoff et al., 2014). After selection of the transformed sporelings on 0.5 µM chlorsulfuron and 10 mg/L hygromycin and 100 mg/L Cefotaxime, resistant plants were genotyped for Mp*B-GATA1* polymorphisms by PCR and Sanger sequencing (Table S1). Two mutant alleles, Mp*b-gata1-1* and Mp*bgata1-2* were used for further analyses (Figure S2).

The *Arabidopsis thaliana gata hexuple* (*gata hex*) mutants were generated by CRISPR/Cas9-mediated gene disruption of *GATA15* and *GATA16* in the previously published *gata quad* (Ranftl et al., 2016). Two gRNAs targeting *GATA15* and *GATA16*, respectively, were cloned into pHEE401E and transformed into *gata quad* by floral dipping (Table S1) (Clough & Bent, 1998; Wang et al., 2015). T1 seedlings were identified on ½ MS agar plants containing 25 mg/L hygromycin, plants were analysed through a combination of high-resolution melt analysis with the Precision Melt Analysis Software, PCR and Sanger sequencing (BioRad) (Samarut et al., 2016). An estimation of InDels was performed with the heterozygous T1 plants using Tracking of Indels by Decomposition (Brinkman & van Steensel, 2019). The analysis resulted in the identification of two *gata15* and nine *gata16* alleles in the *gata quad* background. Since, among 234 plants analyzed, no hexuple mutant could be identified directly in *gata quad*, two quintuple mutants, carrying a frame-shift mutation before and in the zinc-finger-encoding domains of *GATA15* (*gata15-1*) and *GATA16* (gata16-1), respectively, were crossed to obtain a *gnc gnl gata15 gata16 gata17 gata17l* hexuple mutant (*gata hex*) (Figure S3).

To obtain MpB-GATA1ox, the coding sequence of *MpB-GATA1* (Mp7g03490) was amplified from complementary DNA (cDNA) for Gateway-cloning and subcloned via pDONR221 (Thermo Fisher Scientific) into pMpGWB224 (Ishizaki et al., 2016).

2.7 | Quantitative reverse transcription-PCR (qRT-PCR)

For qRT-PCR, total RNA was extracted using the NucleoSpin RNA kit (Macherey-Nagel). cDNA was generated through reverse transcription using oligo(dT) primers and RevertAid reverse transcriptase (Thermo Fisher Scientific). qRT-PCR was carried out using Takyon No ROX SYBR MasterMix (Eurogentec) in a CFX96 or CFX384 Real-Time System Cycler (BioRad) with the following protocol: step 1, 95°C for 3:00 min; step 2, 95°C, 0:10 min; step 3, 55°C, 0:30 min followed by 39 repeats of steps 2 and 3. Mp*ACT7* (*Mp6g11010*) was used as housekeeping gene and normalised expression was calculated according to the $2^{-\Delta\Delta Ct}$ method as described and is shown relative to the wild type (Livak & Schmittgen, 2001).

2.8 | RNA-seq

For transcriptomic analysis comparing the two *Mpb-gata1* alleles to the BoGa wild type, gemmae of BoGa, *Mpb-gata1-1* and *Mpb-gata1-2* were grown on ½ Gamborg's B5 medium for 3 weeks under long day conditions (60 µmol m^{-2} s^{-1}) at 22°C in MLR-351 plant growth

chambers (Sanyo Denki). Three entire plants of each genotype were sampled and homogenised together and were considered a biological replicate, three biological replicates were used for the experiment. Plants were frozen and ground in liquid nitrogen and RNA was extracted using the NucleoSpin RNA kit (Macherey-Nagel). The quality and quantity of the extracted RNA was validated with a Bioanalyzer 2100 Microfluidics System using the RNA Nano chip (Agilent). Library construction for RNA-seq was performed using the TruSeq stranded mRNA Kit with TruSeq RNA Adapter Box Set B (Illumina) and protocol 1000000040498v00 (Elution-2-Frag-Prime step: 4 min at 94°C; 15 cycles enrichment). Paired-end sequencing (2 × 75 bp, single-index) was carried out on a HiSeq. 2500 (Illumina) with HiSeq Rapid SBS Kit v2 (2 × 50 cycles), HiSeq PE Rapid Cluster Kit v2 and onboard clustering (Illumina). The reads were mapped onto the Marchantia genome v5.1 (http://marchantia.info) using the CLC Genomics Workbench (Qiagen). For the discovery of differentially expressed genes (DEGs), genes with a false discovery rate (FDR) $p \leq 0.05$ were considered, no fold change filter was applied.

For the HL intensity RNA-seq experiment, *Marchantia polymorpha* lines BoGa, Mp*b-gata1-1*, Mp*b-gata1-2*, and Mp*B-GATA1ox-3* were grown for 2 weeks on ½ Gamborg's B5 medium under standard conditions, as described above. Following this initial growth period, plants were shifted to the sun simulators of the Research Unit: Environmental Simulation at the Helmholtz Zentrum München (Neuherberg) and grown for adaptation for 1 week at 20°C–22°C under constant light with 60–70 µmol m^{-2} s^{-1} PPFD (low-light [LL]; GL). The light treatments consisted of either 6 h of HL (t6HL) with 1200 µmol m^{-2} s^{-1} PPFD at 20°C–22°C (t6HL) or 6 h LL with 60–70 µmol m^{-2} s^{-1} PPFD at 20°C–22°C (t6LL). Respectively, one-third of the plants were harvested and immediately frozen in liquid nitrogen before the shift to HL (t0), 6 h after the shift to HL (t6HL) or after an additional 6 h in LL (t6LL). Three biological replicates were prepared from each genotype at each time point and condition, each replicate consisting of three entire plants. RNA was extracted using the NucleoSpin RNA kit (Macherey-Nagel), the quality and quantity of the extracted RNA was checked on a Bioanalyzer 2100 Microfluidics System using the RNA Nano Chip (Agilent). Library construction was performed using the TruSeq stranded mRNA Kit with TruSeq RNA Adapter Box Set A and B and the protocol 1000000040498v00 (Elution-2-Frag-Prime step: 4 min at 94°C; 15 cycles enrichment) (Illumina). Paired-end sequencing (2 × 100 bp, single index) was carried out on a NovaSeq. 6000 with NovaSeq. 6000 S1 Reagent Kit v1.5 (200 cycles) at the NGS Core Facility Helmholtz Zentrum München. Reads were mapped onto the Marchantia genome v6.1 (http://marchantia.info) using the CLC Genomics Workbench (Qiagen). For the discovery of DEGs, genes with a FDR $p \leq 0.05$ were considered. No fold change filter was applied. Heat maps were generated using the Cluster 3.0 and Treeview tools (de Hoon et al., 2004; Saldanha, 2004).

For the corresponding HL intensity RNA-seq experiment with Arabidopsis, seeds of the *Arabidopsis thaliana* lines Col-0, *gata hex* and GNLox were plated on ½ Murashige & Skoog medium and stratified for 3 days in the dark at 4°C. Plants were grown under constant LL at 60–70 µmol m^{-2} s^{-1} and 20°C–22°C in the sun simulators of the Research Unit: Environmental Simulation at the Helmholtz Zentrum München for 1 week, before exposing them to a 6 h HL stress or a corresponding LL mock control treatment, as described above. Three biological replicates of 30 seedlings, each, were prepared for each genotype, time point and treatment. Library construction was performed using the Illumina stranded mRNA Kit with IDT for Illumina RNA UD Indexes Set A and the protocol 1000000124518v01 with 11 cycles of enrichment (Illumina). Single-end sequencing (1 × 100 bp, dual-index) was performed on a NovaSeq. 6000 with NovaSeq. 6000 S1 Reagent Kit v1.5 (100 cycles) at the IMGM Laboratories (Martinsried). Reads were mapped onto the Arabidopsis TAIR10 genome using the CLC Genomics Workbench (Qiagen). Genes with a FDR $p \leq 0.05$ were considered as DEGs. Heat maps were generated using the Cluster 3.0 and Treeview tools (de Hoon et al., 2004; Saldanha, 2004). DEG sets were compared with UpSetR (Conway et al., 2017). Data from the RNA-seq experiments have been deposited in NCBI-GEO (https://www.ncbi.nlm.nih.gov/geo/) under accession series number GSE208563.

2.9 | Ortholog identification for comparative gene expression analysis

The limited annotation of the Marchantia genome made it necessary to identify *Arabidopsis thaliana* orthologs of *Marchantia polymorpha* genes. To this end, we used Orthofinder2 with default settings and *Marchantia polymorpha* genome version v6.1 and *Arabidopsis thaliana* genome version TAIR10 as inputs (Emms & Kelly, 2019). Marchantia genes that did not yield an Orthofinder2 result, were annotated by running a DIAMOND search against the TAIR10 genome using default settings (Buchfink et al., 2015). In the latter case, the Arabidopsis Best Hit gene from the DIAMOND analysis was considered an ortholog to the corresponding Marchantia input gene.

2.10 | Gene ontology (GO) enrichment analysis

GO enrichment analysis for *Marchantia 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, Meng, Jin, & Gao, 2020). GO terms with $p \leq 0.05$ were considered statistically significant. The *Arabidopsis thaliana* datasets were analyzed using PANTHER with default settings (Mi et al., 2021).

2.11 | Chlorophyll quantification

For chlorophyll quantification, chlorophylls were extracted as described with slight modifications (Hu et al., 2013). Shortly, 30–100 mg of Marchantia frozen thallus tissue was homogenised using steel beads in the TissueLyser II (Qiagen). Ground material was

incubated in 80% acetone for 30 min in the dark and centrifuged for 15 min at 16 000*g*. The supernatant was mixed (1:1) ratio with methanol (vol/vol) and incubated for 30 min in the dark. The mixture was again centrifuged for 15 min at 16 000*g* and the absorbance of the supernatant was measured at 470, 645, and 663 nm in a Tristar 5 plate reader (Berthold). Chlorophyll and carotenoid contents were calculated as described: Chl_a (mg/mg FW) = [12.25A_{663nm} − 2.79 A_{645nm}] × V/W; Chl_b (mg/mg FW) = [21.5A_{645nm} − 5.10A_{663nm}] × V/W; Chl_{a+b} (mg/mg FW) = [7.15A_{663nm} + 18.71 A_{645nm}] × V/W; C_{x+c} (mg/mg FW) = [[1000A_{470nm} − 1.63Chl_a − 85.0Chl_b]/198] × V/W (Hu et al., 2013).

2.12 | 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.

3 | RESULTS

3.1 | Bryophytes encode B-GATAs with HAN- and LLM-domains

Previous analyses had revealed the existence of HAN- and LLM-domain-containing B-GATAs in the bryophytes *Marchantia polymorpha* and *Physcomitrium patens* (Schwechheimer et al., 2022). To provide an overview of the plant B-GATA family, we generated a phylogenetic tree of the B-GATAs of the streptophyte alga *Chara braunii* (3 B-GATAs), the bryophytes *Marchantia polymorpha* (1 B-GATA) and *Physcomitrium patens* (4 B-GATAs), the lycophyte *Selaginella moellendorffii* (3 B-GATAs), the fern *Salvinia cucullata* (5 B-GATAs), and the angiosperm *Arabidopsis thaliana* (11 B-GATAs) (Figure 1). The analysis showed that LLM-domain-containing B-GATAs are present in all species analyzed, including the three B-GATAs from *Chara braunii* (Figure 1 and Table S3). The HAN-domain, however, could not be found in *Chara braunii*, selected here as a representative algae, but was present in all land plant species examined (1). This supported the previously made conclusion that the HAN-domain is specific to land plant GATAs (Figure 1) (Schwechheimer et al., 2022). B-GATA factors can be identified based on their zinc-finger DNA-binding domain, and we also found B-GATAs devoid of the LLM- or the HAN-domain in *Salvinia cucullata* (Scf718000008998) and *Arabidopsis thaliana* (AtGATA29), in addition to a previously reported *Arabidopsis thaliana* B-GATA with a degenerate LLM-domain (AtGATA23) (Behringer et al., 2014). Importantly, whereas *Arabidopsis thaliana* and other land plant genomes contain either HAN- or LLM-domain B-GATAs, the B-GATAs from the bryophytes *Marchantia polymorpha* and *Physcomitrium patens* contain a HAN-, as well as an LLM-domain (Figure 1) (Schwechheimer et al., 2022).

3.2 | Functional analysis of *Marchantia polymorpha* MpB-GATA1

Marchantia polymorpha encodes for a single B-GATA factor, MpB-GATA1 (Figure 1a). To examine MpB-GATA1 function, we generated a loss-of-function mutant by CRISPR/Cas9-based genome editing in the BoGa ecotype and identified two loss-of-function alleles, Mp*b-gata1-1* and Mp*b-gata1-2*, with 796 and 791 base pair deletions encompassing the zinc-finger domain and resulting in frame-shift mutations (Figure 2a,b and Figure S2) (Sugano et al., 2014). We also obtained Mp*B-GATA1* overexpression lines by expressing the Mp*B-GATA1* cDNA under control of the constitutive *MpEF1a* promoter fragment (Althoff et al., 2014). Amongst the overexpression lines, we initially identified Mp*B-GATA1ox-3*, although the expression strength exceeded the expression of the wild type gene, analysed from plant material of the same age, only threefold (Figure S4). An apparent trait of the Mp*b-gata1* alleles and the overexpression lines was their reduced thallus size when compared to the wild type (Figure 2b,c and 2e). Furthermore, the loss-of-function mutants displayed an increase in the number of gemma cups per thallus area, while gemma cup formation was strongly reduced in Mp*B-GATA1ox-3* (Figure 2b–d and 2f). Even after prolonged growth, gemma cups were rarely observed in Mp*B-GATA1ox-3* plants (Figure 2c,d). We later identified other overexpression lines through independent transformations, and despite the differences in their expression strength, they displayed the same defects as the initially characteriser Mp*B-GATA1ox-3* (Figure S4).

3.3 | Loss of Mp*B-GATA1* results in dramatic gene expression changes

To understand the molecular consequences of the loss of Mp*B-GATA1*, we performed an RNA-seq experiment with 3-week-old plants grown under long day conditions (16 h light/8 h dark) at 60 µmol s^{-1} m^{-2} white light at 22°C. The analysis identified, respectively, 1742 and 1699 down- and upregulated genes (FDR $p \leq 0.05$) in both alleles when compared to the wild type (Table S3). GO enrichment analyses identified terms related to terpene biosynthesis, phosphatidylglycerol metabolic processes and carotene biosynthesis, but also photosynthesis-related categories as the most strongly differentially regulated GO term categories among the downregulated genes (Figure 3a). Among the upregulated genes, many GO terms referred to basic cell biological processes like Golgi function and transport, secretion and cell adhesion (Figure 3a).

Defects in chlorophyll biosynthesis and greening are the most apparent defects of *Arabidopsis thaliana gnc gnl* mutants (Bastakis et al., 2018; Bi et al., 2005). These have previously been characterised in detail and could be explained by the downregulation of chlorophyll biosynthesis genes of the tetrapyrrole biosynthesis pathway (Bastakis et al., 2018). When we examined the expression of tetrapyrrole biosynthesis genes in the Marchantia Mp*b-gata1* mutants, we found a substantial overlap in the identity and directionality of differential

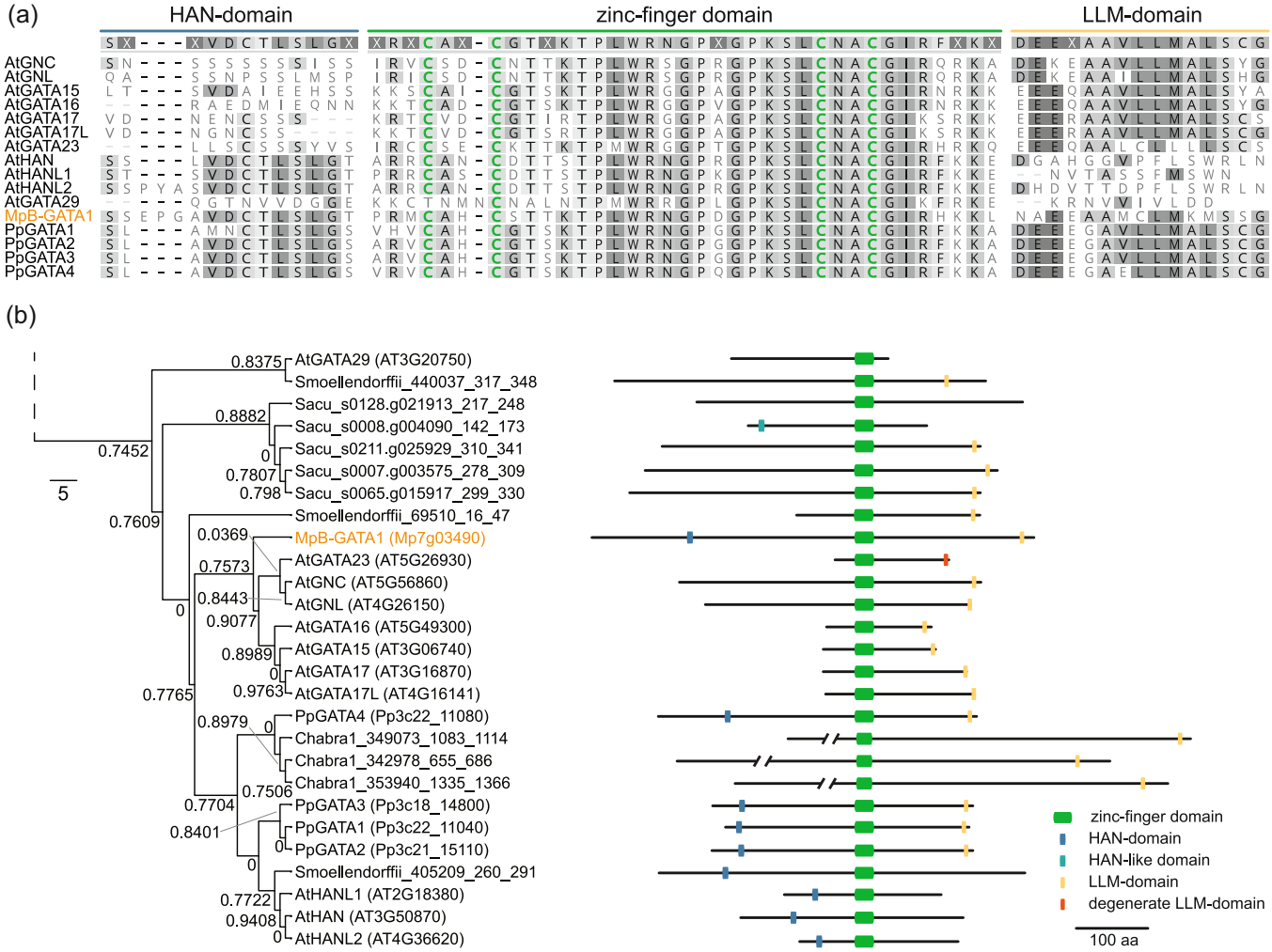

FIGURE 1 Bryophytes contain B-GATAs with a HAN- and an LLM-domain. (a) MUSCLE protein sequence alignment of the HAN-domain, the zinc-finger domain and the LLM-domain of B-GATA factors from *Arabidopsis thaliana* (At), *Marchantia polymorpha* (Mp), and *Physcomitrium patens* (Pp). (b) Phylogenetic tree and schematic representation of the B-GATA factors from *Selaginalla moellendorfii* (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. *Marchantia polymorpha* MpB-GATA1 is highlighted in yellow. The scale bar represents a distance of five amino acid substitutions. Numbers refer to branch support after an aLRT SH-like test. The phylogeny shown here is extracted from a phylogenetic tree of all GATAs of the respective species. The full phylogeny is included as Figure S1. LLM, leucine-leucine-methionine.

regulation between the respective Marchantia and Arabidopsis gene sets (Figure S5 and Table S4). In line with their pale-green appearance, we found that the levels of chlorophyll a and chlorophyll b, and consequently of total chlorophylls, were reduced in both alleles but not in the wild type (Figure S5B–D). We thus concluded that differential regulation of chlorophyll biosynthesis genes may be an evolutionarily conserved B-GATA gene function.

With regard to the differences in gemma cup formation observed between the mutants, we scored the RNA-seq data for expression of the MYB transcription factor *GEMMA CUP-ASSOCIATED MYB1* (Mp*GCAM1*), an essential regulator of gemma cup formation (Yasui et al., 2019). In line with the observed phenotypes (Figure 2), Mp*GCAM1* expression was up- and downregulated in Mp*b-gata1* and Mp*B-GATA1ox-3*, respectively (Figure S6).

One striking observation from the RNA-seq experiment was the differential regulation of 88 genes related to photosynthesis (Figure 3b). Importantly, these included 19 strongly upregulated orthologues of the *Arabidopsis thaliana ELIP*-encoding genes (Figure 3b). ELIPs are thought to prevent photooxidative damage to the photosynthetic machinery by facilitating excess energy dissipation by binding free chlorophyll molecules (Montane & Kloppstech, 2000). While ELIPs have similar structure and similar chlorophyll-binding capacities to light harvesting complexes (LHCs), they differ from LHCs in their transient expression upon exposure to HL intensities (Hutin et al., 2003). Based on this observation, we

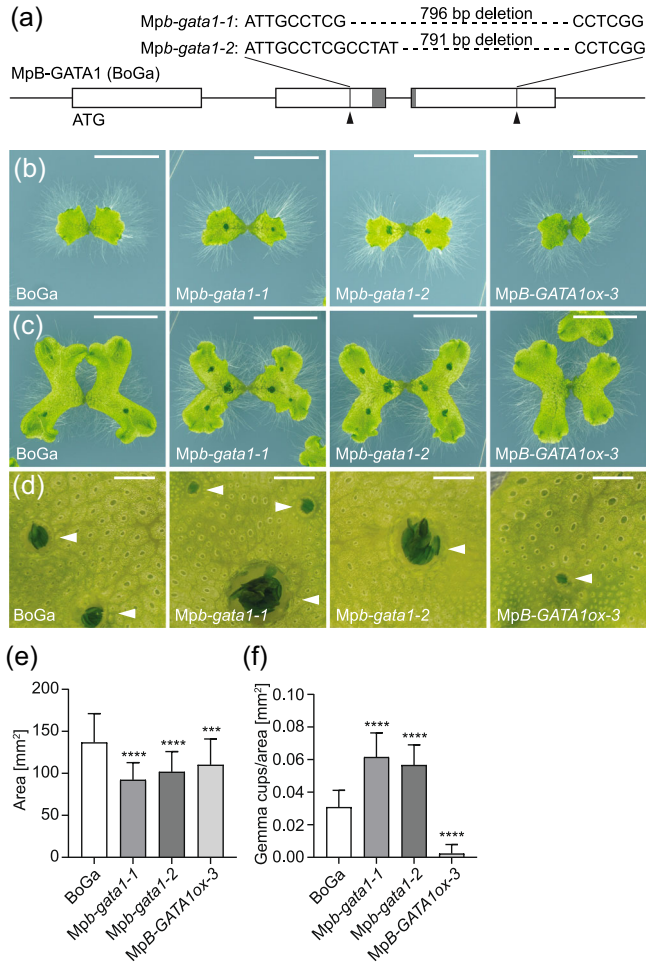

FIGURE 2 Characterisation of Mp*b-gata1* mutants and overexpressors. (a) Schematic representation of the Mp*B-GATA1* gene and delineation of the deletions present in Mp*b-gata1-1* and Mp*b-gata1-2*. Boxes correspond to exons, the line marks the introns; the zinc-finger-encoding region, interrupted by an intron, is shown by a grey box. Arrowheads mark the positions of the guide RNAs used for genome editing. (b–d) Representative photographs of two (b) and 3 weeks old (c) Marchantia plants and their thallus areas (d) of the specified genotypes. Arrowheads mark gemma cups. Scale bars = 10 mm (b) and (c); scale bar = 1 mm (e). (e, f) Graphs displaying the quantification of the thallus area (e) and gemma cup density (f), as determined from 3-week-old plants of the specified genotypes. Student's *t*-test: ***$p \leq 0.001$; ****$p \leq 0.0001$, $n = 35$. [Color figure can be viewed at wileyonlinelibrary.com]

speculated that the Marchantia Mp*b-gata1* mutants may display a constitutive HL stress response.

3.4 | Marchantia Mp*b-gata1* mutants display a HL stress response phenotype

To analyze HL stress responses in Marchantia and the Mp*b-gata1* mutants at the molecular level, we performed an RNA-seq experiment using 3-week-old plants of the BoGa wild type and the two Mp*b-gata1* mutant alleles grown in LL (60 µmol m^{-2} s^{-1} PPFD) and after a 6 h exposure to HL stress (1200 µmol m^{-2} s^{-1} PPFD) or a corresponding LL (60 µmol m^{-2} s^{-1} PPFD) mock treatment (Figure 4a). Following HL stress exposure, the wild type as well as the Mp*b-gata1* mutants had obvious stress symptoms, a kind of wilting of the parts of the thallus tips, which were located most proximal to the light source (Figure 4b).

In the BoGa wild type, we identified 7520 genes that were differentially regulated after HL exposure (FDR $p \leq 0.05$) but not when exposed to the LL mock treatment (BoGa t6HL vs. BoGa t6LL; Figure 4c and Table S5). Of these 7520 HL-regulated genes, 2952 genes (39%; grey-shaded intersection) were already differentially regulated in Mp*b-gata1* grown in LL mock conditions (Mp*b-gata1* t6LL vs. BoGa t6LL; Figure 4c and Table S5). When we examined the expression of these 2952 genes at the quantitative level, we found that 1935 (65%; central intersection) of these genes were still HL stress-regulated in Mp*b-gata1* (Figure 4c and Table S5). A large fraction of the HL stress-responsive genes (71%; 1109 and 984 of 2952) were regulated in the same direction in the wild type after HL stress and LL grown mutant (Figure 4d,e and Table S5). A total of 4923 genes (65%) of the HL stress-induced genes from the wild type (7520) and the Mp*b-gata1* mutant (5963) were HL stress-induced in the wild type and the Mp*b-gata1* mutants (Figure 4c and 4f; Table S5). These findings are in line with the notion that HL stress-responsive gene expression is substantially, but at the quantitative level only partially differentially regulated in the mutants (Figure 4c and Table S5).

Only a subset of 249 and 414 genes among the 2952 differentially regulated genes (21%) were antagonistically regulated between the LL-exposed Mp*b-gata1* mutants and the HL-stressed BoGa wild type (Figure 4d,e and Table S5). These DEGs from the HL stress treatment may correspond to heat stress-responsive genes rather than (only) HL stress-responsive genes because the HL stress treatment was unavoidably associated with local temperature increases, although the experimental set-up had allowed a stable temperature at the chamber level.

3.5 | HL stress differentially affects air pore size in *Marchantia polymorpha MpB-GATA1* genotypes

One visible trait discernable in HL stressed Marchantia was an increase in air pore size when the plants had been grown for 2 weeks in (moderate) HL conditions (400–500 µmol s^{-1} m^{-2}; Figure 5). Air pores are structures at the surface of the thallus that allow efficient gas exchange within the thallus tissue. An increase in air pore size may thus be a physiological response that allows better venting of the plants in HL. Closer inspection revealed a slight reduction of the air pore size in LL-grown Mp*b-gata1* mutants, when compared to the wild type, and a strong increase in air pore size in the MpB-GATA1 overexpression line Mp*B-GATA1ox3* (Figure 5). In all genotypes, but very prominently in Mp*B-GATA1ox3*, air pore size was increased after

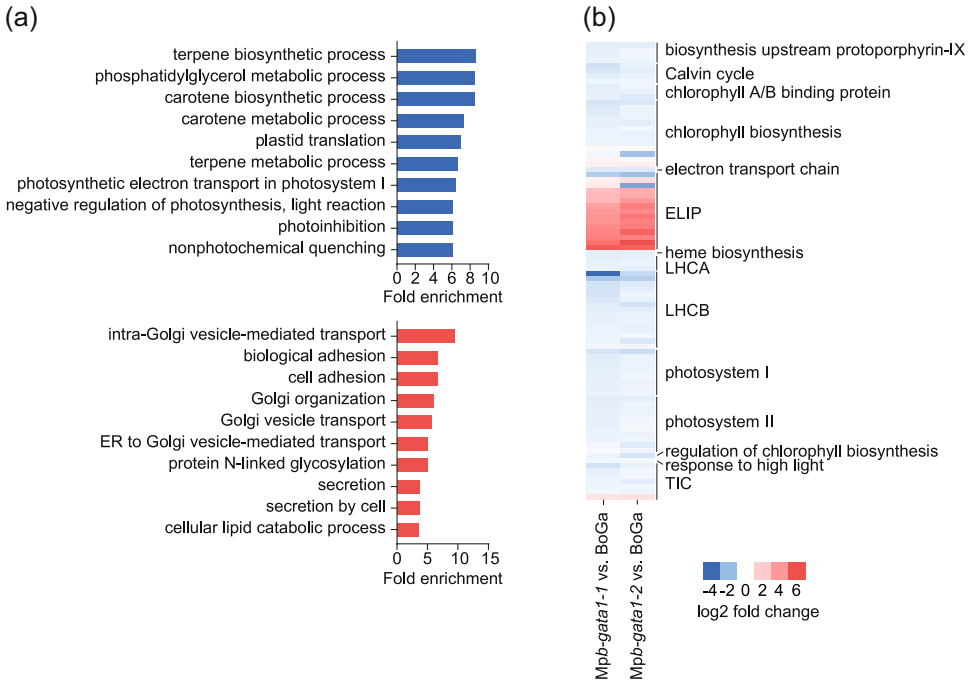

FIGURE 3 Photosynthesis associated-genes are misregulated in Mp*b-gata1* mutants. (a) GO enrichment analysis of down- and upregulated genes from transcriptomic data of Mp*b-gata1* mutants compared to BoGa shown in blue and red, respectively. (b) Heatmap showing expression of photosynthesis-associated genes from both mutant lines. ELIP, early light induced protein; GO, Gene Ontology; LCHA/LHCB, light-harvesting complex; TIC, translocon on the inner chloroplast membrane. [Color figure can be viewed at wileyonlinelibrary.com]

1 week of growth in (moderate) HL conditions (Figure 5b–d). We concluded that *M. polymorpha* responds to high light by increasing air pore size. Additionally, air pore size seemed to correlate with overexpression or loss of Mp*B-GATA1*, suggesting that MpB-GATA1 might quantitatively modulate air pore size in response to HL.

3.6 | *Arabidopsis thaliana gata hexuple* mutants display a HL stress-response phenotype

We next examined whether also *Arabidopsis thaliana* mutants displayed a HL stress response. We had previously generated a quadruple mutant of the *Arabidopsis thaliana* LLM-domain B-GATA factors combining strong alleles in the genes *GNC*, *GNL*, *GATA17* and *GATA17L* (*gata quad)* (Ranftl et al., 2016). For the purpose of this study, we introduced, additionally, predicted loss-of-function mutations in *GATA15* and *GATA16*, for which only partial knock-out mutations had been identified in previous analyses, by genome editing. We identified a *gata15* and a *gata16* allele in *gata* quad, both harbouring one-base pair insertions upstream of the GATA zinc-finger (Figure S3). From these *gata15 gata quad* and *gata16 gata quad* quintuple mutants, we subsequently obtained a hexuple mutant (*gata hex*) through genetic crosses.

Using *gata hex*, we performed a HL stress experiment, equivalent to the experiment with Marchantia, as described above (Figure 6a,b). In the wild type (Col-0), we identified 11 914 genes (Col-0 t6HL vs. Col-0 t6LL) that were differentially regulated after HL exposure (FDR $p \leq 0.05$) but not when exposed to a LL mock treatment (Figure 6c and Table S6). Of these 3500 HL-regulated genes from the wild type (Col-0 t6HL vs. Col-0 t6LL), 2973 genes (25%; light grey-shaded intersection) were already differentially regulated in *gata hex* grown in LL mock conditions (*gata hex* t6LL vs. Col-0 t6LL). When we examined the expression of these 2973 genes at the quantitative level (*gata hex* t6HL vs. *gata hex* t6LL), we found that 1860 (62%) of these genes were still HL-regulated in *gata hex*, in line with the notion that HL stress-responsive gene expression was substantially, but at the quantitative level only partially, affected in *gata hex* (Figure 6c,d and Table S6).

A total of 6696 genes, thus the majority of HL-regulated genes from the wild type (11 914; 56%) and the *gata hex* mutant (7458; 89%) were HL stress-induced in wild type and *gata hex* (Figure 6c and 6f; Table S6), and we concluded that HL stress-induced gene expression is quantitatively compromised in the Arabidopsis B-GATA mutants.

A subset of 611 and 789 genes among the 11 914 differentially regulated genes was antagonistically between the LL-exposed *gata hex* mutants and the HL-stressed wild type (Figure 6e and Table S6). Similarly to the Marchantia experiment, we reasoned also here that the respective genes may be regulated by exposure to local heat stress, a condition experienced by the HL stressed seedlings but not the LL mock control.

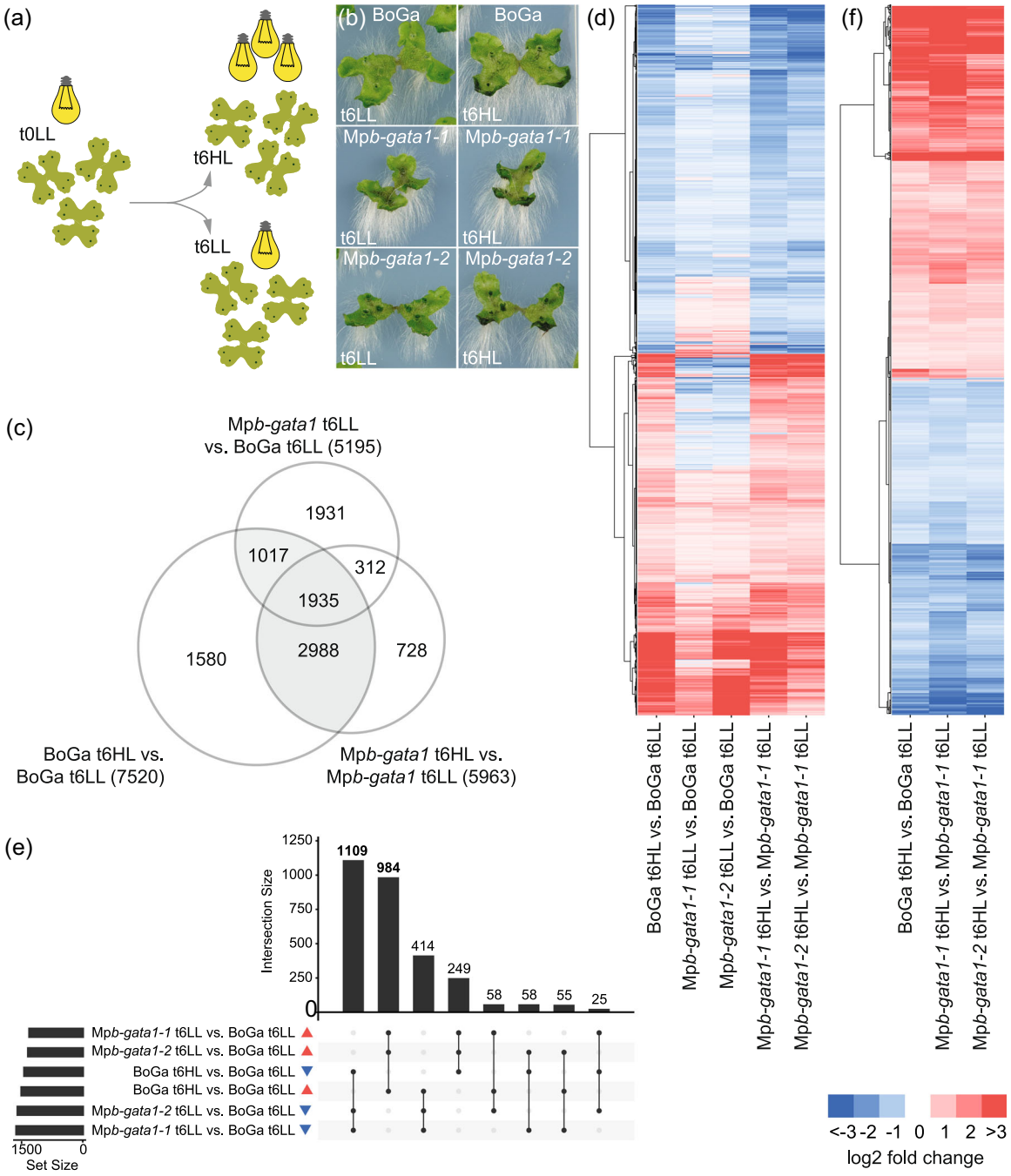

FIGURE 4 *Marchantia polymorpha* Mp*b-gata1* mutants exhibit a constitutive high-light (HL) stress response. (a) Schematic representation of the Marchantia HL stress experiment. Three-week-old low-light (LL) grown *Marchantia polymorpha* BoGa, Mp*b-gata1-1* and Mp*b-gata1-2* were subjected to either 6 h of HL stress (t6HL; 1200 µmol s^{-1} m^{-2}) or 6 h of a corresponding LL mock treatment (t6LL; 60 µmol s^{-1} m^{-2}). (b) Representative photographs of BoGa, Mp-*gata1-1* and Mp*b-gata1-2* plants at t6LL and t6HL. Plants exposed to HL stress show wilting in the areas most proximal to the light source. (c) Venn diagram displaying the subsets of HL stress-regulated genes in BoGa (BoGa t6HL vs. BoGa t6LL) and Mp*b-gata1* (Mp*b-gata1* t6HL vs. Mp*b-gata1* t6LL), as well as of genes differentially regulated in the Mp*b-gata1* mutant in LL conditions (Mp*b-gata1* t6LL vs. BoGa t6LL). (d) Heatmap displaying the differential expression of 1935 differentially expressed genes from the central intersection of the Venn diagram in (c). (e) UpSetR diagram displaying the direction of transcriptional regulation of 2952 differentially expressed genes that are HL stress-induced in the wild type (BoGa t6HL vs. BoGa t6LL) and differentially expressed between Mp*b-gata1* and the BoGa wild type in LL conditions (Mp*b-gata1* t6LL vs. BoGa t6LL). (f) Heatmap displaying the differential expression of 4923 differentially expressed genes induced after HL stress in BoGa and Mp*b-gata1* (BoGa t6HL vs. BoGa t6LL and Mp*b-gata1* t6HL vs. Mp*b-gata1* t6LL). [Color figure can be viewed at wileyonlinelibrary.com]

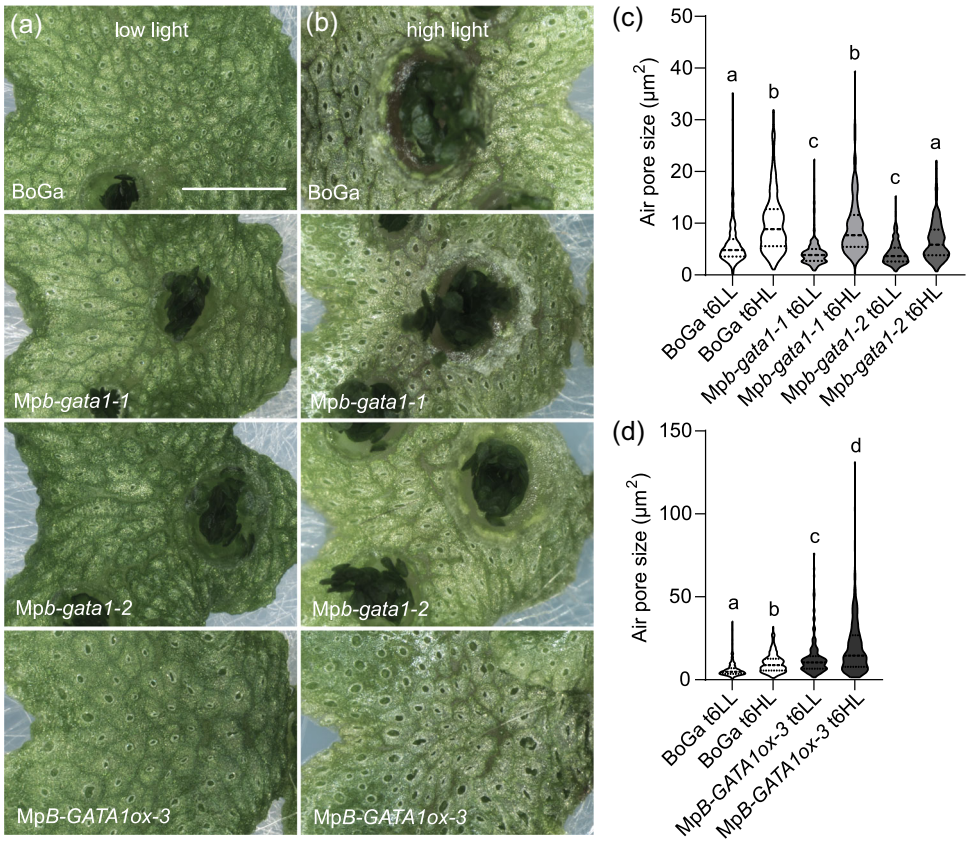

FIGURE 5 Air pore size in *Marchantia polymorpha* is modulated by high-light (HL) and Mp*B-GATA1*. (a) Representative photographs of 3-week-old plants grown in regular growth conditions (low-light [LL]) and (b) grown for 1 week in regular light conditions before shift to moderate HL (400–500 µmol s^{-1} m^{-2}). Scale bar = 2 mm. (c, d) Graphs displaying air pore sizes of LL and HL grown plants as shown in (a) and (b). The BoGa wild type data are identical between (c) and (d). Datasets with no statistical difference after one-way ANOVA and Tukey's HSD post hoc test fall into one group and are labelled with identical letters. ANOVA, analysis of variance. [Color figure can be viewed at wileyonlinelibrary.com]

3.7 | Comparative transcriptomics uncovers a strong conservation in the differential regulation of photosynthesis-related genes between the *Marchantia polymorpha* and *Arabidopsis thaliana*

We had originally identified *ELIP* genes from Marchantia as being strongly differentially regulated in the Mp*b-gata1* mutants, compared to the wild type, when grown in normal (LL) conditions (Figure 3b). We found that also in normal (LL) grown *Arabidopsis thaliana*, the two *ELIP* genes of these species were strongly upregulated when compared to the wild type (Figure S7). To examine a possible conservation of the gene expression defects between the mutants from the two species, we extended our analysis to all genes associated with photosynthesis and chlorophyll biosynthesis (Figure S7). *Arabidopsis thaliana* genes related to photosynthesis were identified based on their functional annotation in TAIR (Berardini et al., 2015), while *Marchantia polymorpha* genes were identified as orthologs to the Arabidopsis genes using Orthofinder2 (Emms & Kelly, 2019). This comparative analysis uncovered, besides the already mentioned differential regulation of *ELIP* genes, a striking similarity in the identity of differentially regulated genes, in low- as well as in HL conditions, from the light harvesting complex, both photosystems, the electron transport chain as well as from ATP synthase complex (Figure 7). We concluded that, at least with regard to this transcriptional output, mutants of *B-GATA* genes from both species have an evolutionarily conserved function in the regulation of photosynthesis and HL stress.

4 | DISCUSSION

Here, we present the first genetic characterisation of a GATA transcription factor from a *Marchantia polymorpha*, the B-GATA Mp*B-GATA1*. In *Arabidopsis thaliana*, LLM-domain B-GATAs play an important role in chlorophyll biosynthesis through the regulation of a broad number of genes from the tetrapyrrole biosynthesis pathway (Bastakis et al., 2018; Zubo et al., 2018). Based on our analysis of gene expression data from Marchantia Mp*b-gata1* mutants, we could provide evidence that this role is conserved between Arabidopsis and Marchantia and may thus represent an ancestral function of the B-GATAs.

We also noted a strong upregulation of 16 of the differentially expressed 18 *ELIP* genes in Marchantia Mp*b-gata1* mutants grown in normal (LL) conditions. In *Arabidopsis thaliana*, the two *ELIP* genes are

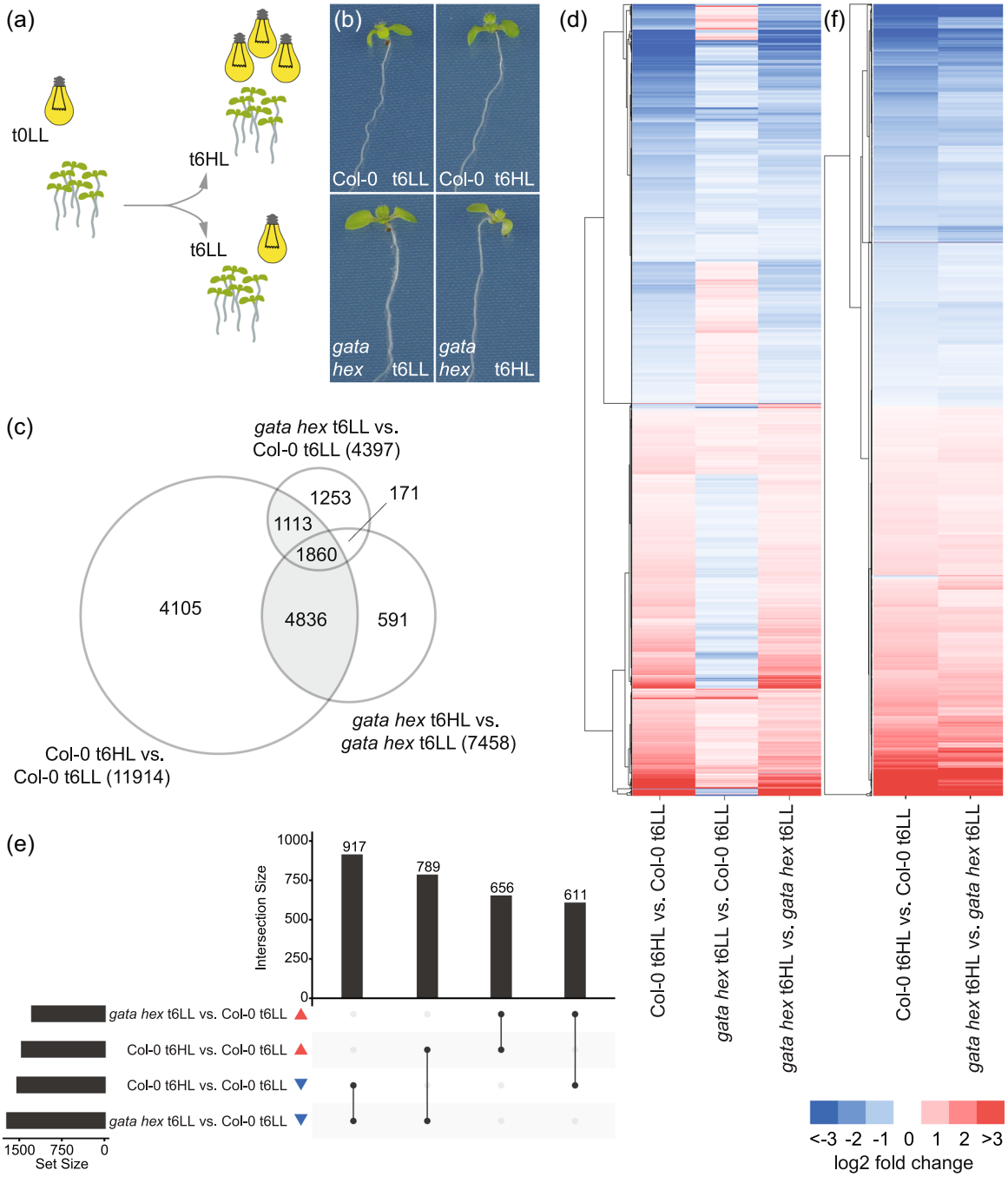

FIGURE 6 *Arabidopsis thaliana gata hex* exhibits a high-light (HL) stress response phenotype. (a) Schematic representation of the HL stress experiment. One-week-old *Arabidopsis thaliana* Col-0 and *gata hex* were subjected to either 6 h of HL (t6HL; 1200 µmol s^{-1} m^{-2}) or 6 h of corresponding low-light mock treatment (t6LL; 60 µmol s^{-1} m^{-2}). (b) Representative photographs of Col-0 and *gata hex* seedlings at t6LL and t6HL. Plants at t6HL show wilting in the areas most proximal to the light source. (c) Venn diagram displaying the differentially expressed gene subsets Col-0 t6HL versus Col-0t 6LL, *gata hex* t6LL versus Col-0 t6LL, and *gata hex* t6HL versus *gata hex* t6LL. (d) Heatmap displaying the differential expression of 1860 genes from the central intersection of the Venn diagram in (C). (e) UpSetR diagram displaying the direction of transcriptional regulation of the differential expressed genes from Col-0 t6HL versus Col-0 t6LL and *gata hex* t6LL versus Col-0 t6LL. (f) Heatmap displaying the differential expression of 2973 differentially expressed genes shared between Col-0 t6HL versus Col-0 t6LL and *gata hex* t6HL versus *gata hex* t6LL. [Color figure can be viewed at wileyonlinelibrary.com]

part of the photoprotection response to HL stress and they act as scavengers for unbound chlorophyll molecules (Hutin et al., 2003). Instigated by this observation, we tested the hypothesis that Mp*b-gata1* mutants display HL stress in normal (LL) conditions. In a dedicated RNA-seq experiment, we found a substantial differential regulation (39%) of genes in Mp*b-gata1* mutants grown in normal (LL) conditions that are regulated in the wild type after HL stress. This suggests that MpB-GATA1 is a repressor of HL stress-responsive

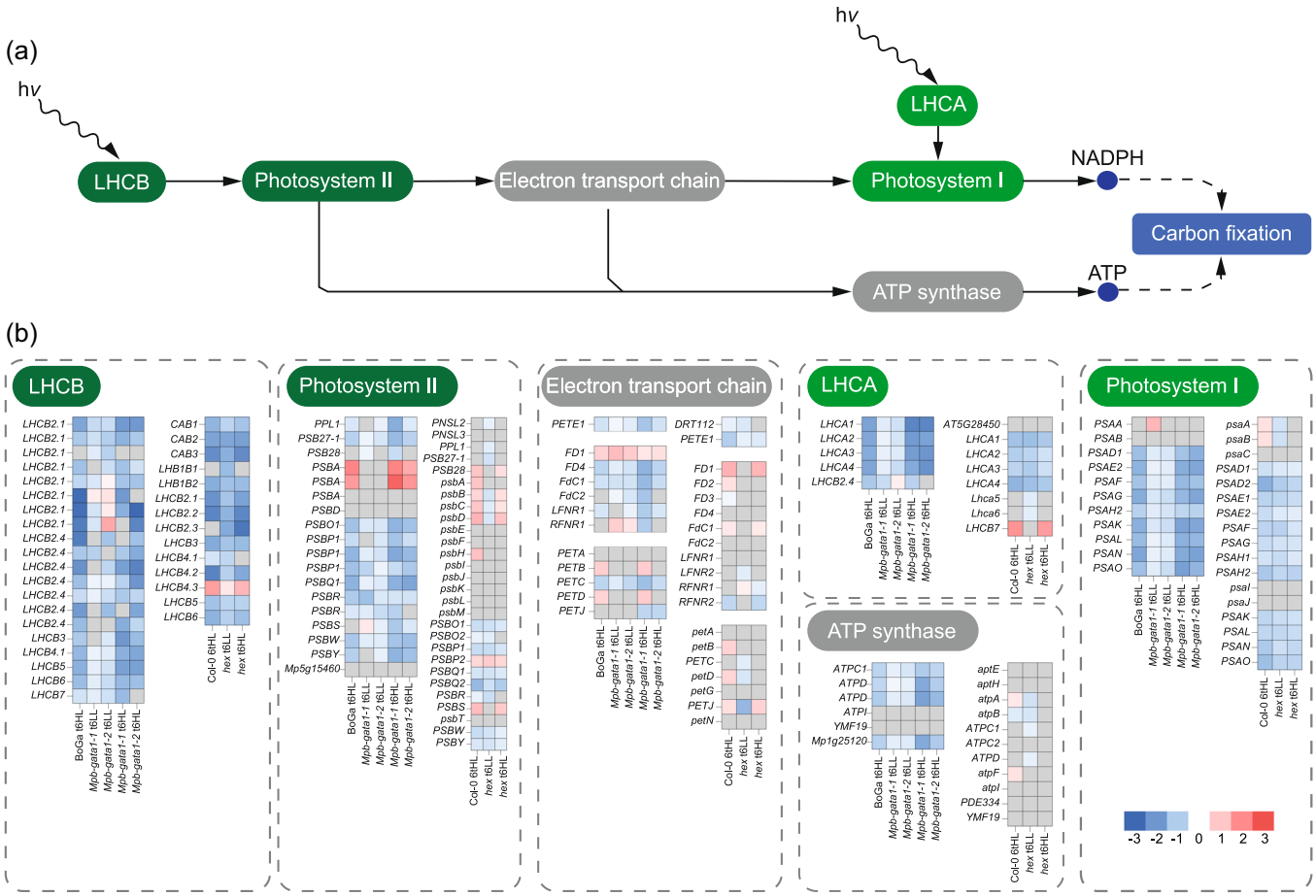

FIGURE 7 Photosynthesis-associated genes are misregulated in mutants of B-GATAs from *Marchantia polymorpha* and *Arabidopsis thaliana*. (a) Schematic representation of the photosynthesis pathway and (b) heatmaps displaying the differential expression of genes in the *Marchantia polymorpha* and *Arabidopsis thaliana* genotypes based on the data obtained in the high-light stress experiment. Grey boxes indicate no differential expression. [Color figure can be viewed at wileyonlinelibrary.com]

gene expression in the wild type. Since we also observed that HL stress responsive gene expression is, with regard to the gene identities, quantitatively as well as qualitatively only partially impaired in the Mp*b-gata1* mutants, the existence of other regulators of HL stress must be postulated.

We subsequently examined the conservation of the transcriptional HL response in the *Arabidopsis thaliana* wild type and *gata hex*, deficient in all six canonical LLM-domain B-GATA factors. There, we confirmed the differential expression of *ELIP* genes after HL stress in the wild type, their upregulation in LL-grown *gata hex* and found that a fraction (25%) of the HL-regulated genes from the wild type was differentially regulated in LL-grown *gata hex*. We concluded that the role of B-GATAs in the regulation of HL responsive gene expression may be a conserved and ancestral function.

We further found that the differential regulation of the entire set of photosynthesis-related genes was differentially regulated in a highly comparable manner between Marchantia and Arabidopsis. In the absence of genetic studies, which are hindered in Marchantia by the absence of relevant mutants or other molecular analysis, it is at present difficult to judge whether the different observations with regard to the role of B-GATAs in chlorophyll biosynthesis, HL-induced gene expression or the expression of photosynthesis-related genes are interrelated or subject to a direct regulation by the GATAs.

Besides the regulation of greening and photosynthesis-related aspects, we also noted two morphological defects in Mp*b-gata1* mutants. Most prominently, Mp*b-gata1* mutants and Mp*B-GATA1* overexpressors showed, respectively, an increase and a decrease in the formation of gemma cups on the surface of their thalli. This phenotype was maintained after extended growth periods, arguing that the defect was not merely due to a delay in this developmental process, and accompanied by a reduced and increased expression of Mp*GCAM1*, a critical regulator of gemma cup formation (Yasui et al., 2019). Second, we observed an increase in the size of air pores in Marchantia plants that had been grown for an extended period under HL stress. Air pore size was strongly increased in the Mp*B-GATA1ox* overexpression lines, and slightly decreased in the Mp*b-gata1* mutants. Air pores likely aid in the aeration of the tissue, which may be an important attribute in plants experiencing HL stress environments. Thus, we propose a dual function of Mp*B-GATA1* in the regulation of abiotic stress responses and developmental

processes. Defects in gemma cup- and air pore formation have also been observed in *Marchantia polymorpha* overexpressors of the auxin corepressor *TOPLESS* and plants deficient in auxin signalling (Flores-Sandoval et al., 2015; Kato et al., 2015). However, the phenotypic alterations were distinct from the ones reported here. Interestingly, in *Arabidopsis thaliana*, the formation of stomata, which regulate gas exchange in vascular plants development—but notably not the regulation of stomata size—is under control of LLM-domain B-GATAs (Klermund et al., 2016). Thus, B-GATAs may be required for the formation of structures with an analogous function in Marchantia and Arabidopsis.

Our previously published phylogenomic analysis had revealed, in all plant species including algae and bryophytes, a conservation of all four GATA transcription factor classes, classes A–D, that had been described in vascular plants, e.g. *Arabidopsis thaliana* and rice (Reyes et al., 2004; Schwechheimer et al., 2022). Besides the Zn-finger, other protein domains are characteristic for individual family members, including the HAN- and the LLM-domain in the case of MpB-GATA1 from *Marchantia polymorpha*, or HAN and GNC or GNL from *Arabidopsis thaliana* (Schwechheimer et al., 2022). The B-GATA family in higher plants is comprised of many members, 11 in *Arabidopsis thaliana*, and its members generally contain either a HAN- or an LLM-domain. Also in the genome of the lycophyte *Selaginella moellendorffii*, we identified B-GATAs with a HAN- or an LLM-domain. In contrast, the B-GATA family from the bryophytes *Marchantia polymorpha* and *Physcomitrium patens* are small, with one and four members, respectively. Whereas the LLM-domain can already be found in algal B-GATAs, the HAN-domain has, at present, only and exclusively been identified in land plant B-GATAs. Intriguingly, bryophyte B-GATAs contain a HAN- as well as an LLM-domain, an arrangement never observed in a vascular plant or lycophyte species.

Since, across all species, the HAN-domain is exclusively found in land plants and in the context of B-GATAs, arguing that it was invented during land plant evolution and present in a common ancestor of the extant land plants. In this ancestor, the HAN- and LLM-domains may have, originally, been present on the same GATA protein. This arrangement may then have been maintained in bryophytes but not in vascular plants or lycophytes where the family was subdivided, e.g. to allow for its subfunctionalization. It could be hypothesised that the HAN- and the LLM-domains are functionally interdependent, which could be fulfilled in the case of the bryophyte B-GATAs by the presence of the two domains on the same protein, and may be achieved in the case of the vascular plant and lycophyte B-GATAs through heterodimerization between LLM- and the HAN-domain B-GATA family members. Based on existing data, there is, however, no indication for a shared biological role between HAN- and LLM-domain B-GATAs in land plants. An alternative explanation could thus be that the two domains are functionally independent and allow B-GATA transcription factors to perform additional regulatory functions, e.g., through interactions with other, additional proteins.

In their conquest of terrestrial habitats, the algal ancestors of extant embryophytes had to adjust to a broad range of new environmental stimuli and abiotic stresses. High irradiances and UV light are major challenges for terrestrial photosynthetic organisms (Becker & Marin, 2009; Delwiche & Cooper, 2015). Air does not filter light to the same degree as water does, leading to spectral differences to which the common ancestor of land plants was subjected to. Also, the light intensity on land fluctuates drastically during the day and can easily reach intensities capable of causing damage to the photosynthetic apparatus. Thus, in the course of terrestrialization, plants had to develop several strategies to deal with HL stress. When subjected to HL stress, plants reduce the production of photosynthesis-related proteins like light-harvesting complexes and photosystem components and decrease the biosynthesis of chlorophylls, while simultaneously promoting their degradation to protect the photosynthetic machinery. We demonstrated that loss-of-function mutants of LLM-domain containing B-GATAs in both *Marchantia polymorpha* and *Arabidopsis thaliana*, two species separated by over 400 million years of distinct evolution, exhibit HL stress phenotypes under regular growth conditions in LL. We speculate that the B-GATAs act as repressors or regulators of the HL stress response and that this role is evolutionary conserved. Thus, the B-GATAs acted as important players in the course of the water-to-land transition and the adaptation to new light stresses.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Deutsche Forschungsgemeinschaft through grants-in-aid DFG SCHW751/9-2 and DFG SCHW751/19-1 through the DFG priority programme 2237: MAdLand—Molecular Adaptation to Land: plant evolution to change. We are grateful for technical support and advice in RNA-seq from Christine Wurmser (Technical University of Munich, Freising, Germany). Open Access funding enabled and organized by Projekt DEAL.

ORCID

Peter Schröder http://orcid.org/0000-0002-0679-218X
Nora Gutsche http://orcid.org/0000-0002-8306-3087
Jana Barbro Winkler http://orcid.org/0000-0002-7092-9742
Boris Hedtke http://orcid.org/0000-0002-1170-1246
Bernhard Grimm http://orcid.org/0000-0002-9730-1074

REFERENCES

Alboresi, A., Gerotto, C., Giacometti, G.M., Bassi, R. & Morosinotto, T. (2010) *Physcomitrella patens* mutants affected on heat dissipation clarify the evolution of photoprotection mechanisms upon land colonization. *Proceedings of the National Academy of Sciences*, 107(24), 11128–11133. Available from: https://doi.org/10.1073/pnas.1002873107

Althoff, F., Kopischke, S., Zobell, O., Ide, K., Ishizaki, K., Kohchi, T. et al. (2014) Comparison of the MpEF1α and CaMV35 promoters for application in *Marchantia polymorpha* overexpression studies. *Transgenic Research*, 23(2), 235–244. Available from: https://doi.org/10.1007/s11248-013-9746-z

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215(3),

403–410. Available from: https://doi.org/10.1016/S0022-2836(05)80360-2

Anisimova, M. & Gascuel, O. (2006) Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology*, 55(4), 539–552. Available from: https://doi.org/10.1080/10635150600755453

Bastakis, E., Hedtke, B., Klermund, C., Grimm, B. & Schwechheimer, C. (2018) LLM-domain B-GATA transcription factors play multifaceted roles in controlling greening in arabidopsis. *The Plant Cell*, 30(3), 582–599. Available from: https://doi.org/10.1105/tpc.17.00947

Becker, B. & Marin, B. (2009) Streptophyte algae and the origin of embryophytes. *Annals of Botany*, 103(7), 999–1004.

Behringer, C., Bastakis, E., Ranftl, Q.L., Mayer, K.F.X. & Schwechheimer, C. (2014) Functional diversification within the family of B-GATA transcription factors through the leucine-leucine-methionine domain. *Plant Physiology*, 166(1), 293–305. Available from: https://doi.org/10.1104/pp.114.246660

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. et al. (2015) The arabidopsis information resource: making and mining the "gold standard" annotated reference plant genome. *Genesis*, 53(8), 474–485. Available from: https://doi.org/10.1002/dvg.22877

Bi, Y.M., Zhang, Y., Signorelli, T., Zhao, R., Zhu, T. & Rothstein, S. (2005) Genetic analysis of arabidopsis GATA transcription factor gene family reveals a nitrate-inducible member important for chlorophyll synthesis and glucose sensitivity. *The Plant Journal*, 44(4), 680–692. Available from: https://doi.org/10.1111/j.1365-313X.2005.02568.x

Brinkman, E.K. & van Steensel, B. (2019) Rapid quantitative evaluation of CRISPR genome editing by TIDE and TIDER. *Methods in Molecular Biology*, 1961, 29–44. Available from: https://doi.org/10.1007/978-1-4939-9170-9_3

Buchfink, B., Xie, C. & Huson, D.H. (2015) Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12(1), 59–60. Available from: https://doi.org/10.1038/nmeth.3176

Buschmann, H., Holtmannspötter, M., Borchers, A., O'Donoghue, M.T. & Zachgo, S. (2016) Microtubule dynamics of the centrosome-like polar organizers from the basal land plant *Marchantia polymorpha*. *New Phytologist*, 209(3), 999–1013. Available from: https://doi.org/10.1111/nph.13691

Clough, S.J. & Bent, A.F. (1998) Floral dip: a simplified method for agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, 16(6), 735–743.

Conway, J.R., Lex, A. & Gehlenborg, N. (2017) UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, 33(18), 2938–2940. Available from: https://doi.org/10.1093/bioinformatics/btx364

de Hoon, M.J.L., Imoto, S., Nolan, J. & Miyano, S. (2004) Open source clustering software. *Bioinformatics*, 20(9), 1453–1454. Available from: https://doi.org/10.1093/bioinformatics/bth078

Delwiche, C.F. & Cooper, E.D. (2015) The evolutionary origin of a terrestrial flora. *Current Biology*, 25(19), R899–R910.

Demmig-Adams, B. & Adams, W.W. (1992) Photoprotection and other responses of plants to high light stress. *Annual Review of Plant Physiology and Plant Molecular Biology*, 43(1), 599–626. Available from: https://doi.org/10.1146/annurev.pp.43.060192.003123

Eberhard, S., Finazzi, G. & Wollman, F.A. (2008) The dynamics of photosynthesis. *Annual Review of Genetics*, 42, 463–515. Available from: https://doi.org/10.1146/annurev.genet.42.110807.091452

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797. Available from: https://doi.org/10.1093/nar/gkh340

Emms, D.M. & Kelly, S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20(1), 238. Available from: https://doi.org/10.1186/s13059-019-1832-y

Flores-Sandoval, E., Eklund, D.M. & Bowman, J.L. (2015) A simple auxin transcriptional response system regulates multiple morphogenetic processes in the liverwort *Marchantia polymorpha*. *PLoS Genetics*, 11(5), e1005207. Available from: https://doi.org/10.1371/journal.pgen.1005207

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321. Available from: https://doi.org/10.1093/sysbio/syq010

Harari-Steinberg, O., Ohad, I. & Chamovitz, D.A. (2001) Dissection of the light signal transduction pathways regulating the two early light-induced protein genes in arabidopsis. *Plant Physiology*, 127(3), 986–997.

Hu, X., Tanaka, A. & Tanaka, R. (2013) Simple extraction methods that prevent the artifactual conversion of chlorophyll to chlorophyllide during pigment isolation from leaf samples. *Plant Methods*, 9(1), 19. Available from: https://doi.org/10.1186/1746-4811-9-19

Huang, J., Zhao, X. & Chory, J. (2019) The arabidopsis transcriptome responds specifically and dynamically to high light stress. *Cell Reports*, 29(12), 4186–4199. Available from: https://doi.org/10.1016/j.celrep.2019.11.051

Hutin, C., Nussaume, L., Moise, N., Moya, I., Kloppstech, K. & Havaux, M. (2003) Early light-induced proteins protect arabidopsis from photooxidative stress. *Proceedings of the National Academy of Sciences*, 100(8), 4921–4926. Available from: https://doi.org/10.1073/pnas.0736939100

Ishizaki, K., Nishihama, R., Yamato, K.T. & Kohchi, T. (2016) Molecular genetic tools and techniques for *Marchantia polymorpha* research. *Plant and Cell Physiology*, 57(2), 262–270. Available from: https://doi.org/10.1093/pcp/pcv097

Kato, H., Ishizaki, K., Kouno, M., Shirakawa, M., Bowman, J.L., Nishihama, R. et al. (2015) Auxin-mediated transcriptional system with a minimal set of components is critical for morphogenesis through the life cycle in *Marchantia polymorpha*. *PLoS Genetics*, 11(5), e1005084. Available from: https://doi.org/10.1371/journal.pgen.1005084

Klermund, C., Ranftl, Q.L., Diener, J., Bastakis, E., Richter, R. & Schwechheimer, C. (2016) LLM-Domain B-GATA transcription factors promote stomatal development downstream of light signaling pathways in *Arabidopsis thaliana* hypocotyls. *The Plant Cell*, 28(3), 646–660. Available from: https://doi.org/10.1105/tpc.15.00783

Landry, L.G., Chapple, C. & Last, R.L. (1995) Arabidopsis mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-b injury and oxidative damage. *Plant Physiology*, 109(4), 1159–1166. Available from: https://doi.org/10.1104/pp.109.4.1159

Lefort, V., Longueville, J.E. & Gascuel, O. (2017) SMS: smart model selection in PhyML. *Molecular Biology and Evolution*, 34(9), 2422–2424. Available from: https://doi.org/10.1093/molbev/msx149

Lentjes, M.H., Niessen, H.E., Akiyama, Y., de Bruïne, A.P., Melotte, V. & van Engeland, M. (2016) The emerging role of GATA transcription factors in development and disease. *Expert Reviews in Molecular Medicine*, 18, e3. Available from: https://doi.org/10.1017/erm.2016.2

Li, Z., Wakao, S., Fischer, B.B. & Niyogi, K.K. (2009) Sensing and responding to excess light. *Annual Review of Plant Biology*, 60, 239–260. Available from: https://doi.org/10.1146/annurev.arplant.58.032806.103844

Livak, K.J. & Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. *Methods*, 25(4), 402–408. Available from: https://doi.org/10.1006/meth.2001.1262

Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albou, L.P., Mushayamaha, T. et al. (2021) PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API.

Nucleic Acids Research, 49(D1), D394–D403. Available from: https://doi.org/10.1093/nar/gkaa1106

Montané, M.H. & Kloppstech, K. (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. Available from: https://doi.org/10.1016/s0378-1119(00)00413-3

Naito, T., Kiba, T., Koizumi, N., Yamashino, T. & Mizuno, T. (2007) Characterization of a unique GATA family gene that responds to both light and cytokinin in *Arabidopsis thaliana*. *Bioscience, Biotechnology, and Biochemistry*, 71(6), 1557–1560. Available from: https://doi.org/10.1271/bbb.60692

Nawy, T., Bayer, M., Mravec, J., Friml, J., Birnbaum, K.D. & Lukowitz, W. (2010) The GATA factor HANABA TARANU is required to position the proembryo boundary in the early arabidopsis embryo. *Developmental Cell*, 19(1), 103–113. Available from: https://doi.org/10.1016/j.devcel.2010.06.004

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. Available from: https://doi.org/10.3389/fpls.2016.01950

Ranftl, Q.L., Bastakis, E., Klermund, C. & Schwechheimer, C. (2016) LLM-domain containing B-GATA factors control different aspects of cytokinin-regulated development in *Arabidopsis thaliana*. *Plant Physiology*, 170(4), 2295–2311. Available from: https://doi.org/10.1104/pp.15.01556

Reyes, C., Muro-Pastor, M.I. & Florencio, F.J. (2004) The GATA family of transcription factors in arabidopsis and rice. *Plant Physiology*, 134(4), 1718–1732. Available from: https://doi.org/10.1104/pp.103.037788

Richter, R., Behringer, C., Müller, I.K. & Schwechheimer, C. (2010) The GATA-type transcription factors GNC and GNL/CGA1 repress gibberellin signaling downstream from DELLA proteins and PHYTOCHROME-INTERACTING FACTORS. *Genes & Development*, 24(18), 2093–2104. Available from: https://doi.org/10.1101/gad.594910

Saldanha, A.J. (2004) Java treeview--extensible visualization of microarray data. *Bioinformatics*, 20(17), 3246–3248. Available from: https://doi.org/10.1093/bioinformatics/bth349

Samarut, É., Lissouba, A. & Drapeau, P. (2016) A simplified method for identifying early CRISPR-induced indels in zebrafish embryos using high resolution melting analysis. *BMC Genomics*, 17, 547. Available from: https://doi.org/10.1186/s12864-016-2881-1

Scazzocchio, C. (2000) The fungal GATA factors. *Current Opinion in Microbiology*, 3, 126–131.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. (2012) NIH image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), 671–675. Available from: https://doi.org/10.1038/nmeth.2089

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. Available from: https://doi.org/10.1146/annurev-arplant-072221-092913

Sugano, S.S., Shirakawa, M., Takagi, J., Matsuda, Y., Shimada, T., Hara-Nishimura, I. et al. (2014) CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* L. *Plant and Cell Physiology*, 55(3), 475–481. Available from: https://doi.org/10.1093/pcp/pcu014

Tanabata, T., Shibaya, T., Hori, K., Ebana, K. & Yano, M. (2012) SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis. *Plant Physiology*, 160(4), 1871–1880. Available from: https://doi.org/10.1104/pp.112.205120

Tian, F., Yang, D.C., Meng, Y.Q., Jin, J. & Gao, G. (2019) PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Research*, 48(D1), D1104–D1113. Available from: https://doi.org/10.1093/nar/gkz1020

Tikkanen, M., Mekala, N.R. & Aro, E.M. (2014) Photosystem II photoinhibition-repair cycle protects photosystem I from irreversible damage. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1837(1), 210–215. Available from: https://doi.org/10.1016/j.bbabio.2013.10.001

Trojak, M. & Skowron, E. (2017) Role of anthocyanins in highlight stress response. *World Scientific News*, 81, 150–168.

Tzvetkova-Chevolleau, T., Franck, F., Alawady, A.E., Dall'Osto, L., Carrière, F., Bassi, R. et al. (2007) The light stress-induced protein ELIP2 is a regulator of chlorophyll synthesis in *Arabidopsis thaliana*. *The Plant Journal*, 50(5), 795–809. Available from: https://doi.org/10.1111/j.1365-313X.2007.03090.x

Wang, Z.P., Xing, H.L., Dong, L., Zhang, H.Y., Han, C.Y., Wang, X.C. et al. (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in arabidopsis in a single generation. *Genome Biology*, 16(1), 144. Available from: https://doi.org/10.1186/s13059-015-0715-0

Whipple, C.J., Hall, D.H., DeBlasio, S., Taguchi-Shiobara, F., Schmidt, R.J. & Jackson, D.P. (2010) A conserved mechanism of bract suppression in the grass family. *The Plant Cell*, 22(3), 565–578. Available from: https://doi.org/10.1105/tpc.109.073536

Yasui, Y., Tsukamoto, S., Sugaya, T., Nishihama, R., Wang, Q., Kato, H. et al. (2019) GEMMA CUP-ASSOCIATED MYB1, an ortholog of axillary meristem regulators, is essential in vegetative reproduction in *Marchantia polymorpha*. *Current Biology*, 29(23), 3987–3995. Available from: https://doi.org/10.1016/j.cub.2019.10.004

Zhao, Y., Medrano, L., Ohashi, K., Fletcher, J.C., Yu, H., Sakai, H. et al. (2004) HANABA TARANU is a GATA transcription factor that regulates shoot apical meristem and flower development in arabidopsis. *The Plant Cell*, 16(10), 2586–2600. Available from: https://doi.org/10.1105/tpc.104.024869

Zubo, Y.O., Blakley, I.C., Franco-Zorrilla, J.M., Yamburenko, M.V., Solano, R., Kieber, J.J. et al. (2018) Coordination of chloroplast development through the action of the GNC and GLK transcription factor families. *Plant Physiology*, 178(1), 130–147. Available from: https://doi.org/10.1104/pp.18.00414

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Schröder, P., 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*, 46, 2376–2390. https://doi.org/10.1111/pce.14629