RESEARCH

Open Access

Natural genetic variation in *GLK1*-mediated photosynthetic acclimation in response to light



Abstract

Background GOLDEN-like (GLK) transcription factors are central regulators of chloroplast biogenesis in Arabidopsis and other species. Findings from Arabidopsis show that these factors also contribute to photosynthetic acclimation, e.g. to variation in light intensity, and are controlled by retrograde signals emanating from the chloroplast. However, the natural variation of GLK1-centered gene-regulatory networks in Arabidopsis is largely unexplored.

Results By evaluating the activities of GLK1 target genes and GLK1 itself in vegetative leaves of natural Arabidopsis accessions grown under standard conditions, we uncovered variation in the activity of GLK1 centered regulatory networks. This is linked with the ecogeographic origin of the accessions, and can be associated with a complex genetic variation across loci acting in different functional pathways, including photosynthesis, ROS and brassinosteroid pathways. Our results identify candidate upstream regulators that contribute to a basal level of GLK1 activity in rosette leaves, which can then impact the capacity to acclimate to different environmental conditions. Indeed, accessions with higher GLK1 activity, arising from habitats with a high monthly variation in solar radiation levels, may show lower levels of photoinhibition at higher light intensities.

Conclusions Our results provide evidence for natural variation in GLK1 regulatory activities in vegetative leaves. This variation is associated with ecogeographic origin and can contribute to acclimation to high light conditions.

Keywords Natural variation, Arabidopsis thaliana, GLK1, Acclimation

Background

Photosynthetic activity is essential for the plant to capture energy from the sunlight and convert it to chemical energy needed for its activities, including growth and reproduction. At the same time, the plant also needs to

*Correspondence: Jose M. Muino jose.muino-acuna@bfr.bund.de Kerstin Kaufmann

kerstin.kaufmann@hu-berlin.de

¹ Plant Cell and Molecular Biology, Institute of Biology, Humboldt-

Universität zu Berlin, Philippstr. 13, 10115 Berlin, Germany

² Current Address: German Federal Institute for Risk Assessment (BfR), German Centre for the Protection of Laboratory Animals (Bf3R), Max-Dohrn-Straße 8-10, 10589 Berlin, Germany

³ Plant Molecular Biology, Faculty of Biology, Ludwig-Maximilians-University Munich, Planegg-Martinsried, Munich, Germany control the imbalances between redox reactions created during photosynthesis, e.g. to avoid the damage produced by reactive oxygen species (ROS) [1–4]. In nature, these processes need to be coordinated under changing and fluctuating environmental conditions [5]. For example, northern European habitats have more dramatic seasonal changes in light and temperature conditions than Mediterranean climatic regions. Indeed, plants show a large variation in photosynthetic acclimation responses depending on their natural ecogeographic habitat range [5–7]. Although many studies have focused on how plants acclimate to a particular environment, little is known about the molecular mechanisms underlying the plasticity of acclimation responses to different environments.

Plants are able to acclimate photosynthetic activity to dynamic light conditions [4, 5], including under-using



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicate otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

their photosynthetic capacity to avoid photooxidative stress [8]. They employ mechanisms that enable them to cope with excessive light or imbalance in light-dependent reactions and Calvin–Benson–Bassham (CBB) cycle, including thermal dissipation of excitation energy. Non photochemical quenching (NPQ) is for example associated with activity of the xanthophyll cycle and protonation of PSII antenna proteins [8]. Retrograde signaling pathways arising in the chloroplast evoke nuclear gene expression responses to acclimate photosynthetic activity and alleviate photooxidative stress [9, 10]. A number of metabolites including heme and carotenoids, along with reactive oxygen species (ROS) and cytosolic sugar levels have been proposed to act in retrograde signaling [9, 11].

GOLDEN2-LIKE 1 (GLK1) and GLK2 TFs are main regulators of chloroplast biogenesis and photosynthetic activity in land plants [12–15]. In Arabidopsis, GLK TFs promote the expression of many nuclear-encoded photosynthetic genes that are associated with chlorophyll biosynthesis and light-harvesting antenna proteins [16]. They bind and potentially regulate genes involved in acclimation to high light conditions [17, 18]. The *glk1 glk2* double mutant exhibits a reduction in NPQ [18], consistent with lowered Chl *b* levels.

The brassinosteroid (BR) pathway was shown to regulate GLK1 activity during chloroplast biogenesis. The BR-activated transcription factor BRASSINAZOLE-RESISTANT 1 (BZR1) and its interaction partner PHY-TOCHROME INTERACTING FACTOR 4 (PIF4) directly repress the expression of *GLK1* [19]. At posttranslational level, the light-dependent GSK-3-like kinase BRASSINOSTEROID INSENSITIVE 2 (BIN2) phosphorylates GLK1, thereby promoting protein stability [20]. BIN2 is also known to phosphorylate the GLK1 repressor BZR1 resulting in reduced activity [21].

GLK1 target gene networks have been characterized previously [16, 17, 22]. This offers an interesting resource and starting point to investigate natural variation in transcriptional regulatory networks underlying photosynthetic acclimation and its plasticity under different local habitat environments. Arabidopsis with its wide ecogeographical habitat range and genomic resources provides an excellent model system.

Results

Gene expression diversity in the GLK1 regulatory network across *Arabidopsis* accessions

To understand the natural diversity and potential adaptation of the nuclear gene regulatory network controlling photosynthetic activity to different natural environments, we characterized the expression variation of GLK1 candidate target genes across 727 *Arabidopsis* accessions collected from a wide range of habitats. We previously performed chromatin immunoprecipitation followed by high throughput sequencing (ChIP-seq) for GLK1 [17]. As expected, those data confirmed that GLK1 directly binds to genes with central roles in chlorophyll biosynthesis and photosynthesis activity, including genes needed for acclimation to high-light levels. To be conservative, we only used the top 100 GLK1 binding sites (FDR $< 7.2^{*}10^{-11}$), and we defined a gene as a direct GLK1 target when its DNA-binding site resides within the region from 1 kb upstream to 1 kb downstream of the gene. This resulted in 136 candidate GLK1 target genes. To study their expression diversity across a large group of Arabidopsis accessions, we re-analyzed publicly available transcriptome data (RNA-seq) generated from rosettes of 727 Arabidopsis accessions collected just before bolting and grown under standard greenhouse long day conditions [23]. These accessions were collected from a global Arabidopsis habitat range (see Suppl. Fig. 1a). Figure 1a shows the expression pattern of these 136 GLK1 potential target genes across the 727 accessions. We observed three main clusters of genes (Fig. 1a, Table S1): cluster G1 (n=68) was significantly enriched in photosynthesis and related Gene Ontology (GO) terms (Suppl. Fig. 2), G2 (n = 63) showed a small (non-significant) enrichment of genes involved in circadian rhythm (FDR < 0.14, Suppl. Fig. 2). Cluster G3 contained only 5 genes and we did not test for GO overrepresentation due to low sample size. When testing for the enrichment of transcriptional regulators using ShinyGO v0.77 [24], G1 genes were mostly enriched in targets of SQUAMOSA-BOX BINDING PROTEIN-LIKE 7 (SPL7) (Suppl. Fig. 2). SPL7 is a regulator of copper signaling that also acts upon ROS-detoxifying enzyme activities in plastids [25]. Genes in cluster G2 were most strongly enriched in CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) target genes (Suppl. Fig. 2). Together, the analyses suggest that cluster G2 represents potential target genes of GLK1 not directly related to photosynthetic activity. Indeed, GLK1 has been reported to also be involved in other processes such as anthocyanin biosynthesis and flowering [26, 27].

Interestingly, the clustering algorithm identifies 3 different groups of *Arabidopsis* accessions with specific gene expression patterns (Fig. 1a, Table S2): accessions in group A1 (n=215) show a much lower expression of genes from group G1 (mainly photosynthetic genes) and of *GLK1* itself (Fig. 1b) compared to A2 (n=479) accessions. Meanwhile, candidate GLK1 target genes from the G2 group did not show a strongly reduced expression like the G1 genes. Group A3 (n=33), which includes *Col*-0, was very similar in expression to A2 accessions, although with lower expression levels of *GLK2*. Genes of other known proteins affecting *GLK1* expression also show distinctive expression patterns in these accession



Fig. 1 Gene expression diversity of GLK1 candidate target genes across 727 Arabidopsis accessions. **A** Heatmap showing the relative gene expression of the candidate target genes (n = 136) of the top 100 GLK1 binding events across 727 Arabidopsis accessions. Rows represent genes and column accessions. Three clusters of accessions and three clusters of genes were identified and labeled in the heatmap depending on their particular expression patterns. **B** Expression of *GLK1* in the three clusters identified in Fig. 1A. The expression is significantly different (t-test; p-value< 2*10⁻¹⁶) between A1 vs A2, and A1 vs A3, and slightly significant (p-value< 0.03) between A2 vs A3. **C** Distribution of the monthly standard deviation (SD) of solar radiation of the coordinates from the regions where the accessions were collected. Groups A1 and A2 were significantly different (p-value< 10⁻⁷; t-test). **D** Map representing the location of the 727 *Arabidopsis* accessions. Only a subsection of the map is plotted, see Suppl. Fig. 1 for a complete map. A1 accessions are plotted in blue, A2 in red, and A3 in green

groups (Suppl. Fig. 3): While GENOMES UNCOUPLED 1 (GUN1) shows an expression pattern that is essentially opposite to GLK1 in A1 vs. A2, GLK2 and PIF4 show a pattern similar to GLK1. Figure 1d shows the distribution of accessions in Europe, see Suppl. Fig. 1a for their global distribution. We focused our further analysis on clusters A1 and A2, since they show the most distinct gene expression profiles and contain most of the Arabi*dopsis* accessions. The A1 accessions were collected from regions enriched in countries such as GBR and USA, while A2 accessions were enriched e.g. in Germany, Sweden, and especially Spain (Supp. Fig. 1a). This indicates that they may represent adaptation to different climatic environments. To further investigate this, we downloaded environmental variable values associated with the places of collection from WorldClim (version 2.0) [28]. Among the variables studied, we found that the intermonthly standard variation of solar radiation (p-value $<1.1*10^{-7}$; t-test) was the most significantly different average between groups A1 and A2. Accessions from group A1 originate mainly from environments with lower inter-monthly variation in solar radiation (Fig. 1c). See Table S3 for p-values of other variables studied.

Next, we aimed to characterize the general differences in gene expression patterns of the A1 and A2 accession groups. When analyzing the expression of all nuclear-encoded genes using the DESeq2, an R package for RNA-seq data analysis, we identified 1179 genes as significantly differentially expressed in A1 compared to A2 accessions (FDR < 0.01 and abs $\log_2 FC > 2$; Fig. 2a). 824 of those genes were more active in accession group A1. On the other hand, 366 genes were more strongly expressed in the accession group A2. Together, these analyses suggest that the A1 and A2 accession groups show clear differences in genome-wide gene activities. To better characterize these genes, we performed GO enrichment analysis (Fig. 2b-c). Here, we found that genes related to reactive oxygen species (ROS) signaling are highly expressed in A1 (so linked with low GLK1 expression), while genes involved in defense responses are more strongly expressed in accession group A2 (so associated with high GLK1 activity), which is in line with



Fig. 2 Differential gene expression signature in clusters A1 vs A2. **A** Scatterplot showing the average expression of all genes in cluster A1 vs A2. In blue are indicated differentially expressed genes (FDR < 0.01 & abs log2FC > 2), and in red the potential direct targets of GLK1 represented in Fig. 1A. **B** GO analysis (Biological process, ShinyGO) of genes with a significantly increased expression in A1 vs A2. **C**) GO analysis (Biological process, ShinyGO) of genes with a significantly increased expression in A1 vs A2. **C**) GO analysis (Biological process, ShinyGO) of genes with a significantly increased expression in A2 vs A1

a role of GLK1 in disease resistance based on previous results [29].

Next, we studied the phylogenetic relationship among these accessions. The 1001 genomes consortium defined several admixture groups among the Arabidopsis accessions depending on their genome sequence similarity [30]. Indeed, we see that group A1 is enriched in the groups "Western Europe", "Germany" and "Italy Balkan Caucasus", while group A2 is enriched in the admixture groups "Central Europe", "South-Sweden" and "Spain" (Suppl. Fig. 1c), which indicates that different genetic structure is associated with the two accession groups.

In summary, we have identified at least two groups of *Arabidopsis* accessions with distinct transcriptome profiles related to *GLK1* activity. These two groups were collected from eco-geographic locations with different environmental properties, with solar radiation being the most statistically different (see Table S3 for the list and significance of all variables studied). These accessions represent different genetic structures, as they belong to different admixture groups.

The genetic structure behind different *GLK1* activities across *Arabidopsis* accessions

Since we observed a different proportion of admixture groups in our *Arabidopsis* accession groups, we decided to characterize their genetic differences in more detail. We utilized the Single Nucleotide Polymorphisms (SNPs) and short insertions/deletions obtained from the 1001 Arabidopsis genomes project [30] to perform genome-wide association analysis (GWAS) to identify SNPs separating groups A1 and A2 (Chi-square test, PLINK software; Suppl. Fig. 4). A large number of SNPs (n=77,424) were significantly associated with the trait at study (p-value<10⁻¹⁰; Table S4), indicating a complex genetic structure behind the phenotype studied. We filtered SNPs located outside exons or with a frequency

lower than 25% in A1 or A2 accession groups. Among the remaining 6918 SNPs, 115 SNPs were estimated by Variant Effect Prediction (VEP) to have a "HIGH" impact on their protein activity. Among them, 98 were predicted to introduce a new stop codon in the protein sequence. Figure 3 shows the genetic structure of the A1 and A2 groups, for SNPs predicted to have a "HIGH" or "MOD-ERATE" impact on protein function.

Genetic variation with a "HIGH" impact on protein activity was found for genes that potentially could affect the GLK1 gene regulatory network. For example, FTSH PROTEASE 12 (FTSH12) serves as an import motor involved in the ATP-dependent translocation of preproteins through the TIC channel and is important for the retrograde/anterograde signaling pathways [31]. The mutation of this gene in Col-0 is reported to be embryo lethal [32]. However, many of our *Arabidopsis* accessions carry a premature stop codon in FTSH12 without causing lethality which indicates a different gene-regulatory re-wiring that can compensate for this mutation. The mutation is causing a premature stop codon in the third exon of all three described isoforms of the gene. Whether this stop codon entirely abolishes translation of a fulllength protein needs to be experimentally tested in future research. Another interesting example is BRASSINOS-TEROID INACTIVATOR 2 (BIA2), a gene that controls brassinosteroid homeostasis. Its overexpression is associated with change in expression of genes involved in photosynthesis (including GLK1), carbohydrate metabolism, and circadian rhythms [33]. Also affecting the brassinosteroid pathway, PLANT U-Box 41 (PUB41) acts in the degradation of BZR1 [34]. Another gene affected encodes the non-expressor of pathogenesis-related (NPR) transcription factor BLADE-ON-PETIOLE 2 (BOP2) that is regulated by ROS signaling, controls boundary establishment in development and acts in plant defense [35]. Other genes harboring mutations in subsets of the



Fig. 3 Genetic variation behind the different GLK1 activity. Heatmap showing SNPs among the Arabidopsis accessions studied. Only SNPs differently associated with A1 and A2 accession groups and with a predicted high or moderate impact on gene function are shown. Genes predicted to have a high impact and mentioned in the manuscript are labeled on the right side

accessions are e.g. *GLK1* itself and *STROMAL ASCOR-BATE PEROXIDASE* (*SAPX*) that mediates ROS scavenging in the chloroplast stroma [36].

In summary, we identified a complex genetic structure behind these two distinct clusters of Arabidopsis accessions. Focusing on SNPs predicted to cause a change in protein activity (i.e. by introducing premature stop codons), we determined candidate genes that may contribute to natural variation in the levels of GLK1 activity and thereby differential regulation of its target networks.

Gene-regulatory network analysis

To better understand the differences in gene activities observed between A1 and A2 accession groups, we predicted the gene regulatory network (GRN). Based on the analysis of genetic polymorphisms, we focused on genes involved in GO categories related to chlorophyll metabolism, brassinosteroid signaling and retrograde signaling (see Methods; n=216 genes) using GeneNet [37] for the group of A1 and A2 accessions independently. We did not preselect for TFs when running GeneNet as many regulators of the chlorophyll metabolism are not only TFs, but also e.g. transporters and posttranscriptional or posttranslational regulators, such as RNA binding proteins. GeneNet infers large-scale gene association networks using graphical Gaussian models that represent dependencies among genes calculated based on partial correlations. Partial correlations measure the dependence of two variables given a set of confounding variables (other genes in this context) and therefore provide a closer measure of direct effects.

We initially used the GRN of A2 as a reference as it is more similar at the *GLK1* target expression level to *Col*-0 compared to A1 (Fig. 1a). Among the 216 genes

considered, 1394 interactions were found significant in the A2 accessions (p-value<0.05; Table S5). To predict genes regulating GLK1 or genes acting in the same regulatory module, we selected the top 36 gene interactions in the neighborhood of GLK1 predicted by GENENET (p-value< 0.01), and plotted these interactions in a graphical network (Fig. 4; see Suppl. Fig. 5 for the GRN based on A1 for comparison). GLK1 was predicted to be associated with (in order of importance) GLK2, LZF1, GNC, BEH1, and BCM1. Indeed, GLK2 may cross-regulate with GLK1 as they bind each other's promoters [17, 22]. LZF1/ BBX22 is a B-box protein involved in the positive regulation of photomorphogenesis [38]. Other B-box TFs are regulated by GLK1 [17, 27]. GNC and GLK1 have overlapping functions coordinating chloroplast development [18]. BEH1 is a TF homolog of BEZ1 controlling brassinosteroid signaling and photosynthetic activity [39]. Lastly, BCM1 encodes a Mg-dechelatase that catalyzes chlorophyll *a* degradation and is regulated by GLK1 [40].

In a next step, we determined the GRN for the set of A1 accessions (Table S6). In order to identify the main differences in the interactions of these two networks, we plotted the interaction strength predicted by GENENET in the A1 vs. A2 GRNs, for all interactions found significant (*p*-value<0.01) in the A2 GRN. Among all the interactions, the links between BIA2-CYP722A1, EXO-THE1 and ELIP1-ELIP2 were the most different. For the BIA2-CYP722A1 interaction, this could be explained by the fact that our GWAS analysis detected an SNP causing

a new stop codon in BIA2 with different frequencies in the A1 and A2 accessions. The cell-wall localized protein EXORDIUM (EXO) and the receptor kinase THESEUS1 (THE1) control cell growth and act in the brassinosteroid pathway [41, 42]. ELIP1 and ELIP2 are two chlorophyll a/b binding proteins that have been suggested to play roles in protection from photo-oxidative stress [43], and thus may contribute to environmental acclimation pathways converging on GLK1 and GLK2 activities. In summary, by comparing the gene regulatory network centered on GLK1 activity in A1 vs A2, we revealed multiple links with specific genes in the brassinosteroid pathway, with a candidate gene affecting the natural variation in the levels of GLK1 activity being BIA2. Indeed, previously it was found that BIA2 overexpression (BIA2-OE) leads to elevated GLK1 activity (FDR < 4*10^-22) and GLK2 (FDR < 1.36*10^-4) [33]. Additionally, we found that the bia2 loss-of function alleles in natural accessions are correlated with low GLK1 activity (Suppl. Fig. 6a). A comparison of transcriptome levels in accession cluster A1 and A2 showed that accessions with high expression of BIA2 (A1) show a BIA2-OE transcriptome signature (Suppl. Fig. 6b).

GLK1 expression variation is linked with fluctuations in solar radiation

Given the difference in genetic structure and photosynthesis-related expression patterns between the two groups of Arabidopsis accessions studied here, and their



Fig. 4 Gene regulatory network differences. A Predicted GRN around *GLK1* in the A2 group of accessions. B Interaction strength predicted in the A2 accessions (y-axis) versus A1 accessions (x-axis) for all possible interactions of the genes depicted in **A**. The continuous black line is the diagonal, dotted black lines represent a difference of 0.2 points from the diagonal line. Red points indicate the interactions whose strength is the difference of 0.2 points indicated by the dotted black lines

association with different natural environments, we speculated that these genetic and expression changes may represent an adaptation to different environmental conditions. Indeed, these two subpopulations are significantly associated with different distributions of inter-monthly solar radiation variation. The average solar radiation in a particular month depends on the day length and the solar intensity during the day. Countries with a variable day length (e.g. Sweden), or with a variable solar intensity across the year (e.g. Spain) have a more variable solar radiation than other countries, such as the GBR. Indeed, long or short-day conditions are reported to impact the gene activities associated with photosynthetic pathways [44]. Also, excess light can result in the production of ROS, and plants have evolved the mechanism of non-photochemical quenching (NPQ) to dissipate excess light energy as heat to protect the photosynthetic machinery (for review, see [45]). Rungrat et al. showed that Arabidopsis adapted to environments with different solar and temperature levels by adjusting their NPQ activity [46]. They profiled 284 Arabidopsis accessions in two sets of conditions reflecting seasonal and regional (coastal vs. inland) climates, by simulating specific gradients in light intensity and temperature. The maximum light intensity was $150 \,\mu\text{mol}\,\text{m}^{-2}\,\text{s}^{-1}$ (coastal) or $300 \,\mu\text{mol}\,\text{m}^{-2}\,\text{s}^{-1}$ (inland), and photosynthetic parameters were measured by pulse amplitude modulation (PAM) fluorometry. Among the profiled accessions, 12 belong to our group A1, and 32 to the A2 group.

Comparing photosynthetic parameters across conditions and accession groups (Fig. 5a) revealed that maximum quantum efficiency of photosystem II as measured by dark-adapted Fv/Fm (QY-max) is higher under



Fig. 5 Photosynthetic activity measures across Arabidopsis accessions. The two groups of accessions studied have different photosynthetic activity. **A** Average photosynthetic value (column) among the accession groups and environmental conditions considered (rows). **B** Distributions for QY_max, the distributions for A2 accessions are not significantly different between the conditions studied (t-test, *p*-value<0.543). A1 accessions are significantly different when growing in the two environments (t-test, *p*-value<0.010). **C** Distributions for NPQ_Lss, the distributions for A1 accessions are not significantly different between the conditions are significantly different when growing in the two environments (t-test, *p*-value<0.010). **C** Distributions for NPQ_Lss, the distributions for A1 accessions are not significantly different between the conditions studied, however, A2 accessions are significantly different when growing in different environments considered (t-test, *p*-value<2.6*10⁻⁵). **D** Distributions for the parameter QY_Lss in the A1 and A2 accession groups for the two different sets of conditions

inland conditions than under coastal conditions for accession group A1 (t-test, *p*-value<0.010). In contrast, QY-max of A2 accessions is not significantly different between inland and coastal conditions (p-value<0.543) (Fig. 5b). In line with this, A2 accessions display a significant increase in the photoprotection parameter NPQ_Lss (steady state non-photochemical quenching) (t-test, *p*-value<2.6*10–5, Fig. 5c), comparing "coastal" to "inland" conditions. There are no significant differences in NPQ Lss for the A1 accessions when comparing these two environmental conditions (p-value<0.5). It is tempting to speculate that the A2 accessions, natural from countries like Spain and Sweden with high variation in solar radiation levels, show a stronger adaptive capacity in protection against photoinhibition, while photosynthetic efficiency (QY_Lss) is similar in A1 and A2 (Fig. 5d). In contrast to that, the A1 accessions, natural from low variation solar radiation environments like GBR, may not have the capacity to modulate protection against photoinhibition in adaptation to seasonal light differences.

Discussion

The acclimation of photosynthetic activities to variable environmental regimes requires regulatory adjustments at post-translational, post-transcriptional and transcriptional levels, and the coordination of gene activities in both plastid and nuclear genomes. GLK transcription factors function not only as central regulators of chloroplast biogenesis, but have also been found to act in response to abiotic and biotic environmental factors. Here, we showed that the levels of GLK1 activity in rosette leaves vary across natural Arabidopsis accessions, and that this variation is linked with eco-geographic environmental factors, especially seasonal fluctuations in solar radiation levels. GWAS and gene-regulatory network analyses predicted brassinosteroids and ROS signaling as potential upstream pathways modulating GLK1 activity in the different accession groups.

Acclimation to varying light conditions is mediated by a combination of short-term metabolic shifts and posttranslational regulation of the photosynthetic apparatus. Prolonged exposure to high-light conditions is mediated by additional adjustment in the photosynthetic status, e.g. mediated by changes in enzyme and antenna protein composition, and stomatal aperture [47]. This in turn results in changes in the maximum photosynthetic rate (Pmax). These acclimation responses require retrograde signaling from the plastids to the nucleus. Different types of retrograde signals mediating plant acclimation to environmental conditions have been proposed, including ROS signaling and cytosolic sugar levels [10]. GLK1 is a target of GUN1-mediated retrograde signaling [48]. This has been suggested to help the plants protect against photo-oxidative damage. At the posttranslational level, GLK1 was found to be regulated by retrograde signals that influence protein stability [49]. More recently, it was shown that BIN2-mediated phosphorylation of GLK1 promotes chloroplast biogenesis by enhancing GLK1 protein stability and activity in the light [20]. This complexity of transcriptional and posttranscriptional regulation makes GLK1 a candidate factor for mediating acclimation of photosynthetic activity to different habitats. Our findings suggest that GLK1 activity is genetically variable across Arabidopsis accessions, and that the variation is mediated by variants in different genes that contribute to fine tuning GLK1 expression. These genes include known regulatory pathways linked with GLK1, for example brassinosteroid signaling genes such as BIA2, a regulator of brassinosteroid homeostasis [33]. Another factor that is associated with variation in GLK1 activity is BLADE-ON-PETIOLE 1 (BOP1), which was also shown to be regulated by brassinosteroids. Interestingly, BOP1 is a TGACG sequence-specific binding protein (TGA) TF and shows conserved protein residues that are known to be regulated in other TGA TFs by the intracellular redox status [50], thus providing a potential link with ROS signaling. The high number of genetic variants potentially contributing to variation in GLK1 gene activities in rosette leaves across the studied accessions suggest frequent and evolutionary dynamic adaptation of this regulatory system to local habitats. Indeed, we identified a different ecogeographic distribution of accessions with low vs. high GLK1 levels. The complexity of genetic variation suggests multiple origins of genetic variation contributing to the diversity of GLK1 upstream pathways. To which extent this is linked with coupled genetic variation in the downstream pathways of GLK1 will be an interesting question for future research.

Conclusions

In summary, our findings suggest that variation in GLK1 regulatory pathways, such as the brassinosteroid pathway, mediates differences in the regulation of basal *GLK1* levels in different Arabidopsis accessions. Together with variation in downstream networks, this results in different adaptive potential to environmental conditions, in particular fluctuating light activities, by elevating or lowering the responsiveness of GLK1 target genes. Future experiments should test the roles of the identified candidate TFs and pathways, in particular brassinosteroid and ROS signaling in the regulation of GLK1 responsiveness in different natural environments. This will also shed

light on the interplay between biogenic and operational retrograde signaling mechanisms converging on this important transcription factor.

Methods

RNA-seq data analysis

Fastq files from publicly available RNA-seq data were downloaded from Sequence Read Archive (SRA; https:// www.ncbi.nlm.nih.gov/sra) with ID PRJNA319904. Individual fastq files were trimmed with Trimmomatic v0.36 (default parameters), and mapped to the TAIR9 Arabidopsis genome sequence and the Araport11 gene annotation using STAR v2.7.2b with options --outFilterMultimapNmax 2 --outMultimapperOrder Random --alignIntronMax 3000 --outSAMstrandField intron-Motif --outFilterIntronMotifs RemoveNoncanonical. Mapped reads were assigned to genes (intron/exon) using featureCounts from the package subread v1.6.4 with parameters -s 2 -p -C -M -t gene -g gene_id. Non protein-encoding genes or genes encoded in the organelle genomes were filtered out at this step. Next, individual sequencing libraries with less than 10⁶ mapped reads were filtered out. The remaining libraries were combined depending on the Arabidopsis accession used. This resulted in raw read counts for 727 Arabidopsis accessions and 27,445 genes. Next, the data was normalized using the function varianceStabilizingTransformation from the R package Deseq2 v1.34.0. When needed, relative expression per gene was calculated as the expression of a particular gene in one particular Arabidopsis accession minus the average expression of that gene across all accessions.

The function *pheatmap* from the R package *pheatmap* v1.0.12 was used to cluster the relative expression of the candidate GLK1 targets across all Arabidopsis accessions studied. For this, the top 100 GLK1 binding sites were obtained from [17]. Candidate GLK1 targets were defined as protein-coding genes with a top 100 GLK1 binding sites in the 1 kb upstream, inside, and 1 kb downstream region of the gene. This resulted in 136 genes.

Climate data

Average monthly climate data were downloaded from the WorldClim v2.0 database using a spatial resolution of 30s (~1 km²) across the years 1970–2000 [28]. The R package RGDAL v1.6–7 was used to obtain the climate data from the location where the different *Arabidopsis* accession were collected as reported by the 1001Genome project. The average climate values were obtained by calculating the average value across the 12 months of the year, and the standard variation was obtained by calculating the standard deviation across the 12 months of the year.

GWAS analysis

SNP data for the *Arabidopsis* accession studied were downloaded from the 1001 Genome project v3.1, file name: 1001genomes_snp-short-indel_only_ACGTN.vcf. gz which only contains genetic variants located in the nuclear genome. GWAS analysis (Chi-square test for the A1 versus A2 groups) was performed with the *Plink* software v1.90b6.21, with parameters: *---assoc -allow-no-sex --pfilter 0.05.* In addition, only variants with a *p*-value lower than 10⁻¹⁰, with a frequency bigger than 25% in the A1 or A2 *Arabidopsis* accession groups as estimated by Plink were retained for further analysis, and plots. The impact of SNPs was predicted using Variant Effect Prediction (VEP) provided by EnsemblPlants (https://plants.ensembl.org/Arabidopsis_thaliana/Tools/VEP). This tool uses the TAIR10 gene annotation.

Gene regulatory network reconstruction

Gene expression profiles were obtained using the analysis described before. Only genes belonging to the next gene ontologies and their children were used: "chlorophyll metabolic process" (GO:0015994), "chloroplast-nucleus signaling" (GO:0010019), "brassinosteroid homeostasis" (GO:0010268) and "brassinosteroid mediated signaling pathway" (GO:0009742). The list of genes belonging to these ontologies was downloaded from the TAIR website (https://www.arabidopsis.org). This results in 214 genes. Gene regulatory networks were estimated for the A1 and A2 accession groups independently using the function ggm. estimate.pcor of the R package GeneNet with parameters method="dynamic". Only edges with a p-value associated lower than 0.05 were considered. To predict GLK1 regulatory network (Fig. 4), we selected genes in the GLK1 network as the genes with a p-value<0.01 to be associated with GLK1. To provide more depth to the network, genes directly associated with the previously selected GLK1-associated network were also selected. This was done 5 steps upstream of GLK1, which resulted in 36 genes. For these genes, the association was plotted in Cytoscape v3.10.

Re-analysis of photosynthesis activity

Photosynthesis activity data for 271 Arabidopsis ecotypes and genotypes were downloaded from the supplementary material of [46]. Only data from 'leaf 14' were used. Among their profiled accessions, 12 belong to our group A1, and 32 to the A2 group of *Arabidopsis* accessions.

Abbreviations

TF	transcription factor
GRN	gene regulatory network
GWAS	genome-wide association mapping

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s12870-024-04741-1.

Additional file 1: Table S1. List of gene TAIR IDs in clusters G1, G2 and G3. Table S2. List of Arabidopsis natural accessions for clusters A1, A2 and A3. Table S3. Mean difference (t-test) between the environmental variables (WorldClim 2.0) of the place of collection of the accession cluster A1 versus A2. Table S4. Different association (Chi-square) of genetic variants between cluster A1 versus A2. Table S5. Gene regulatory network estimated by Genenet for accession cluster A1. Table S6. Gene regulatory network estimated by Genenet for accession cluster A2.

Additional file 2: Suppl. Fig 1. Spatial distribution of the studied Arabidopsis accessions. A) World map plotting the location from where accessions were collected. A1 accessions are plotted in blue, A2 in red, and A3 in green. B) Proportion of country of origin is plotted for accessions belonging to clusters A1 and A2. C) Proportion of admixture groups as identified by the 1001 Genomes project [30]. Suppl. Fig 2. GO analysis among genes classified in the G1 and G2 groups. GO analysis was performed with ShinyGO v0.77. The GO (Biological processes) enrichment for G1 genes is shown in A, and for G2 genes is shown in B. ShinyGO can also test the enrichment among the Plant Gene Set Annotation database [51] and other databases. The enrichment using all databases in ShinyGO v 0.77 for G1 genes is shown in C and for G2 genes in D. Suppl. Figure 3. Expression of known GLK1 regulators in different Arabidopsis accession groups. The expression distribution of GLK2 (A), HY5 (B), PIF4 (C) and GUN1 (D) is shown for each group of accessions defined in Fig. 1a. Suppl. Figure 4. A Distribution of -log 10 p-values obtained by PLINK when testing (Chi-square test) for association between the group of accessions A1 versus A2 for each chromosome. Typically, a fixed p-value threshold of 5*-8 is widely used to identify association between genetic variants and a trait of interest [52]. B Quantile-Quantile (Q-Q) plot of the observed versus expected p-values. Red dashed line indicates the diagonal. Suppl. Figure 5. Predicted gene regulatory network around GLK1 estimated for the A1 group of accessions as described for Fig. 4A. Suppl. Figure 6. A Relation of GLK1 and BIA2 expression, as well as impact of mutation causing stop codon in the BIA2 gene. B Expression signature of accessions with high BIA2 expression fits with the observed consequence of BIA2 overexpression in Col-0.

Acknowledgements

The authors wish to thank Christian Schmitz-Linneweber for valuable comments on an earlier version of the manuscript. We acknowledge support by the Open Access Publication Fund of Humboldt-Universität zu Berlin.

Authors' contributions

J.M. performed the data analyses. J.M. and K.K. wrote the main manuscript text and prepared the figures. C.G. and T.K. edited the manuscript. All authors reviewed the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. K.K. wishes to thank the DFG for funding (project no. 270050988 (C05); 458750707). T.K. wishes to thank the DFG for funding (project no. 270050988 (C01)). J.M. wishes to thank the DFG for funding (project no. 407463262).

Availability of data and materials

Datasets and scripts are available upon request.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 10 October 2023 Accepted: 10 January 2024 Published online: 05 February 2024

References

- Kono M, Terashima I. Long-term and short-term responses of the photosynthetic electron transport to fluctuating light. J Photochem Photobiol B Biol. 2014;137:89–99.
- Kono M, Noguchi K, Terashima I. Roles of the cyclic Electron flow around PSI (CEF-PSI) and O2-dependent alternative pathways in regulation of the photosynthetic Electron flow in short-term fluctuating light in Arabidopsis thaliana. Plant Cell Physiol. 2014;55:990–1004.
- Pospíšil P. Production of reactive oxygen species by photosystem II as a response to light and temperature stress. Front Plant Sci. 2016;7:232551.
- Vialet-Chabrand S, Matthews JSA, Simkin AJ, Raines CA, Lawson T. Importance of fluctuations in light on plant photosynthetic acclimation. Plant Physiol. 2017;173:2163–79.
- Tanaka Y, Adachi S, Yamori W. Natural genetic variation of the photosynthetic induction response to fluctuating light environment. Curr Opin Plant Biol. 2019;49:52–9.
- Demmig-Adams B, Adams WW. Photoprotection in an ecological context: the remarkable complexity of thermal energy dissipation. New Phytol. 2006;172:11–21.
- Van RR, Aarts MGM, Harbinson J. Natural genetic variation for acclimation of photosynthetic light use efficiency to growth irradiance in Arabidopsis. Plant Physiol. 2015;167:1412–29.
- Li Z, Wakao S, Fischer BB, Niyogi KK. Sensing and responding to excess light. Annu Rev Plant Biol. 2009;60:239–60.
- Chan KX, Phua SY, Crisp P, McQuinn R, Pogson BJ. Learning the Languages of the Chloroplast: Retrograde Signaling and Beyond. Annu Rev Plant Biol. 2016;67:25–53.
- Richter AS, Nägele T, Grimm B, Kaufmann K, Schroda M, Leister D, et al. Retrograde signaling in plants: A critical review focusing on the GUN pathway and beyond. Plant Commun. 2023;4:100511.
- 11. Chi W, Feng P, Ma J, Zhang L. Metabolites and chloroplast retrograde signaling. Curr Opin Plant Biol. 2015;25:32–8.
- Fitter DW, Martin DJ, Copley MJ, Scotland RW, Langdale JA. GLK gene pairs regulate chloroplast development in diverse plant species. Plant J. 2002;31:713–27.
- Rossini L, Cribb L, Martin DJ, Langdale JA. The maize Golden2 gene defines a novel class of transcriptional regulators in plants. Plant Cell. 2001;13:1231–44.
- 14. Yasumura Y, Moylan EC, Langdale JA. A conserved transcription factor mediates nuclear control of organelle biogenesis in anciently diverged land plants. Plant Cell. 2005;17:1894–907.
- Yelina NE, Frangedakis E, Wang Z, Schreier TB, Rever J, Tomaselli M, et al. Streamlined regulation of chloroplast development in the liverwort Marchantia polymorpha. bioRxiv. 2023.01.23.525199.
- Waters MT, Wang P, Korkaric M, Capper RG, Saunders NJ, Langdale JA. GLK transcription factors coordinate expression of the photosynthetic apparatus in Arabidopsis. Plant Cell. 2009;21:1109–28.
- Atanasov V, Schumacher J, Muiño JM, Larasati C, Wang L, Kaufmann K, et al. Arabidopsis BBX14 is involved in high light acclimation and seedling development. Plant J. 2023. https://doi.org/10.1111/tpj.16597.
- Zubo YO, Blakley IC, Franco-Zorrilla JM, Yamburenko MV, Solano R, Kieber JJ, et al. Coordination of chloroplast development through the action of the GNC and GLK transcription factor families. Plant Physiol. 2018;178:130.
- Oh E, Zhu JY, Wang ZY. Interaction between BZR1 and PIF4 integrates brassinosteroid and environmental responses. Nat Cell Biol. 2012;14:802–9.
- Zhang D, Tan W, Yang F, Han Q, Deng X, Guo H, et al. A BIN2-GLK1 signaling module integrates Brassinosteroid and light signaling to repress chloroplast development in the dark. Dev Cell. 2021;56:310–324.e7.
- 21. Wang ZY, Bai MY, Oh E, Zhu JY. Brassinosteroid signaling network and regulation of photomorphogenesis. Annu Rev Genet. 2012;46:701–24.

- 22. Tu X, Ren S, Shen W, Li J, Li Y, Li C, et al. Limited conservation in crossspecies comparison of GLK transcription factor binding suggested widespread cistrome divergence. Nat Commun. 2022;13:7632.
- Kawakatsu T, Huang SS, Jupe F, Sasaki E, Schmitz RJ, Urich MA, et al. Epigenomic diversity in a global collection of Arabidopsis thaliana accessions. Cell. 2016;166:492–505.
- 24. Ge SX, Jung D, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 2020;36:2628–9.
- Araki R, Mermod M, Yamasaki H, Kamiya T, Fujiwara T, Shikanai T. SPL7 locally regulates copper-homeostasis-related genes in Arabidopsis. J Plant Physiol. 2018;224–225:137–43.
- Li Y, Lei W, Zhou Z, Li Y, Zhang D, Lin H. Transcription factor GLK1 promotes anthocyanin biosynthesis via an MBW complex-dependent pathway in Arabidopsis thaliana. J Integr Plant Biol. 2023;65:1521–35.
- Susila H, Nasim Z, Gawarecka K, Jung JY, Jin S, Youn G, et al. Chloroplasts prevent precocious flowering through a GOLDEN2-LIKE–B-BOX DOMAIN PROTEIN module. Plant Commun. 2023;4:100515.
- Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol. 2017;37:4302–15.
- Savitch LV, Subramaniam R, Allard GC, Singh J. The GLK1 "regulon" encodes disease defense related proteins and confers resistance to fusarium graminearum in Arabidopsis. Biochem Biophys Res Commun. 2007;359:234–8.
- Alonso-Blanco C, Andrade J, Becker C, Bemm F, Bergelson J, Borgwardt KMM, et al. 1,135 genomes reveal the global pattern of polymorphism in Arabidopsis thaliana. Cell. 2016;166:481–91.
- Mielke K, Wagner R, Mishra LS, Demir F, Perrar A, Huesgen PF, et al. Abundance of metalloprotease ftsh12 modulates chloroplast development in arabidopsis thaliana. J Exp Bot. 2021;72:3455–73.
- 32. Meinke D, Muralla R, Sweeney C, Dickerman A. Identifying essential genes in Arabidopsis thaliana. Trends Plant Sci. 2008;13:483–91.
- Zhang Z, Xu L. Arabidopsis BRASSINOSTEROID INACTIVATOR2 is a typical BAHD acyltransferase involved in brassinosteroid homeostasis. J Exp Bot. 2018;69:1925–41.
- Kim EJ, Lee SH, Park CH, Kim SH, Hsu CC, Xu S, et al. Plant u-box40 mediates degradation of the brassinosteroid-responsive transcription factor bzr1 in arabidopsis roots. Plant Cell. 2019;31:791–808.
- Khan M, Xu H, Hepworth SR. BLADE-ON-PETIOLE genes: setting boundaries in development and defense. Plant Sci. 2014;215–216:157–71.
- Kameoka T, Okayasu T, Kikuraku K, Ogawa T, Sawa Y, Yamamoto H, et al. Cooperation of chloroplast ascorbate peroxidases and proton gradient regulation 5 is critical for protecting Arabidopsis plants from photooxidative stress. Plant J. 2021;107:876–92.
- Opgen-Rhein R, Strimmer K. From correlation to causation networks: a simple approximate learning algorithm and its application to highdimensional plant gene expression data. BMC Syst Biol. 2007;1:37.
- Chang CSJ, Maloof JN, Wu SH. COP1-mediated degradation of BBX22/ LZF1 optimizes seedling development in arabidopsis. Plant Physiol. 2011;156:228–39.
- Otani Y, Kawanishi M, Kamimura M, Sasaki A, Nakamura Y, Nakamura T, et al. Behavior and possible function of Arabidopsis BES1/BZR1 homolog 2 in brassinosteroid signaling. Plant Signal Behav. 2022;17:2084277.
- Yamatani H, Ito T, Nishimura K, Yamada T, Sakamoto W, Kusaba M. Genetic analysis of chlorophyll synthesis and degradation regulated by BALANCE of CHLOROPHYLL METABOLISM. Plant Physiol. 2022;189:419–32.
- Schröder F, Lisso J, Lange P, Müssig C. The extracellular EXO protein mediates cell expansion in Arabidopsis leaves. BMC Plant Biol. 2009;9:1–12.
- Guo H, Ye H, Li L, Yin Y. A family of receptor-like kinases are regulated by BES1 and involved in plant growth in Arabidopsis thaliana. Plant Signal Behav. 2009;4:784–6.
- Heddad M, Adamska I. Light stress-regulated two-helix proteins in Arabidopsis thaliana related to the chlorophyll a/b-binding gene family. Proc Natl Acad Sci U S A. 2000;97:3741–6.
- Baerenfaller K, Massonnet C, Hennig L, Russenberger D, Sulpice R, Walsh S, et al. A long photoperiod relaxes energy management in Arabidopsis leaf six. Curr Plant Biol. 2015;2:34–45.
- Ruban AV. Nonphotochemical chlorophyll fluorescence quenching: mechanism and effectiveness in protecting plants from Photodamage. Plant Physiol. 2016;170:1903–16.
- Rungrat T, Almonte AA, Cheng R, Gollan PJ, Stuart T, Aro EM, et al. A genome-wide association study of non-photochemical quenching in

- Dietz KJ. Efficient high light acclimation involves rapid processes at multiple mechanistic levels. J Exp Bot. 2015;66:2401–14.
- Martin G, Leivar P, Ludevid D, Tepperman JM, Quail PH, Monte E. Phytochrome and retrograde signalling pathways converge to antagonistically regulate a light-induced transcriptional network. Nat Commun. 2016;7:11431.
- Tokumaru M, Adachi F, Toda M, Ito-Inaba Y, Yazu F, Hirosawa Y, et al. Ubiquitin-proteasome dependent regulation of the GOLDEN2-LIKE 1 transcription factor in response to plastid signals. Plant Physiol. 2017;173:524.
- Hepworth SR, Zhang Y, McKim S, Li X, Haughn GW. BLADE-ON-PETIOLE– dependent signaling controls leaf and floral patterning in Arabidopsis. Plant Cell. 2005;17:1434.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.