Natural variation in *Arabidopsis thaliana* identifies GlyoxalaseI;2 as key for the detoxification of glucose-derived reactive carbonyl species

Manuel Balparda[1*], Aylin Kerim[1], Maroua Bouzid[1], Clarisa E. Alvarez[2], Saleh Alseekh[3-4], Alisdair R. Fernie[3-4], Veronica G. Maurino[1*]

[1]Molecular Plant Physiology, Institute of Cellular Molecular Botany (IZMB), University of Bonn, Kirschallee 1, 53115 Bonn, Germany

[2]Centro de Estudios Fotosintéticos y Bioquímicos (CEFOBI-CONICET), University of Rosario, Suipacha 570, 2000 Rosario, Argentina.

[3]Max-Planck-Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

[4]Center for Plant Systems Biology and Biotechnology, 4000 Plovdiv, Bulgaria

[*]Corresponding authors: Manuel Balparda, mbalpard@uni-bonn.de and Veronica G. Maurino, vero.maurino@uni-bonn.de

ORCID: Manuel Balparda 0000-0001-5771-4790; Saleh Alseekh 0000-0003-2067-5235; Alisdair Fernie 0000-0001-9000-335X; Veronica G. Maurino 0000-0002-3124-1451.

Summary: Natural variation in the expression of the glyoxalase system in *Arabidopsis thaliana* ecotypes indicate that GlyoxalaseI;2 is mostly responsible for the detoxification of glucose-derived reactive carbonyl species.

Key words: glyoxalase system, natural variation, oxidative stress, reactive carbonyl species, scavenging systems

Abstract

Reactive carbonyl species (RCS) are highly toxic molecules produced during normal metabolism and increased under conditions of oxidative stress. Here we show that *Arabidopsis thaliana* ecotypes exhibit natural variation in the degree of detoxification of RCS. We have isolated the ecotype IP-Pal-0 as an ecotype that has evolved a higher resistance to the toxic effects of glucose-derived RCS by acquiring a higher activity of the glyoxalase system through increased expression of some of its components. In particular, Viridiplantae-specific GLXI;2, GLXII;4 and GLXII;5 isoforms are highly expressed when plants are grown in the presence of 2-keto-D-glucose (KDG; glucosone) or methylglyoxal. We found that specific motif/*cis*-regulatory elements of the Col-0 and IP-Pal-0 GLXI;2 promoter regions may be involved in the differences in GLXI;2 gene expression associated with KDG detoxification. IP-Pal-0 GLXI;2 contains two different amino acids compared to Col-0, but these do not affect the basic kinetics of the protein. Interestingly, we found that the simultaneous change of these amino acids also occurs together in the GLXI proteins of some other organisms, suggesting a convergence in the simultaneous change of both amino acid residues. Our study of natural variants of *A. thaliana* suggests that the Viridiplantae-specific isoforms of the glyoxalase system are involved in the detoxification of glucose-derived RCS, particularly KDG.

Introduction

The study of plant natural ecotypes is constantly providing new insights into the biological mechanisms of the adaptation to the environment. By studying genes associated with environmental changes, it is possible to identify those responsible for stress tolerance in the natural variants of *Arabidopsis thaliana*. Ecotypes of this model plant have been used to link genes to cold (Barah *et al.*, 2013), drought (Bouzid *et al.*, 2019), and heat (Barah *et al.*, 2013) stresses, and even genes associated with viral infections (Prasch and Sonnewald, 2013).

Under environmental stresses and normal growing conditions, plants can produce unwanted small molecules that affect their physiology and reproductive performance (Hudig *et al.*, 2018). Two classic examples of these compounds are reactive oxygen species (ROS) and reactive carbonyl species (RCS), which are highly toxic to the cell due to their reactivity with macromolecules (Mano *et al.*, 2019, Mansoor *et al.*, 2022). RCS commonly found in biological systems include methylglyoxal (MGO), glyoxal (GO), 2-keto-D-glucose (KDG; glucosone) and 3-deoxyglucosone (Semchyshyn, 2014, Balparda *et al.*, 2023). These RCS are mainly

produced as byproducts of sugar metabolism through the oxidation of D-glucose (Thornalley *et al.*, 1999) and the cleavage of glycated proteins (Nagai *et al.*, 2002). Several abiotic stresses have been reported to increase RCS levels. In *Cucurbita maxima*, exposure to high NaCl concentrations increased MGO content in plants (Hossain *et al.*, 2009). Drought stress has been shown to increase MGO content in *Vigna radiata* and *Brassica juncea* (Veena *et al.*, 1999, Nahar *et al.*, 2017). Similarly, plants exposed to toxic metals have increased levels of RCS (Hasanuzzaman *et al.*, 2017).

Because of the harmful effects of RCS, plants seek to avoid and regulate high concentrations of these compounds and have therefore developed scavenging systems to convert RCS into harmless products (Hudig *et al.*, 2018). The main pathway for the conversion of RCS such as MGO and KDG into harmless metabolites is the glyoxalase (GLX) system, which consists of two enzymes, S-D-lactoylgluathione lyase (glyoxalase I, GLXI) and S-2-hydroxyacylgluathione hydrolase (glyoxalase II, GLXII), which work in sequence (Thornalley, 1990, Schmitz *et al.*, 2017, Balparda *et al.*, 2023).

GLXI uses as substrate a hemithioacetal formed by the non-enzymatic reaction of an RCS with glutathione (GSH), resulting in the formation of a glutathione thioester in which GSH is covalently bound to the RCS (Thornalley, 2003). GLXII then catalyses the conversion of the glutathione thioester to a (D)-α-hydroxy acid derivative which can be reintegrated into cellular metabolism, releasing GSH. Through this pathway, MGO is converted into D-lactate, which can be further converted into pyruvate by the enzymatic activity of D-lactate dehydrogenase (Engqvist *et al.*, 2009, Welchen *et al.*, 2016). GLXI;2 can also act on KDG, which produces D-gluconic acid (Balparda *et al.*, 2023).

In *A. thaliana*, three gene loci encode active GLXI protein isoforms: At1g67280 (*GLXI;1*), At1g11840 (*GLXI;2*) and At1g01180 (*GLXI;3*) (Schmitz *et al.*, 2017). GLXI is a lyase that depends on a divalent metal ion as cofactor. The plant GLXI homologues are classified into two groups of different evolutionary origin and structural properties. One group, represented by *A. thaliana* GLXI;1 and GLXI;2, is formed by isoforms only found in viridiplantae and bacteria (Balparda *et al.*, 2023). The viridiplantae-specific isoforms are monomers containing two structural domains that form two distinct concavities: one forming the active site and the other one forming a cryptic site of unknown function (Ridderstrom and Mannervik, 1996, Frickel *et al.*, 2001, Turra *et al.*, 2015, Balparda *et al.*, 2023). GLXI isoforms of this group are classified as Ni^{2+}-type GLXI, although they prefer to use Ni^{2+} when MGO-GSH is the substrate and Mn^{2+} when GO-GSH is the substrate (Schmitz *et al.*, 2017). The other group, represented

by Arabidopsis GLXI;3, is conserved across eukaryotic kingdoms (Schmitz *et al.*, 2018, Balparda *et al.*, 2023). These proteins have a single domain and form homodimers with two functionally equivalent active sites in Animalia and Viridiplantae, whereas other groups such as Rhodophyta, Phaeophyta, Diatoms, Protozoa and Fungi contain both single-domain and two-domain proteins that also form two functionally equivalent active sites (Balparda *et al.*, 2023). *A. thaliana* GLXI;3 is classified as a Zn^{2+}-type GLXI although it strongly prefers Mn^{2+} for catalysis (Ridderstrom and Mannervik, 1996, Frickel *et al.*, 2001, Schmitz *et al.*, 2017),

In *A. thaliana*, three gene loci encode homologs of GLXII: At3g10850 (GLXII;2), At1g06130 (GLXII;4), and At2g31350 (GLXII;5) (Schmitz *et al.*, 2017). GLXII possesses a Fe^{3+}/Zn^{2+} binuclear metal center and is active as a dimer on a wide range of glutathione thioesters (Thornalley, 1990, Ridderstrom and Mannervik, 1996, Bito *et al.*, 1997, Marasinghe *et al.*, 2005). Like GLXI, plant GLXII isoforms have two evolutionary origins; whereas GLXII;2 is found in all eukaryotic kingdoms, GLXII;4, and GLXII;5 are found only in Viridiplantae and Proteobacteria (Balparda *et al.*, 2023).

Genetic modification of components of the glyoxalase system in crop plants has been shown to increase resistance to various abiotic and biotic stresses, such as increased salt tolerance in transgenic tobacco and brown mustard (Veena *et al.*, 1999, Singla-Pareek *et al.*, 2008), and to pathogen attack, heat, drought and salinity in transgenic rice (Saxena *et al.*, 2011, Gupta *et al.*, 2018, Yan *et al.*, 2018). However, the use of genetically modified organisms in agriculture, particularly in Europe, is restricted to commercial crops by EU Directive 2015/412 (https://eur-lex.europa.eu/eli/dir/2015/412/oj). Therefore, the exploration of natural genetic variation, particularly of the GLX system in ecotypes, is a promising approach that may lead to crop improvement (Faralli *et al.*, 2019, Schulman *et al.*, 2020).

A. thaliana ecotypes grow naturally in different habitats and have thus adapted to different environmental conditions, including light intensity, temperature, humidity, water availability and soil properties (Weigel, 2012). Several reports have suggested important phenotypic plasticity between Arabidopsis ecotypes (Mitchell-Olds, 1995, Kliebenstein *et al.*, 2001, Koornneef *et al.*, 2004). Therefore, analysis of the genetic capacity within components of the GLX system among Arabidopsis ecotypes could help to dissect the genetic and physiological determinants of the GLX system.

In this work, we analyzed 20 *A. thaliana* ecotypes and studied their growth characteristics in the presence of MGO and KDG. Among them, the ecotype IP-Pal-0 showed a better growth performance under KDG and MGO than Col-0, which was chosen as a control. Using

complementary molecular and physiological analyses, we conclude that the higher expression of the Viridiplantae-specific GLXI;2 isoform in IP-Pal-0 is most likely related to its better growth performance in the presence of KDG.

Results

Growth of *A. thaliana* ecotypes in the presence of reactive carbonyl species

To investigate whether Arabidopsis ecotypes show differences in the degree of detoxification of RCS, we performed toxicity analysis with 20 selected ecotypes from 5 different locations (Fig. 1A). We analyzed plant development after germination in the presence of 0.2 mM (KDG; or 0.5 mM MGO. The concentration of the RCS was selected based on our previous work on the Col-0 ecotype (Schmitz *et al.*, 2017, Balparda *et al.*, 2023). It is worth mentioning that the selected KDG concentration is not highly harmful to the development of Col-0 under the conditions of our analysis, while the same MGO concentration already strongly affected plant development (Balparda *et al.*, 2023).

Treatment with MGO resulted in a significant reduction in primary root length after 14 days of growth in all ecotypes studied compared to growth in media without MGO (Fig. 1B and D). Interestingly, IP-Pal-0 was the least affected ecotype, followed by Ts-5. When grown on media containing MGO, IP-Pal-0 showed a 35% reduction in its primary root length compared to the growth on control media (Fig. 1B and D). In addition, the rosettes of IP-Pal-0 were larger and greener than those of Col-0 (Fig. 1D). In the presence of 0.2 mM KDG, the ecotypes Col-0, Liarum, and T-510 showed no differences in primary root length after 14 days of growth compared to growth in media without KDG (Fig. 1B) (Balparda *et al.*, 2023). Other ecotypes showed reduced primary root length in the presence of KDG, with the exception of IP-Pal-0 and Ts-5, which showed significantly increased primary root length (Fig. 1B and D). In particular, IP-Pal-0 showed 50% longer primary root than Col-0 (Fig. 1B and D). This analysis indicates that IP-Pal-0 has increased tolerance to the toxicity of KDG and MGO.

We recently showed that the presence of glucose in the growth media induces the accumulation of KDG in *A. thaliana* tissues (Balparda *et al.*, 2023). KDG is produced during cellular metabolism by the oxidation of glucose or the degradation of glycosylated proteins (Bean and Hassid, 1956, Baute, 1984, Semchyshyn, 2014). Therefore, we investigated whether the growth of Col-0 and IP-Pal-0 is affected by the presence of glucose. Similar to KDG, we observed that IP-Pal-0 had a longer main root and larger and greener rosettes than Col-0 in media containing

2 and 4% glucose (Fig. 2). The magnitude of the effect is positively correlated with glucose concentration, suggesting that the reduction in plant development in media containing glucose *in vivo* is at least partly due to the production of KDG and likely other RCS, such as MGO, from glucose.

Transcriptional analysis of GLXI genes in *A. thaliana* ecotypes

In the IP-Pal-0 ecotype, the toxicity analysis showed an increased tolerance to KDG and to a lesser extent, to MGO. Knowing that KDG is specifically metabolized by GLXI;2 (Balparda *et al.*, 2023) and that MGO is metabolized by all GLXI isoforms, we further analyzed the expression of all GLXI isoforms in the selected *A. thaliana* ecotypes. We conducted a qRT-PCR analysis of roots and shoots of seedlings grown with and without KDG in the growth media.

We found that the expression of the GLXI isoforms in the MS control media is similar in all analyzed ecotypes, except for *GLXI;1*, which is reduced in Taz-0 with respect to Col-0 in roots and is 2-fold higher in IP-Pal-0 than in Col-0 in shoots; and *GLXI;2*, which is approximately 2-fold higher in IP-Pal-0 than in Col-0 in both roots and shoots (Fig. 3).

In comparison to Col-0, the expression of *GLXI;2* and *GLXI;3* is highly enhanced in both roots and shoots of IP-Pal-0 in media containing KDG (Fig. 4). In the roots, the KDG/MS relative expression of *GLXI;2* is enhanced by about 3.7-fold, while that of *GLXI;3* is enhanced by about 2.3-fold (Fig. 4A) in comparison to the relative expression in Col-0. In shoot, the relative KDG/MS expression of *GLXI;2* is increased 9.4-fold, whereas that of GLXI;3 is increased 3.4-fold in comparison to Col-0 (Fig. 4B). *GLXI;1* expression is also increased in shoot of IP-Pal-0 grown with KDG compared to Col-0; relative KDG/MS expression of *GLXI;1* is increased 4-fold. Other ecotypes showed changes in the expression of some isoforms, but these are minor compared with those of IP-Pal-0.

Overall, qRT-PCR analysis showed increased transcription of GLXI isoforms, particularly GLXI;2, in IP-PAL-0 when exposed to high concentrations of KDG.

Metabolism of KDG by protein extracts of *A. thaliana* ecotypes

Considering the particularly high *GLXI;2* expression in IP-Pal-0 growing in media containing KDG, we further analyzed the total GLXI activity in whole seedlings of the selected ecotypes

growing in the presence of 0.2 mM KDG. The free GSH concentration was determined after incubation of protein extracts with different KDG concentrations and a fixed concentration of GSH as we previously described (Turra *et al.*, 2015, Balparda *et al.*, 2023). We found that KDG was similarly metabolized by all ecotypes, with the exception of IP-Pal-0, which showed a statistically significant enhanced metabolization of KDG (Fig. 5). This result together with the better growth of IP-Pal-0 in KDG-containing media (Fig. 1C and D) suggest a higher KDG scavenging capacity of IP-Pal-0 compared to the other ecotypes.

Comparison of GLXI;2 from IP-Pal-0 and Col-0

The higher GLXI activity of the protein extracts from IP-Pal-0 compared to Col-0 is most likely due to an increased amount of GLXI;2 protein caused by the observed increased expression of GLXI;2 in IP-Pal-0 (Fig. 3). Nevertheless, we cannot exclude the possibility that single nucleotide polymorphisms (SNPs) in the coding sequence of IP-Pal-0 GLXI;2 can be responsible for the production of a more active protein than that of Col-0. We analyzed the genomic sequence of IP-Pal-0 *GLXI;2* obtained from the 1001 Genomes Project and found some missing nucleotides in the coding region. We thus amplified and sequenced the *GLXI;2* cDNA from IP-Pal-0 leaves. We detected two SNPs in two codons that result in two different amino acids in GLXI;2 of IP-PAL-0 compared to Col-0 (Fig. 6A). Amino acid residue 229 of Col-0 GLXI;2 is a Lys, while that of IP-PAL-0 is an Asn; and amino acid residue 278 of Col-0 is a Phe, while IP-PAL-0 is an Ile.

GLXI;2 is a monomeric two-domain protein, in which both structural domains contribute to the formation of two non-equivalent concavities. One of this, is a cryptic site (site 1, Fig. 6B) that cannot bind a metal ion, and the other is the functional metal-dependent active site (site 2, Fig. 6B) (Turra *et al.*, 2015). We comparatively modeled the three-dimensional structure of GLXI;2 from Col-0 and IP-Pal-0 and analyzed the localization of amino acids 229 and 278. We observed that both amino acids are located in the structural domain B region that contribute to the formation of the concavity bearing the cryptic site 1 and are thus far away from the amino acids that participate in metal-ion binding (His89, Glu137, Gln150, and Glu201) in the active site (Turra *et al.*, 2015, Balparda *et al.*, 2023) (Fig. 6B).

To empirically analyse whether the presence of Asn at position 229 and/or Ile at position 278 induces changes in the enzymatic properties of Col-0 GLXI;2, we generated the K229N, F278I and K229N-F278I mutants by site-directed mutagenesis of Col-0 GLXI;2. Previously, we indicated that the activity of recombinant GLXI;2 cannot be assessed spectrophotometrically

using KDG as substrate, but can be assessed using MGO or GO (Schmitz *et al.*, 2017, Balparda *et al.*, 2023). We thus analyzed the recombinant proteins, Col-0 GLXI;2 WT, K229N, F278I and K229-NF278I using MGO and GO as substrates and the best metal cofactor for each substrate (Table 1). The kinetic parameters measured, turn-over rate (k_{cat}), substrate affinity (K_m), and catalytic efficiency (k_{cat}/K_m) obtained were highly similar to those previously published (Schmitz *et al.*, 2017). We observed no significant differences in any of these parameters between Col-0 GLXI;2 WT and the mutant variants (simulating GLXI;2 from IP-Pal-0; Table 1), indicating that the presence of a Lys or Asn at position 229 and a Phe or Ile at position 278, or both simultaneously, does not alter the basic kinetic properties of the enzyme under the conditions of our assay. Analysis of the behavior of the proteins in native PAGE showed that all proteins have similar mobilities (Suppl. Fig. S1), indicating that the mutations do not alter the basic oligomeric organization.

In order to analyse whether these two SNPs also occur in other GLXI, we compared GLXI sequences of different evolutionary origin (Supplementary Table 1) (Balparda *et al.*, 2023). We found that the amino acids of the IP-Pal-0 GLXI;2 SNPs were not present at the homologous position in any of the other Viridiplantae-specific GLXI;1 or GLXI;2 (Fig. 6C; Supplementary Table 1). Interestingly, all sequenced Viridiplantae Zn^{2+}-type GLXI, represented by *A. thaliana* GLXI;3, contain the same amino acid (Ile) as the SNP at position 278 of IP-Pal-0 GLXI;2 (Fig. 6C; Supplementary Table 1). GLXI from other organisms almost always have different amino acids at positions 229 and 278 from those found in Viridiplantae (Fig. 6C; Supplementary Table 1)

Analysis of *GLXI;2* promoter regions of IP-Pal-0 and Col-0

Since the enzymatic properties of GLXI;2 of Col-0 and IP-Pal-0 (Col-0 K229N-F278I) are very similar (Table 1), we propose that the high expression of GLXI;2 observed in IP-Pal-0 in the presence of KDG is most likely responsible for the enhanced growth of this ecotype. We therefore analyzed and compared the promoter regions of Col-0 and IP-Pal-0 for the presence of different cis-regulatory elements (CREs) or of conserved CREs in different positions. Using the online database *Plant CARE*, we identified around 27 motifs present in both ecotypes (Fig. 7A) (Supplementary Table 2). We found that some of the motifs share similar positions in Col-0 and IP-Pal-0, but others exist in very different positions in the region from -1250 to -800 (Supplementary Table 3). When we compared the sequences of the identified CREs between both ecotypes, we detected ecotype-specific motifs that exist only in one of the ecotypes such

as ERE (ATTTCATA) and MYC (CAATTG) that are only found in Col-0, and ABRE (ACGTG; CACGTG; GCAACGTGTC) only found in IP-Pal-0 (Fig. 7B). This analysis suggests that specific motif/*cis*-regulatory elements may be involved in the differences in *GLXI;2* gene expression associated with KDG detoxification in Col-0 and IP-Pal-0.

Transcriptional analysis of GLXII genes in *Arabidopsis thaliana* ecotypes

In Arabidopsis, KDG is metabolized by the sequential action of GLXI;2, and GLXII;4 or GLXII;5 (Balparda *et al.*, 2023). To investigate whether the expression of second-step genes of the metabolic pathway is also deregulated in IP-Pal-0 growing in the presence of KDG, we performed qRT-PCR analysis of all GLXII genes using the same ecotypes in which we had previously analyzed GLXI expression.

We found that the expression of the GLXII isoforms in the MS control media is similar in all ecotypes analyzed, except for *GLXII*,*4* and *GLXII*,*5*, which are approximately 0.6-fold higher in shoot of IP-Pal-0 than in Col-0 (Fig. 8).

In the presence of KDG, the expression of *GLXII;2* was not changed in the root of IP-Pal-0 compared with Col-0, whereas that of *GLXII;4* and *GLXII;5* was greatly increased (Fig. 9A). The expression of *GLXII;4* is approximately 13-fold higher in IP-Pal-0 than in Col-0, whereas the expression of *GLXII;5* is approximately 12-fold higher (Fig. 9A). In the shoot, the expression of *GLXII;2* and *GLXII;*5 is 0.5-fold higher in IP-Pal-0 than in Col-0, whereas that of *GLXII;4* is 3.3-fold higher (Fig. 9B). As in the case of *GLXI;1* expression, other ecotypes also showed specific changes in the expression of some isoforms, in most cases in both organs studied (Fig. 9B).

Overall, and in agreement with our previous results (Balparda *et al.*, 2023), the relative level of *GLXII* expression in IP-PAL-0 compared with Col-0 indicates that *GLXII;4* and *GLXII;5* are mainly involved in the second step of KDG detoxification.

Metabolite profile analysis in seedlings of Col-0 and IP-Pal-0 grown in MS and KDG

We hypothesize that if IP-Pal-0 grows better than Col-0 in the presence of KDG, and exhibits an enhanced expression of *GLXI;2*, increased total GLXI activity, and enhanced expression of *GLXII,4* and *GLXII;5*, the relative KDG levels in tissues of IP-Pal-0 growing in the presence of KDG should be lower than in tissues of Col-0 growing under the same conditions. To explore

this, we conducted a metabolite profile analysis via GC-MS of seedling extracts grown in the presence of 0.2 mM KDG and in MS as control condition (Supplementary data 1). We run a principal component analysis (PCA) to reduce the dimensionality of the dataset and visualize sample grouping (Fig. 10A and Suppl. Fig. S2). The first axis of the PCA (principal component 1; PC1) explains 46.3% of the variance and it separates the Col-0 and IP-Pal-0 ecotypes. The second axis of the PCA (PC2) explains 20.1% of the variance and it separates the growth conditions (MS vs KDG in growth media). The PCA identified a distinction between Col-0 in the presence of KDG and in MS. However, IP-Pal-0 grown in MS and KDG showed similar relative metabolite levels. The PCA biplot also includes the loadings (Fig. 10A), where the large loadings indicate that a variable has a strong effect on a PC. Our results indicate that KDG and other sugars such as fructose and glucose, as well as proline, show a strong loading on PC2 in the direction of Col-0 grown in a media containing KDG. These results suggest that Col-0, when fed with KDG, cannot completely metabolize it and also accumulate metabolites known to be enhanced in response to abiotic stress.

Next, we examined the relative KDG content in Col-0 and IP-Pal-0 obtained by GC-MS (Fig. 10B). We found that Col-0 grown in the presence of KDG accumulated about 2.5-fold higher amounts of KDG than the other plants analyzed. Interestingly, IP-Pal-0 grown in MS or in the presence of KDG showed no differences in relative KDG levels, indicating that enhanced GLXI;2, and GLXII;4 and GLXII;5 activities are most probably responsible for the rapid detoxification of this RCS. Gluconate, the product of glyoxalase system activity on KDG, accumulated to similar levels in Col-0 and IP-Pal-0 grown in the presence of KDG (Supplementary data 1). In plants grown in the presence of KDG, these levels were 3.6-fold higher in Col-0 and 3.4-fold higher in IP-Pal-0.

Discussion

The Arabidopsis ecotypes studied in this work originate from different geographical regions (Fig. 1A) and may therefore have been subject to different selective pressures. As a result, the ecotypes may have developed adaptive responses to environmental stresses, which we have exploited here to analyse the activity of the glyoxalase system, a scavenging system for RCS. After growing the ecotypes in the presence of two glucose-derived RCS, MGO and KDG, which are specifically produced under oxidative stress, the IP-Pal-0 ecotype showed an increased tolerance to the toxic effects of both RCS, especially KDG, compared to most of the other ecotypes (Fig 1B-D). A second ecotype, Ts-5, also showed a similar behavior, but to a

much lesser extent. It is interesting to note that both ecotypes originate from very restricted specific regions in Spain (Fig. 1A).

Here we show that in IP-Pal-0, specifically the GLXI;2 and GLXII;4 and GLXII;5 isoforms are highly expressed in root and shoot when plants are grown in the presence of the RCS, specially of KDG (Fig. 4 and Fig. 8). We also show that protein extracts from IP-Pal-0 have a higher total GLXI activity than the other ecotypes analyzed (Fig. 5). In addition, IP-Pal-0 does not accumulate KDG when grown on media containing this RCS, whereas Col-0 does (Fig. 10B). These results suggest that the expression of GLXI;2 and GLXII;4 and GLXII;5 is enhanced by KDG, although it cannot be excluded that the high level of KDG in the growth media further increases the production of other RCS and their derivatives, and thus the plant also responds to them by increasing its scavenging capacity. It is worth mentioning that although GLXI;2 is specifically involved in the detoxification of KDG (Balparda *et al.*, 2023), it can also act on the hemithioacetal formed by MGO and GSH (Schmitz *et al.*, 2017). In any case, the high expression levels of GLXI;2 and GLXII;4 and GLXII;5 in response to KDG in the growth media are consistent with previous work showing that these isoforms are specifically involved in the detoxification of KDG (Balparda *et al.*, 2023).

Apart from the high expression of GLXI;2 in IP-Pal-0, other possibility is that this isoform is more active than the isoform from Col-0. We found that the sequence of IP-Pal-0 GLXI;2 contains two different amino acids compared to Col-0. We produced the single and double mutants of the Col-0 protein to mimic that of IP-Pal-0 and found no differences in the basic kinetic behavior between the proteins (Table 1). Mapping the mutations in the modelled structure of GLXI;2 revealed that both residues are located far away from the active site in the part of domain B involved in the formation of the cryptic site (Fig. 6B). It has been proposed that this cryptic site has lost its catalytic capacity due to the accumulation of mutations that do not allow the binding of the metal required for catalysis (Turra *et al.*, 2015, Schmitz *et al.*, 2018). Interestingly, we found that the specific combination of amino acids found in IP-Pal-0 GLXI;2 (Asn instead of Lys 229, and Ile instead of Phe 278) does not occur in GLXI proteins of any other organism, but that all sequenced Viridiplantae Zn^{2+}-type GLXI contain the Ile at position 278 (Fig. 6C; Supplementary Table 1), indicating convergent evolution. In contrast to Ni^{2+}-type GLXI, Viridiplantae Zn^{2+}-type GLXI are monomeric proteins with two structural domains that form two similar active sites (Balparda *et al.*, 2023), so the presence of Ile at this position does not appear to be detrimental to the catalytic function. Importantly, we found that specific motif/cis-regulatory elements of the Col-0 and IP-Pal-0 GLXI;2 promoter regions may be involved in the differences in GLXI;2 gene expression associated with KDG detoxification.

In the future, this knowledge can be further exploited to identify promoter regions important for the enhancement of this trait.

Our study of natural variants of *A. thaliana* indicates that the Viridiplantae-specific isoforms of the glyoxalase system are involved in the detoxification of glucose-derived RCS, in particular KDG. These results allow us to propose a strategy to improve plant resistance to abiotic stresses that generate glucose-derived RCS by overexpressing GLXI;2 and GLXII;4 or GLXI;2 and GLXII;5 in the appropriate cell compartments. Further study of this type of natural variation promises to shed light on the adaptive mechanisms of scavenging systems.

Materials and Methods

Growth conditions

A. thaliana ecotypes listed in Table 1 were kindly provided by Prof. Juliette de Maux (University of Cologne). All seedlings were grown under a long-day photoperiod of 16 hours of light and 8 hours of darkness on soil or sterile media at 22 °C and 55% relative humidity. A combination of Osram HO 80W/840 and HO 80W/830 bulbs with a light intensity of 120 μE^{-2} s^{-1} was used.

Sterilized seeds of each *A. thaliana* line were placed on plates with 0.5X MS (Duchefa; Murashige and Skoog, 1962) alone or either containing 0.2 mM KDG (Sigma-Aldrich, USA) or 0.5 mM MGO (Sigma Aldrich, USA) and vernalized at 4 °C for 48 hours in the dark. Growth experiments were also conducted in media plates with 0.5X MS either containing 2 % glucose or 4 % glucose. Per ecotype and RCS treatment, 15 seeds were arranged on a plate in a line. After vernalization at 4 °C for 48 h, the plates were placed vertically in a growth cabinet (Percival Scientific Inc., USA) and kept under long-day conditions. Root length was recorded for every seedling via the usage of the image processing program ImageJ (Schneider *et al.*, 2012).

RNA isolation from plant tissue and synthesis of complementary cDNA

Total RNA was extracted from roots and shoots from plants grown in 0.5X MS and 0.5X MS supplemented with 0.2 mM KDG for 14 days. RNA was extracted from rosette leaves using the Universal RNA Purification Kit (Roboklon GmbH (Germany) following the manufacturer's instructions. In 30 µl of RNase-free dH2O, isolated RNA was eluted. The NanoDrop Nabi™

Ultraviolet-Visible (UV/Vis) (MicroDigital Co.Ltd, South Korea) and subsequent agarose gel electrophoresis were used to determine the purity of the recovered RNA. The isolated total RNA was stored at -20 °C until use.

After the removal of possible genomic DNA contaminations with an Ambion DNA-free DNA removal kit (Thermo Fisher, USA), 2 µg RNA was transcribed into the cDNA (cDNA) with RevertAid reverse transcriptase (Thermo Fisher, USA) using an oligo(dT) primer according to the manufacturer's instructions. The RNA was mixed with 1 µl of oligo(dT) primer and heated to 65 °C for 10 minutes. The RNA was then mixed with 2 µl of dNTPs, 0.8 µl of RevertAid H Minus M-MuLV Reverse Transcriptase (Thermo Fisher, USA), 4 µl of 5X RevertAid H Minus First Strand cDNA Synthesis Kit reaction buffer (Thermo Fisher, USA), and 10 µl of water. The combination was incubated for 60 minutes at 42° C and then inactivated for 10 minutes at 70° C. The cDNA was stored at −20° C until use.

Quantification of gene expression via qRT-PCR

Quantitative real-time PCR (qRT-PCR) analysis was performed using the KAPA SYBR FAST qPCR kit (KAPA Biosystems Inc., Switzerland) to check the gene expression levels of *GLXI;1*, *GLXI;2*, *GLXI;3, GLXII;2, GLXII;4 and GLXII;5* in the cDNA samples of the *A. thaliana* ecotypes. The qRT-PCRs were conducted in the Mastercycler® ep realplex (Eppendorf, Germany), and cycle threshold (Ct) values for each sample and gene of interest (GOI) were obtained from the realplex software belonging to the same company. The housekeeping gene *Actin2* (ACT2), served as an internal normalization.

Protein extracts and GLXI activity measurements

Whole seedlings grown for 14 days in 0.5X MS or with 0.2 mM KDG were harvested and stored at -80° C until use. For protein extraction, 200 mg of plant material was ground in the presence of liquid N_2 with an extraction buffer containing 100 mM Tris-HCl, pH 8.0; 100 mM NaCl; 0.5% (v/v) Triton X-100; 2 mM PMSF; and 1% (w/v) PVP40. The samples were centrifuged at 20,000 g for 10 minutes at 4 °C, and the supernatant was transferred to a fresh tube.

To determine the total enzymatic activity of GLXI in seedling protein extracts, KDG (2-keto-D-glucose; Sigma-Aldrich, USA) was incubated with GSH for 10 min before the start of the

experiment to allow synthesis of the hemithioacetal. Variable concentrations of KDG were used (0, 0.25, 0.5, 1 and 2 mM), while the GSH concentration remained at 1 mM. Then 20 µg of protein extract was added to the reaction mixture and after 30 min incubation the absorbance was measured at 412 nm. Finally, 1 mM DNTB (5,5-dithiol-bis(2-nitrobenzoic acid; Geno Technologies Inc., USA), which reacts with GSH, was added to the solution. After incubation for 10 min, the absorbance was read again at 412 nm. Assays were performed on a Synergy HT plate reader (Biotek Instruments Inc, USA) at 25°C. All assay reagents were dissolved in 0.1 M MOPS pH 7.0. A calibration curve using GSH concentrations of 0.2, 0.3, 0.4, 0.6, 0.8, 1.2 and 2 mM in 0.1 M MOPS pH 7.0 was used to determine the GSH concentration. Each assay was performed with a minimum of three biological replicates and a minimum of three technical replicates.

Expression and purification of GLXI;2

For heterologous protein expression, the pET16b expression constructs containing the GLXI;2 variants (GLXI;2-Phe278 and GLXI;2-Ile278) were transformed into *E. coli* Rosetta2 (DE3) pLysS. The transformed cells were grown to an OD_{600} of 0.4–0.6 in 400 ml of LB medium at 37°C and with continuous shaking using ampicillin (50 µg/ml) and Chloramphenicol (34 µg/ml) as selection markers. Adding 1 mM IPTG to the culture and leaving the cells to incubate for 4 hours at 37°C with constant shaking enabled the expression of the heterologous protein to be induced. Pellets were generated after the cellular culture was collected at 6000 *g* for 15 min, which were kept at 20 °C until required.

Pellets were thawed on ice, resuspended in 20 mM Tris-HCl (pH 8.0) containing 500 mM NaCl, 5 mM imidazole, 1 mM phenylmethylsulfonyl fluoride, and a spatula tip of lysozyme, subsequently sonicated and centrifuged at 14 000 g for 20 min at 4°C. Through the use of gravity flow immobilized metal ion chromatography and nickel nitrilotriacetic acid agarose (Ni-NTA Agarose; ThermoFisher, USA), the supernatant was used to purify the proteins. A 20 mM Tris-HCl buffer containing 500 mM NaCl and 5 mM imidazole was utilized to pre-equilibrate the column. Following supernatant loading, the columns were washed four times with 500 mM NaCl in increasing amounts of imidazole (5, 40, and 60 mM) in 20 mM Tris-HCl (pH 8.0). In the final step, 1 ml of 20 mM Tris-HCl, 500 mM NaCl, and 200 mM imidazole elution buffer was utilized to elute the proteins. Finally, protein concentration and buffer exchange (20 mM Tris-HCl, pH 7.5) was performed using Vivaspin 30K columns (Sartorius, Germany, Göttingen, Germany).

Gel electrophoresis

Native-PAGE was performed using 10% (w/v) polyacrylamide gels according to Laemmli (1970) and the SERVA Marker Mix for Blue/Clear Native PAGE (SERVA Electrophoresis GmbH; Heidelberg, Germany). Proteins were visualized by staining with Coomassie Brilliant Blue.

Measurement of GLXI;2 enzymatic activities and determination of kinetic constants

The GLXI;2 activities were monitored spectrophotometrically with a Spark® (Tecan) in 96 well UV-Star Microplates (Greiner Bio-one). The standard assay media contained 100 mM MOPS pH 7, 3.5 mM MGO or GO and 1.7 mM GSH together with 0.5 mM of the metal ion (Ni^{2+} for the case of MGO and Mn^{2+} for GO; Schmitz et al., 2017). The spontaneous reactivity of the reactive carbonyl species with GSH, resulting in the formation of hemithioacetal (HTA), the natural substrate of GLXI, is required for the conversion of MGO or GO by GLXI. The determination of the kinetic constants proceeded the same way as Schmitz et al. (2017). All measurements were performed in at least three independent experiments.

Homology modeling of GLXI;2 3D structure

The AlphaFold protein structure prediction server ColabFold (Jumper *et al.*, 2021) as used to generate *Arabidopsis thaliana* GLXI;2 3D models using the primary amino acid sequence as input. Images of the models and distances between amino acids were prepared using PyMOL v.2.5.4 by Schrödinger (https://pymol.org/).

Metabolite abundance analysis by GC-MS

Plant samples for gas chromatography-mass spectrometry analysis were harvested after being in 0.5 mM MS or 0.2 mM KDG for 14 days. The same equipment configuration and methodology as stated in Lisec et al. 2006 were used for extraction and analysis by gas chromatography-mass spectrometry (GC-MS). Shortly, 200 µl of chloroform and 300 µl of methanol were combined with 300 µl of water to homogenize 10 mg of frozen material for 15 min at 70°C. The polar fraction was dried under vacuum, and the residue was derivatized for 120 min at 37°C (in 40 µl of 20 mg ml^{-1} methoxyamine hydrochloride (Sigma-Aldrich, cat. no.

593-56-6) in pyridine followed by a 30 min treatment at 37°C with 70 µl of *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA reagent; Macherey-Nagel, cat. no. 24589-78-4). An autosampler Gerstel Multi-Purpose system (Gerstel GmbH & Co. KG, Mülheim an der Ruhr, Germany) was used to inject 1 µl of the samples in splitless mode to a chromatograph coupled to a time-of-flight mass spectrometer system (Leco Pegasus HT TOF-MS; Leco Corp., St Joseph, MI, USA). Helium was used as carrier gas at a constant flow rate of 2 ml s^{-1} and GC was performed on a 30 m DB-35 column (capillary column, 30 m length, 0.32 mm inner diameter, 0.25 µm film thickness, PN: G42, Agilent). The injection temperature was 230°C and the transfer line and ion source were set to 250°C. The initial temperature of the oven (85°C) increased at a rate of 15°C min^{-1} up to a final temperature of 360°C. After a solvent delay of 180 s, mass spectra were recorded at 20 scans s^{-1} with m/z 70–600 scanning range. Chromatograms and mass spectra were evaluated using Chroma TOF 4.5 (Leco) and TagFinder 4.2 software. Based on a computation of the retention index with a variance of no more than 5%, metabolites were annotated and compared to the reference data from the Golm Metabolome Database (http://gmd.mpimp-golm.mpg.de; (Luedemann *et al.*, 2008)). The results were statistically analyzed using Student's t-test, using a threshold of $P<0.05$ between the samples of the control and treated plants (Supplemental Dataset).

Sequence analysis of Col-0 and IP-Pal-0 GLXI;2

The *GLXI;2* (AT1G11840) cDNA sequences in Col-0 were obtained from the TAIR database (https://www.arabidopsis.org/). In the case of IP-PAL-0, cDNA synthesized from RNA from a plant tissue extraction was used. The cDNA was sequenced from the start (ATG) to the end of the gene (stop codon) with Macrogen, Inc (South Korea).

The coding sequences obtained were translated into protein sequences using the EMBOSS Transeq online tool (https://www.ebi.ac.uk/Tools/st/emboss_transeq/). Finally, the protein sequences were aligned using the Clustal Omega (1.2.4) online tool (https://www.ebi.ac.uk/Tools/msa/clustalo/). The MEGA7 software was used to perform the multiple alignment of GLXI sequences of Figure 6C.

Promoter analysis

To analyze the natural variation in the GLX1;2 promoter of the Col-0 and IP-Pal-0 ecotypes, we used the "Pseudogenomes Download" tool from the website https://tools.1001genomes.org/

to download the sequences of the 5'UTR of the AT1G11840 gene from 1.5 kb upstream to the last nucleotide before the ATG. To do this, we selected both ecotypes and used the following query "Chr1:3993657..3996044" to define the position of interest. We then used the online database and analysis tool PlantCARE (Lescot *et al.*, 2002) to identify the cis-acting regulatory elements in the DNA sequences of the Col- and IP-Pal-0 promoters. We then used TBtools (Chen *et al.*, 2020) to visualize the identified cis-regulatory-elements and R scripts to compare presence/absence and positions of the identified motifs between the two ecotypes.

Statistical analysis

P values were obtained using one-way ANOVA with Dunnett's test for the cases of Figures 1, 3, 4, 8, 9 and 10. P and the unpaired t-test for the cases of Figures 2 and 5. For the case of Figure 10, the prcomp function of the basic R package stat was used to perform the principal component analysis, and the R package factoextra was used to visualize the results.

Legend to figures

Figure 1. Effect of KDG and MGO on the development of *A. thaliana* ecotypes. (A) Geographical origin of the ecotypes analyzed. For those from Spain and Sweden, the localisation shows the exact coordinates provided by 1001genomes (https://1001genomes.org). (B) Main root length after 14 days of germination on media containing 0.5 mM MGO relative to the length in control media (MS). (C) Main root length after 14 days of germination on media containing 0.2 mM KDG relative to the length in MS. The graphs show the means and standard errors of the measurements of 3 independent experiments (n=3) of 20 roots each. Significant differences were analyzed by one-way ANOVA. An asterisk (*) indicates a significant difference ($P<0.001$) between an ecotype and Col-0 by Dunnett's multiple comparison test. (D) Images of seedling of selected ecotypes at 14 d after germination in different growth media. Scale bar = 1 cm.

Figure 2. Development of seedlings of Col-0 and IP-PAL-0 in the presence of 2% and 4% glucose in the growth media. (A) Images of seedlings 14 d after germination in different growth media. Scale bar = 1 cm. (B) Main root length after 14 days of germination on media containing 2% or 4% glucose compared to the length in control media (MS). The graphs show the means and standard errors of the measurements of 3 independent experiments (n=3) of 20

roots each. An asterisk (*) indicates a significant difference (P <0.001) between Col-0 and IP-PAL-0 determined by unpaired t-test.

Figure 3. Transcriptional analysis of the GLXI isoforms in selected *A. thaliana* ecotypes grown on MS. Expression levels of GLXI isoforms in root (A) and shoot (B) of 14-week-old seedlings of selected ecotypes grown on MS relative to those measured in Col-0. One-way ANOVA was used to test for significant differences. An asterisk (*) indicates a significant difference (P<0.001) between an ecotype and Col-0 by Dunnett's multiple comparison test.

Figure 4. Transcriptional analysis of the GLXI isoforms in selected *A. thaliana* ecotypes grown on media with and without KDG. Expression levels of GLXI isoforms in root (A) and shoot (B) of 14-week-old seedlings of selected ecotypes grown in 0.2 mM KDG relative to those measured in seedlings grown in MS. One-way ANOVA was used to test for significant differences. An asterisk (*) indicates a significant difference (P<0.001) between an ecotype and Col-0 by Dunnett's multiple comparison test.

Figure 5. GLXI activity using KDG in protein extracts from seedlings of selected ecotypes. The concentration of free GSH was determined after incubation of protein extracts with different concentrations of KDG. The initial amount of free GSH when no KDG was added to the reaction media was assumed to be 100%. Values are mean ± SD (n=3). An asterisk (*) indicates a significant difference between ecotype and Col-0 values at each KDG concentration (P<0.05), determined by unpaired t-test.

Figure 6. Analysis of GLXI sequences and structures. (A) Alignment of the primary amino acid sequences of GLXI;2 of IP-Pal-0 and Col-0. (B) Model structures of Col-0 (green) and IP-Pal-0 (light blue) showing the localization of the amino acids at positions 229 and 278 and those involved in ion and GSH at the binding functional active site (site 2) and the equivalents in the cryptic site (site 1). (C) Alignment of 17 selected Ni^{2+}-type and Zn^{2+}-type GLXI sequences from different organisms showing differential conservation of the amino acids at position 229 and 278 (corresponding to the *A. thaliana* Col-0 sequence). The analysis was conducted with the MEGA7 software. All 72 sequences analysed are listed in Supplementary Table 1.

Figure 7. Cis-regulatory elements in Col-0 and IP-PAL-0. (A) Promoter analysis for the presence of cis-regulatory elements 1.5 kb upstream of the 5'UTR (2.4 kb upstream of the transcription initiation site) of *GLX1;2* in the *A. thaliana* ecotypes Col-0 and IP-Pal-0. Cis-regulatory elements were mapped using TBtools (Chen *et al.*, 2020). (B) Unique cis-regulatory elements of *GLX1;2* identified in the ecotypes Col-0 and IP-Pal-0.

Figure 8. Transcriptional analysis of the GLXII isoforms in selected *A. thaliana* ecotypes grown on MS. Expression levels of GLXII isoforms in root (A) and shoot (B) of 14-week-old seedlings of selected ecotypes grown on MS relative to those measured in Col-0. One-way ANOVA was used to test for significant differences (P<0.0001). An asterisk (*) indicates a significant difference (P<0.001) between an ecotype and Col-0 by Dunnett's multiple comparison test.

Figure 9. Transcriptional analysis of the GLXII isoforms in selected *A. thaliana* ecotypes grown on media with and without KDG. Expression levels of GLXI isoforms in root (A) and shoot (B) of 14-week-old seedlings of selected ecotypes grown in 0.2 mM KDG relative to those measured in seedlings grown in MS. One-way ANOVA was used to test for significant differences (P<0.0001). An asterisk (*) indicates a significant difference (P<0.001) between an ecotype and Col-0 by Dunnett's multiple comparison test.

Figure 10. Analysis of the metabolite relative contents. (A) 2D Principal Component Analysis (PCA) biplot of variables in black labels (metabolites measured in seedlings grown in media without (MS) or with 0.2 mM KDG and individuals in colored points (*A. thaliana* ecotypes Col-0 and IP-Pal-0). Ellipses group the ecotypes per growth medium. The first axis explains 46.3% of the variance and the second axis explains 20.1%. (B) Relative amount of KDG in seedlings of Col-0 and IP-Pal-0 grown in media without (MS) and with 0.2 mM KDG.

Supplementary material

Supplementary Figure 1. Native PAGE of GLXI;2. Two µg of recombinant GLXI;2 of Col-0 WT, and the mutants K229N, F278I, K229N-F278I were analysed in a 10% (w/v) native PAGE and visualized with Coomassie Brilliant blue. MWM: SERVA Marker Mix for Blue/Clear Native PAGE (SERVA Electrophoresis GmbH; Heidelberg, Germany).

Supplementary Figure 2. Screen plot used for the PCA analysis to determine the number of factors to be retained.

Supplementary Data 1. Relative amounts of metabolites determined by GC-MS. Six biological replicates (BR1-BR6) were analyzed from each line.

Supplementary Table 1. GLXI sequences used for the analysis of the amino acids at positions 229 and 278 corresponding to the *A. thaliana* GLXI;2 Col-0 sequence. The sequences were extracted from the evolutionary analysis of GLXI presented in Supplementary Fig. 2 of Balparda et al. 2023.

Supplementary Table 2. Cis-regulatory elements identified in Col-0 and IP-Pal-0. The PlantCARE (Lescot *et al.*, 2002) tool was used. The positions are from 0 to -2400 bp in the 5' direction (0 is in the 3' end, which is the position of the last nucleotide before the ATG, and -2400 is in the 5' end).

Supplementary Table 3. Differences in the position of the identified cis-regulatory elements in Col-0 and IP-Pal-0. The positions are from 0 to -2400 bp in the 5' direction (0 is in the 3' end which is the position of the last nucleotide before the ATG and -2400 is in the 5' end).

Authorship

V.G.M. conceived and designed research. V.G.M and M.Ba. supervised data analysis and interpretation. M.Ba., A.K., M.Bo., C.E.A., and S.A. performed research and acquired the raw data. A.R.F. contributed with analytical tools. M.Ba., M.Bo. and V.G.M. wrote the manuscript. All authors contributed to the generation of the figures and the revision of the manuscript. All authors accepted the final version of the manuscript.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft through grant MA2379/21-1 to V.G.M., by the University of Bonn through an Argelander Grant to M.Ba and to M.Bo, and by the European Union's Horizon 2020 research and innovation programme, project PlantaSYST (SGA-CSA No. 739582 under FPA No. 664620), and by the European Regional Development Fund through the Bulgarian "Science and Education for Smart Growth" Operational Programme grant BG05M2OP001-1.003-001-C01 to A.R.F.

Conflict of interest statement

None declared.

Bibliography

Balparda, M., Schmitz, J., Duemmel, M., Wuthenow, I.C., Schmidt, M., Alseekh, S., Fernie, A.R., Lercher, M.J. and Maurino, V.G. (2023) Viridiplantae-specific GLXI and GLXII isoforms co-evolved and detoxify glucosone in planta. *Plant Physiol*, **191**, 1214-1233.

Barah, P., Jayavelu, N.D., Rasmussen, S., Nielsen, H.B., Mundy, J. and Bones, A.M. (2013) Genome-scale cold stress response regulatory networks in ten Arabidopsis thaliana ecotypes. *BMC Genomics*, **14**, 722.

Baute, M.-A.B., R. (1984) Occurrence among macrofungi of the bioconversion of glucosone to cortalcerone. *Phytochemistry*, **23**, 271-274.

Bean, R.C. and Hassid, W.Z. (1956) Enzymatic Oxidation of Glucose to Glucosone in a Red Alga. *Science*, **124**, 171-172.

Bito, A., Haider, M., Hadler, I. and Breitenbach, M. (1997) Identification and phenotypic analysis of two glyoxalase II encoding genes from Saccharomyces cerevisiae, GLO2 and GLO4, and intracellular localization of the corresponding proteins. *J. Biol. Chem.*, **272**, 21509-21519.

Bouzid, M., He, F., Schmitz, G., Häusler, R.E., Weber, A.P.M., Mettler-Altmann, T. and De Meaux, J. (2019) Arabidopsis species deploy distinct strategies to cope with drought stress. *Ann. Bot.*, **124**, 27-40.

Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y. and Xia, R. (2020) TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular plant*, **13**, 1194-1202.

Engqvist, M., Drincovich, M.F., Flugge, U.I. and Maurino, V.G. (2009) Two D-2-hydroxy-acid dehydrogenases in Arabidopsis thaliana with catalytic capacities to participate in the last reactions of the methylglyoxal and beta-oxidation pathways. *J. Biol. Chem.*, **284**, 25026-25037.

Faralli, M., Matthews, J. and Lawson, T. (2019) Exploiting natural variation and genetic manipulation of stomatal conductance for crop improvement. *Curr. Opin. Plant Biol.*, **49**, 1-7.

Frickel, E.M., Jemth, P., Widersten, M. and Mannervik, B. (2001) Yeast glyoxalase I is a monomeric enzyme with two active sites. *J. Biol. Chem.*, **276**, 1845-1849.

Gupta, B.K., Sahoo, K.K., Ghosh, A., Tripathi, A.K., Anwar, K., Das, P., Singh, A.K., Pareek, A., Sopory, S.K. and Singla-Pareek, S.L. (2018) Manipulation of glyoxalase pathway confers tolerance to multiple stresses in rice. *Plant Cell Environ*, **41**, 1186-1200.

Hasanuzzaman, M., Nahar, K., Hossain, M.S., Mahmud, J.A., Rahman, A., Inafuku, M., Oku, H. and Fujita, M. (2017) Coordinated Actions of Glyoxalase and Antioxidant Defense Systems in Conferring Abiotic Stress Tolerance in Plants. *Int J Mol Sci*, **18**.

Hossain, M.A., Hossain, M.Z. and Fujita, M. (2009) Stress-induced changes of methylglyoxal level and glyoxalase I activity in pumpkin seedlings and cDNA cloning of glyoxalase I gene. *Australian Journal of Crop Science*, **3**, 53.

Hudig, M., Schmitz, J., Engqvist, M.K.M. and Maurino, V.G. (2018) Biochemical control systems for small molecule damage in plants. *Plant signaling & behavior*, **13**, e1477906.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P. and Hassabis, D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583-589.

Kliebenstein, D.J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J. and Mitchell-Olds, T. (2001) Genetic control of natural variation in Arabidopsis glucosinolate accumulation. *Plant Physiol.*, **126**, 811-825.

Koornneef, M., Alonso-Blanco, C. and Vreugdenhil, D. (2004) Naturally occurring genetic variation in Arabidopsis thaliana. *Annu. Rev. Plant Biol.*, **55**, 141-172.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P. and Rombauts, S. (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.*, **30**, 325-327.

Luedemann, A., Strassburg, K., Erban, A. and Kopka, J. (2008) TagFinder for the quantitative analysis of gas chromatography--mass spectrometry (GC-MS)-based metabolite profiling experiments. *Bioinformatics*, **24**, 732-737.

Mano, J., Biswas, M.S. and Sugimoto, K. (2019) Reactive Carbonyl Species: A Missing Link in ROS Signaling. *Plants (Basel)*, **8**.

Mansoor, S., Ali Wani, O., Lone, J.K., Manhas, S., Kour, N., Alam, P., Ahmad, A. and Ahmad, P. (2022) Reactive Oxygen Species in Plants: From Source to Sink. *Antioxidants (Basel)*, **11**.

Marasinghe, G.P., Sander, I.M., Bennett, B., Periyannan, G., Yang, K.W., Makaroff, C.A. and Crowder, M.W. (2005) Structural studies on a mitochondrial glyoxalase II. *J. Biol. Chem.*, **280**, 40668-40675.

Mitchell-Olds, T. (1995) The molecular basis of quantitative genetic variation in natural populations. *Trends Ecol. Evol.*, **10**, 324-328.

Nagai, R., Unno, Y., Hayashi, M.C., Masuda, S., Hayase, F., Kinae, N. and Horiuchi, S. (2002) Peroxynitrite induces formation of N(epsilon)-(carboxymethyl) lysine by the cleavage of Amadori product and generation of glucosone and glyoxal from glucose: novel pathways for protein modification by peroxynitrite. *Diabetes*, **51**, 2833-2839.

Nahar, K., Hasanuzzaman, M., Alam, M.M., Rahman, A., Mahmud, J.A., Suzuki, T. and Fujita, M. (2017) Insights into spermine-induced combined high temperature and drought tolerance in mung bean: osmoregulation and roles of antioxidant and glyoxalase system. *Protoplasma*, **254**, 445-460.

Prasch, C.M. and Sonnewald, U. (2013) In silico selection of Arabidopsis thaliana ecotypes with enhanced stress tolerance. *Plant signaling & behavior*, **8**, e26364.

Ridderstrom, M. and Mannervik, B. (1996) The primary structure of monomeric yeast glyoxalase I indicates a gene duplication resulting in two similar segments homologous with the subunit of dimeric human glyoxalase I. *Biochem. J.*, **316 (Pt 3)**, 1005-1006.

Saxena, M., Deb Roy, S., Singla-Pareek, S.-L., Kumar Sopory, S. and Bhalla-Sarin, N. (2011) Overexpression of the glyoxalase II gene leads to enhanced salinity tolerance in Brassica juncea. *The Open Plant Science Journal*, **5**.

Schmitz, J., Dittmar, I.C., Brockmann, J.D., Schmidt, M., Hudig, M., Rossoni, A.W. and Maurino, V.G. (2017) Defense against Reactive Carbonyl Species Involves at Least Three Subcellular Compartments Where Individual Components of the System Respond to Cellular Sugar Status. *Plant Cell*, **29**, 3234-3254.

Schmitz, J., Rossoni, A.W. and Maurino, V.G. (2018) Dissecting the Physiological Function of Plant Glyoxalase I and Glyoxalase I-Like Proteins. *Front Plant Sci*, **9**, 1618.

Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**, 671-675.

Schulman, A.H., Oksman-Caldentey, K.M. and Teeri, T.H. (2020) European Court of Justice delivers no justice to Europe on genome-edited crops. *Plant Biotechnol. J.*, **18**, 8-10.

Semchyshyn, H.M. (2014) Reactive carbonyl species in vivo: generation and dual biological effects. *TheScientificWorldJournal*, **2014**, 417842.

Singla-Pareek, S.L., Yadav, S.K., Pareek, A., Reddy, M.K. and Sopory, S.K. (2008) Enhancing salt tolerance in a crop plant by overexpression of glyoxalase II. *Transgenic Res.*, **17**, 171-180.

Thornalley, P.J. (1990) The glyoxalase system: new developments towards functional characterization of a metabolic pathway fundamental to biological life. *Biochem. J.*, **269**, 1-11.

Thornalley, P.J. (2003) Glyoxalase I--structure, function and a critical role in the enzymatic defence against glycation. *Biochem. Soc. Trans.*, **31**, 1343-1348.

Thornalley, P.J., Langborg, A. and Minhas, H.S. (1999) Formation of glyoxal, methylglyoxal and 3-deoxyglucosone in the glycation of proteins by glucose. *Biochem. J.*, **344**, 109-116.

Turra, G.L., Agostini, R.B., Fauguel, C.M., Presello, D.A., Andreo, C.S., Gonzalez, J.M. and Campos-Bermudez, V.A. (2015) Structure of the novel monomeric glyoxalase I from Zea mays. *Acta Crystallogr D Biol Crystallogr*, **71**, 2009-2020.

Veena, Reddy, V.S. and Sopory, S.K. (1999) Glyoxalase I from Brassica juncea: molecular cloning, regulation and its over-expression confer tolerance in transgenic tobacco under stress. *Plant J.*, **17**, 385-395.

Weigel, D. (2012) Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiol.*, **158**, 2-22.

Welchen, E., Schmitz, J., Fuchs, P., Garcia, L., Wagner, S., Wienstroer, J., Schertl, P., Braun, H.P., Schwarzlander, M., Gonzalez, D.H. and Maurino, V.G. (2016) D-Lactate dehydrogenase links methylglyoxal degradation and electron transport through cytochrome C. *Plant Physiol.*

Yan, G., Xiao, X., Wang, N., Zhang, F., Gao, G., Xu, K., Chen, B., Qiao, J. and Wu, X. (2018) Genome-wide analysis and expression profiles of glyoxalase gene families in Chinese cabbage (Brassica rapa L). *PloS one*, **13**, e0191159.

Table 1. Kinetic parameters of Col-0 GLXI;2 WT and K229N, F278I, and K229N-F278I mutants using MGO and Ni^{2+}, or GO and Mn^{2+}. The kinetic constants were calculated by nonlinear regression analysis. Values are means ± SE (*n*=3) of independent enzyme preparations. Testing for significant differences was performed by t-test ($P < 0.05$).

Col-0 GLXI;2	**MGO**			**GO**		
	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)
WT	297 ± 14	7.5 ± 0.9	25.3 ± 1.7	411 ± 18	2.8 ± 0.3	6.8 ± 0.6
K229N	295 ± 23	9.1 ± 1.7	26.4 ± 2.2	415 ± 17	2.4 ± 0.3	6.5 ± 0.5
F278I	311 ± 18	8.4 ± 1.2	27.0 ± 2.4	403 ± 12	2.5 ± 0.4	6.2 ± 0.3
K229N-F278I	302 ± 15	8.1 ± 1.1	25.8 ± 2.1	407 ± 15	2.6 ± 0.3	6.7 ± 0.5

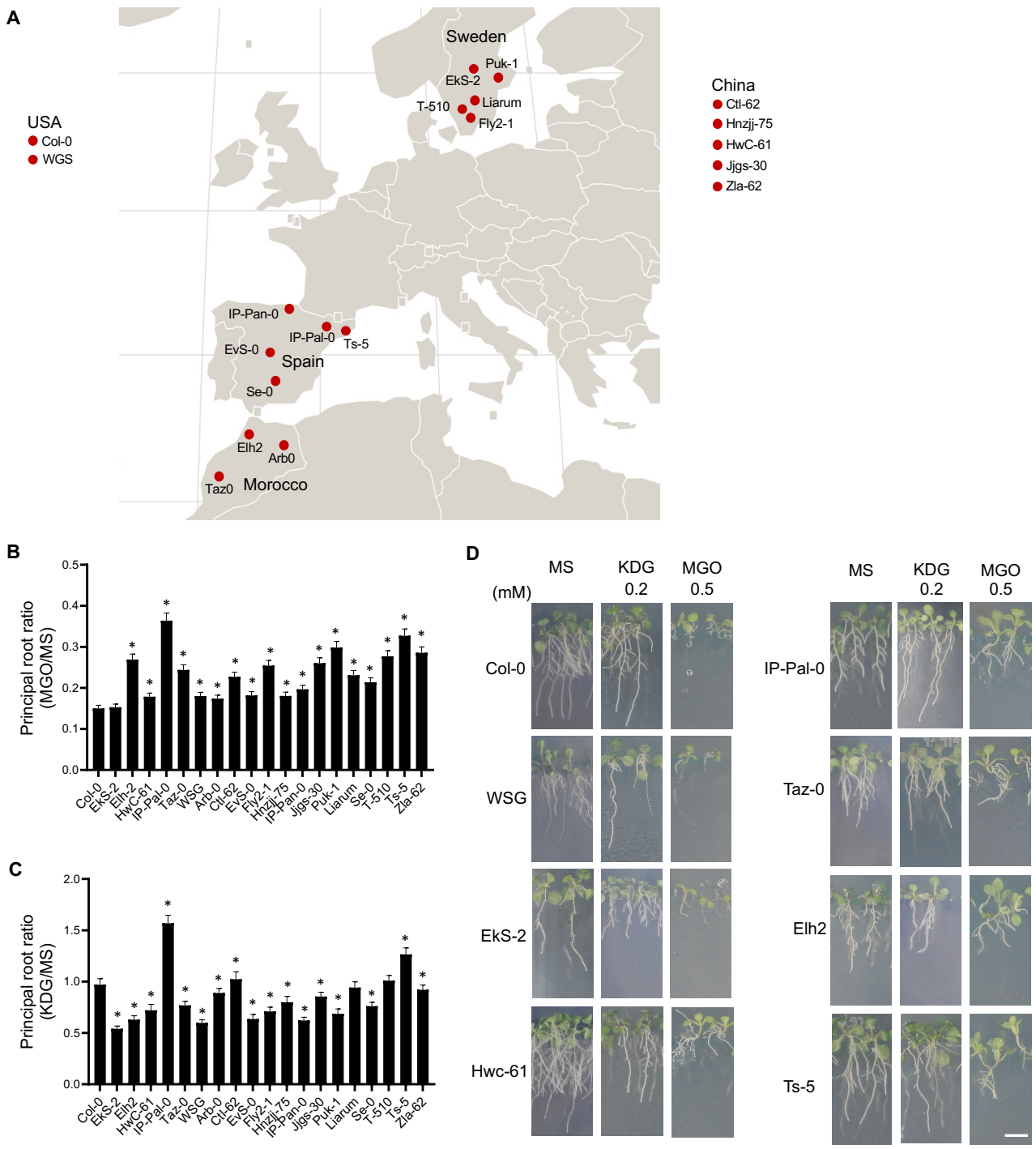

Figure 1. Effect of KDG and MGO on the development of *A. thaliana* ecotypes. (A) Geographical origin of the ecotypes analyzed. For those from Spain and Sweden, the localisation shows the exact coordinates provided by 1001genomes (https://1001genomes.org). (B) Main root length after 14 days of germination on media containing 0.5 mM MGO relative to the length in control media (MS). (C) Main root length after 14 days of germination on media containing 0.2 mM KDG relative to the length in MS. The graphs show the means and standard errors of the measurements of 3 independent experiments (*n*=3) of 20 roots each. Significant

differences were analyzed by one-way ANOVA. An asterisk (*) indicates a significant difference ($P<0.001$) between an ecotype and Col-0 by Dunnett's multiple comparison test. (D) Images of seedling of selected ecotypes at 14 d after germination in different growth media. Scale bar = 1 cm.

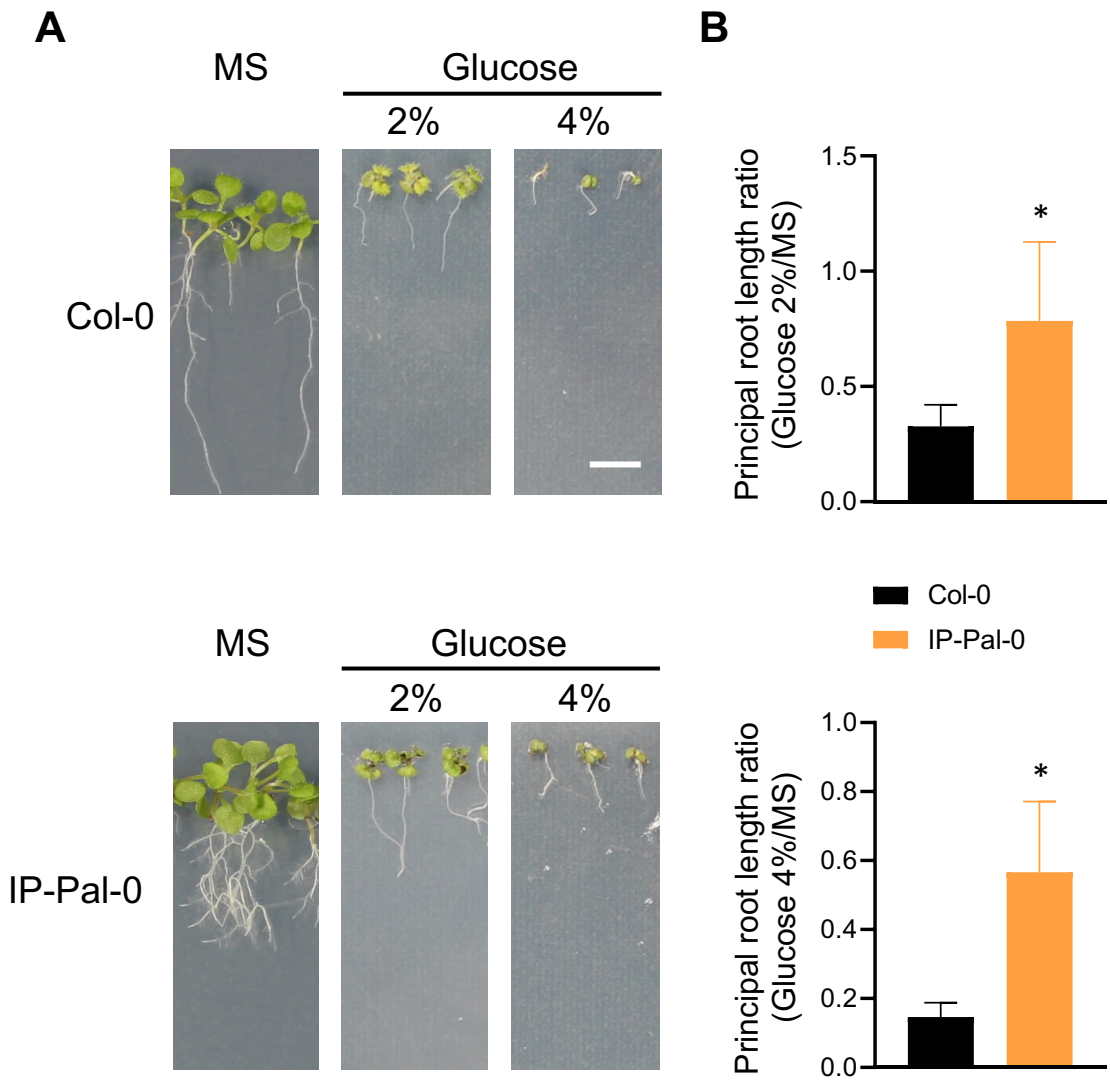

Figure 2. Development of seedlings of Col-0 and IP-PAL-0 in the presence of 2% and 4% glucose in the growth media. (A) Images of seedlings 14 d after germination in different growth media. Scale bar = 1 cm. (B) Main root length after 14 days of germination on media containing 2% or 4% glucose compared to the length in control media (MS). The graphs show the means and standard errors of the measurements of 3 independent experiments (*n*=3) of 20 roots each. An asterisk (*) indicates a significant difference (P <0.001) between Col-0 and IP-Pal-0 determined by unpaired t-test.

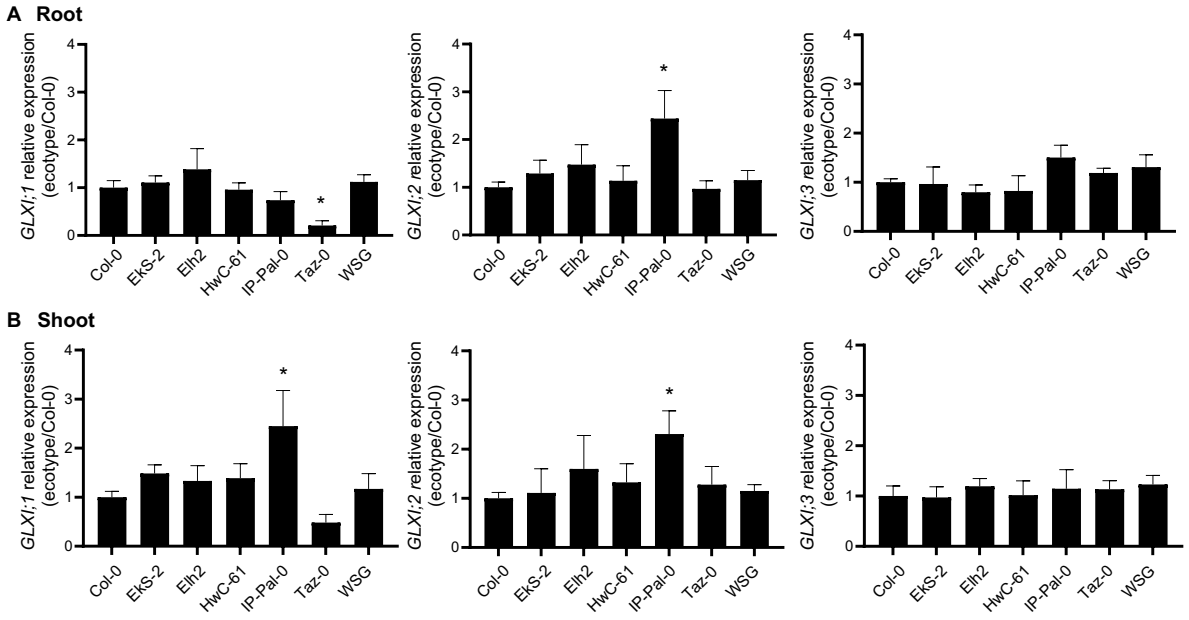

Figure 3. Transcriptional analysis of the GLXI isoforms in selected *A. thaliana* ecotypes grown on MS. Expression levels of GLXI isoforms in root (A) and shoot (B) of 14-week-old seedlings of selected ecotypes grown on MS relative to those measured in Col-0. One-way ANOVA was used to test for significant differences. An asterisk (*) indicates a significant difference (P<0.001) between an ecotype and Col-0 by Dunnett's multiple comparison test.

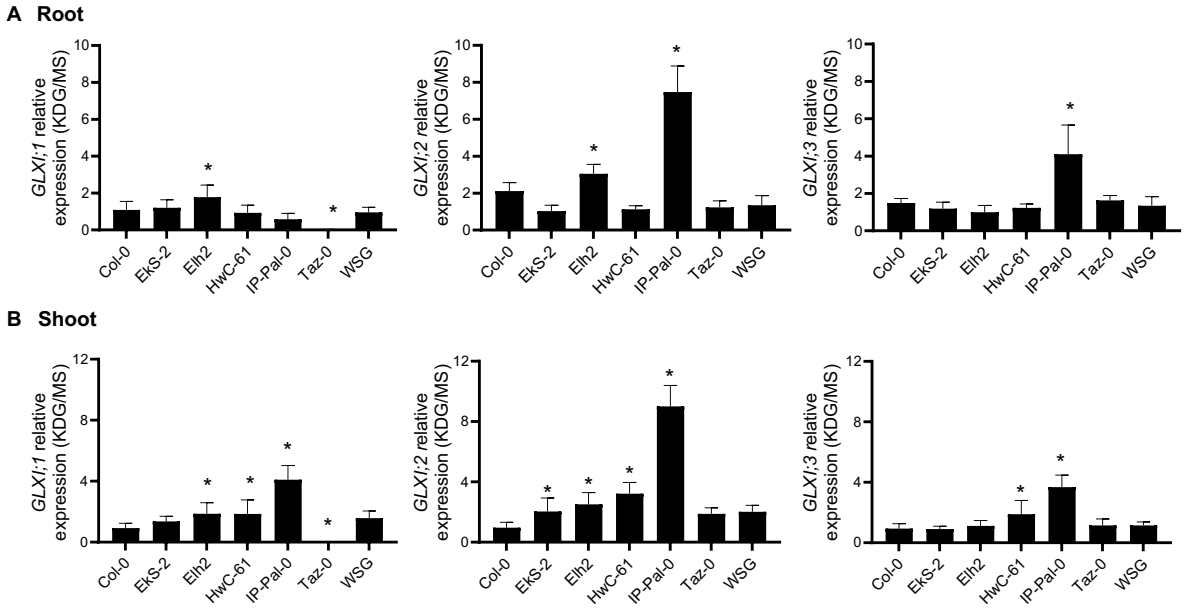

Figure 4. Transcriptional analysis of the GLXI isoforms in selected *A. thaliana* ecotypes grown on media with and without KDG. Expression levels of GLXI isoforms in root (A) and shoot (B) of 14-week-old seedlings of selected ecotypes grown in 0.2 mM KDG relative to those measured in seedlings grown in MS. One-way ANOVA was used to test for significant differences. An asterisk (*) indicates a significant difference (P<0.001) between an ecotype and Col-0 by Dunnett's multiple comparison test.

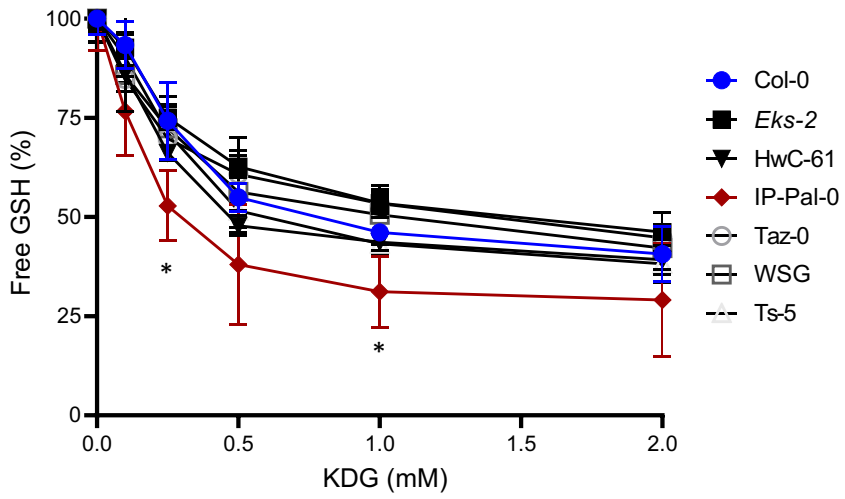

Figure 5. GLXI activity using KDG in protein extracts from seedlings of selected ecotypes. The concentration of free GSH was determined after incubation of protein extracts with different concentrations of KDG. The initial amount of free GSH when no KDG was added to the reaction media was assumed to be 100%. Values are mean ± SD (*n*=3). An asterisk (*) indicates a significant difference between ecotype and Col-0 values at each KDG concentration ($P<0.05$), determined by unpaired t-test.

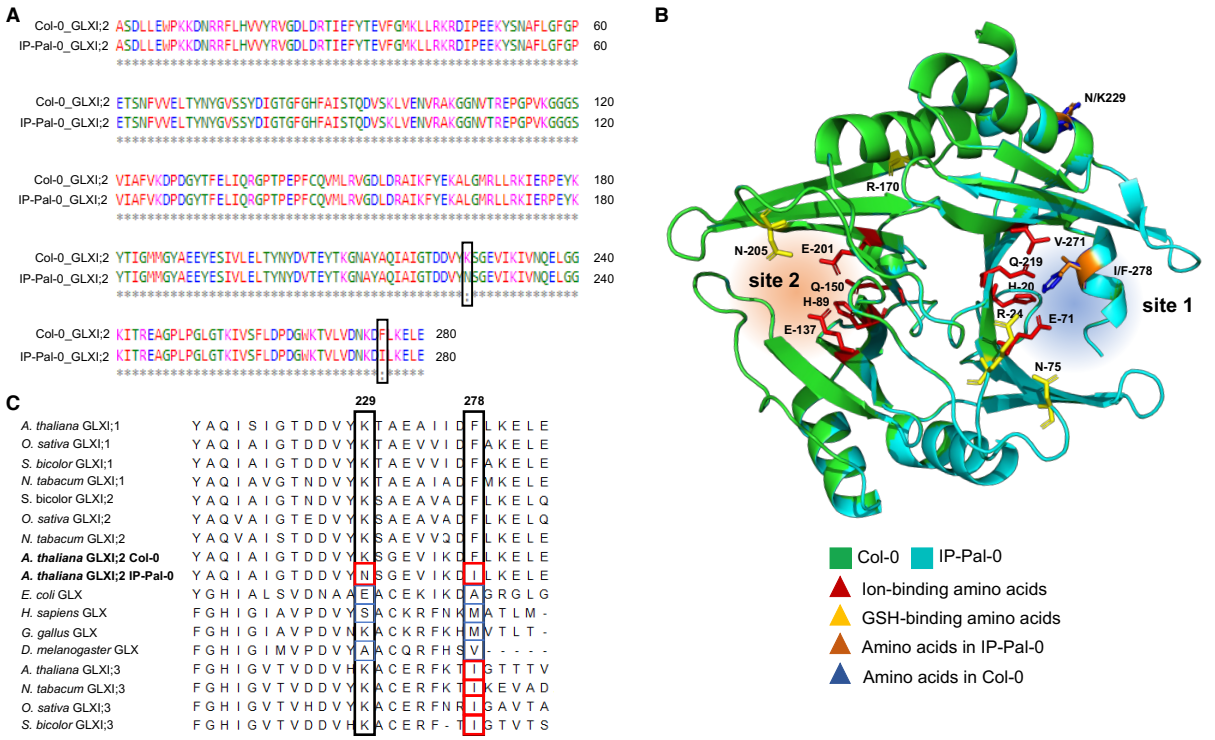

Figure 6. Analysis of GLXI sequences and structures. (A) Alignment of the primary amino acid sequences of GLXI;2 of IP-Pal-0 and Col-0. (B) Model structures of Col-0 (green) and IP-Pal-0 (light blue) showing the localization of the amino acids at positions 229 and 278 and those involved in ion and GSH at the binding functional active site (site 2) and the equivalents in the cryptic site (site 1). (C) Alignment of 17 selected Ni^{2+}-type and Zn^{2+}-type GLXI sequences from different organisms showing differential conservation of the amino acids at position 229 and 278 (corresponding to the *A. thaliana* Col-0 sequence). The analysis was conducted with the MEGA7 software. All 72 sequences analysed are listed in Supplementary Table 1.

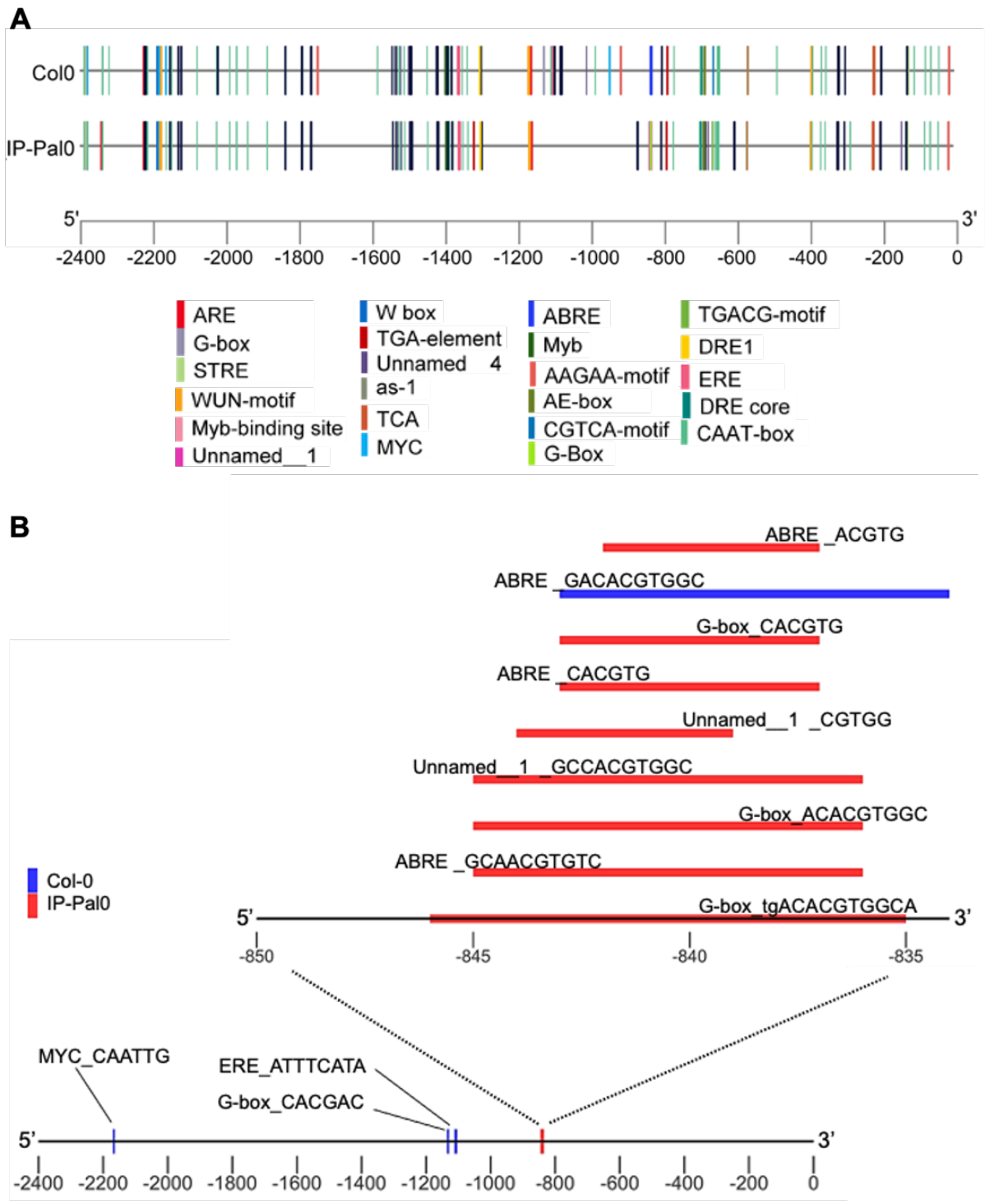

Figure 7. Cis-regulatory elements in Col-0 and IP-PAL-0. (A) Promoter analysis for the presence of cis-regulatory elements 1.5 kb upstream of the 5'UTR (2.4 kb upstream of the transcription initiation site) of *GLX1;2* in the *A. thaliana* ecotypes Col-0 and IP-Pal-0. Cis-regulatory elements were mapped using TBtools (Chen *et al.*, 2020). (B) Unique cis-regulatory elements of *GLX1;2* identified in the ecotypes Col-0 and IP-Pal-0.

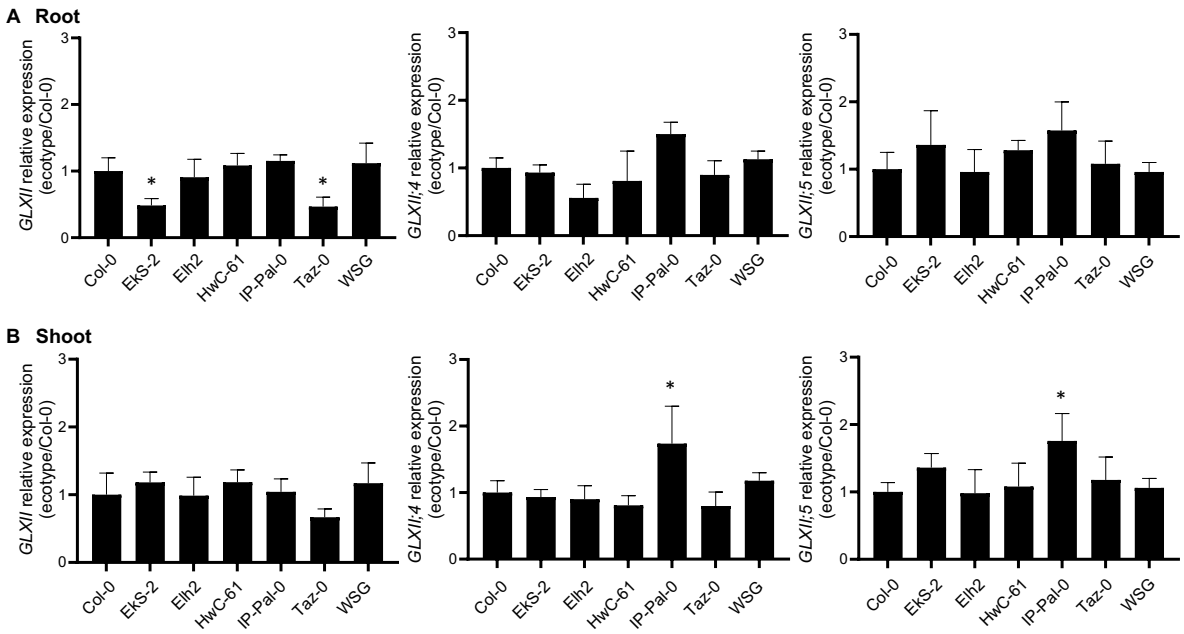

Figure 8. Transcriptional analysis of the GLXII isoforms in selected *A. thaliana* ecotypes grown on MS. Expression levels of GLXII isoforms in root (A) and shoot (B) of 14-week-old seedlings of selected ecotypes grown on MS relative to those measured in Col-0. One-way ANOVA was used to test for significant differences (P<0.0001). An asterisk (*) indicates a significant difference (P<0.001) between an ecotype and Col-0 by Dunnett's multiple comparison test.

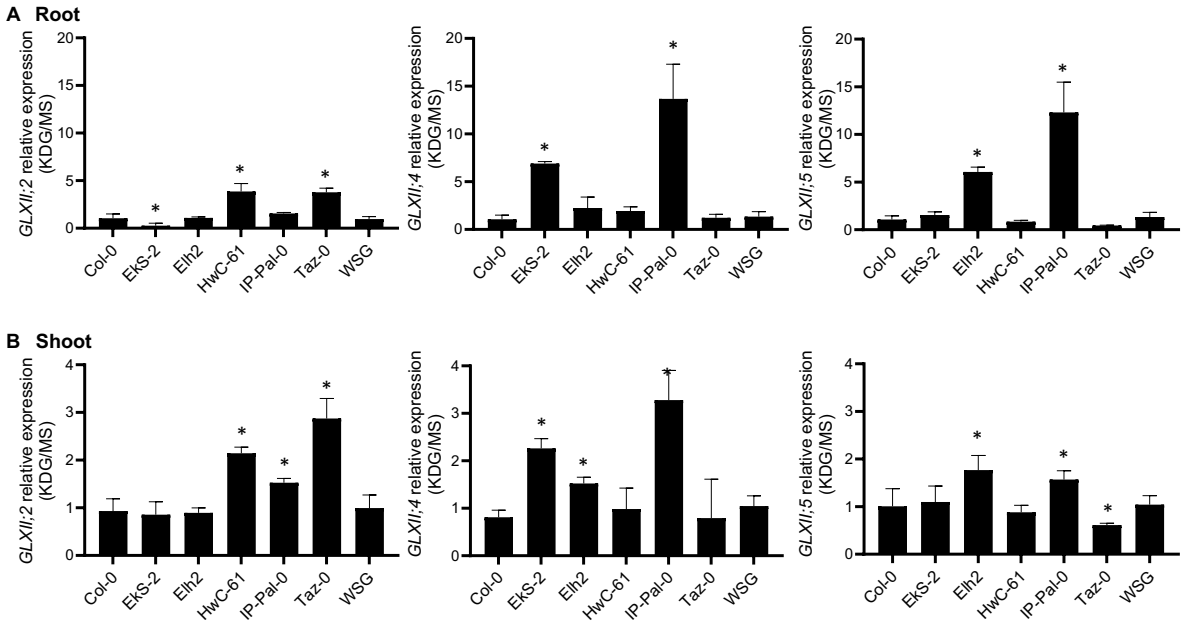

Figure 9. Transcriptional analysis of the GLXII isoforms in selected *A. thaliana* ecotypes grown on media with and without KDG. Expression levels of GLXI isoforms in root (A) and shoot (B) of 14-week-old seedlings of selected ecotypes grown in 0.2 mM KDG relative to those measured in seedlings grown in MS. One-way ANOVA was used to test for significant differences (P<0.0001). An asterisk (*) indicates a significant difference (P<0.001) between an ecotype and Col-0 by Dunnett's multiple comparison test.

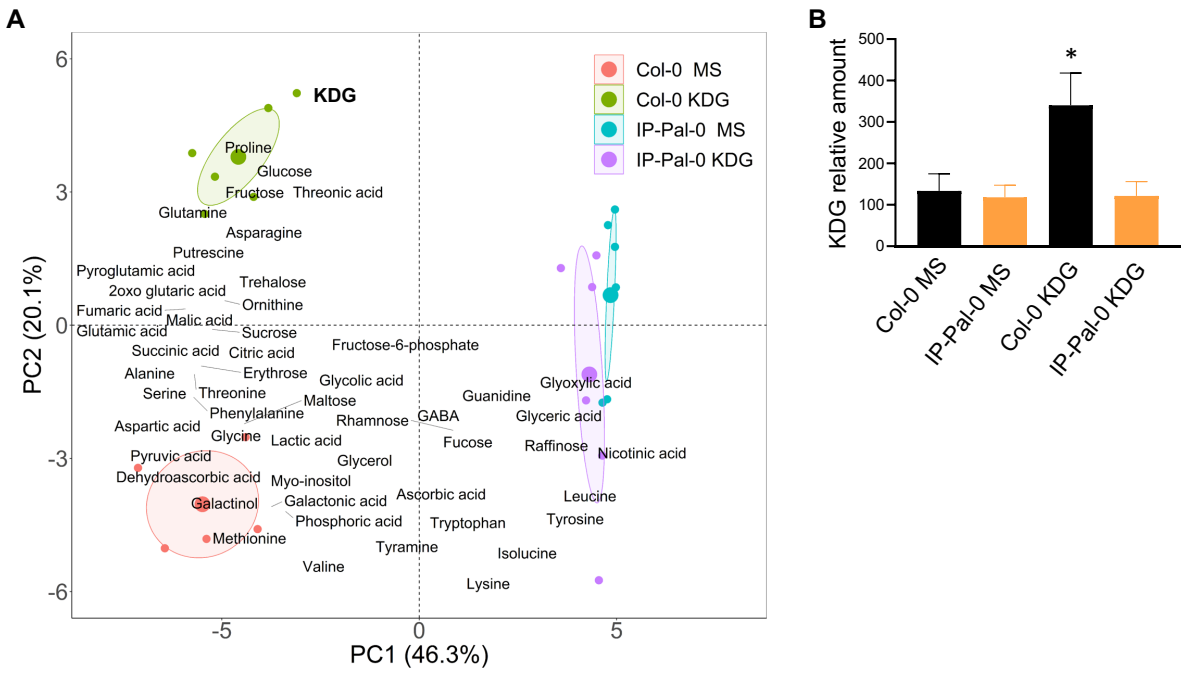

Figure 10. Analysis of the metabolite relative contents. (A) 2D Principal Component Analysis (PCA) biplot of variables in black labels (metabolites measured in seedlings grown in media without (MS) or with 0.2 mM KDG and individuals in colored points (*A. thaliana* ecotypes Col-0 and IP-Pal-0). Ellipses group the ecotypes per growth medium. The first axis explains 46.3% of the variance and the second axis explains 20.1%. (B) Relative amount of KDG in seedlings of Col-0 and IP-Pal-0 grown in media without (MS) and with 0.2 mM KDG.