Plant Soil
https://doi.org/10.1007/s11104-026-08980-z

RESEARCH ARTICLE

Molecular and physiological mechanisms of the cadmium response in seedlings of two *Theobroma cacao L.* genotypes

P. Delgadillo-Durán · F. M. Menéndez-Burns · C. Rodríguez-Medina · A. C. Montenegro · A. Istvan · M. J. Guiltinan · R. Yockteng · S. N. Maximova

Received: 14 January 2026 / Accepted: 10 August 2026
© The Author(s) 2026

Abstract

Aims Understanding the mechanisms of cadmium (Cd) accumulation in cacao plants is critical to mitigating health risks associated with Cd exposure through chocolate consumption. A general and genotype-specific molecular and physiological responses were characterized.

Methods Seedlings of cacao genotypes PA121 and TSH660 were exposed to 0 and 10 ppm Cd in hydroponic conditions. Samples were collected at 0, 24, and 48 h (RNAseq) and at 60 days post-treatment (ICP-OES). Gene expression profiles were compared using differential gene expression (DEG) and gene ontology analyses. Gas exchange (Gs) and abscisic acid (ABA) measurements were conducted on greenhouse-grown PA121 seedlings.

Results Both shared and genotype-specific changes were observed in genes associated with detoxification, ROS, and hormone metabolism, notably ABA-related genes. In leaves, the low-accumulator exhibited fewer DEGs than the high-accumulator, whereas in roots, both genotypes displayed similar expression profiles. The high-accumulating genotype exhibited fewer statistically significant expression changes in key hormone biosynthesis genes. In roots, Cd transport family genes, including *ZIP/IRT* and *NRAMP,* were downregulated. The high-accumulator genotype exhibited significant downregulation of one *ZIP* and one *HMA5* ortholog, whereas the response of the low-accumulator genotype included upregulation of a different *HMA5* ortholog. Additionally, high-accumulator seedlings exposed to 12 ppm Cd reduced Gs without significant changes in photosynthesis, while ABA concentrations increased in roots at 36 h after Cd exposure.

Conclusions Together, these results support a model in which Cd induces ABA-mediated root signaling

Responsible Editor: Guillaume Echevarria.

P. Delgadillo-Durán and F. M. Menéndez-Burns contributed equally to the manuscript.

The authors responsible for the distribution of materials integral to the findings presented in this article, in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plphys/pages/General-Instructions), are Siela Maximova and Roxana Yockteng.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11104-026-08980-z.

P. Delgadillo-Durán · A. C. Montenegro · R. Yockteng (✉)
Km 14 Vía Mosquera. Centro de Investigación Tibaitatá, Corporación Colombiana de Investigación Agropecuaria - Agrosavia, Mosquera, Colombia
e-mail: ryockteng@agrosavia.co

F. M. Menéndez-Burns · M. J. Guiltinan · S. N. Maximova (✉)
Department of Plant Science, The Pennsylvania State University, University Park, PA, USA
e-mail: snm104@psu.edu

Published online: 08 September 2026

that reduces stomatal conductance and transpiration-driven mass flow, while suppressing the expression of active Cd transporters, thereby limiting Cd uptake.

Keywords Cadmium · Theobroma cacao · ABA · RNAseq · Gas exchange · Heavy metals

Introduction

Cadmium (Cd) is an environmental toxin present in many human food sources and is associated with several adverse health outcomes in humans (Charkiewicz et al. 2023). Cd occurs naturally in different types of rocks, usually associated with carbonates and phosphorites, and is present in soils at concentrations ranging from 0.01 to 1 mg/kg, with a global mean of approximately 0.36 mg/kg (Kubier et al. 2019). Among heavy metals, Cd exhibits the highest mobility in soils (Kubier et al. 2019). Due to its physicochemical similarity to essential elements, Cd can enter and accumulate in plant tissues at high rates through uptake mechanisms normally used for essential macro- and micronutrients (Lux et al. 2011; Kubier et al. 2019).

Theobroma cacao L. tree, the seeds of which are the primary ingredient of chocolate, is the focus of intense research because of its capacity to accumulate Cd. Similar to other plants, cacao trees uptake Cd from soil following the physical principles controlling water and solute movement (Barber 1995), with competing elemental abundance and solubility and pH acting as key determinants in bioavailability to plants (Smolders 2001; Wade et al. 2022; Barati et al. 2023).

F. M. Menéndez-Burns · M. J. Guiltinan · S. N. Maximova
Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, USA

C. Rodríguez-Medina
Centro de Investigación Palmira, Corporación Colombiana de Investigación Agropecuaria - Agrosavia, Diagonal a La Intersección de La Carrera 36 A Con Calle 23, Valle del Cauca, Colombia

A. Istvan
Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA

R. Yockteng
Institut de Systématique, Évolution, Biodiversité-UMR-CNRS 7205, National Museum of Natural History, Paris, France

At the molecular level, several metal transporters involved in diffusion-based uptake have been functionally characterized, including members of the NRAMP family in cacao TcNRAMP5 and TcNRAMP6 (Ullah et al. 2018); and the HMA family member TcHMA3 (Moore et al. 2020), which mediate Zn^{2+} and Cd^{2+} transmembrane transport, the former proposed to function at the rhizosphere and the latter in subcellular compartments. Additionally, the ZIP family transporter TcIRT1 has been shown to facilitate the uptake of both Cd^{2+} and Fe^{2+} (Dashner 2024).

As a consequence of Cd accumulation in cacao seeds, chocolate has been identified as a major contributor to dietary Cd exposure in the European population (EFSA 2012). Chocolate products with higher cocoa solids content tend to exhibit higher Cd concentrations, although levels vary depending on the geographical origin of the cacao (Pedron et al. 2025). However, Cd accumulation varies significantly among cacao genotypes (Lewis et al. 2018; Engbersen et al. 2019; Moore et al. 2020; Vanderschueren et al. 2021). Furthermore, differing Cd isotope profiles in seeds and xylem (Barati et al. 2023) as well as differing variances between seed, leaf, and roots (Lewis, et al. 2018) suggest that Cd uptake, translocation, and compartmentalization are likely governed by distinct mechanisms. Likewise, this variability indicates that the mechanisms governing Cd uptake and accumulation, as well as the associated molecular responses and physiological adaptations, differ among genotypes.

Plants in general are affected by Cd, but in rare cases display tolerance to its deleterious effects. Sensitivity to Cd is typically characterized by pronounced growth inhibition and strong molecular responses (Rui et al. 2016; Feng et al. 2018). Cd tolerance, in contrast, refers to the ability to maintain growth despite high Cd exposure without necessarily hyperaccumulating the metal (Ernst et al. 2008). For food crops, an ideal phenotype would combine normal growth on Cd-contaminated soils with the ability to exclude Cd from edible tissues (Wang et al. 2021).

At a physiological level, cacao plants exposed to Cd exhibit reduced photosynthetic performance and biomass accumulation, often accompanied by decreased water use (Barroso et al. 2023; Vera Pinargote et al. 2025). In one field trial, photosynthetic efficiency was associated with reduced Cd uptake and growth tolerance to Cd, suggesting that this trait may be useful for selecting cacao genotypes with

enhanced Cd avoidance capacity (Vera Pinargote et al. 2025). Additional physiological traits of interest include the preferential accumulation of Cd in lignified tissues, which generally follow the pattern roots > stems > leaves > nibs (Vanderschueren et al. 2021; Blommaert et al. 2022).

Other hallmark physiological responses to Cd reported in plants, but not yet corroborated in cacao, include early signaling, such as a rapid spike in calcium levels (Zhang et al. 2023), increases in reactive oxygen species (ROS) (Garnier et al. 2006; Chmielowska-Bąk et al. 2014; Pérez-Chaca et al. 2014; Liu et al. 2024), and antioxidant activation, which includes the regulation of glutathione-S transferase (*GST*) (Herbette et al. 2006; Liu et al. 2017, 2024; Wang et al. 2023a), and modulation of abscisic acid (ABA) (Hsu and Kao 2003; Stroiński et al. 2013; Liu et al. 2024), ethylene (Rodriguez-Serrano et al. 2009; Chmielowska-Bak et al. 2013; Schellingen et al. 2014; Liu et al. 2024), and indole-3-acetic acid (IAA) (Chmielowska-Bąk et al. 2014).

We therefore investigated the molecular mechanisms underlying high or low Cd accumulation in cacao by identifying associated candidate genes and characterizing the molecular components involved in Cd uptake, translocation, and tolerance. To this end, two genotypes were selected from a large Cd-uptake screening: a low-accumulating genotype, TSH660 (TSH660LA), and a high-accumulating genotype, PA121 (PA121HA). These contrasting genotypes were subjected to transcriptomic analysis. Additionally, we have developed a structured working model that integrates our transcriptomic results with existing knowledge from cacao and other plant species, providing insights into the molecular basis of Cd accumulation in cacao.

Materials and methods

Plant material and growth conditions

Plants used for measuring Cd concentration in cacao tissues and growth parameters, and RNAseq analysis were grown in the AGROSAVIA Research Center located in Palmira, Valle del Cauca, Colombia (3° 31′ 12″ N, 76° 19′ 50″ W at ~ 100 m.a.s.l.), with an average annual temperature of 26 °C and relative humidity of 65%. Plants were grown from open-pollinated seeds obtained from two clonally propagated mother plants. Genotypes TSH660 and PA121 were selected based on a preliminary screening experiment conducted at AGROSAVIA's Palmira Research Station under similar conditions. Open-pollinated seeds from the mother plants of both genotypes were germinated in sand, grown for two months in a greenhouse, and transferred to 150 L containers with half-strength Hoagland's solution (Hoagland and Arnon 1950) (7.5 mM N, 3.5 mM K, 2.5 mM Ca, 0.5 mM P, 0.1 mM S, 1 mM Mg, 1.01 mM Fe, 0.5 mM B, 0.5 mM Mn, 0.05 mM Zn, 0.02 mM Cu, and 0.01 mM Mo). For transcriptome analysis, seedlings were grown in three individual 150 L tanks, each housing one plant of each genotype, totaling three plants per genotype. After four weeks, $Cd(NO_3)_2$ was added at a concentration of 10 ppm Cd (~ 0.089 mM Cd), increasing total nitrogen (N) by 0.027 mM. After two months of Cd exposure, Cd concentration in leaves and roots, and plant growth parameters were measured.

For gas exchange and phytohormone quantification, plants were grown at The Pennsylvania State University, University Park, PA, USA, in a greenhouse with controlled conditions: 16-h average PPFD of 116.2 $\mu mol \bullet m^{-2} \bullet s^{-1}$ (SD = 10 $\mu mol \bullet m^{-2} \bullet s^{-1}$), temperature of 29.82 °C (SD = 5.9 °C), and humidity at 73.84% RH (SD = 15.43% RH). Forty-eight seeds from open-pollinated PA121 mother plants were obtained from the USDA, ARS Tropical Agriculture Research Station in Mayagüez, Puerto Rico. They were germinated in 32-Star Deep Plug Trays (T.O. Plastics, Clearwater, MN) filled with sand and misted for 2 min every 8 min. Plants were transferred after a week to 7 in D40H Deepots™ (Stuwe & Sons, Tangent, Oregon) with two drippers (flow rate of 22 mL $minute^{-1}$), and fertigated with a modified Hoagland's solution (16 mM N, 6 mM K, 4 mM Ca, 2 mM P, 1 mM S, 1 mM Mg, 0.00537 mM Fe, 0.05 mM Cl, 0.025 mM B, 0.002 mM B, 0.002 mM Zn, 0.0005 mM Cu, 0.00350 mM Mo) twice a day at 8 a.m. and 6 p.m. for 2 min each time for 3 months before 48 plants were randomly assigned to 3 treatment levels ($n = 16$) of Cd^{2+} ion (added $CdCl_2$): ~ 0.055 mM (6.13 ppm), ~ 0.1096 mM (12.26 ppm) or without Cd^{2+} treatment. The higher Cd concentration was applied to compensate for continuous exposure to liquid media in the initial experiment conducted in Colombia. These experiments

were limited to genotype PA121 due to its availability in Puerto Rico, as well as regulations restricting the movement of plant material between the U.S.A. and Colombia, and logistical constraints related to equipment and personnel.

Cd analysis

Cd concentrations were measured at the AGROSAVIA Research Center, Tibaitatá (Mosquera, Colombia) using inductively coupled plasma optical emission spectrometry (ICP-OES Thermo Scientific ICAP 6500) following a modified method of the US EPA (US Environmental Protection Agency 1996, 2007). Calibration curves were performed using the standard reference material from the National Institute of Standards and Technology, a solution of 1000 mg/L Cd prepared by digesting Cd $(NO_3)_2$ in 0.5 mol/L HNO_3 (Certipur®). Calibration curves were performed using reference cacao material from genotype WEPAL IPE 965 (WEPAL 2019) provided by Wageningen Agricultural University (Netherlands).

Tissue sampling, RNA purification, RNA-Seq library preparation, and sequencing

Leaves at developmental stage C (Fister et al. 2016) were excised from plants once before the application of Cd (time point 0) and then 24 and 48 h after the application of Cd (AC), then immediately frozen in liquid nitrogen, and stored at −80 ºC. Root tissues were sampled similarly from the same plants at 0 and 48 h. One hundred mg of tissue from each sample was ground in liquid nitrogen using a mortar and pestle and used to purify total RNA with the PureLink™ Plant RNA Reagent, following the protocol of Yockteng et al. (2013). RNA concentrations were determined spectrophotometrically using a NanoDrop® 2000 (ThermoScientific). RNA was separated on a 1% agarose gel and stained with GelRed® (Biotium) to confirm RNA integrity. Barcoded libraries were made from 1 µg of RNA using the TruSeq RNA Library Prep Kit v2, following the manufacturer's Instructions (Illumina, San Diego, CA). The quality of each library was assessed using a high-sensitivity DNA kit on an Agilent TapeStation 4200 (Agilent Technologies, Santa Clara, CA). In total, 18 libraries from leaf tissue and 12 from root tissue were constructed and sequenced using a 150 bp paired-end strategy on an Illumina HiSeq X system, dedicating one lane to each replicate tank, targeting a 300 bp insert size. Sequencing was performed at Macrogen Inc. (Seoul, South Korea).

Transcriptome assembly and read mapping

Adapters and low-quality reads were removed using Trimmomatic (Bolger et al. 2014), and the quality of the remaining reads was verified using FastQC (Andrews 2010). Filtered reads were mapped using Hisat2 (Kim et al. 2019) to a GFF annotation containing the last exon of each gene model of Criollo B97 from Belize V2 reference genome (Argout et al. 2017) (Figure S1A-P). In the study, we present the transcript annotations using NCBI's RefSeq for that genome. Fractional counts were calculated using the *featureCounts* package (Liao et al. 2014). Genes with less than 10 normalized reads across samples were removed.

Log2 fold change calculation, filtering, and size factor estimation

DEGs were identified using functions from the package DESeq2 (Michael et al. 2022) and an ad hoc R function created to minimize type I error and maximize DEG discovery. The function had three main capabilities: 1) filtering genes based on mapping events across libraries, as performed by the edgeR package (Robinson et al. 2010); 2) utilizing the row variance for the independent filtering statistic as suggested previously (Bourgon et al. 2010); 3) performing steps one and two and normalization using gene dispersion values only using libraries included in the comparison. An adjusted p-value cut-off of 0.05 was applied.

The gene expression model included "genotype" and "time after Cd treatment (AC)" as the main effects and a blocking factor to account for tank-to-tank variation. The interaction between time AC and genotype was examined. Since each plant was continuously sampled throughout the experiment, each plant is nested within the interaction of the "genotype" and "block" terms. PA121 and 0 h were used as references. Another set of models without "genotype" was applied to compute log2-fold changes of combined genotypes (model matrix and contrast weights; Table S1 A-C).

Ortholog classification, pathway mapping, and subcellular localization prediction

DEGs were reciprocally blasted using CRB-BLAST (Aubry et al. 2014) to the coding sequences (CDS) of the *A. thaliana* genome (arabidopsis.org) using three E-value cut-offs: 10^{-2}, 10^{-3}, and 10^{-5}. Gene families were then identified using PlantTribes2 (Wafula et al. 2023). KEGG pathway mapping (Kanehisa et al. 2017) was performed using orthogroups identified by KOFAM KOALA (Aramaki et al. 2020). Subcellular localization was predicted using DeepLoc2 (Thumuluri et al. 2022).

Gene ontology term enrichment analysis and clustering

Gene Ontology (GO) terms (Ashburner et al. 2000; Aleksander et al. 2023) were assigned using Blast2GO (Götz et al. 2008). Overrepresentation analysis (ORA) was done using Fisher's hypergeometric test in the "Fast GSEA" package (Korotkevich et al. 2019) against reference genes with at least one normalized count across their respective tissue-specific libraries. GO term semantic similarity indices and clustering were performed using "SimplifyEnrichment" (Gu and Hübschmann 2022), the Rel method (Schlicker et al. 2006), and the binary-cut algorithm (Mezentsev 2017). Graphs were produced with R packages (Wickham 2009; Gu et al. 2016).

Photosynthesis and gas exchange measurements

Photosynthesis (A) and gas exchange (Gs) were measured using a LI-6400 XT photosynthesis system with a Leaf Fluorescence Chamber (LCF) (Li-Cor ®, Lincoln, NE) in survey mode. The system had a constant flow of 500 µmol s^{-1}, complete desiccant bypass, reference CO_2 of 400 µmol mol^{-1}, and light set to match the outside of the LFC at 116.2 µmol•m^{-2}•s^{-1}. Measurements were taken at 8:00 a.m. and 1:00 p.m. for 2 days following the onset of treatment.

Sample size was estimated using Cohen's η (F-value, (Cohen 1988), implemented with the pwr package in R, and based on standard deviations reported for stomatal conductance (Gs) and photosynthesis (A) data in Cd-treated cacao plants (Barroso et al. 2023). Data collected each day were processed, parsed, and merged using a custom R script prior to analysis. The objective was to compare Gs and A after 2 days of Cd treatment. Accordingly, F-statistics were calculated using the anova function (stats package) on data from the last day of treatment, applied to linear models constructed with the lm function, using the formulas Gs ~ Cd level + position or A ~ Cd level + position. The "position" term was included as a clocking factor and represents individual tray positions on the greenhouse table. Post hoc comparisons were conducted using Tukey's HSD test, implemented with the TukeyHSD function on aov objects (stats package) generated from the same data and model. A similar analytical approach was applied to the 12 ppm Cd-treated plants, using models defined as Gs ~ time + position or A ~ time + position, which were fitted to assess temporal trends across measurements.

Hormone quantification

Phytohormone extraction was performed using a protocol adapted from previously published methods (Trapp et al. 2014; Floková et al. 2020). Plants were uprooted, and their roots washed in water, after which fully developed E-stage leaves (Fister et al. 2016) and adventitious roots, branching from the taproot, were excised and immediately placed on dry ice. One hundred mg of frozen tissue per sample was ground in liquid nitrogen and placed in a 2 mL centrifuge tube. One mL of extraction buffer (7:3 parts methanol to water with 5 nM of labeled internal standards) was added, samples were vortexed and left to shake for 30'/4 °C. Tubes were centrifuged at 16,000x*g* for 5'/4 °C, after which the supernatant was freeze-dried by centrifugal evaporation. Samples were than resuspended in fresh extraction buffer without internal standards using sonication in an ice bath for 10 min, centrifuged at 16,000x*g* for 5'/4 °C, and ABA concentration in the supernatant was measured by The Huck Institutes' Metabolomics Core Facility (RRID: SCR_023864) using a Sciex ZenoTOF 7600 LC–MS.

Results

Genotypic variation in Cd uptake in roots and leaves of hydroponically grown cacao seedlings

The Cd concentrations in tissues of 2 month-old plants treated with 10 ppm Cd were significantly higher than those of the controls, both in leaves (one-sided t-test, $t=4.2205$, $df=19$, $p=2.315\times10^{-4}$) and in roots (one-sided t-test, $t=16.459$, $df=20$, $p=2.05\times10^{-13}$) (Table 1) (Fig. 1A, B). On average, Cd concentration in the roots was higher than in leaves, regardless of treatment (Table 1).

In Cd-treated PA121 roots, the mean Cd concentration was ~ 1.6-fold higher than the average recorded for Cd-treated TSH660 roots (Table 1) (Fig. 1A). The average concentration in Cd-treated PA121 leaves was ~ 6.4-fold higher than the average recorded for Cd-treated TSH660 leaves (Table 1) (Fig. 1B). Cd concentrations in tissues from the two genotypes grown at 0 ppm Cd were not significantly different (Table 1). These results demonstrate that under our experimental conditions, progeny from open-pollinated PA121 accumulated more Cd in leaf and root tissues than those from TSH660. We designated the progeny from PA121 as the high accumulator, PA121(HA), and the progeny from TSH660 as the low accumulator, TSH660(LA).

Exposure to 10 ppm Cd significantly stunted plant growth, resulting in similar overall biomass between the two genotypes (Table 1; Figs. 1C and D). Under control conditions (0 ppm Cd), PA121 exhibited significantly greater leaf mass than TSH660 (one-sided $t=3.9361$, $df=2.8835$, $p=0.0157$) and marginally greater root mass (one-sided $t=1.8287$, $df=3.952$, $p=0.07115$).

Differential gene expression after cadmium treatment

The 18 transcriptome libraries averaged over 16 million reads per sample, with an average mapping rate of 85.7%. Differentially expressed genes (DEGs) were obtained from 14 linear contrasts (Table S1B-C). The number of DEGs from the main contrasts, the percent of upregulated DEGs of these contrasts, and selected intersections are represented in Fig. 2. Differential expression results are presented in Table S2. After 48 h of Cd treatment, we identified 4185 shared DEGs in roots (shared response, Fig. 2A). The shared response in leaves had significantly fewer DEGs than in roots; in leaves, 1162 at 24 h AC and 1999 at 48 h AC (Fig. 2A). In the shared response, the number of downregulated DEGs was higher in the roots (2344/4185 DEGs or 56%). Conversely, upregulation was higher in leaves (1127/1999 DEGs or 56.4%), deviating significantly from a 1:1 expectation ($X^2=82.21$, $df=1$, $p\text{-value}<2.2\times10^{-16}$, Fig. 2C).

PA121(HA) had 2297 DEGs in roots 48 h AC, and 1538 and 604 in leaves 48 h and 24 h AC, respectively (Fig. 2A). The number of downregulated DEGs was greater than expected in roots, while in leaves, upregulated DEGs were greater than expected ($X^2=112.45$, $df=1$, $p\text{-value}<2.2\times10^{-16}$). On the other hand, TSH660(LA) had 2396 DEGs in roots 48 h AC, slightly higher than in PA121(HA) (Fig. 2A). In leaves, TSH660(LA) showed 180 DEGs

Table 1 Mean Cd concentrations and dry weight of plants treated with Cd

Tissue (Leaves v Roots)	Genotype (PA121 v TSH660) (N)	Cd treatment (10 ppm vs 0 ppm)	Mean Cd measured (ppm) (95% C.I) by genotype and treatment	Mean Cd measured (ppm) by treatment (95% C.I)	Mean dry weight (g) (95% C.I) by genotype and treatment	Mean dry weight (g) (95% C.I) by genotype
Leaves	PA121 ($N=3$)	0 ppm	0.3 (C.I. 0.27)	0.43 (C.I. 0.3)	25.3 (C.I. 6.7)	18.3 (C.I. 9)
	TSH660 ($N=3$)		0.5 (C.I. 0.98)		11.2 (C.I. 13.9)	
	PA121 ($N=9$)	10 ppm	97.4 (C.I. 38.6)	52.3 (C.I. 25.7)	7.3 (C.I. 2)	6.7 (C.I. 1.2)
	TSH660 ($N=11$)		15.3 (C.I. 13)		6.2 (C.I. 1.9)	
Roots	PA121 ($N=3$)	0 ppm	7.5 (C.I. 22)	4.6 (C.I. 6.7)	7.6 (C.I. 5)	5.97 (C.I. 2.8)
	TSH660 ($N=3$)		1.7 (C.I. 1.9)		4.3 (C.I. 5.7)	
	PA121 ($N=10$)	10 ppm	1416.3 (C.I. 118.8)	1143.1 (C.I. 144.2)	2.4 (C.I. 1)	2.5 (C.I. 0.5)
	TSH660 ($N=11$)		895.1 (C.I. 122.8)		2.6 (C.I. 0.5)	

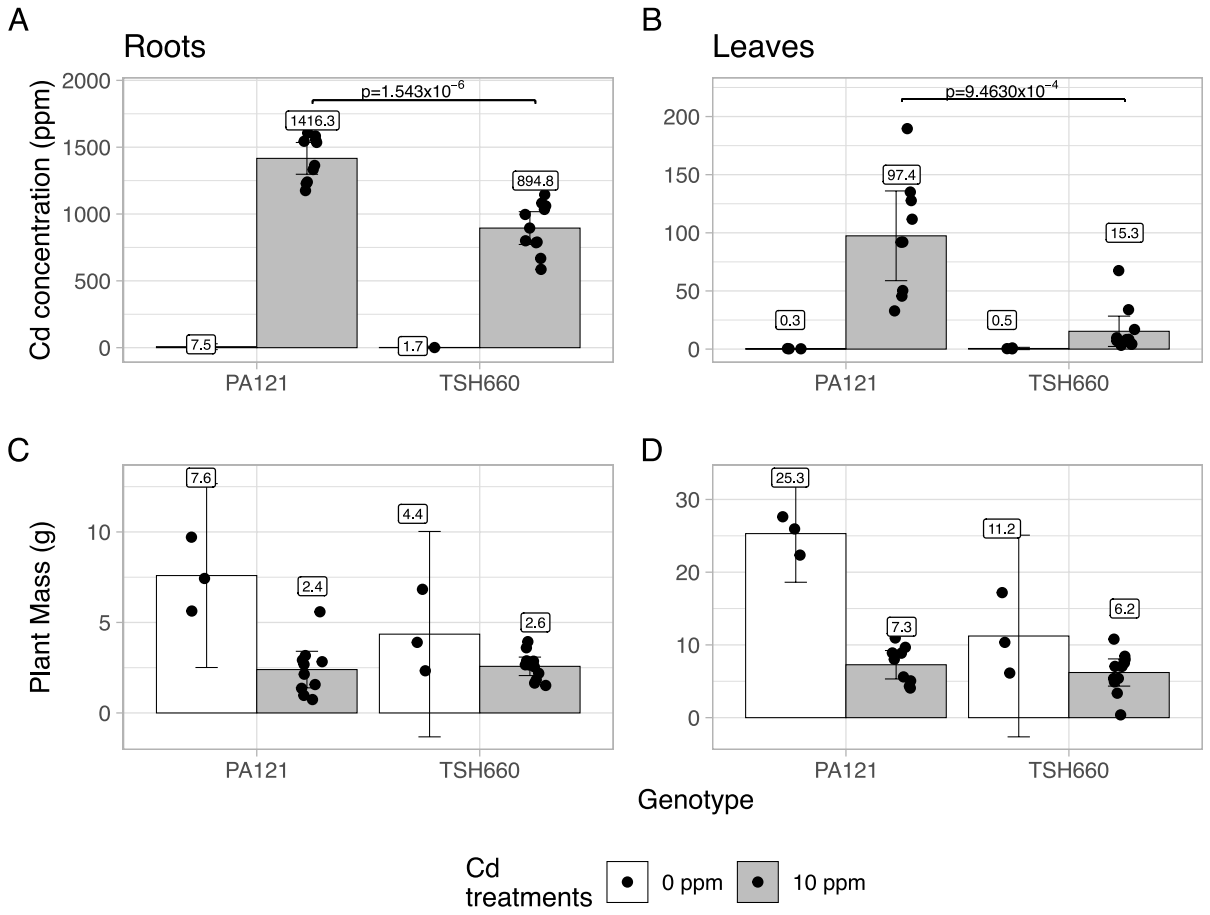


Fig. 1 Cadmium accumulation and growth responses in cacao genotypes after chronic Cd exposure. Cacao seedlings were grown with 0 or 10 ppm Cd for 2 months. **A–B** Cd concentrations in root and leaf tissues of PA121 (high accumulator, HA) and TSH660 (low accumulator, LA). **C–D** Root and leaf dry biomass under control (0 ppm) and Cd treatment (10 ppm). Error bars indicate 95% CI, and labels indicating means are placed above each column. Only for plants treated with 10 ppm Cd, a two-sided t-test was performed to compare the mean Cd concentration across genotypes; the p-values are included above each comparison. For the rest, we relied on the 95% CI. See Table 1 for the sample size in each group, as well as the exact value of the 95% CI

at 48 h and 57 DEGs at 24 h, fewer than PA121(HA), which showed 547 and 1358 DEGs at 24 h and 48 h, respectively (Fig. 2A). Nevertheless, the number of downregulated DEGs in TSH660(LA) roots still deviated significantly from a 1:1 expectation ($X^2 = 11.849$, df = 1, p-value = 0.0005771) (Fig. 2C).

GO term enrichment analysis

To gain insight into the mechanisms activated during Cd-response, enriched Gene Ontology (GO) terms from ten predefined sets of DEGs were identified (Table 2, Table S3). In roots, the enriched GO terms in these sets of DEGs were combined, resulting in a total of 199 unique Biological Process (BP), 90 Molecular Function (MF), and 41 Cellular Component (CC) GO terms. In leaves, there were 63 BP, 18 MF, and 7 CC unique GO terms. After calculating semantic similarities and removing GO terms with zero similarity, 37 BP, 27 MF, and 7 CC remained in roots, and 14 BP in leaves. BP GO was prioritized because it exhibited the most enriched GO terms and provided the most informative GO categories. BP GO produced eight clusters, with at least two GO terms per cluster, a relaxed criterion that ensured maximal retention of GO terms (Fig. 3, Table 3). The

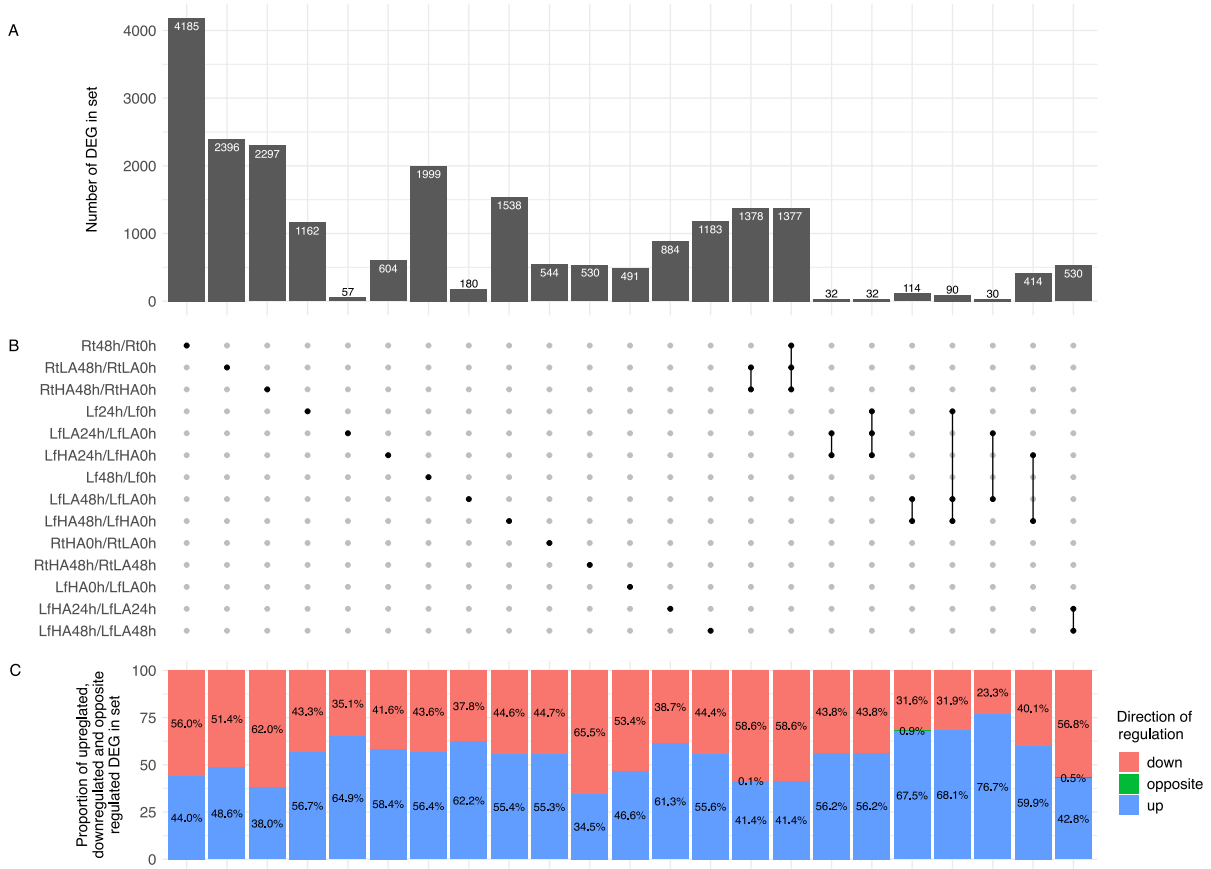


Fig. 2 Differentially expressed genes (DEGs) in cacao seedlings in response to Cd treatment. RNA-seq analysis of root (Rt) and leaf (Lf) tissues from PA121 (HA) and TSH660 (LA) at 0 h, 24 h, and 48 h. **A** Number of DEGs identified in each pairwise comparison. **B** The intersection matrix shows shared and unique DEGs across comparisons. **C** Proportion of DEGs classified as upregulated, downregulated, or oppositely regulated in each comparison

Table 2 GO term enrichment results. The column (DEG Set) states the comparison, the second column lists the number of DEGs, followed by the number of enriched BP GO in roots, and the third column lists the enriched BP GO in leaves. The 10 sets included are those that had enriched BP GO (non-zero)

N°	DEG Set	DEG / Enriched BP GO Roots	DEG / Enriched BP GO Leaves
1	Tc48h/0 h (shared response)	4185 / 120	1999 / 30
2	PA121(HA) 48 h/0 h	2297 / 86	1538 / 16
3	TSH660(LA) 48 h/0 h	2396 / 64	180 / 11
4	Intersection of 2 & 3	1378 / 51	114 / 9
5	Complements TSH660(LA)	1018 / 4	66 / 0
6	Complements PA121(HA)	919 / 0	1424 / 9

clustering analysis for MF is presented in Figure S2. A second hypergeometric test compared GO term names in each cluster to the names within their ontology, identifying enriched keywords used to label each cluster (Table 3, Fig. 3A, B).

Cd activation of indole acetic acid, abscisic acid, melatonin, strigolactone, and ethylene biosynthetic and signaling pathways in roots

DEGs contributing to each enriched BP GO term in the "blue light-activated abscisic acid, defense, and ion response pathway" cluster were classified into orthologous protein families and mapped to KEGG pathways. Gene families identified using KOFAM were consolidated with those identified using PlantTribes2. Genes were mapped to the biosynthetic steps

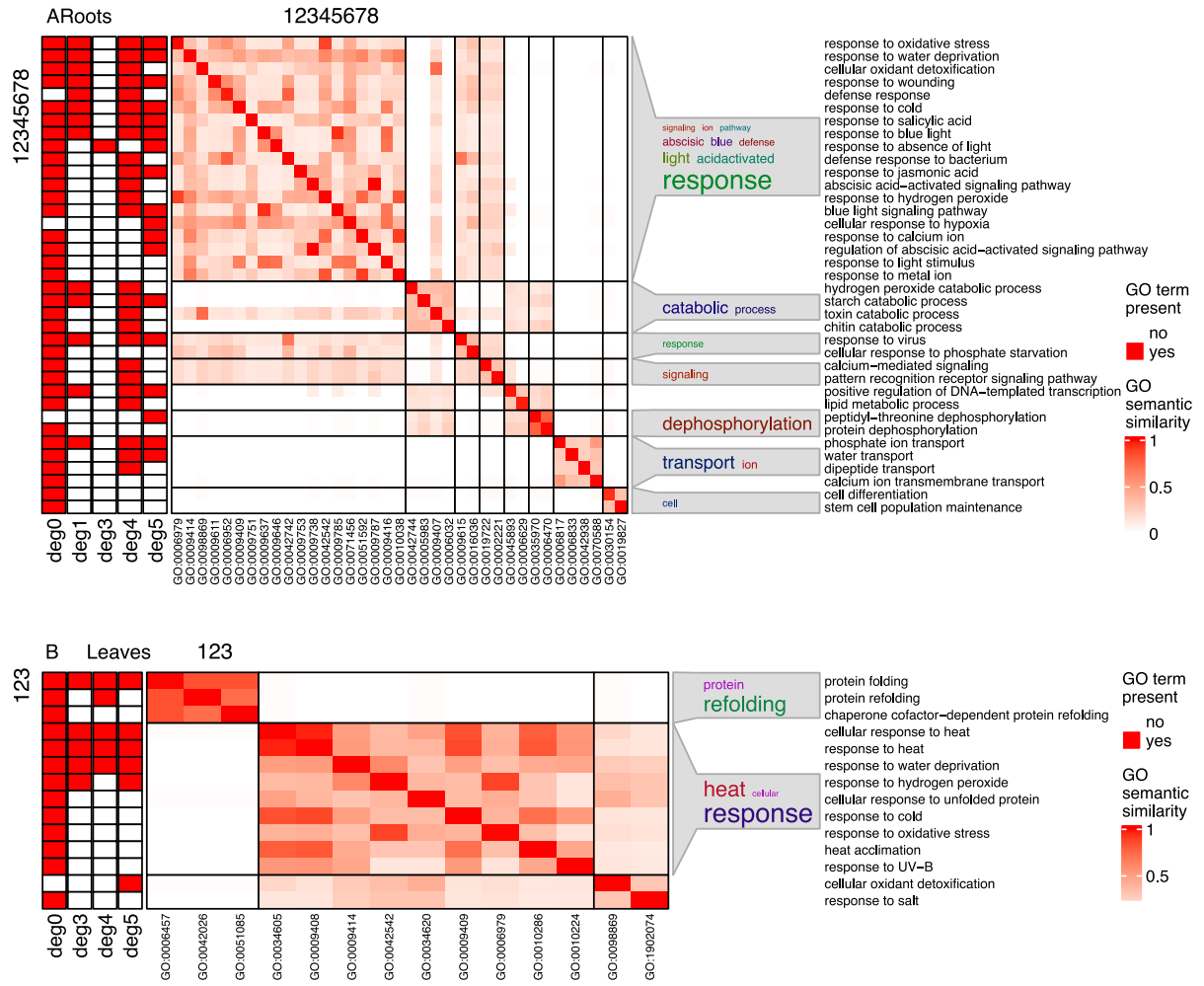


Fig. 3 Enriched gene ontology (GO) terms in cacao roots and leaves after Cd treatment. Clustered semantic similarity matrices of enriched GO terms in (**A**) roots and (**B**) leaves. Each cell represents the similarity between pairs of GO terms (0 = unrelated, 1 = identical). GO term names and IDs are displayed within the matrices

and signal transduction components of various phytohormones, including indole-3-acetic acid (IAA or Auxin), abscisic acid (ABA), strigolactones, melatonin (Mt), ethylene, and strictosidine (Fig. 4). Mapping results are presented in Table S4.

Indoles

In the indole biosynthetic pathway (Fig. 4A and D), two orthologs of a terminal amino acid transferase (*TAA*) gene family, which deaminases Trp to indole-3-pyruvic acid (IPA), XM_007019556.2 and XM_007019553.2, were downregulated. The first one was significantly downregulated 21-fold (−4.39 $\log_2$-fold) when combining both genotypes, and the second one was downregulated 1.7-fold (−0.8 $\log_2$-fold) only in PA121(HA) (Fig. 4A). Additionally, seven downstream signaling components of the IAA, members of the *AUX/IAA* gene family, were upregulated, except for XM_007034894.2, an ortholog of *AtIAA14,* which was downregulated ~ fourfold (Table S2 and Fig. 4D). A *SNAT* gene that acetylates serotonin (XM_007045624.2) was upregulated 3.11-fold (1.64 $\log_2$-fold) at 48 h AC in TSH660(LA) (Fig. 4A). One candidate of a caffeoyl-methyltransferase (*COMT*, XM_018121419.1) was downregulated 2.5-fold (−1.33 $\log_2$-fold) in both genotypes, while another (XM_018120814.1) was upregulated 43.1-fold (5.4

Table 3 GO term clusters

Clusters in Roots	Cluster Name	Number (GO terms)	Genotype
1	Blue light abscisic activated defense, and ion response signaling pathway	19	Both
2	Catabolic process	4	Both
3	Response	2	Both
4	Unnamed	2	Both
5	Signaling	2	HA only (PA121)
6	Ion transport	4	Both
7	Dephosphorylation	2	Both
8	Cell	2	
Clusters in Leaves	Cluster Name		
1	Protein refolding	3	Both
2	Cellular heat response	9	Both
3	Cell detoxification	2	LA only (TSH660)

log_2-fold) in TSH660(LA). A marked difference was observed between TSH660(LA) and PA121(HA) strictosidine synthase (*SS*) genes, *SS1* and *SS2*. Changes in expression occurred only in TSH660(LA). The *SS1* gene XM_018120572.1 was downregulated 2.54-fold (−1.3 log_2-fold) and three genes, XM_018126632.1, XM_018127382.1, and XM_007046480.2, were upregulated 3.11-fold (1.6 log_2-fold), fourfold (2 log_2-fold), and 4.6-fold (2.2 log_2-fold) respectively.

Apocarotenoids

Several key genes encoding enzymes central to the ABA biosynthetic pathway were strongly upregulated in response to Cd treatment (Fig. 4B and F). The shared response identified a 5.9-fold (2.56 log_2-fold) upregulation of an ortholog to β-carotene hydrolase 1 (*BCH1*), XM_007038743.2, and a 1.76-fold (0.81 log_2-fold) upregulation of an ortholog to zeaxanthin epoxidase (*ZEP*), XM_007028495.2. Two orthologs of 9-cis epoxy dioxygenase (*NCED*), XM_007022118.2 and XM_018118862.1, were upregulated 7.5-fold (2.9 log_2-fold) and 8.6-fold (3.1 log_2-fold), respectively. Three xanthoxin dehydrogenases (*XDH*) were upregulated, and three were downregulated. Notably, XM_007027846.2 was upregulated 13.5-fold (3.21 log_2-fold) in TSH660(LA). In the indole aldehyde oxygenase (*IAO*) family, XM_007015509.2 was upregulated 2.3-fold (1.22 log2-fold), while XM_018127404.1 was downregulated 6.5-fold (−2.7 log_2-fold). One ABA-hydroxylase (*ABAH*) candidate gene, XM_007016306.2, was upregulated 20.2-fold (4.34 log2-fold) in both genotypes, and another, XM_007042365.2, 8.6-fold (3.11 log_2-fold) only in PA121(HA). On the other hand, one *ABAH*, XM_018116623.1, was downregulated 4.8-fold (−2.26 log_2-fold) in PA121(HA) (Fig. 4B and F). Furthermore, signaling components downstream of ABA, the ABA receptor gene family *PYL*, XM_007008792.2, XM_007024407.2, XM_007026527.2, and XM_018127083.1, were downregulated 4.92-fold (−2.3 log_2-fold), 31.2-fold (~−5 log_2-fold), 4.1-fold (−2.0 log_2-fold), and 3.4-fold (−1.8 log_2-fold), while one, XM_007051517.2, was upregulated 2.9-fold (1.5 log_2-fold) (Fig. 4F).

Finally, a key step in strigolactone synthesis, the *CCD8* gene, also known as *MAX4* (XM_007052456.2), was upregulated 2.7-fold (1.43 log_2-fold) (Fig. 4B). Furthermore, two downstream receptors of strigolactones (Chen et al. 2022), one annotated as a D14, ortholog of *AtKAI2*, XM_018128515.1, was downregulated 1.9-fold (~0.9 log_2-fold), and another *MAX2* ortholog, XM_018123475.1, was upregulated 1.9-fold (0.9 log_2-fold) (Table S2).

Ethylene

The aminocyclopropane oxidase (*ACO*) XM_007019139.2 was downregulated 6.45-fold (−2.69 log_2-fold) in the shared response. The *ACO* XM_007026145.2 was upregulated

2.12-fold (1.08 log_2-fold) in PA121(HA) and the *ACO* XM_007012805.2 was upregulated 2.43-fold (1.28 log_2-fold) in TSH660(LA) (Fig. 4C and E).

Catabolism cluster, including regulation of GST, lignin-forming peroxidases, ROS modulation, and sugar-mediated stress response in roots

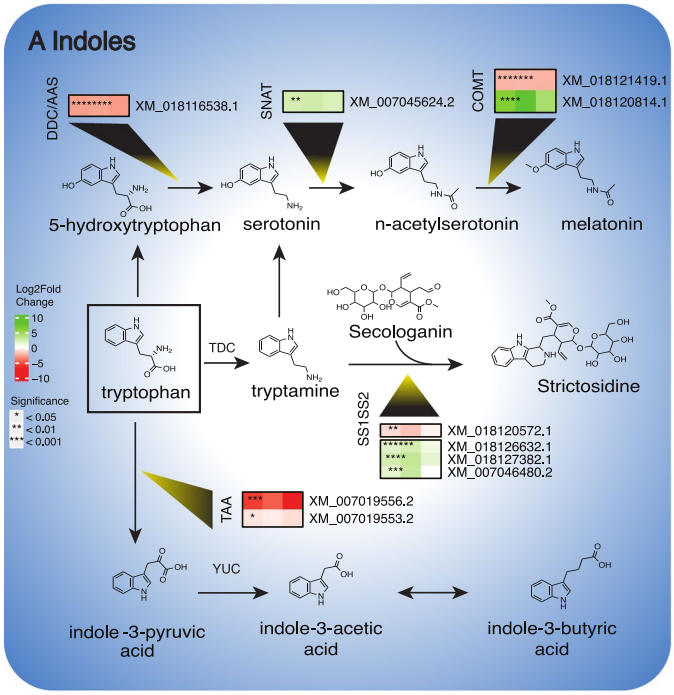
The DEGs in the second-largest cluster, "catabolic process" included enzymes with catabolic functions, such as peroxidases (PER), glutathione S-transferases (GST), and enzymes involved in carbohydrate degradation, as well as chitinases (Fig. 5A). The PER included 31 genes annotated as lignin-forming peroxidases (*LFP*) (Fig. 5A). Of those, 25 were downregulated. In addition, 20 genes were genotype-specific to TSH660(LA) and 26 to PA121(HA), with an overlap of 17 genes between both genotypes. One *LFP*, XM_007045286.2, was strongly upregulated 175.7-fold (7.5 log_2-fold) in TSH660(LA) (Fig. 5A). In contrast to most other *LFPs*, XM_007011153.2 and XM_007011167.2 exhibited similar responses in roots and leaves (Fig. 5A). The next largest family in this cluster was the GST, which contained 21 genes (Fig. 5A). In the shared response, 15 were downregulated (Fig. 5A). In the genotype-specific response, only ten genes were significant for TSH660(LA) and 13 for PA121(HA), with an overlap of seven genes between genotypes (Fig. 5A). There were ten carbohydrate catabolism genes, ten genotype-specific to PA121(HA), and eight genotype-specific to TSH660(LA) (Fig. 5A). The two DEGs that were exclusive to PA121(HA) were an alpha-glucan water dikinase (*GWD*) gene, XM_007027585.2, and a phosphoglucan phosphatase *LSEX4*, XM_007011190.2 (Fig. 5A). Interestingly, we identified an alpha-amylase that was upregulated only in leaves of PA121(HA) (Fig. 5A).

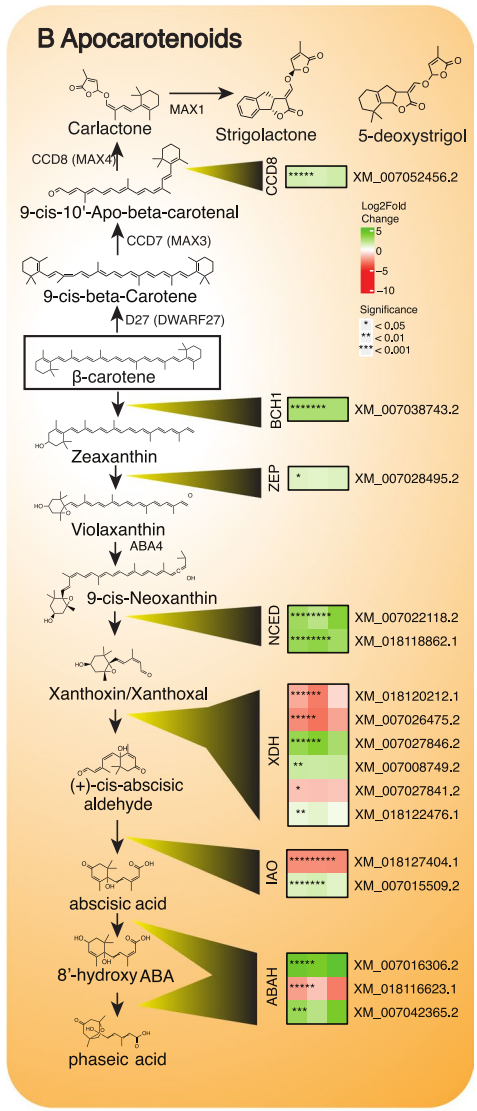
Finally, 17 chitinases in the catabolism cluster are represented in the shared response: 12 are specific to PA121(HA), and nine are specific to TSH660(LA), with an overlap of seven DEGs (Fig. 5A). One marked difference between genotypes was the 6.7-fold (−2.8 log_2-fold) downregulation of XM_007033739.2, an *EP3* chitinase in PA121(HA) roots (Fig. 5A). In leaves of PA121(HA) the endochitinases ortholog (*CHT1*), XM_007041606.2 was 15.8-fold (3.98 log_2-fold) upregulated, the endochitinases *EP3*, XM_018119383.1 was 76.7-fold (6.3 log_2-fold) upregulated, and the chitinase-2 (*CHT2*), XM_018127043.1, was 15.4-fold (3.9-log_2fold) upregulated (Fig. 5A).

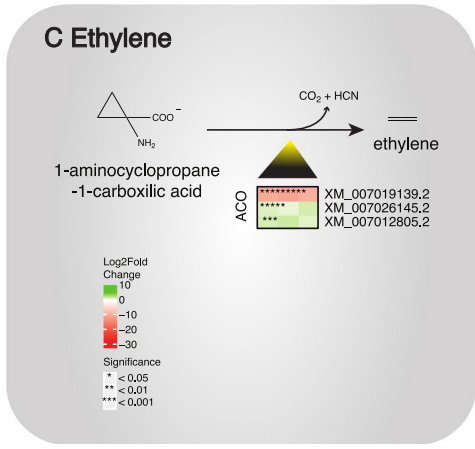
Ion transport clusters indicating regulation of nitrogen, sulfur, and phosphate transporters

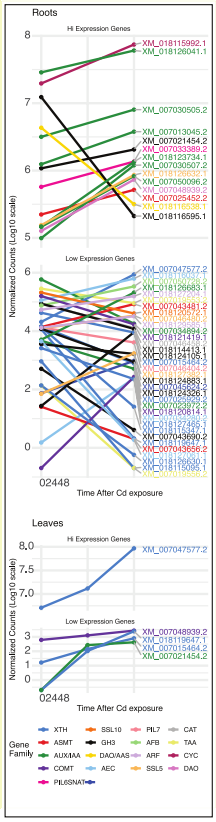
The ion transport cluster contains six sulfur (S), 21 nitrogen (N), and seven phosphate (P) transporters. This cluster also includes 16 calcium (Ca) transporters, 10 aquaporins, four annexins, and five receptor class proteins (Fig. 5B). Nitrogen transporters fell into two orthogroups of *NRT1/PTR* gene families, one containing 19 members (*NTR/PTR1*), of which eight were downregulated (Fig. 5B). *NTR1/PTR1* had 15 genotype-specific DEGs to PA121(HA), 13 to TSH660(LA), with an overlap of ten (Fig. 5A). The six S transporters belong to a single family (*SULTR*), with four of them being downregulated. Additionally, four were genotype-specific to each, with an overlap of two (Fig. 5B). P transporters had 2 DEGs in the *PHO1* gene family and five in the *IPT* family (Fig. 5B). In the *IPTs*, TSH660(LA) had two downregulated genotype-specific genes. In PA121(HA), while five *IPT* genes were significantly downregulated, one *IPT*, XM_018122229.1, was upregulated 3.9-fold (1.97 log_2-fold) (Fig. 5B). Membrane intrinsic proteins (aquaporins) have five tonoplast intrinsic proteins (*TIP*), four plasma membrane intrinsic proteins (*PIP*), and one NOD25 intrinsic protein (*NIP*). One *PIP* gene, XM_007016678.2 (*TcPIP2-2*), exhibited a relatively high twofold (1 log_2-fold) upregulation, exclusive to TSH660(LA). More Ca transporters were regulated in PA121(HA), which had seven genotype-specific DEGs compared to four in TSH660(LA). Despite this, one member of the *CAX2* gene family, XM_007016045.2, was 15-fold (3.99-log2-fold) upregulated specifically in TSH660(LA). This contrasted with the 1.3-fold upregulation in PA121(HA), which did not pass the significance cut-off (Fig. 5B). Interestingly, in leaves, this gene was genotype-specific to PA121(HA) and was downregulated 5.4-fold (−2.42 log_2-fold) (Fig. 5B).

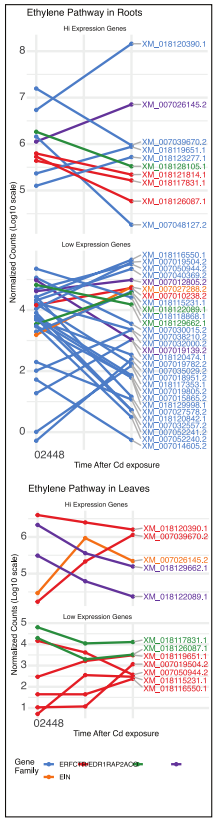
The potential overlap between the phosphate starvation response (PSR) in *A. thaliana* (Wang et al. 2023b) and the ion transport cluster in the Cd

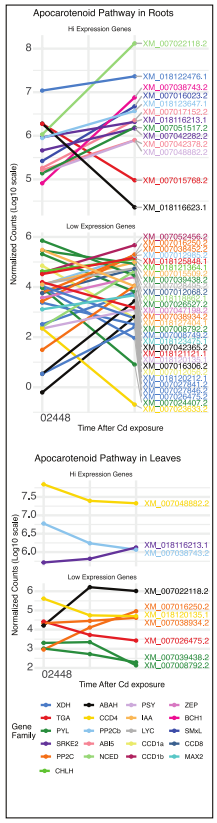

A Indoles
DDC/AAS
XM_018116538.1
SNAT
XM_007045624.2
COMT
XM_018121419.1
XM_018120814.1
5-hydroxytryptophan
serotonin
n-acetylserotonin
melatonin
Log2Fold Change
Significance
tryptophan
TDC
tryptamine
Secologanin
Strictosidine
SS1SS2
XM_018120572.1
XM_018126632.1
XM_018127382.1
XM_007046480.2
TAA
XM_007019556.2
XM_007019553.2
YUC
indole-3-pyruvic acid
indole-3-acetic acid
indole-3-butyric acid


B Apocarotenoids
Carlactone
MAX1
Strigolactone
5-deoxystrigol
CCD8 (MAX4)
CCD8
XM_007052456.2
9-cis-10'-Apo-beta-carotenal
CCD7 (MAX3)
9-cis-beta-Carotene
D27 (DWARF27)
β-carotene
BCH1
XM_007038743.2
Zeaxanthin
ZEP
XM_007028495.2
Violaxanthin
ABA4
9-cis-Neoxanthin
NCED
XM_007022118.2
XM_018118862.1
Xanthoxin/Xanthoxal
XDH
XM_018120212.1
XM_007026475.2
XM_007027846.2
XM_007008749.2
XM_007027841.2
XM_018122476.1
(+)-cis-abscisic aldehyde
IAO
XM_018127404.1
XM_007015509.2
abscisic acid
8'-hydroxy ABA
ABAH
XM_007016306.2
XM_018116623.1
XM_007042365.2
phaseic acid


C Ethylene
COO⁻
NH₂
CO₂ + HCN
ethylene
1-aminocyclopropane -1-carboxilic acid
ACO
XM_007019139.2
XM_007026145.2
XM_007012805.2
Log2Fold Change
Significance


D Indoles
Roots
Leaves
Normalized Counts (Log10 scale)
Time After Cd exposure
Gene Family


E Ethylene
Ethylene Pathway in Roots
Ethylene Pathway in Leaves
Normalized Counts (Log10 scale)
Time After Cd exposure
Gene Family


F Apocarotenoids
Apocarotenoid Pathway in Roots
Apocarotenoid Pathway in Leaves
Normalized Counts (Log10 scale)
Time After Cd exposure
Gene Family

◂**Fig. 4** Expression of hormone biosynthetic and signaling genes from the *blue light abscisic acid-activated defense and ion response signaling* cluster. **A** Indole pathway genes. **B** Apocarotenoid pathway genes. **C** Ethylene biosynthesis genes. **D**–**F** Signaling components associated with indole, apocarotenoid, and ethylene pathways. Heatmaps display $\log_2$-fold changes (48 h/0 h) in root and leaf tissues for the combined genotypes (Both), TSH660(LA), and PA121(HA) from left to right. Time-course expression plots (0, 24, 48 h) are shown for selected gene families in roots and leaves. Asterisks denote significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Gene IDs are displayed in the figure; full gene names are provided in Supplementary Table S2

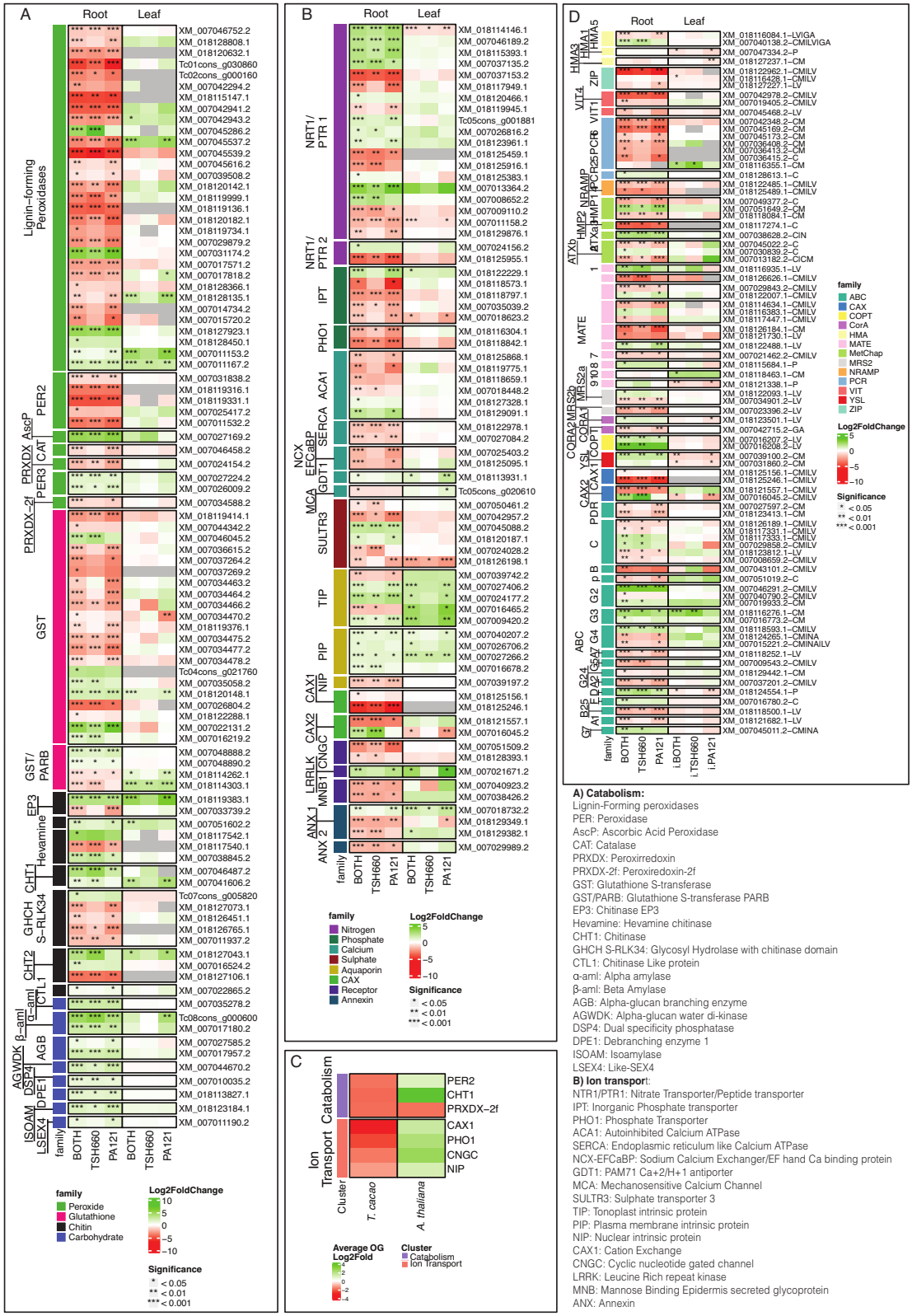
response in cacao was determined using the PSR genes from Table S1 in Z. Wang et al. (2023a, b). In each gene family identified using PlantTribes2, genes from either *A. thaliana* or cacao had to all be regulated in the same manner (all upregulated or downregulated, Table S5). The $\log_2$-fold averaged by gene families is plotted in Fig. 5C. In the ion transport cluster, all *CAX1*, *PHO1*, *CNGC*, and *NIP* were downregulated in cacao and upregulated in *A. thaliana* (Fig. 5C). In the catabolism cluster, the peroxidases-2 (*PER2*) and chitinases-1 (*CHT1*) were all downregulated in the cacao, while upregulated in the *A. thaliana*, except for the peroxiredoxin-2f gene family, which was downregulated in both (Fig. 5C).

Metal transporter gene families regulated by Cd in cacao

By searching the cacao genome for 16 gene superfamilies involved in metal or coordination compound transmembrane transport, 82 DEGs encoding metal transporters or metal detoxification genes were identified, 75 in roots and 12 in leaves. Their subcellular localization was then predicted (Fig. 5D, Table S6). Seven were leaf-specific and 70 root-specific (Fig. 5D). Our results indicated that A Cd/Zn transporting PIB-2 *HMA* gene (XM_018127237.1) was downregulated 1.7-fold (−0.77 $\log_2$-fold) in the leaves of the PA121(HA) 48 h AC (Fig. 5D). This protein is predicted to be localized to the cytosol membrane (CM). *TcHMA1* (XM_007047334.2) was also downregulated 2.1-fold (1.1 $\log_2$-fold) in PA121(HA) leaves at 48 h AC. In roots, the orthologs to HMA5 and RAN1 included two distinct members. The first one, *TcHMA5* (XM_018116084.1), predicted to localize to the CM, was downregulated 2.3-fold (−1.22 $\log_2$-fold) in PA121(HA) at 48 h AC. Conversely, *TcRAN1* (XM_007040138.2), predicted to localize to the chloroplast membrane, was upregulated threefold (1.6 $\log_2$-fold) in TSH660(LA) at 48 h AC (Fig. 5D). *TcNRAMP6* (XM_018122485.1) was downregulated 3.4-fold (−1.77 $\log_2$-fold) in roots of both genotypes 48 h AC. *TcNRAMP5* (XM_018125489) was downregulated 9.64-fold (−3.26 $\log_2$-fold) in the roots and was genotype-specific to TSH660(LA) 48 h after Cd (Fig. 5D). The ZIP gene, *TcIRT1* (XM_018122962.1), was strongly downregulated in response to Cd treatment in both genotypes, 76.79-fold (−6.27 $\log_2$-fold) in PA121(HA) and 39.24-fold (−5.29 $\log_2$-fold) in TSH660(LA). Additionally, PA121(HA) had one *ZIP* family member, XM_018127227.1, downregulated 2.93-fold (−1.55 $\log_2$-fold) in roots 48 h AC. Plant Cd resistance (*PCRs*) were mostly downregulated in roots, with one member (XM_018116355.1) being upregulated in the leaves of PA121(HA) (Fig. 5D). The vacuolar iron transporter (*VIT*), XM_007042978.2, was 20-fold (−4.32 $\log_2$-fold) downregulated, and one Yellow Stripe-Leaf (*YSL*), XM_007039100.2, was upregulated in roots of both genotypes while being downregulated in the leaves of PA121(HA) (Fig. 5D).

GO term clusters and DEGs mapping to TCA, carbohydrate, and fatty acid metabolic pathways

KEGG pathway analysis mapped 561 genes from Cellular Heat Response (CHR) leaf cluster (Fig. 3B) to 236 different pathways and clustered once more using the binary-cut algorithm (Table S7 and Figure S3). After this, thirty-nine genes with shared pathways were selected and used to create ad hoc groups related to carbohydrate metabolism (CHO), fatty acid (FA) metabolism, the tricarboxylic acid cycle (TCA), and terpenoid biosynthesis (Fig. 6). Twenty-eight of the 39 genes were significantly upregulated in PA121(HA) (Fig. 6). One DEG shared by both genotypes, a cupredoxin ascorbate oxidase (XM_007016622.2), was upregulated 10.3-fold (3.4 $\log_2$-fold) at 48 h AC. In the terpenoid group, an *NCED* (XM_007022118.2) was upregulated 5.6-fold (2.5 $\log_2$-fold), and a purple acid phosphatase ortholog (XM_018116461.1) was downregulated 2.6-fold (1.4 $\log_2$-fold); both genes are related to ABA signaling. The shared response in the carbohydrate cluster included a cell wall invertase 1, XM_018122432.1, which was upregulated sevenfold (2.8 log2-fold), and a beta-amylase, XM_018117250.1, downregulated 4.7-fold



A) Catabolism:
Lignin-Forming peroxidases
PER: Peroxidase
AscP: Ascorbic Acid Peroxidase
CAT: Catalase
PRXDX: Peroxiredoxin
PRXDX-2f: Peroxiredoxin-2f
GST: Glutathione S-transferase
GST/PARB: Glutathione S-transferase PARB
EP3: Chitinase EP3
Hevamine: Hevamine chitinase
CHT1: Chitinase
GHCH S-RLK34: Glycosyl Hydrolase with chitinase domain
CTL1: Chitinase Like protein
α-aml: Alpha amylase
β-aml: Beta Amylase
AGB: Alpha-glucan branching enzyme
AGWDK: Alpha-glucan water di-kinase
DSP4: Dual specificity phosphatase
DPE1: Debranching enzyme 1
ISOAM: Isoamylase
LSEX4: Like-SEX4

B) Ion transport:
NTR1/PTR1: Nitrate Transporter/Peptide transporter
IPT: Inorganic Phosphate transporter
PHO1: Phosphate Transporter
ACA1: Autoinhibited Calcium ATPase
SERCA: Endoplasmic reticulum like Calcium ATPase
NCX-EFCaBP: Sodium Calcium Exchanger/EF hand Ca binding protein
GDT1: PAM71 Ca+2/H+1 antiporter
MCA: Mechanosensitive Calcium Channel
SULTR3: Sulphate transporter 3
TIP: Tonoplast intrinsic protein
PIP: Plasma membrane intrinsic protein
NIP: Nuclear intrinsic protein
CAX1: Cation Exchange
CNGC: Cyclic nucleotide gated channel
LRRK: Leucine Rich repeat kinase
MNB: Mannose Binding Epidermis secreted glycoprotein
ANX: Annexin

◂**Fig. 5** Expression changes in root-enriched functional categories under Cd treatment. **A** Catabolism and detoxification gene families. **B** Ion and macronutrient transport gene families. **C** Average $\log_2$ fold change of shared gene families (orthogroups) between the *Arabidopsis thaliana* phosphate starvation responses and the cacao Cd response. **D** Metal-transporting gene families. Heatmaps display $\log_2$ fold changes (48 h/0 h) in root and leaf tissues for combined genotypes (Both), TSH660 (LA), and PA121 (HA). Gene family names and abbreviations are defined; full gene IDs are provided in Supplementary Table S2. Asterisks denote significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(2.2 $\log_2$-fold). All DEGs were specific to PA121(HA) in the fatty acid cluster. Eceriferum had two members: one downregulated (XM_018115329.1) threefold (−1.6 $\log_2$-fold) and another (XM_007008667.2) upregulated 10.7-fold (3.4 $\log_2$-fold). Cell wall invertases (*CWIN*), raffinose synthase (*RS*), and galactinol transferase (*GalT*) were upregulated in PA121(HA) leaves. One *CWIN*, XM_018117430.1, was downregulated in TSH660(LA) roots.

Genes with convergent and divergent gene expression patterns

In this analysis, convergent genes were defined as those that were differentially expressed at 0 h but not at 48 h. Conversely, divergent genes were not differentially expressed at 0 h but became so at 48 h AC. In roots, 322 convergent genes and 308 divergent genes were identified. In the divergent group, 45 genes became divergent due to changes in expression in PA121(HA) and 54 genes due to changes in expression in TSH660(LA). The convergent and divergent genes in leaves were similarly identified, as 182 and 867 DEGs, respectively. In leaves, 338 genes became convergent due to changes in expression specific to PA121(HA) and six due to changes in expression specific to TSH660(LA). The convergent and divergent genes were cross-referenced to genes contributing to each semantic similarity cluster (Table 3). The list of genes is provided in Table S8. The expression changes for the top genes in both groups were plotted by tissue in Figure S4.

Cd treatment in PA121 reduced gas exchange in cacao leaves and increased ABA concentration in roots

To test whether ABA-related transcriptional changes under Cd exposure translated into physiological responses, we measured stomatal conductance (Gs), photosynthesis (A), and ABA concentrations in cacao seedlings under 0, 6, and 12 ppm Cd^{2+}. Cd treatment had a significant effect on stomatal conductance (one-way ANOVA, $p = 0.0192$), accounting for 20.8% of its variation (Fig. 7A); however, photosynthesis was unaffected ($p = 0.6151$) (Fig. 7B). Tukey's post hoc test indicated that plants treated with 12 ppm Cd had significantly lower Gs than the control group at the end of the experiment (difference = −0.0119, p-adj. = 0.0145). No significant differences were observed between plants treated with 6 ppm Cd and the controls (difference = −0.0051, p-adj. = 0.4242) or between plants treated with 12 ppm and 6 ppm (difference = −0.0067, p-adj. = 0.2363) (Fig. 7A). In plants treated with 12 ppm, a significant difference in Gs was observed between the morning of the first day and the afternoon of the second day (difference = −0.0190, p = 0.0499) (Fig. 7C). Like before, no significant changes in A were detected (Fig. 7D). After 36 h, the mean ABA concentration in roots was significantly higher in the 12 ppm group (1.95 nM) compared to the control group (0.21 nM) (t = 4.0822, df = 2.0528, $p = 0.0264$) (Fig. 7E). No significant difference between the treated and control groups was observed in leaf ABA concentrations (t = 0.3749, df = 3.0921, $p = 0.366$) (Fig. 7F).

Discussion

Theobroma cacao exhibits both physiological and molecular adaptations to cadmium (Cd) exposure. TSH660(LA) genotype inherently absorbs significantly less Cd in its roots, approximately 63% of the amount absorbed by PA121(HA) genotype, and translocates markedly less to its leaves, at only ~ 15.7% of the PA121(HA) levels (Fig. 1A-B). In contrast, PA121(HA) tolerated a higher Cd uptake and translocation, as well as a more robust activation of stress-response genes in leaves. These divergent strategies are reflected at the transcriptomic level: TSH660(LA) shows fewer differentially expressed genes (DEGs) in its leaves, consistent with an avoidance mechanism. PA121(HA), on the other hand, exhibits a substantially higher number of DEGs in its leaves (Fig. 2A),

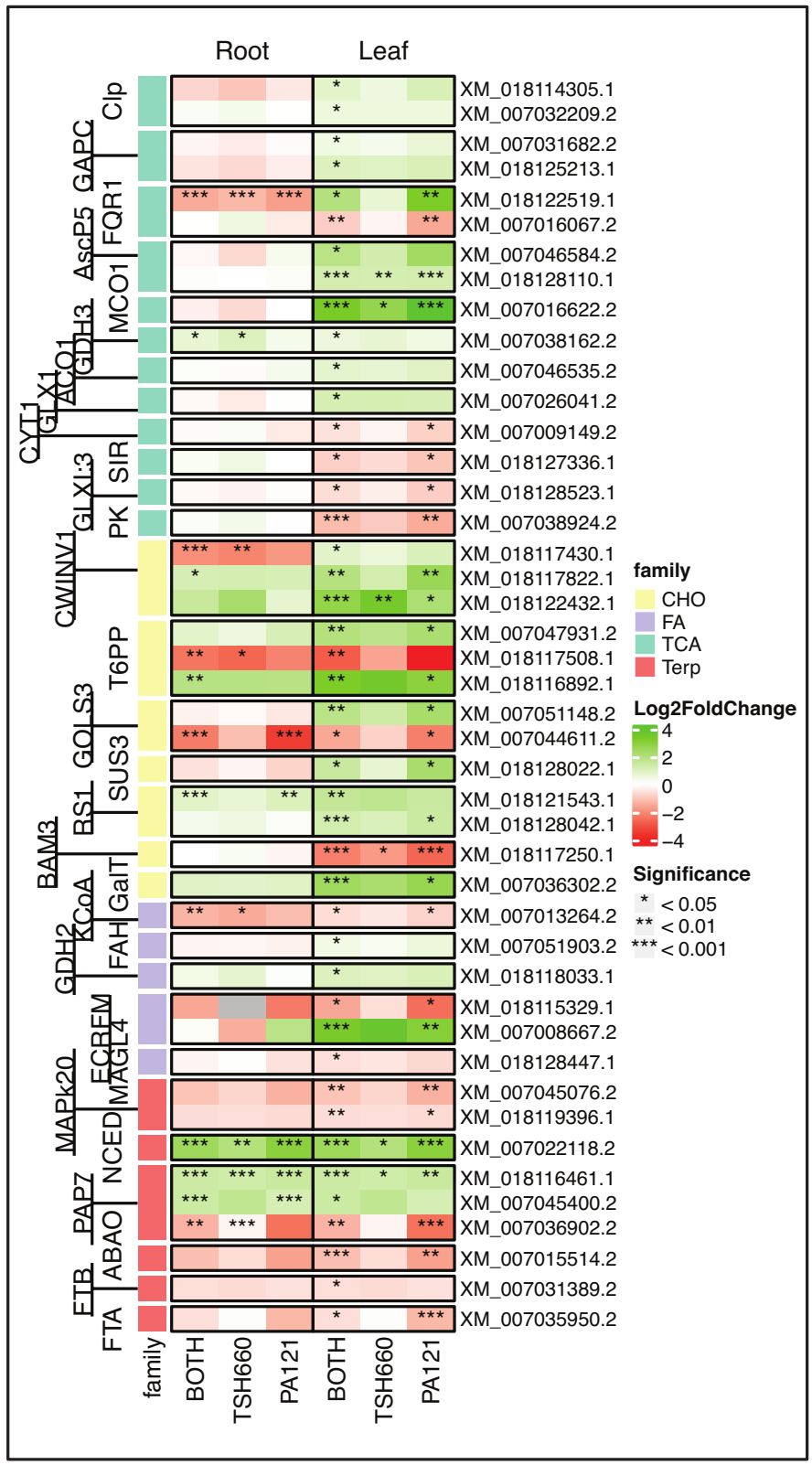


Fig. 6 Expression changes in leaf-enriched metabolism pathways in cacao seedlings exposed to Cd. Heatmaps show $\log_2$ fold changes (48 h/0 h) in root and leaf tissues for genes in the tricarboxylic acid (TCA) cycle, carbohydrate metabolism (CHO), fatty acid metabolism (FA), and terpenoid metabolism (Terp). Data are shown for combined genotypes (Both), TSH660 (LA), and PA121 (HA). Gene family abbreviations and representative gene IDs are displayed in the figure; full gene names are provided in Supplementary Table S2. Asterisks denote significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

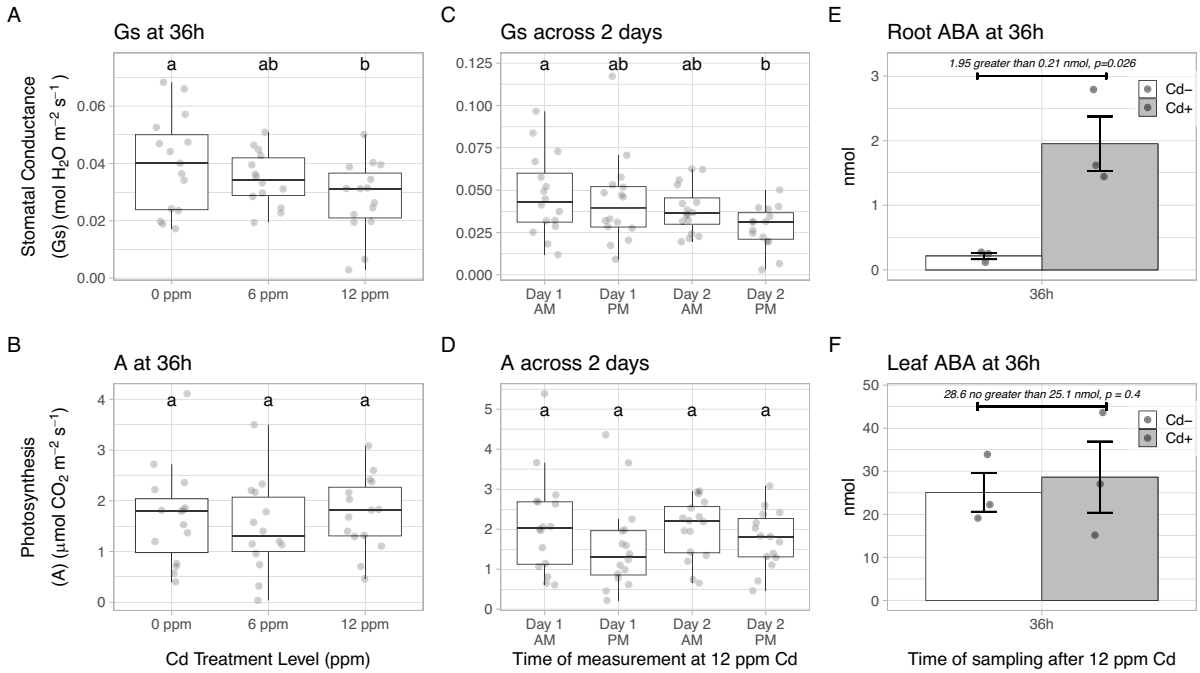


Fig. 7 Physiological and hormonal responses of cacao seedlings to Cd treatment. **A**–**B** Stomatal conductance (Gs) and photosynthetic rate (A) at 36 h under 0, 6, and 12 ppm Cd. **C**–**D** Time-course of Gs and A over two days (AM and PM measurements) at 12 ppm Cd. **E**–**F** Abscisic acid (ABA) concentrations in root and leaf tissues after 36 h. Cd exposure significantly increased ABA in roots ($p = 0.026$) but not in leaves ($p = 0.4$). Data represent means ± SE ($n = 16$ for Gs and A and $n = 3$ for phytohormone quantification)

with enrichment in genes associated with heat stress responses (Fig. 3B), including heat shock proteins typically activated under abiotic and biotic stress conditions.

Comparative transcriptomic responses to cadmium in cacao roots and leaves

At 48 h after Cd exposure, the number of DEGs was nearly twice as high in cacao roots (4185) as in leaves (1999) (Fig. 2B). In leaves, PA121(HA) had 1538 DEGs compared to 180 in TSH660(LA), where in roots, both genotypes showed a similar number of DEGs (Fig. 2B). The higher number of DEGs in cacao roots aligns with findings in *Solanum lycopersicum* (Chen et al. 2023), *Brassica juncea* (Thakur et al. 2019), *Nicotiana rustica* (Zhang et al. 2021), and *A. thaliana* (Herbette et al. 2006), but not in *N. tabacum* (Zhang et al. 2021). Both cacao genotypes exhibited a higher proportion of downregulated genes in roots and more upregulated genes in leaves; a pattern not consistently observed in the species cited above.

Despite differences in gene expression and Cd accumulation in leaves and roots, the mechanisms of Cd response in both genotypes share common features. These include a stress-detection and fast-acting signaling phase, followed by hormone-mediated signaling. These cause transcriptional changes that may result in reducing Cd toxicity, such as divalent cation uptake downregulation (Fig. 5D); N and P uptake regulation (Fig. 5C); regulation of chelating molecules (Fig. 5A); and other adaptations related to water use, which include reduced stomatal conductance (Fig. 7A-C), and transcriptional changes that may be related to lowering the plants water potential and increasing leaf thickness (Fig. 5 and 6). We also anticipate an adverse effect on growth (Fig. 1C-D). Figure 8 integrates the data and the proposed model.

Signaling components activated by cadmium

Although *T. cacao* is increasingly recognized as a model system for tropical perennial trees, the mechanisms underlying Cd sensing remain poorly

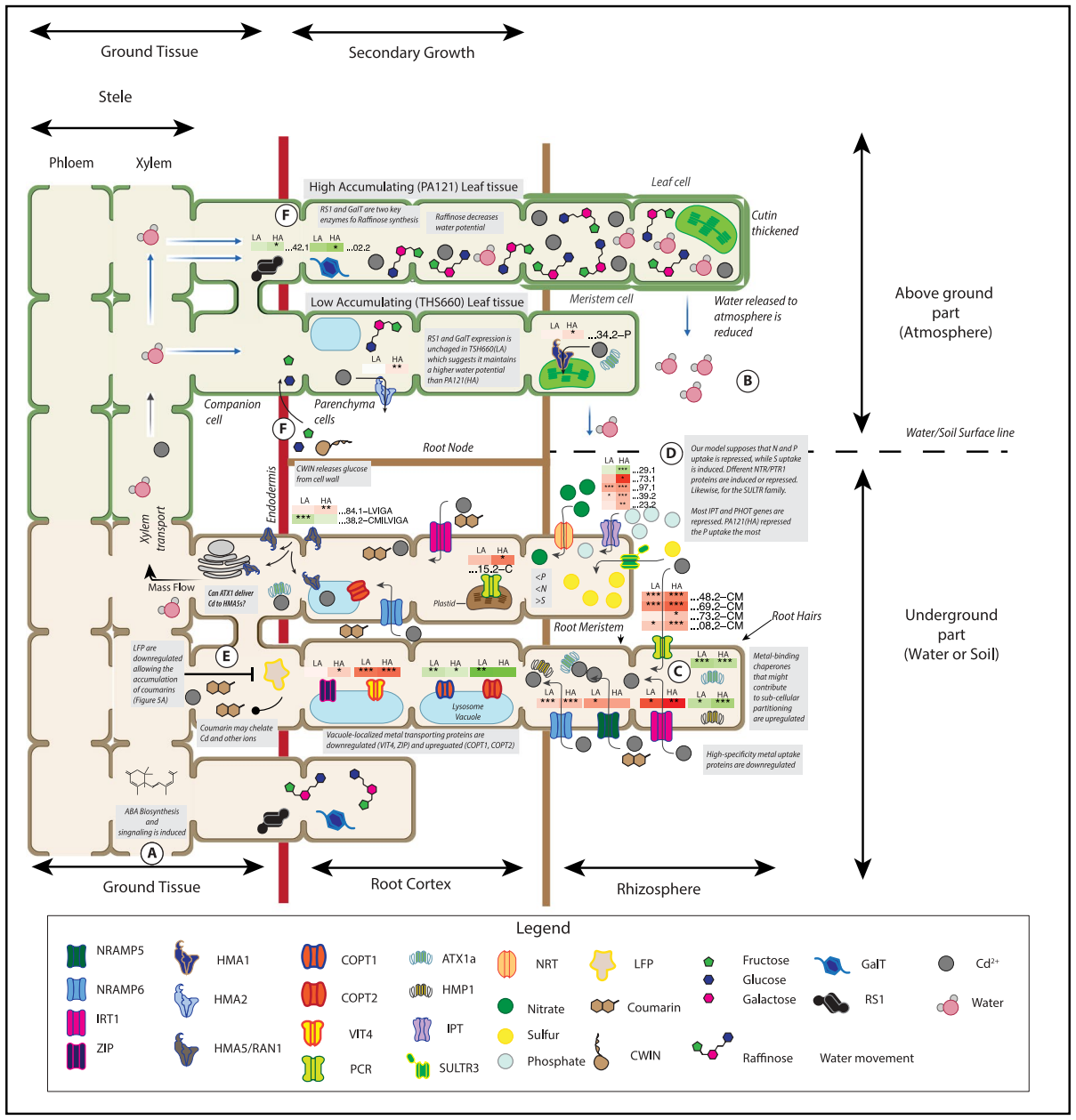


Fig. 8 Working model of Cd responses in cacao seedlings. The schematic illustrates physiological and molecular processes from the rhizosphere to the atmosphere, with tissue zones indicated (rhizosphere, root cortex, ground tissue, stele, and aboveground tissues). The high accumulator, PA121(HA), and the low accumulator, TSH660(LA), are contrasted. Key processes include: **A** stress signaling associated with abscisic acid (ABA) and related molecules, **B** reduced stomatal conductance limiting transpiration-driven mass flow and Cd uptake, **C** altered nutrient and ion transporter activity (e.g., NRAMP5, IRT1, HMA2, NRAMP6, HMA5, ZIP, HMA1), **D** changes in macronutrient uptake (reduced phosphorus and nitrogen, increased sulfur), **E** enhanced chelation and sequestration mediated by compounds such as coumarins, and (**F**) shifts in carbohydrate metabolism and raffinose accumulation affecting water potential. Embedded heatmaps show $\log_2$ fold changes (48 h/0 h) in root and leaf tissues, with significance indicated (* $p<0.05$, ** $p<0.01$, *** $p<0.001$). Full gene names are provided in Figs. 5 and 6. Transmembrane protein subcellular localizations follow the predictions made by DeepLoc2 (Supplementary Table S6)

understood. We hypothesize that ROS and Ca^{2+} signaling occur in cacao plants, and, similar to *A. thaliana*, early responses to Cd involve a rapid calcium spike (Zhang et al. 2023), followed by an increase in reactive oxygen species (ROS), a pattern also observed in other plant species (Chmielowska-Bąk et al. 2014). Following this initial sensing event, hormone signaling can induce changes in gene expression. Consistent with this hypothesis, the largest cluster of GO terms in roots comprised 19 biological process (BP) categories, including 767 differentially expressed genes (DEGs) involved in key biosynthetic and signaling pathways for ABA, IAA, strigolactone, and ethylene (Fig. 4). Likewise, ABA concentration increased in the roots of PA121(HA) (Fig. 7E). Plants of the genotype TSH660(LA) were not available at the time of the hormone quantification. While acknowledging this limitation, our findings establish a foundation for future investigations into hormone signaling under Cd stress in cacao, including studies across different genotypes and experiments employing hormone inhibitors or ABA-related mutants, and add support to the body of evidence of the role of ABA and Ethylene in response to Cd exposure (Liu et al. 2024).

Effects of cadmium on water-use efficiency

Cd is thought to alter water-use efficiency (WUE), prompting investigations into the link between drought tolerance and Cd accumulation in cacao (Ortiz-Álvarez et al. 2023). This relationship is plausible because reduced stomatal conductance lowers water loss to the atmosphere, thereby slowing mass flow and slowing down solute uptake (Marschner and Rengel 2011), including Cd (Page and Feller 2015). In a subsequent physiological experiment conducted independently in PA121(HA), Cd significantly reduced stomatal conductance (Fig. 7A, C). This result agrees with previous publications reporting a negative effect of Cd on transpiration in cacao (dos Santos et al. 2020; Barroso et al. 2023). Both studies also reported a reduction in photosynthesis, suggesting another unintended effect of forced stomatal closure. Under our experimental conditions, we did not detect any trend in photosynthetic rate aside from a higher rate in the morning (Fig. 7B, D), likely due to the much shorter treatment duration.

To further explore the link between Cd exposure and WUE, we analyzed a previously published transcriptome data of cacao drought response (CDR) (Kulesza et al. 2024),, where, similarly to cacao Cd response, we observed a correlation between the log2-fold changes of ABA-related genes, including *LEA* and *PYL*. However, some differences emerged. For instance, Cd affected the expression of nine *IAA/ARP* genes, while only six were differentially expressed in the CDR; of those, only three showed similar regulation patterns. A comparable trend was observed in the ethylene response factor (*ERF*) gene family. Additionally, transcriptional changes associated with reduced water availability were reflected in the largest GO term cluster in leaves, identified as "cellular response to heat" (Table 3). Within this cluster, induction of Eceriferum gene expression in leaves was observed, indicative of potential enhanced cutin biosynthesis and cuticular thickening, processes that likely contribute to reduced water loss due to transpiration. Furthermore, key enzymes for raffinose synthesis, identified as carbohydrate metabolism in leaves (Fig. 6), galactinol synthase (*GOLS1*), raffinose synthase (*RS*), and galactosyltransferase (*GalT*) (Elsayed et al. 2014), were significantly upregulated. Seven additional *GalT* and *RS* genes (Table S2) exhibited similar expression patterns to those observed in the cacao drought response.

Downregulation of IRT and NRAMP transporters in response to cadmium

Stagnation of mass flow alone may not be sufficient to limit Cd uptake, as most micronutrients (e.g., Fe, Mn, Zn, Cu) are predominantly acquired via active transport, enabling efficient assimilation even at low external concentrations. *TcIRT1*, a key gene in Strategy I Fe uptake (Dashner 2024), was strongly downregulated in roots: 48-fold in PA121(HA) and 22.72-fold in TSH660(LA) (Fig. 5D). Similarly, its Arabidopsis ortholog (*AtIRT1*) was induced under Fe deficiency but was downregulated when Fe deficiency was combined with 90 µM Cd (Connolly et al. 2003), and after Cd exposure alone (Herbette et al. 2006). Together, these findings suggest a regulatory mechanism in which the expression of Cd-permeable transporters, such as *AtIRT1* and *TcIRT1*, is attenuated to limit Cd uptake. ABA has also been linked to *IRT1* downregulation under cadmium exposure in *A. thaliana*,

suggesting a hormonal role in restricting Cd uptake (Shen et al. 2022).

TcNRAMP6 and *TcNRAMP5*, which mediate transmembrane transport of Zn and Cd (Ullah et al. 2018), exhibited differential expression. *TcNRAMP6* was upregulated in both genotypes, whereas *TcNRAMP5* was downregulated in TSH660 (LA) (Fig. 5D). These patterns suggest that the suppression of specific NRAMP transporters may help limit Cd uptake. Evidence from Arabidopsis NRAMP knockout lines supports this hypothesis (Cailliatte et al. 2009). Similarly, Cd xylem translocation was reduced in *O. sativa* RNAi lines (Ishimaru et al. 2012). A comparable trend occurs in *Nicotiana tabacum*, which accumulates less Cd than *Nicotiana rustica* and downregulates *NtNRAMP2* and *NtNRAMP6* in response to Cd exposure (Zhang et al. 2021). Although further functional validation in cacao is required, repression of Cd-permeable transporters appears to be a conserved mechanism across species for restricting Cd accumulation.

Cadmium translocation and partitioning by the HMA gene family

Among the heavy metal ATPases in cacao, only *TcHMA3* has been functionally characterized to date (Moore et al. 2020). In our study, *TcHMA3* is strongly downregulated in the roots of PA121(HA) after 48 h of Cd exposure (Fig. 5D). This transcriptional response contrasts with *AtHMA3*, which was upregulated after 30 h of exposure to 50 µM Cd (Herbette et al. 2006), but resembles the behavior of *AtHMA2* and *AtHMA4*, both of which are repressed by the transcription factors *AtMYC2* and *AtMYB43* following Cd exposure (Zheng et al. 2022; Cao et al. 2024). Given that *AtHMA2* and *AtHMA4* function as xylem loaders (Hussain et al. 2004) and *AtHMA3* mediates vacuolar sequestration (Morel et al. 2008), the divergent regulation observed in cacao warrants further investigation into the evolutionary dynamics and functional specialization of the HMA family in this species.

In addition, *TcHMA5* (XM_007040138.2), an ortholog to the ethylene-coupled Cu transporter *AtRAN1* (Li et al. 2017), was upregulated in the roots of TSH660(LA) (Fig. 5D), coinciding with changes in ethylene-related gene expression and signaling. This pattern suggests that TcHMA5 may function similarly to RAN1 and is consistent with its increased expression during leaf maturation (Kulesza et al. 2024). Furthermore, *TcHMA1,* whose Arabidopsis ortholog, *AtHMA1*, contributes to Zn detoxification (Kim et al. 2009), was downregulated in PA121(HA) leaves. These observations highlight genotype-specific regulation of HMA transporters under Cd stress.

Chelation- and nutrient-mediated processes in cadmium detoxification

Chelation, sequestration, and precipitation are critical processes in Cd detoxification. Nitrogen (N) and phosphorus (P) can directly precipitate (Van Belleghem et al. 2007), whereas sulfur (S) acts as a reducing agent in molecules such as glutathione S-transferases (GSTs), which mobilize Cd bound to glutathione (GSH) and detoxify ROS (Flores et al. 2000). In our study, we observed transcriptional changes in genes encoding for N, P, and S transporters (Fig. 5B-C). Changes in abundance or chelation were not tested directly, so we limit our discussion to their transcriptional behavior. Notably, given lignin's capacity for metal chelation, the downregulation of genes for lignin-forming proteins (Fig. 5A) is intriguing. *LFP* suppression has been associated with increased levels of monolignols, such as coumarins (Abdur Rahim et al. 2022). This raises an important question regarding the role of monolignols in the Cd response: whether they act as direct coordination compounds facilitating Cd mobilization or serve as precursors required for mobilizing other elements (Tsai and Schmidt 2017).

Concluding remarks

Our results suggest that no single factor predominates in explaining differences in Cd uptake in cacao. However, we detected distinct differentially expressed genes (DEGs) that provide a foundation for future functional validation. The proposed model integrates observations from cacao with evidence from other species into a cohesive hypothesis involving water-use efficiency, specialized elemental transport, chelation, and detoxification. Several hypotheses are proposed for future research that may advance the development of climate-resilient, low-Cd cacao varieties. Importantly, this work represents the first published whole-transcriptome study of Cd response for this species.

In conclusion, this study reveals a coordinated physiological and molecular response in cacao aimed at mitigating Cd uptake and toxicity. This response involves hormonal signaling, reduced mass flow, modulation of nutrient and ion transport, and activation of detoxification pathways. Future research is needed to validate this model by investigating cellular redox dynamics, monolignol-mediated chelation, root structural adaptations, and controlled phytohormone or inhibitor applications. Given the substantial overlap between drought and Cd responses, integrating not only both stressors (Ortiz-Álvarez et al. 2023), but also multiple omics approaches, may accelerate the development of climate-resilient, low-Cd cacao varieties. Finally, the molecular features identified here represent promising targets for marker-assisted selection and gene-editing strategies in cacao improvement.

Acknowledgements The authors would like to acknowledge the Huck Institutes' Metabolomics Core Facility (RRID: SCR_023864) for the use of the Sciex ZenoTOF 7600 LC-MS and Sergei Koshkin for helpful discussions, the team at USDA, ARS Tropical Agriculture Research Station in Mayagüez, Puerto Rico, for their service in providing cacao pods for our experiment, and Dr. Naomi Altman for supporting the statistical model design. They would also like to acknowledge the use of BioRender® template for Figure 8, and OpenAI's ChatGPT and Grammarly for writing refinement and grammatical editing.

Author contributions PD-D designed the research, contributed new laboratory methods and assays, assembled the transcriptome, analyzed data, and wrote the paper. FMM-B, MG, and SM designed the physiological experiments. FMM-B performed and analyzed the physiological experiments, contributed new analytical methods for phytohormone purification, contributed to the transcriptome assembly, contributed new analytical and computational tools, analyzed data, prepared all figures, and wrote the paper. IA and NC contributed to the transcriptome assembly and analyzed data. MG and SM participated in the interpretation of the results and contributed to writing the manuscript. ACM designed the research. CR-M performed and set up greenhouse experiments to measure growth, Cd concentration, RNA purification, and sampled material. RY designed the research, collected samples, analyzed data, and contributed to writing the manuscript.

Funding This work was funded by the "Ministerio de Ciencia, tecnologia e Innovación de Colombia" Agreements TV-18 and TV-19 through the "Corporación Colombiana de Investigación Agropecuaria" (AGROSAVIA), under the project: "Selección de parentales de cacao por resistencia a enfermedades y estrés hídrico, productividad, compatibilidad, menor absorción de cadmio y calidad", ID 1000429. It was also funded by the Huck Institutes of the Life Sciences, the Endowed Program in the Molecular Biology of Cacao and USDA Hatch appropriations under Project #PEN05003 and Accession #7007428 and Project #PEN4879 and Accession #7005892.

Data availability The reads used in the current study are available at GenBank (https://www.ncbi.nlm.nih.gov/genbank/) under the BioProject ID: PRJNA943175. The Licor 6400XT output and cleaned data used for gas exchange analyses are available on Zenodo (https://doi.org/10.5281/zenodo.17316758.). The scripts and analytical workflows used for data cleaning, statistical modeling, and figure generation are available in the associated GitHub repository at https://github.com/FranciscoMenendez/cacao-cadmium-response. The sequences obtained in this project belong to the Colombian State, as the country of origin of the genetic material, in accordance with the Convention on Biological Diversity, the Nagoya Protocol, and Andean Decision 391.

Declarations

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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 indicated 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/.

References

Abdur Rahim M, Atikur Rahman M, Korea Md Mizanur Rahman S, Song C, Li X, Jia B, Liu L, Wei P, Aamir Manzoor M, Wang F, et al (2022) Comparative Transcriptomics Unveil the Crucial Genes Involved in Coumarin Biosynthesis in Peucedanum praeruptorum Dunn. https://doi.org/10.3389/fpls.2022.899819

Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, Feuermann M, Gaudet P, Harris NL, Hill DP et al (2023) The Gene Ontology knowledgebase in 2023. Genetics 224:1–14

Andrews S (2010) FastQC: A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H (2020) KofamKOALA: KEGG

Ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics 36:2251–2252
Argout X, Martin G, Droc G, Fouet O, Labadie K, Rivals E, Aury JM, Lanaud C (2017) The cacao Criollo genome v2.0: an improved version of the genome for genetic and functional genomic studies. BMC Genomics 18:1–9
Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al (2000) Gene ontology: tool for the unification of biology. Nat Genet 25:25–29
Aubry S, Kelly S, Kümpers BMC, Smith-Unna RD, Hibberd JM (2014) Deep evolutionary comparison of gene expression identifies parallel recruitment of trans-factors in two independent origins of C4 photosynthesis. PLoS Genet. https://doi.org/10.1371/journal.pgen.1004365
Barati E, Moore RET, Ullah I, Kreissig K, Coles BJ, Dunwell JM, Rehkämper M (2023) An investigation of zinc isotope fractionation in cacao (*Theobroma cacao* L.) and comparison of zinc and cadmium isotope compositions in hydroponic plant systems under high cadmium stress. Sci Rep. https://doi.org/10.1038/s41598-023-30899-z
Barber SA (1995) Soil nutrient bioavailability: a mechanistic approach, 2nd ed. John Wiley & Sons
Barroso JP, de Almeida A-AF, do Nascimento JL, Oliveira BRM, dos Santos IC, Mangabeira PAO, Ahnert D, Baligar VC (2023) The damage caused by Cd toxicity to photosynthesis, cellular ultrastructure, antioxidant metabolism, and gene expression in young cacao plants are mitigated by high Mn doses in soil. Environ Sci Pollut Res 30:115646–115665
Blommaert H, Aucour AM, Wiggenhauser M, Moens C, Telouk P, Campillo S, Beauchêne J, Landrot G, Testemale D, Pin S et al (2022) From soil to cacao bean: Unravelling the pathways of cadmium translocation in a high Cd accumulating cultivar of Theobroma cacao L. Front Plant Sci 13:1–19
Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120
Bourgon R, Gentleman R, Huber W (2010) Independent filtering increases detection power for high-throughput experiments. Proc Natl Acad Sci U S A 107:9546–9551
Cailliatte R, Lapeyre B, Briat J-F, Mari S, Curie C (2009) The NRAMP6 metal transporter contributes to cadmium toxicity. Biochem J 422:217–228
Cao L, Liu L, Zhang C, Ren W, Zheng J, Tao C, Zhu W, Xiang M, Wang L, Liu Y et al (2024) The MYC2 and MYB43 transcription factors cooperate to repress HMA2 and HMA4 expression, altering cadmium tolerance in *Arabidopsis thaliana*. J Hazard Mater. https://doi.org/10.1016/j.jhazmat.2024.135703
Charkiewicz AE, Omeljaniuk WJ, Nowak K, Garley M, Nikliński J (2023) Cadmium toxicity and health effects—a brief summary. Molecules. https://doi.org/10.3390/molecules28186620
Chen J, Nelson DC, Shukla D (2022) Activation mechanism of strigolactone receptors and its impact on ligand selectivity between host and parasitic plants. J Chem Inf Model 62:1712–1722
Chen L, Wu M, Jin W, Lei T, Li Y, Wu X, Fu S (2023) Gene identification and transcriptome analysis of cadmium stress in tomato. Front Sustain Food Syst 7:1–11
Chmielowska-Bak J, Lefèvre I, Lutts S, Deckert J (2013) Short term signaling responses in roots of young soybean seedlings exposed to cadmium stress. J Plant Physiol 170:1585–1594
Chmielowska-Bąk J, Gzyl J, Rucińska-Sobkowiak R, Arasimowicz-Jelonek M, Deckert J (2014) The new insights into cadmium sensing. Front Plant Sci 5:1–13
Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd ed. Lawrence Erlbaum Associates, Publishers
Connolly EL, Campbell NH, Grotz N, Prichard CL, Guerinot ML (2003) Overexpression of the FRO2 ferric chelate reductase confers tolerance to growth on low iron and uncovers posttranscriptional control. Plant Physiol 133:1102–1110
Dashner ZS (2024) Novel biotechnological approaches to understand and reduce cadmium in the iron uptake pathway of theobroma cacao. Doctoral Dissertation. The Pennsylvania State University, University Park
dos Santos MLS, de Almeida AAF, da Silva NM, Oliveira BRM, Silva JVS, Junior JOS, Ahnert D, Baligar VC (2020) Mitigation of cadmium toxicity by zinc in juvenile cacao: physiological, biochemical, molecular and micromorphological responses. Environ Exp Bot 179:104201
EFSA (2012) Cadmium dietary exposure in the European population. EFSA J 10:2551
Elsayed AI, Rafudeen MS, Golldack D (2014) Physiological aspects of raffinose family oligosaccharides in plants: protection against abiotic stress. Plant Biol 16:1–8
Engbersen N, Gramlich A, Lopez M, Schwarz G, Hattendorf B, Gutierrez O, Schulin R (2019) Cadmium accumulation and allocation in different cacao cultivars. Sci Total Environ 678:660–670
Ernst WHO, Krauss GJ, Verkleij JAC, Wesenberg D (2008) Interaction of heavy metals with the sulphur metabolism in angiosperms from an ecological point of view. Plant Cell Environ 31:123–143
Feng J, Jia W, Lv S, Bao H, Miao F, Zhang X, Wang J, Li J, Li D, Zhu C et al (2018) Comparative transcriptome combined with morpho-physiological analyses revealed key factors for differential cadmium accumulation in two contrasting sweet sorghum genotypes. Plant Biotechnol J 16:558–571
Fister AS, Shi Z, Zhang Y, Helliwell EE, Maximova SN, Guiltinan MJ (2016) Protocol: transient expression system for functional genomics in the tropical tree *Theobroma cacao* L. Plant Methods 12:1–13
Floková K, Shimels M, Andreo Jimenez B, Bardaro N, Strnad M, Novák O, Bouwmeester HJ (2020) An improved strategy to analyse strigolactones in complex sample matrices using UHPLC-MS/MS. Plant Methods 16:1–17
Flores HE, Vivanco JM, Loyola-Vargas VM (2000) Plant glutathione s-transferases: enzymes with multiple functions in sickness and in health. Trends Plant Sci 5:193–198
Garnier L, Simon-Plas F, Thuleau P, Agnel J-P, Blein J-P, Ranjeva R, Montillet J-L (2006) Cadmium affects tobacco cells by a series of three waves of reactive oxygen

species that contribute to cytotoxicity. Plant Cell Environ 29:1956–1969
Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36:3420–3435
Gu Z, Hübschmann D (2022) Simplify enrichment: a bioconductor package for clustering and visualizing functional enrichment results. Genom Proteom Bioinform. https://doi.org/10.1016/j.gpb.2022.04.008
Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32:2847–2849
Herbette S, Taconnat L, Hugouvieux V, Piette L, Magniette MLM, Cuine S, Auroy P, Richaud P, Forestier C, Bourguignon J et al (2006) Genome-wide transcriptome profiling of the early cadmium response of *Arabidopsis* roots and shoots. Biochimie 88:1751–1765
Hoagland DR, Arnon DI (1950) The water-culture method for growing plants without soil, 3rd ed. Circular 347 of the California Agricultural Experimental Station, Berkeley
Hsu YT, Kao CH (2003) Role of abscisic acid in cadmium tolerance of rice (*Oryza sativa* L.) seedlings. Plant Cell Environ 26:867–874
Hussain D, Haydon MJ, Wang Y, Wong E, Sherson SM, Young J, Camakaris J, Harper JF, Cobbett CS (2004) P-Type ATPase heavy metal transporters with roles in essential zinc homeostasis in *Arabidopsis*. Plant Cell 16:1327–1339
Ishimaru Y, Takahashi R, Bashir K, Shimo H, Senoura T, Sugimoto K, Ono K, Yano M, Ishikawa S, Arao T et al (2012) Characterizing the role of rice NRAMP5 in manganese, iron and cadmium transport. Sci Rep 2:1–8
Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45:D353–D361
Kim YY, Choi H, Segami S, Cho HT, Martinoia E, Maeshima M, Lee Y (2009) AtHMA1 contributes to the detoxification of excess Zn(II) in *Arabidopsis*. Plant J 58:737–753
Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907–915
Korotkevich G, Sukhov V, Sergushichev A (2019) Fast gene set enrichment analysis. bioRxiv. BioAr
Kubier A, Wilkin RT, Pichler T (2019) Cadmium in soils and groundwater: a review. Appl Geochem. https://doi.org/10.1016/j.apgeochem.2019.104388
Kulesza E, Thomas P, Prewitt SF, Shalit-Kaneh A, Wafula E, Knollenberg B, Winters N, Esteban E, Pasha A, Provart N et al (2024) The cacao gene atlas: a transcriptome developmental atlas reveals highly tissue-specific and dynamically regulated gene networks in *Theobroma cacao* L. BMC Plant Biol. https://doi.org/10.1186/s12870-024-05171-9
Lewis C, Lennon AM, Eudoxie G, Umaharan P (2018) Genetic variation in bioaccumulation and partitioning of cadmium in Theobroma cacao L. Sci Total Environ 640–641:696–703
Li W, Lacey RF, Ye Y, Lu J, Yeh KC, Xiao Y, Li L, Wen CK, Binder BM, Zhao Y (2017) Triplin, a small molecule, reveals copper ion transport in ethylene signaling from ATX1 to RAN1. PLoS Genet 13:1–21
Liao Y, Smyth GK, Shi W (2014) FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–930
Liu Y, Yu X, Feng Y, Zhang C, Wang C, Zeng J, Huang Z, Kang H, Fan X, Sha L et al (2017) Physiological and transcriptome response to cadmium in cosmos (*Cosmos bipinnatus* Cav.) seedlings. Sci Rep 7:1–16
Liu J, Yang L, Wen Y, Li Y, Shabala S, Zhang J, Song WY (2024) The roles of abscisic acid and ethylene in cadmium accumulation and tolerance in plants. Plant Soil 505:115–126
Lux A, Martinka M, Vaculík M, White PJ (2011) Root responses to cadmium in the rhizosphere: a review. J Exp Bot 62:21–37
Marschner P, Rengel Z (2011) Nutrient availability in soils. In: Marschner H (ed) Marschner's mineral nutrition of plants. Elsevier Science & Technology, pp 499–522
Mezentsev Y (2017) Binary cut-and-branch method for solving mixed integer programming problems. 2017 Constructive nonsmooth analysis and related topics (Dedicated to the Memory of VF Demyanov), CNSA 2017 - Proceedings 1–3
Michael A, Ahlmann-eltze C, Forbes K, Anders S (2022) Package 'DESeq2 .'
Moore RET, Ullah I, de Oliveira VH, Hammond SJ, Strekopytov S, Tibbett M, Dunwell JM, Rehkämper M (2020) Cadmium isotope fractionation reveals genetic variation in Cd uptake and translocation by Theobroma cacao and role of natural resistance-associated macrophage protein 5 and heavy metal ATPase-family transporters. Hortic Res 7:71
Morel M, Crouzet J, Gravot A, Auroy P, Leonhardt N, Vavasseur A, Richaud P (2008) AtHMA3, a P1B-ATPase allowing Cd/Zn/Co/Pb vacuolar storage in *Arabidopsis*. Plant Physiol 149:894–904
Ortiz-Álvarez A, Magnitskiy S, Silva-Arero EA, Rodríguez-Medina C, Argout X, Castaño-Marín ÁM (2023) Cadmium accumulation in cacao plants (*Theobroma cacao* L.) under drought stress. Agronomy. https://doi.org/10.3390/agronomy13102490
Page V, Feller U (2015) Heavy metals in crop plants: transport and redistribution processes on the whole plant level. Agronomy 5:447–463
Pedron T, Cristine Augusto C, Carvalho Silva G, Cezar Valeriano M, Benicia Mamián-López M, Slaveykova VI, Lemos Batista B (2025) Human health risk assessment and concentration of Al, Mn, Fe, Cu, Zn, Se, As, Cd, Pb, Hg, Rb, and REEs in chocolate. Food Chem Toxicol 206:115770
Pérez-Chaca MV, Rodríguez-Serrano M, Molina AS, Pedranzani HE, Zirulnik F, Sandalio LM, Romero-Puertas MC (2014) Cadmium induces two waves of reactive oxygen species in *Glycine max* (L.) roots. Plant Cell Environ 37:1672–1687
Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140
Rodriguez-Serrano M, Romero-Puertas MC, Pazmino DM, Testillano PS, Risueno MC, del Rio LA, Sandalio LM (2009) Cellular Response of Pea Plants to Cadmium

Toxicity: Cross Talk between Reactive Oxygen Species, Nitric Oxide, and Calcium. Plant Physiol 150:229–243
Rui H, Chen C, Zhang X, Shen Z, Zhang F (2016) Cd-induced oxidative stress and lignification in the roots of two *Vicia sativa* L. varieties with different Cd tolerances. J Hazard Mater 301:304–313
Schellingen K, Van Der Straeten D, Vandenbussche F, Prinsen E, Remans T, Vangronsveld J, Cuypers A (2014) Cadmium-induced ethylene production and responses in *Arabidopsis thaliana* rely on ACS2 and ACS6 gene expression. BMC Plant Biol 14:1–14
Schlicker A, Domingues FS, Rahnenführer J, Lengauer T (2006) A new measure for functional similarity of gene products based on gene ontology. BMC Bioinformatics 7:1–16
Shen C, Yang YM, Sun YF, Zhang M, Chen XJ, Huang YY (2022) The regulatory role of abscisic acid on cadmium uptake, accumulation and translocation in plants. Front Plant Sci 13:1–13
Smolders E (2001) CADMIUM UPTAKE BY PLANTS *. Int J Occup Med Environ Health 14:177–183
Stroiński A, Giżewska K, Zielezińska M (2013) Abscisic acid is required in transduction of cadmium signal to potato roots. Biol Plant 57:121–127
Thakur S, Choudhary S, Bhardwaj P (2019) Comparative transcriptome profiling under cadmium stress reveals the uptake and tolerance mechanism in *Brassica juncea*. J Plant Growth Regul 38:1141–1152
Thumuluri V, Almagro Armenteros JJ, Johansen AR, Nielsen H, Winther O (2022) DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. Nucleic Acids Res 50:W228–W234
Trapp MA, De Souza GD, Rodrigues-Filho E, Boland W, Mithöfer A (2014) Validated method for phytohormone quantification in plants. Front Plant Sci 5:1–11
Tsai HH, Schmidt W (2017) Mobilization of iron by plant-borne coumarins. Trends Plant Sci 22:538–548
Ullah I, Wang Y, Eide DJ, Dunwell JM (2018) Evolution, and functional analysis of natural resistance-associated macrophage proteins (NRAMPs) from *Theobroma cacao* and their role in cadmium accumulation. Sci Rep 8:1–15
US Environmental Protection Agency (1996) Method 3052: Microwave assisted acid digestion of siliceous and organically based matrices. US EPA
US Environmental Protection Agency (2007) Method 6010C: Inductively coupled plasma-atomic emission spectrometry.
Van Belleghem F, Cuypers A, Semane B, Smeets K, Vangronsveld J, D'Haen J, Valcke R (2007) Subcellular localization of cadmium in roots and leaves of *Arabidopsis thaliana*. New Phytol 173:495–508
Vanderschueren R, Argüello D, Blommaert H, Montalvo D, Barraza F, Maurice L, Schreck E, Schulin R, Lewis C, Vazquez JL et al (2021) Mitigating the level of cadmium in cacao products: reviewing the transfer of cadmium from soil to chocolate bar. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2021.146779
Vera Pinargote LD, Maddela NR, Jaimez RE, Corozo-Quiñonez L (2025) Unraveling cadmium stress in cacao (*Theobroma cacao* L): photosynthetic performance and differential accumulation in two new hybrid populations under controlled conditions. S Afr J Bot 185:372–382
Wade J, Ac-Pangan M, Favoretto VR, Taylor AJ, Engeseth N, Margenot AJ (2022) Drivers of cadmium accumulation in *Theobroma cacao* L. beans: a quantitative synthesis of soil-plant relationships across the Cacao Belt. PLoS ONE. https://doi.org/10.1371/journal.pone.0261989
Wafula EK, Zhang H, Von Kuster G, Leebens-Mack JH, Honaas LA, dePamphilis CW (2023) PlantTribes2: tools for comparative gene family analysis in plant genomics. Front Plant Sci. https://doi.org/10.3389/fpls.2022.1011199
Wang B, Wang Y, Yuan X, Jiang Y, Zhu Y, Kang X, He J, Xiao Y (2023a) Comparative transcriptomic analysis provides key genetic resources in clove basil (*Ocimum gratissimum*) under cadmium stress. Front Genet 14:1–16
Wang M, Wang R, Mur LAJ, Ruan J, Shen Q, Guo S (2021) Functions of silicon in plant drought stress responses. Hortic Res 8:1–13
Wang Z, Zheng Z, Zhu Y, Kong S, Liu D (2023b) PHOSPHATE RESPONSE 1 family members act distinctly to regulate transcriptional responses to phosphate starvation. Plant Physiol 191:1324–1343
WEPAL (2019) Wageningen Evaluating Programs Certificate of Analysis (WEPAL) - ISE sample 890. 1–12
Wickham H (2009) Ggplot2: elegant graphics for data analysis. Springer
Yockteng R, Almeida AMR, Yee S, Andre T, Hill C, Specht CD (2013) A method for extracting high-quality RNA from diverse plants for next-generation sequencing and gene expression analyses. Appl Plant Sci 1:1300070
Zhang Y, Chao J, Li X, Zhang C, Khan R, Du S, Xu N, Song L, Liu H, Shi Y (2021) Comparative transcriptome combined with biochemical and physiological analyses provide new insights toward cadmium accumulation with two contrasting *Nicotiana* species. Physiol Plant 173:369–383
Zhang Y, Wang Z, Liu Y, Zhang T, Liu J, You Z, Huang P, Zhang Z, Wang C (2023) Plasma membrane-associated calcium signaling modulates cadmium transport. New Phytol 238:313–331
Zheng P, Cao L, Zhang C, Pan W, Wang W, Yu X, Li Y, Fan T, Miao M, Tang X et al (2022) MYB43 as a novel substrate for CRL4PRL1 E3 ligases negatively regulates cadmium tolerance through transcriptional inhibition of HMAs in *Arabidopsis*. New Phytol 234:884–901

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