Integrative BSA-seq and time-course RNA-seq prioritize candidate loci and genes associated with whole-plant cadmium (Cd) accumulation and tolerance in dark jute (Corchorus olitorius L.)

Jian-tang Xu

xujiantang@fafu.edu.cn

Fujian Agriculture and Forestry University

Shixin Zhu

Fujian Agriculture and Forestry University

Victoria Oluwapelumi Akintayo

Fujian Agriculture and Forestry University

Yaqi Zhao

Fujian Agriculture and Forestry University

Lijun Li

Fujian Agriculture and Forestry University

Han Li

Fujian Agriculture and Forestry University

Wei-qun Zhang

Fujian Agriculture and Forestry University

Aifen Tao

Fujian Agriculture and Forestry University

Pingping Fang

Fujian Agriculture and Forestry University

Jianmin Qi

Fujian Agriculture and Forestry University

Shu-biao Zhang

Fujian Agriculture and Forestry University

Research Article

Keywords: Dark jute, Whole-plant cadmium accumulation, Cadmium tolerance, Time-course RNA-seq, BSA-seq, Candidate genes

Posted Date: June 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9612727/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Plant Biology on July 23rd, 2026. See the published version at https://doi.org/10.1186/s12870-026-09553-z.

Abstract

Background:

Dark jute (*Corchorus olitorius* L.) is a high-biomass fiber crop with potential value for phytoremediation, but the genetic basis of whole-plant cadmium (Cd) accumulation and tolerance remains poorly understood. Identifying loci and candidate genes associated with Cd-related traits can support molecular improvement of dark jute for cultivation in Cd-affected environments.

Results:

In this study, we integrated bulked segregant analysis sequencing (BSA-seq) with time-course RNA sequencing (RNA-seq), Cd phenotyping, and qRT-PCR validation to investigate Cd accumulation and tolerance in dark jute. A Cd-tolerant, low-accumulating accession (CO57) and a Cd-sensitive, high-accumulating accession (CO83) were used to generate an F_2 population for extreme-phenotype bulking. Time-course RNA-seq of the tolerant parent identified 2,830 and 4,525 Cd-responsive differentially expressed genes in roots and aboveground tissues, respectively, with 1,207 shared genes. Functional enrichment showed that aboveground responses were mainly associated with plastid and photosynthesis-related processes, whereas root responses involved redox regulation, cell wall and cytoskeleton remodeling, and metal-stress-related functions. For BSA-seq, 50 extremely low-accumulating and 50 extremely high-accumulating F_2 individuals were pooled, yielding 625,354 high-quality SNPs. ΔSNP-index analysis detected five candidate intervals spanning 7.13 Mb and containing 662 annotated genes. By integrating positional evidence, Cd-responsive expression patterns, and functional annotations, seven candidate genes were prioritized, including genes encoding ABC transporter-like proteins, NAC-domain regulators, and two PLAC8-domain proteins annotated as Cd-resistance-related proteins. qRT-PCR analysis supported the RNA-seq expression patterns of the prioritized candidates.

Conclusions:

These results define candidate genomic intervals and Cd-responsive genes associated with whole-plant Cd accumulation and tolerance in dark jute. The study provides a genetic and transcriptomic framework for future functional validation, marker development, and phytoremediation-oriented improvement of dark jute.

1. Introduction

Elevated Cadmium (Cd) in agricultural environments arises from multiple anthropogenic and geogenic sources, including mining and smelting activities, atmospheric deposition, and the long-term use of Cd-bearing phosphate fertilizers, which can increase Cd inputs to arable soils over time [1, 2]. Because

dietary exposure to Cd is largely driven by plant-derived foods, reducing Cd accumulation in crops and crop-based products is a priority for both sustainable agriculture and public health [3].

In plants, Cd disrupts growth and development by impairing photosynthesis, disturbing nutrient homeostasis, and triggering oxidative damage and cellular dysfunction [4]. Cd-induced reactive oxygen species (ROS) can damage nucleic acids and proteins [5], and Cd stress has been associated with reductions in photosynthetic indices, including photosynthetic rate and intracellular CO_2 concentration [6]. Consistent with these effects, Cd can impede plant growth and development even at low concentrations [7].

At the molecular level, Cd enters root cells largely via transport system for essential divalent cations (e.g., Zn^{2+}, Fe^{2+}, Ca^{2+}), and is subsequently regulated through chelation, sequestration, and transport process [4]. Natural variation for Cd accumulation and tolerance is widespread, but these traits are typically quantitative, environmentally sensitive, and controlled by multiple loci [3, 4, 8].

High-biomass crops are attractive candidates for reducing Cd bioavailability through phytoremediation and phytoextraction, where whole-plant metal accumulation integrates tissue concentration with biomass production [9, 10]. Exploiting plant diversity remains important for improving heavy metal (HM) accumulation capacity and identifying effective hyperaccumulators [11, 12].

Dark jute (*Corchorus olitorius* L.) is a major bast-fiber crop with substantial biomass potential and expanded genomic resources that facilitate locus discovery and molecular breeding [13, 14]. Recent studies support the capacity of jute to tolerate HM stress and accumulate metals [15], and integrative genetic dissection has demonstrated that mapped loci can be connected to candidate genes using combined genomic and transcriptomic evidence [16].

From a crop improvement perspective, breeding for Cd-related traits requires resolving two partially independent components: whole-plant Cd accumulation which reflects uptake, internal transport, and retention across organs; and Cd tolerance, reflecting the ability to maintain growth and limit toxicity under Cd exposure [3, 4]. Bulked Segregant Analysis (BSA), and sequencing-based BSA approaches (BSA-seq) enable efficient mapping of loci by comparing pooled DNA from phenotypic extremes [17, 18], but candidate intervals often contain many genes, requiring additional evidence to prioritize candidate genes. Transcriptome profiling provides complementary support by identifying stress-responsive pathways and genotype-specific programs, and integrating mapping-by-sequencing with RNA-seq can improve candidate resolution by intersecting mapped regions with differentially expressed genes (DEGs) and enriched pathways [3, 4, 16, 19, 20]. Such integrative strategies have successfully identified candidate loci and genes for Cd accumulation in other crops, including rapeseed [8].

Despite these advances, the genetic architecture underlying whole-plant Cd accumulation and tolerance in dark jute remains insufficiently characterized, and integrative datasets linking mapped loci to time-resolved transcriptional responses are limited. Addressing this gap is important for developing

molecular markers and candidate genes that support Cd-resilient dark jute improvement and broader deployment of high-biomass crops in Cd-impacted production systems [1, 21, 22].

Here, we combined BSA-seq of F_2 bulks defined by whole-plant Cd accumulation with time-course RNA-seq in the tolerant parent to map Cd accumulation-associated loci, characterize tolerance-associated transcriptional responses and prioritize candidate genes, including those related to metal transport. The objectives of this study were to: (i) evaluate Cd tolerance and accumulation differences between contrasting parental lines, (ii) identify candidate genomic intervals associated with whole-plant Cd accumulation using BSA-seq; (iii) characterize tissue-specific and time-dependent transcriptional responses to Cd stress; and (iv) prioritize candidate genes associated with Cd accumulation and tolerance for future functional validation and marker development.

2. Materials and Methods

2.1 Plant material, Cd growing conditions, and experimental design

Seeds of dark jute accessions, CO57 (P1) and CO83 (P2), which were used as parental lines, were obtained from the germplasm collection maintained by the Key Laboratory of Genetics, Breeding and Multiple Utilization of Bast Fiber Crops, Fujian Agriculture and Forestry University (FAFU), Fuzhou, China. The accession names CO57 and CO83 were assigned by our laboratory for germplasm identification and management. These accessions were cultivated germplasm materials and were not collected from the wild; therefore, no specific permission or licence was required for sample collection. The plant materials were authenticated as *Corchorus olitorius* L. by Dr. Xu Jian-tang, FAFU, based on institutional germplasm records and morphological characteristics. No voucher specimen was deposited because the materials used in this study were cultivated accessions from the institutional germplasm collection. The experimental research on plants complied with relevant institutional, national, and international guidelines and legislation.

Seeds were sown in 72-cell trays and grown in the greenhouse of the Field Laboratory of Genetics, Breeding and Multiple Utilization of Bast Fiber Crops, FAFU, under a 12 h light/12 h dark photoperiod at 25 ^{o}C. After two weeks, uniform seedlings at the four-leaf stage were selected and transferred to a 100 mL conical flask containing half-strength of Hoagland solution to allow root recovery from handling stress. The half-strength Hoagland solution was prepared using a modified formulation to avoid Cd^{2+} precipitation, following the hydroponic protocol described previously [23].

To characterize parental responses and determine an appropriate working Cd concentration, seedlings of both parents were exposed to a Cd concentration gradient (30, 50, 70, and 100 μmol/L) supplied as $CdCl_2.2.5H_2O$ in the modified half-strength Hoagland solution. Control (CK) plants were maintained in the same nutrient solution without Cd addition. Treatments were arranged in a completely randomized design with three biological replicates. The nutrient solution was renewed every three days to avoid

microbial growth [24]. Cd tolerance was evaluated using biomass-derived indices, including the tolerance index (TI) and growth inhibition (GGI), calculated from CK and Cd-treated biomass at the replicate levels. Visible toxicity symptoms (e.g., chlorosis and leaf discolouration) were recorded photographically (Fig. S1). Whole-plant Cd accumulation was assessed as Cd concentration on a dry-weight basis (roots and shoots combined) and quantified by Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES) (Section 2.3).

Based on the parental dose-response results, 70 µmol/L Cd^{2+} was selected as the working concentration for population screening. The two parental lines were crossed to generate F_1 plants (Fig. S2) and F_1 individuals were self-pollinated to produce F_2 populations. A total of 300 F_2 seedlings were screened under 70 µmol/L Cd^{2+} using the same hydroponic system. After 7 days of Cd exposure, surviving plants with sufficient biomass were harvested for phenotyping variation, and whole-plant Cd concentration was determined by ICP-OES (Section 2.3). A total of 296 F_2 individuals yielded valid Cd concentration data and were used for downstream analyses.

2.2 Biomass-based indices (TI and GGI)

Fresh weight (FW) and dry weight (DW) were measured for each accession under CK and Cd treatments at 30, 50, 70, and 100 µmol/L. For each biological replicate, the tolerance index (TI) and growth inhibition rate (GGI) were calculated using biomass values under CK and Cd treatments, where Bc represents biomass under CK and Bt represents biomass under Cd treatment. The indices were calculated as follows:

GGI = (Bc-Bt)/Bc (1)

TI = (Bt/Bc) [25] (2)

2.3 Cd quantification using ICP-OES

To quantify whole-plant Cd accumulation, roots of stressed seedlings were immersed in 20 mM Na_2 EDTA for 15–20 minutes to remove Cd adhered to the surface of the root. After which the roots were washed thrice with distilled water and finally with de-ionized water [26]. Whole plants (shoot and root combined) were weighed and dried in a 65 ˚C oven for 3 days until reaching a constant weight. The samples were ground into powder using a hammer cyclone mill at 60 Hz frequency for 90 seconds. For acid digestion, 0.1 g of each powdered sample were weighed in three replicates and put in a 50 mL conical flask. 5 mL of HNO_3 were added to each conical flask, and the digestion was performed on a hot block digester. A glass funnel was placed on each of the 50 mL conical flasks, the samples were heated as follows, 120 ^{o}C for 60 mins, 140 ^{o}C for 60 mins, and 160 ^{o}C for 60 mins and the whole digestion process was done in a ventilator.

After the digestion was completed (a clear colourless liquid), the sample was diluted with distilled water to make 50 mL and 10 mL of the diluent was later used to detect the Cd concentration [27]. Cd concentration was measured using an ICP-OES (Avio 220 Max, PerkinElmer, Waltham, MA, USA) at an

analytical wavelength of 228.802 nm with four technical readings per sample. A standard curve was generated using Cd standard solution containing 5% HNO_3 at 0 mg/L, 5 mg/L, 10 mg/L, 20 mg/L, 30 mg/L, 40 mg/L and 50 mg/L (Fig. S3). Sample analysis followed the Chinese national standard GB 5009.268–2016, issued by National Standard Substances Center of China.

2.4 Time-course RNA-seq design and analysis

The tolerant parental line P1 was grown for three weeks and subjected to Cd stress as described in Section 2.1. To characterize Cd-responsive transcriptional dynamics, root (R) and aboveground tissue (L) samples were collected at 0, 24, 48, and 72 h. The 0 h CK samples were used as baseline controls, and Cd-treated samples were collected at 24, 48, and 72 h after exposure to 70 μmol/L Cd^{2+}. Tissues were immediately frozen in liquid nitrogen and stored at − 80°C until RNA extraction. For each time-course × tissue × treatment combination, three biological replicates were collected and sequenced. RNA-seq libraries were prepared and sequenced by Jisi Huiyuan Biotechnology (Nanjing, China) following standard protocols. Sequencing yield and base-quality metrics for all libraries are summarized in Table S1.

Raw reads were quality-controlled using fastp v0.20.0 with adapter removal, filtering of reads containing more than 5 ambiguous bases (N), and removal of reads in which bases with quality < 5 accounted for more than 50% of the read length (main parameters: -q 5 -n 5). Clean reads were aligned to the dark jute reference genome TC [14] using HISAT2 v2.1.0 with the option --novel-splicesite-outfile, and the resulting SAM files were converted, sorted and indexed as BAM files using samtools v1.9.

Alignment quality, mapping efficiency and coverage of exons and transcripts were assessed using Qualimap v2.2.1 in RNA-seq mode. Transcript assembly and expression quantification were performed with StringTie v2.1.3. For each sample, transcripts were assembled with parameters -f 0.2 -a 30 -c 5 -j 5, and assemblies were merged using the --merge function to obtain an optimized annotation. StringTie was then rerun in quantification mode (-e -p 4) using this annotation to generate gene-level abundance estimates. Read counts were organized using Ballgown (R), and FPKM/TPM values were calculated using in-house Python scripts. Differential expression analysis was performed on raw counts using DESeq2 v1.42.0, with DEGs defined as those with $|\log_2(\text{fold change})| \geq 1$ and FDR ≤ 0.05. The DEGs were annotated by sequence alignment to the Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Non-Redundant Protein Database (NR), Swiss-Prot (swissprot), and InterPro databases using BLAST (v 2.15.0) [28].

2.5 DNA isolation and sequencing

Leaf samples were collected from 5 plants of each parent, P1 and P2. From the F_2 population, 50 plants with high whole-plant Cd concentration and 50 plants with low whole-plant Cd concentration were selected, and a bulk pool was made for each phenotypic trait. Genomic DNA was extracted from each pool using CTAB-based protocol [29]. The concentration and quality of the total genomic DNA were determined using a NanoDrop2000 Spectrophotometer (Thermo Fisher Scientific). DNA libraries (350

bp) for Illumina/BGI sequencing were constructed for the bulk pools and parent. After DNA library construction, sequencing was performed on an Illumina HiSeq XTen/NovaSeq/BGI platform by a commercial service Jisi Huiyuan Biotechnology (Nanjing, China), with 150 bp read lengths. Raw reads (FASTQ) were processed on BMK Cloud platform (www.biocloud.net) to remove adapter contamination and low-quality reads; Q20, Q30, GC content and duplication levels were calculated, and only high-quality clean reads were retained (Table S2).

2.6 Read mapping, variant calling, and BSA-seq

Clean reads were aligned to the dark jute reference genome TC [14] using bwa-mem2 v2.2 (mem -t 4 -M), and the alignments were sorted and indexed with samtools v1.9. Mapping statistics, average sequencing depth and genome coverage at ≥ 1×, ≥ 5× and ≥ 10× were calculated from the sorted BAM files; most samples showed high mapping rates, with mapping rates ranging from 87.28% to 99.20%. Genome coverage at ≥ 10× was higher in the F_2 bulks than in the parental samples (Table S3). SNPs and small InDels were called using GATK v3.8 HaplotypeCaller. Variants within 5 bp of an InDel or within 10 bp of adjacent InDels (bcftools vcfutils.pl varFilter -w 5 -W 10), clustered variants (clusterSize = 2, clusterWindowSize = 5), and those failing QUAL < 30, QD < 2.0, MQ < 40 or FS > 60.0 were removed, with other parameters kept at GATK defaults (Table S4).

Based on the mapping results, structural variants (insertions, deletions, inversions and ectopic rearrangements) were detected using Manta with default parameters. For BSA-seq analysis, SNP-index values were calculated for the high- and low-Cd accumulation bulks. The ΔSNP-index was calculated by subtracting the SNP-index of the low-accumulation bulk from that of the high-accumulation bulk. Genome-wide ΔSNP-index values were examined using a sliding-window approach, and candidate regions were defined as intervals exceeding the 99th percentile threshold. The coding genes in the candidate region were functionally annotated using multiple databases: NR (NCBI non-redundant protein sequences) [30], SwissProt (A manually annotated and reviewed protein sequence database) [30], GO (Gene Ontology) [31], KO (KEGG Ortholog database) [32], COG (Clusters of Orthologous Groups of protein) [33] using BLAST [34] software.

2.7 Integration of BSA-seq and RNA-seq data for candidate-gene prioritization

BSA-seq was used to define genomic intervals associated with segregation for whole-plant Cd accumulation, and time-course RNA-seq in the tolerant parent was used as supportive functional evidence for candidate gene prioritization. Genes located within BSA-seq candidate intervals were extracted from the reference genome annotation. Candidate genes were first prioritized by retaining genes that (i) were located within the BSA-seq intervals, and (ii) carried at least one predicted functional variant (non-synonymous SNP or frameshift InDel) between high accumulation and low accumulation materials and/or had GO/KEGG annotation related to glutathione metabolism and transporter-associated pathways.

Next, RNA-seq expression data (Section 2.4) were then used to refine candidates based on Cd-responsive transcriptional regulation. Differential expression between Cd-treated and CK samples was assessed using DESeq2 ($|\log_2 FC| \geq 1$, FDR ≤ 0.05). Genes were retained as integrated candidates if they were located within the BSA-seq intervals and were significantly differentially expressed in at least one relevant comparison, with expression patterns consistent with Cd response in the tolerant parent. FPKM values were used to summarize expression levels and visualize temporal expression profiles.

2.8 qRT-PCR analysis

Gene-specific qPCR primers for the seven candidate genes and the reference gene were designed using PrimerQuest (Integrated DNA Technologies) based on the *C. olitorius* genome and transcript sequences (Table S5). The *C. olitorius* PP2A catalytic subunit (CoPP2A/PP2Ac) was used as the internal reference gene because of its stable expression reported in jute [35]. For each time-course × treatment × tissue combination, three biological replicates were analyzed, and each biological sample was run with two technical qPCR replicates.

The RNA samples used for qRT-PCR were the same as those used for transcriptome sequencing and cDNA was synthesized using the full-strand gold reverse transcription kit (AT311, China). For each reaction, 1 μg of total RNA was reverse transcribed using Superscript II Reverse Transcriptase (Cat. no. 18064-014, Invitrogen, Thermo Fisher Scientific, USA) and oligo $(dT)_{12-18}$ primers (Invitrogen, Thermo Fisher Scientific, USA), following the manufacturers' instructions. qRT-PCR was performed on a CFX96 real-time PCR detection system (Bio-Rad Laboratories, Hercules, CA, USA) using SYBR Green chemistry. Relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method [36].

2.9 Statistical analysis

All phenotypic measurements were obtained from three biological replicates unless otherwise stated. Statistical analyses were performed using GraphPad Prism v10.1.1. Differences between treatments or accessions were assessed using two-tailed Student's *t*-tests. Statistical significance was defined as $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***). Graphs were generated using GraphPad prism, and TBtools.

3. Results

3.1 Parental responses to Cd concentration gradient

Both parental lines exhibited dose-dependent growth suppression under Cd stress, reflected by decreasing TI and increasing GGI as Cd concentration increased. At 30 μmol/L, inhibition was mild and no significant difference was observed between accessions. With increasing Cd concentration, P2 showed a stronger loss of tolerance, with a pronounced decline in TI (Table S6) and increased GGI (Fig. 1). In contrast, P1 maintained higher tolerance across moderate to high Cd levels, showing consistently higher TI and lower GGI. Because 70 μmol/L produced clear phenotypic differentiation between the parents while avoiding the most severe inhibition observed at 100 μmol/L, it was selected as the working concentration for subsequent stress treatment and F_2 population screening. Cd exposure

also induced distinct visual symptoms in the two parents (Fig. S1). At 30 μmol/L, both lines displayed leaf chlorosis relative to CK. In P1, chlorosis progressively intensified with increasing Cd and was accompanied by gradual senescence and abscission of older leaves. In P2, red venation in mature leaves was evident and intensified with Cd concentration, together with stronger chlorosis and more extensive leaf loss. At 70 and 100 μmol/L, symptoms severity increased in both lines, but P2 consistently showed more pronounced discolouration and defoliation than P1.

GGI were calculated per biological replicate using CK biomass (Bc) and Cd-treated biomass (Bt). Bars represent mean ± SEM. Asterisks indicate significant differences between accessions at each concentration by two-tailed Student's *t*-test: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***).

3.2 Cd accumulation differed between parental lines

Following 7 d Cd exposure, whole-plant Cd concentration (mg/kg DW) increased markedly in both parental lines compared with CK (Fig. 2). In P1, Cd concentration increased from 4.04 mg/kg DW to 328.05 mg/kg DW in Cd treatment, while in P2 it increased from 5.30 mg/kg DW in CK to 584.20 mg/kg DW in Cd treatment. Thus, Cd exposure resulted in ≈ 81-fold and 110-fold increases in P1 and P2, respectively. Under Cd treatment, P2 accumulated significantly more Cd than P1 (1.78-fold higher), confirming clear parental divergence in whole-plant Cd accumulation capacity.

3.3 Segregation in the F_2 population under Cd stress

At 70 μmol/L Cd, a total of 300 F_2 seedlings were screened for whole-plant Cd accumulation, and valid ICP-OES measurements were obtained for 296 individuals (Table S7). Whole-plant Cd concentration (mg/kg DW) showed a continuous, unimodal distribution spanning a broad range (approximately 25.5-853.9 mg/kg DW), with a central tendency of μ = 338 mg/kg DW and dispersion of σ = 149 mg/kg DW (coefficient of variation ≈ 44%). The fitted normal density curve closely tracked the histogram across the central bins, indicating that Cd accumulation behaves as a quantitative trait rather than a discrete, single-gene class. Notably, the distribution exhibited a slight right tail, with a small number of individuals showing markedly elevated Cd concentrations (> 600 mg/kg DW), suggesting transgressive segregation for high accumulation and providing a clear basis for selecting extreme phenotypes.

Normality was formally assessed using the Shapiro-Wilk test ($W = 0.9812$, $p = 0.007095$). Although the p-value indicates a statistically detectable deviation from perfect normality (which is common for large sample sizes), the W statistic close to 1 and the overall histogram shape support an approximately normal, quantitative inheritance pattern for whole-plant Cd accumulation. The two parental lines differed in their positions within the F_2 distribution, with P1 close to the population mean and P2 shifted toward the higher-accumulation side (Fig. 3), consistent with segregating alleles contributing to variation in Cd accumulation.

3.4 Differentially expressed genes in response to Cd stress

To characterize Cd-responsive transcriptional dynamics, time-course RNA-seq was performed in P1 using root (R) and aboveground tissue (L) samples collected at 24, 48, and 72 h after Cd exposure. The 0 h CK samples served as baseline controls for differential expression comparisons with Cd-treated samples at each subsequent time point. Pearson correlation analysis showed strong consistency among biological replicates and clear separation by tissue and treatment/time (Fig. 4).

Differential expression was assessed using DESeq2 (padj ≤ 0.05, $|\log_2 FC| \geq 1$). In roots, 3784 DEGs (1923 up, 1861 down) were detected at 24 h, 4167 (2001 up, 2166 down) at 48 h, and 4102 (1884 up, 2218 down) at 72 h (Table S8). Overall, Cd triggered a consistently large transcriptional response in roots across the time-course, whereas aboveground tissues showed a more dynamic pattern, with a reduction in DEG number at 48 h followed by a marked increase by 72 h.

Overlap analysis indicated both shared and tissue-specific Cd responses. Across tissues, 1207 DEGs were shared between roots and aboveground tissues, while 2623 and 3318 were tissue-specific to roots and aboveground tissues, respectively (Fig. 5A). Time-course Venn analysis further showed that only a small core set of DEGs was shared across consecutive time transitions under Cd stress (roots: 10 shared; aboveground tissues: 34 shared), indicating that most transcriptional changes were stage-specific during the progression of Cd stress (Fig. 5B-C).

(A) Shared and tissue-specific DEGs between root (R) and aboveground tissues (L), (B) Time-course overlap of DEGs in roots (R), and (C) Time-course overlap of DEGs in aboveground tissues (L)

3.5 Functional enrichment analysis of Cd-responsive DEGs

To interpret the biological functions associated with Cd-responsive transcriptional changes, GO and KEGG enrichment analyses were performed separately for aboveground tissues (L) and roots (R) using the union of DEGs identified at 24, 48, and 72 h (Fig. 6; Table S8).

GO enrichment patterns differed markedly between tissues (Fig. 6A-B). In aboveground tissues, enriched GO terms were dominated by chloroplast and photosynthesis-associated cellular component (e.g., plastid/chloroplast, stroma, thylakoid, and photosynthetic membrane), indicating that Cd stress strongly impacted plastid-associated structures and photosynthetic machinery. In contrast, root DEGs showed broader enrichment spanning biological process, cellular component, and molecular function categories, including cell wall organization/biogenesis and secondary cell wall formation, cytoskeleton and microtubule-associated terms, and redox-related processes, together with oxidoreductase activity (including activity acting on metal ions), consistent with structural remodeling and redox regulation in the primary Cd-exposed tissue.

KEGG analysis further supported both shared and tissue-specific responses (Fig. 6C-D; Table S9). Both tissues were enriched for glutathione metabolism and transporters, consistent with detoxification and membrane transport during Cd stress. Root-enriched pathways additionally emphasized transporter-associated related categories, whereas aboveground tissues showed strong enrichment for photosynthesis-related pathways (e.g., photosynthesis and photosynthesis-antenna proteins). Together,

these enrichment patterns suggest a shared detoxification/transport core response to Cd stress, with tissue-specific specialization characterized by root structural/redox responses and aboveground impacts on plastid and photosynthesis-associated functions.

(A) Top 20 enriched GO terms in aboveground tissues (L). (B) Top 20 enriched GO terms in roots (R). (C) Top 20 enriched KEGG pathways in aboveground tissues (L). (D) Top 20 enriched KEGG pathways in roots (R). DEGs were identified by DESeq2 (padj ≤ 0.05, $|\log_2 FC| \geq 1$) at 24, 48, and 72 h after Cd exposure, and DEGs from each tissue were combined as a union set for enrichment analysis. Bars represent $-\log_{10}$ (P-value).

3.6 Identification of candidate intervals associated with whole-plant Cd accumulation by BSA-seq

To map genomic regions associated with whole-plant Cd accumulation, BSA-seq was performed on the two parental lines (P1 and P2) and two F_2 DNA bulks representing extreme phenotypes (high-accumulation, and low accumulation). Pairwise SNP differences among parents and bulks showed clear genetic divergence between the extreme pools (189967 SNPs) and between each bulks and its corresponding parent (Table S10). After filtering, 625354 high quality SNPs were retained for downstream SNP-index analysis (Table S11). Genome-wide ΔSNP-index analysis identified discrete peaks exceeding the 99th percentile threshold (ΔSNP-index cutoff = 0.02), defining five candidate intervals across three chromosomes (Fig. 7; Table S11). These intervals included three regions on Chromosome 3 (GWHBQTL00000003), one region on Chromosome 4 (GWHBQTL00000004), and one region on Chromosome 7 (GWHBQTL00000007), spanning a total of 7.13 Mb and containing 662 annotated genes in the reference genome (Table S11). These candidate intervals were used for subsequent candidate gene prioritization and integration with transcriptomic data.

3.7 Integrative prioritization of candidate genes using BSA-seq and time-course RNA-seq

BSA-seq (ΔSNP-index) identified five candidate intervals associated with whole-plant Cd accumulation across Chromosomes 3, 4, and 7, spanning 7.13 Mb and containing 662 annotated genes (Table S11; Fig. 7). To prioritize candidates within these intervals, we integrated positional evidence from BSA-seq with time-course RNA-seq profiles under Cd treatment.

In a first filtering step, genes located within the BSA intervals were shortlisted based on functional annotations consistent with HM stress responses (e.g., transport, detoxification, stress signaling, and metal binding) and their proximity to association peaks. In a second step, RNA-seq data were used to refine the shortlist by evaluating differential expression between Cd-treated and CK samples across 0, 24, 48 and 72 h in roots (R) and aboveground tissues (L). Genes were retained as key candidates if they (i) showed significant Cd-responsive expression in at least one relevant comparison and/or (ii) exhibited clear time-dependent expression dynamics consistent with Cd response, particularly in roots. Using

these combined criteria, seven genes were prioritized as candidate genes associated with Cd accumulation and tolerance: COS03g_00658, COS03g_00659, COS03g_03287, COS03g_03316, COS03g_03200, COS03g_03180, and COS07g_05235 (Table S12; Fig. 8).

Time-course profiles supported tissue-dependent Cd responses among these candidates (Fig. 8). COS03g_00658 and COS03g_00659 showed consistently higher expression in roots than aboveground tissues across the Cd exposure period, consistent with a root-associated response. COS03g_03316 also displayed higher expression in roots and dynamic changes following Cd exposure. In contrast, COS07g_05235 showed comparatively modest temporal variation between tissues. Collectively, the combined positional mapping and Cd-responsive expression evidence supports these genes as leading candidates for further validation.

To provide additional functional context, conserved protein features were examined using domain-based annotations. Domain architectures supported the presence of motifs and families consistent with stress-response and transport-related functions, including PLAC8 domains in COS03g_00658 and COS03g_00659, ABC transporter-associated features in COS03g_03287 and COS03g_03316, and a NAM/NAC-related domain in COS03g_03200 (Fig. S4). Collectively, the combined mapping-by-sequencing and time-resolved expression evidence highlights these candidates as strong targets for downstream validation and functional characterization.

Expression levels (FPKM) are shown for Cd-treated samples at 24, 48, and 72 h, with 0 h CK samples used as baseline controls. Points represent mean values across biological replicates and error bars indicate ± SE.

3.8 qRT-PCR validation of candidate gene expression

qRT-PCR was used to validate expression patterns for the prioritized candidate genes across the Cd exposure time course. Overall, qRT-PCR trends were consistent with the RNA-seq expression directionality for the tested genes, supporting the reliability of the transcriptomic dataset (Fig. 9A-G). Notably, COS03g_00658 showed stronger induction in P1 than in P2 at 24 h and 72 h, whereas COS03g_00659 displayed higher expression in P2 across Cd-stress time points. The ABC transporter candidate COS03g_03316 showed consistently higher expression in P1 than in P2 across the time course. Other candidates showed weaker or more variable differences between accessions.

A-G represent the seven candidate genes, COS03g_03287, COS03g_03316, COS03g_03200, COS07g_05235, COS03g_03180, COS03g_00658, and COS03g_00659 respectively. Values are mean ± SE. Significant differences were assessed by t-test: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***).

4. Discussion

This study used an integrative genomic and transcriptomic approach to investigate whole-plant Cd accumulation and Cd tolerance-associated responses in dark jute. The continuous distribution of Cd concentration in the F_2 population (Fig. 3) suggests that Cd accumulation is a quantitative trait governed

by multiple loci rather than a single major gene, consistent with reports in other crop species [8, 37]. The identification of five candidate intervals spanning 7.13 Mb across Chromosomes 3, 4, and 7 and containing 662 annotated genes (Table S11) underscores the complex architecture of Cd accumulation and provides a defined genomic framework for downstream functional analysis.

Tissue-resolved transcriptome profiling revealed marked specialization of Cd-responsive programs. In shoots, DEGs were primarily enriched in plastid and chloroplast-associated components, suggesting that aerial tissues largely reflect downstream physiological impacts of Cd stress on photosynthesis and membrane metabolism (Fig. 6A). Cd is well known to disrupt chloroplast ultrastructure, impair thylakoid function, and induce oxidative damage in leaves [4, 38]. Thus, the transcriptional signatures observed in shoots are consistent with structural and metabolic perturbations secondary to Cd exposure.

In contrast, root DEGs were enriched across oxidoreductase activity, membrane transport, cellular recycling, and stress-response categories (Fig. 6B). Because roots represent the primary site of Cd entry, this broader transcriptional activation likely reflects active detoxification and defense. Enhanced oxidoreductase activity and redox regulation are central to limiting metal-induced oxidative stress [39, 40]. The enrichment of glutathione-related pathways across tissues further highlights the importance of thiol-mediated detoxification. Glutathione plays a critical role in Cd chelation, ROS buffering, and the formation of Cd-phytochelatin complexes that enable vacuolar sequestration [41, 42].

Transporter-associated pathways were also significantly enriched (Fig. 6D). Membrane transporters, particularly members of the ATP-binding cassette (ABC) family, are key regulators of HM efflux and compartmentalization [37, 43, 44]. In rice, ABCG36 mediates Cd extrusion from root cells [45], while ABCC-type transporters contribute to vacuolar sequestration of Cd-glutathione complexes in crop species [46]. Within our candidate intervals, COS03g_03287 and COS03g_03316 encode ABC transporter–like proteins. The higher root expression of COS03g_03316 (Fig. 9B) suggests that it may be involved in root Cd transport, redistribution, or sequestration, although functional validation will be required to confirm its role.

Two additional candidates, COS03g_00658 and COS03g_00659, contain PLAC8 motifs characteristic of PCR family proteins (Fig. S4). PCR/PLAC8 proteins have been shown to regulate Cd tolerance and accumulation in several species [47]. Overexpression of *OsPCR1* and *OsPCR3* enhances Cd tolerance and reduces shoot Cd accumulation in rice [48], while *AtPCR2* contributes to Cd resistance in *Arabidopsis thaliana* [49]. The root-biased expression of COS03g_00658 and COS03g_00659 under Cd stress, together with their contrasting regulation between tolerant and sensitive parents, suggests functional divergence between the two PLAC8-domain candidates and support their prioritization for future functional analysis. Notably, COS03g_00658 showed stronger induction in P1, whereas COS03g_00659 was more highly expressed in P2 (Fig. 9F-G).

The NAC transcription factor-like gene COS03g_03200 exhibited relatively high expression levels. NAC family proteins are well-established regulators of abiotic stress responses, coordinating antioxidant

defenses, cell wall modification, and downstream detoxification pathways [50, 51]. Its expression profile suggests that it may function upstream of transporter and detoxification genes in Cd stress signaling.

COS03g_03180, predicted to function extracellularly, displayed higher expression in roots (Fig. 9E). Extracellular or cell wall-associated proteins can contribute to Cd immobilization in the apoplast, thereby limiting symplastic entry [52]. This suggests that COS03g_03180 may contribute to Cd immobilization at the root interface, thereby affecting whole-plant Cd accumulation. In contrast, COS07g_05235 showed no strong tissue bias (Fig. 9D) and may function in general stress signaling rather than Cd-specific detoxification.

Root qRT-PCR validation supported several RNA-seq trends, particularly for COS03g_00658, COS03g_00659, and COS03g_03316. The concordance between transcriptome data and targeted validation strengthens confidence in these genes as core candidates. However, not all genes within the BSA intervals showed strong expression differences, which is expected for quantitative traits such as Cd accumulation. Candidate intervals often contain genes with modest or context-dependent effects that may not exhibit pronounced transcriptional variation under a single experimental condition.

Although this study provides candidate genomic intervals and Cd-responsive genes associated with whole-plant Cd accumulation and tolerance in dark jute, several limitations should be noted. First, Cd accumulation was evaluated as whole-plant Cd concentration, whereas root- and shoot-specific Cd accumulation, translocation factor, bioconcentration factor, and total Cd uptake were not separately assessed. Therefore, the present results should be interpreted as an integrated estimate of whole-plant Cd accumulation rather than a complete description of Cd uptake, transport, and partitioning among organs. Second, the population screening was conducted under hydroponic conditions using a defined Cd concentration. This approach enabled controlled phenotyping and clear differentiation among accessions, but it may not fully represent Cd availability, plant–soil interactions, or environmental variation under field conditions. Validation under Cd-contaminated soil and multi-environment conditions will therefore be necessary. Third, time-course RNA-seq was performed in the tolerant, low-accumulating parent, whereas transcriptome profiling of the sensitive, high-accumulating parent was not included. As a result, the RNA-seq data provide evidence for Cd-responsive expression in the tolerant parent but cannot fully resolve accession-specific transcriptional differences between the two parents. Finally, the prioritized genes were identified based on positional evidence, Cd-responsive expression patterns, functional annotation, and qRT-PCR validation, but their biological functions remain to be experimentally confirmed. Future studies using fine mapping, gene editing, overexpression or complementation analysis, subcellular localization, transporter assays, and field-based phenotyping will be important to verify the roles of these candidate genes in Cd accumulation and tolerance.

Taken together, the integration of BSA-seq and time-course RNA-seq supports a root-centered model of Cd response in dark jute. Roots appear to activate coordinated detoxification, glutathione-linked redox buffering, membrane transport, and possibly apoplastic immobilization, whereas shoots mainly reflect structural and photosynthetic consequences of Cd stress. Among the prioritized genes, COS03g_00658,

COS03g_00659, and COS03g_03316 represent particularly promising candidates for further study because they combine positional evidence, Cd-responsive expression, and functional annotations related to HM tolerance or transport. These findings provide a useful genetic and transcriptomic framework for future functional validation, marker development, and improvement of Cd-related traits in dark jute.

5. Conclusion

This study integrated BSA-seq, time-course RNA-seq, Cd phenotyping, and qRT-PCR validation to investigate whole-plant Cd accumulation and tolerance-associated responses in dark jute. Using two contrasting parental accessions and an F_2 population screened under Cd stress, BSA-seq identified five candidate genomic intervals on Chromosomes 3, 4, and 7, spanning 7.13 Mb and containing 662 annotated genes.

Time-course RNA-seq of the tolerant parent revealed tissue-specific transcriptional responses to Cd stress. Aboveground tissues were mainly associated with plastid- and photosynthesis-related processes, whereas roots showed broader responses involving redox regulation, cell wall and cytoskeleton remodeling, transport, and stress defense. By combining mapped intervals with Cd-responsive expression patterns and functional annotations, seven candidate genes were prioritized, including PLAC8-domain/PCR-family genes, ABC transporter-like genes, and stress-response-related genes.

Overall, these findings provide candidate genomic regions and prioritized genes potentially associated with whole-plant Cd accumulation and tolerance in dark jute. This study establishes a useful genetic and transcriptomic foundation for future fine mapping, functional validation, marker development, and improvement of Cd-related traits in this high-biomass fiber crop.

Abbreviations

ABC
ATP-binding cassette
BSA
bulked segregant analysis
BSA-seq
bulked segregant analysis sequencing
Cd
cadmium
CK
control
COG
Clusters of Orthologous Groups
CTAB
cetyltrimethylammonium bromide

DEG
differentially expressed gene
DNA
deoxyribonucleic acid
DW
dry weight
FDR
false discovery rate
FPKM
fragments per kilobase of transcript per million mapped reads
FW
fresh weight
GGI
growth inhibition rate
GO
Gene Ontology
HM
heavy metal
ICP-OES
inductively coupled plasma optical emission spectroscopy
InDel
insertion/deletion
KEGG
Kyoto Encyclopedia of Genes and Genomes
KO
KEGG Orthology
NAC
NAM, ATAF, and CUC transcription factor family
NR
NCBI non-redundant protein database
PCR
plant cadmium resistance
PLAC8
placenta-specific 8
qRT-PCR
quantitative real-time polymerase chain reaction
RNA
ribonucleic acid
RNA-seq
RNA sequencing

ROS
reactive oxygen species
SEM
standard error of the mean
SNP
single-nucleotide polymorphism
TI
tolerance index
TPM
transcripts per million

Declarations

Ethics approval and consent to participate

The plant materials used in this study were cultivated germplasm accessions obtained from the germplasm collection maintained by the Key Laboratory of Genetics, Breeding and Multiple Utilization of Bast Fiber Crops, Fujian Agriculture and Forestry University (FAFU), Fuzhou, China. The materials were not collected from the wild; therefore, no specific permission or license was required for sample collection. The plant materials were authenticated as *Corchorus olitorius* L. by Dr. Xu Jian-tang, FAFU, based on institutional germplasm records and morphological characteristics. The experimental research on plants complied with relevant institutional, national, and international guidelines and legislation. No ethics approval or consent to participate was required.

Consent for publication

Not applicable.

Availability of data and materials

The sequencing datasets generated and/or analyzed during the current study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1465950. All other data supporting the conclusions of this article are included within the article and its supplementary files.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This study was supported by the National Agri-Industry Technology Research System for Crops of Bast and Leaf Fiber, China (nycytx-19-E06), the Fujian Natural Science Foundation (2022J01584), the Weng

Hongwu Original Scientific Research Foundation, Peking University, and the Chinese Government Scholarship.

Authors' contributions

XJ conceived and designed the study, acquired funding, and supervised the research. ZSX performed the methodology, investigation, data curation, and data analysis. VOA contributed to the methodology, data analysis, validation, and preparation of the original draft. ZY, LLJ, LH, and ZWQ contributed to the methodology, validation, and review and editing of the manuscript. TA, FP, QJ, and ZSB provided resources and research support. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank all individuals and institutions who provided support during this study.

References

1. Kubier A, Wilkin RT, Pichler T. Cadmium in soils and groundwater: A review. Appl Geochem. 2019;108:1–16.
2. Liu X et al. Cadmium pollution from phosphate fertilizers in arable soils and crops: an overview. Front Agricultural Sci Eng, 2019. 6(4).
3. Clemens S, Ma JF. Toxic Heavy Metal and Metalloid Accumulation in Crop Plants and Foods. Annu Rev Plant Biol. 2016;67:489–512.
4. Clemens S. Toxic metal accumulation, responses to exposure and mechanisms of tolerance in plants. Biochimie. 2006;88(11):1707–19.
5. Balestrasse KB, et al. Effect of cadmium stress on nitrogen metabolism in nodules and roots of soybean plants. Funct Plant Biol. 2003;30(1):57–64.
6. Krishnamurti GSR, et al. Kinetics of Cadmium Release from Soils as Influenced by Organic Acids: Implication in Cadmium Availability. J Environ Qual. 1997;26(1):271–7.
7. Feller U, Anders I, Wei S. Distribution and Redistribution of (109)Cd and (65)Zn in the Heavy Metal Hyperaccumulator Solanum nigrum L.: Influence of Cadmium and Zinc Concentrations in the Root Medium. Plants (Basel), 2019. 8(9).
8. Wang H et al. Mapping and Identifying Candidate Genes Enabling Cadmium Accumulation in Brassica napus Revealed by Combined BSA-Seq and RNA-Seq Analysis. Int J Mol Sci, 2023. 24(12).
9. Ali H, Khan E, Sajad MA. Phytoremediation of heavy metals–concepts and applications. Chemosphere. 2013;91(7):869–81.
10. Shehata SM, Badawy RK, Aboulsoud YIE. Phytoremediation of some heavy metals in contaminated soil. Bull Natl Res Centre, 2019. 43(1).
11. Ali SY, Chaudhury S. EDTA-Enhanced Phytoextraction by Tagetes sp. and Effect on Bioconcentration and Translocation of Heavy Metals. Environ Processes. 2016;3(4):735–46.

12. Mahar A, et al. Challenges and opportunities in the phytoremediation of heavy metals contaminated soils: A review. Ecotoxicol Environ Saf. 2016;126:111–21.
13. Zhang L, et al. Reference genomes of the two cultivated jute species. Plant Biotechnol J. 2021;19(11):2235–48.
14. Yang Z, et al. Multi-omics provides new insights into the domestication and improvement of dark jute (Corchorus olitorius). Plant J. 2022;112(3):812–29.
15. Yang Z, et al. Machine learning phenotyping and GWAS reveal genetic basis of Cd tolerance and absorption in jute. Environ Pollut. 2024;362:124918.
16. Akintayo VO, et al. Mapping and identifying candidate genes enabling lead (Pb) accumulation and tolerance in dark jute (Corchorus olitorius L.) using BSA-seq and RNA-seq. Industrial Crops and Products; 2026. p. 239.
17. Michelmore RW, Paran I, Kesseli RV. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci U S A. 1991;88(21):9828–32.
18. Takagi H, et al. QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant J. 2013;74(1):174–83.
19. Guo Z et al. *Identification of candidate genes controlling chilling tolerance of rice in the cold region at the booting stage by BSA-Seq and RNA-Seq.* R Soc Open Sci, 2020. 7(11): p. 201081.
20. Zhou H et al. Combined BSA-Seq and RNA-Seq Analysis to Identify Candidate Genes Associated with Aluminum Toxicity in Rapeseed (Brassica napus L). Int J Mol Sci, 2024. 25(20).
21. Saleem MH et al. Jute: A Potential Candidate for Phytoremediation of Metals-A Review. Plants (Basel), 2020. 9(2).
22. Zulfiqar U, et al. Cadmium Phytotoxicity, Tolerance, and Advanced Remediation Approaches in Agricultural Soils; A Comprehensive Review. Front Plant Sci. 2022;13:773815.
23. Bian JL, et al. Tolerance Mechanism and Cadmium Enrichment Abilities in Two Brassica napus L. Cultivars. Huan Jing Ke Xue. 2020;41(2):970–8.
24. Nath K, Saini S, Sharma YK. Chromium in tannery industry effluent and its effect on plant metabolism and growth. J Environ Biol. 2005;26(2):197–204.
25. Xiao-fei MENG, Jun-xing GJ-mYANG, Jian HU, Guo-di ZHENG. BIAN Jian-lin, LI Yu-feng, CHEN Tong-bin1, LIU Jie, *Mechanism of the two cultivars of rapes with different Pb enrichment ability*. China Environ Sci. 2020;40(10):4479–.
26. Ma J, et al. Impact of foliar application of syringic acid on tomato (Solanum lycopersicum L.) under heavy metal stress-insights into nutrient uptake, redox homeostasis, oxidative stress, and antioxidant defense. Front Plant Sci. 2022;13:950120.
27. Zhang ZH, et al. A multiomics approach reveals the pivotal role of subcellular reallocation in determining rapeseed resistance to cadmium toxicity. J Exp Bot. 2019;70(19):5437–55.

28. Camacho C, et al. ElasticBLAST: accelerating sequence search via cloud computing. BMC Bioinformatics. 2023;24(1):117.
29. Aboul-Maaty NA-F, Oraby HA-S. Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. Volume 43. Bulletin of the National Research Centre; 2019. 1.
30. Deng YY, Wu JLSF, Zhu YP, Zhu YDENGY, Chen YW, He FC, Li J, Maximilian Deng SWU, Li J-YQ, Wu S, Zhu Y, Chen Y, He F, Li J, wu S, Deng YY, Li JQ, Wu SF, Li QQ, Wu S, ZHU Y, Chen Y, HE F, Deng Y, WU S, Deng Y, Deng Y, Lily Y, Deng Y, He F, Deng TY, Zhu Deng Y, ZHU, Chris Deng AY, Deng X, Deng JT, Li Y-YM, Chen ZL, HE Y, Deng J, LI, Lawrence Deng YS, CHEN, ALAN Y. DENG, WU GX, He ML, Jiuqiang Q, Li SF, Wang DYY, WU SF, Zhu YP, Chen YW, He FC. Y.W. CHEN, Li Deng, F Yueqiu He, *Integrated nr database in protein annotation system and its localization.* Computer Engineering, 2006. 32(5): pp. 71–74.
31. Ashburner M, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25–9.
32. Kanehisa M, et al. The KEGG resource for deciphering the genome. Nucleic Acids Res. 2004;32(Database issue):D277–80.
33. Tatusov RL, et al. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 2000;28(1):33–6.
34. Altschul SF, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389–402.
35. Hossain MS, et al. Identification and validation of reference genes for real-time quantitative RT-PCR analysis in jute. BMC Mol Biol. 2019;20(1):13.
36. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402–8.
37. Zhe L, et al. Genome-wide association study of novel genetic loci for cadmium accumulation and germplasm screening for low-cadmium accumulation in common wheat (Triticum aestivum L). Journal of Integrative Agriculture; 2024.
38. DalCorso G, Manara A, Furini A. An overview of heavy metal challenge in plants: from roots to shoots. Metallomics, 2013. 5(9).
39. Gill SS, Tuteja N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol Biochem. 2010;48(12):909–30.
40. Terron-Camero LC, et al. Role of nitric oxide in plant responses to heavy metal stress: exogenous application versus endogenous production. J Exp Bot. 2019;70(17):4477–88.
41. Hasan MK, et al. Responses of Plant Proteins to Heavy Metal Stress-A Review. Front Plant Sci. 2017;8:1492.
42. Verbruggen N, Hermans C, Schat H. Molecular mechanisms of metal hyperaccumulation in plants. New Phytol. 2009;181(4):759–76.

43. He G, et al. Identification of novel heavy metal detoxification proteins in Solanum tuberosum: Insights to improve food security protection from metal ion stress. Sci Total Environ. 2021;779:146197.
44. Song WY, et al. Arsenic tolerance in Arabidopsis is mediated by two ABCC-type phytochelatin transporters. Proc Natl Acad Sci U S A. 2010;107(49):21187–92.
45. Fu S, et al. The ABC transporter ABCG36 is required for cadmium tolerance in rice. J Exp Bot. 2019;70(20):5909–18.
46. Song W-Y et al. *ABC Transporters and Heavy Metals*, in *Plant ABC Transporters*. 2014. pp. 1–17.
47. Guan J, et al. Plant cadmium resistance 10 enhances tolerance to toxic heavy metals in poplar. Plant Physiol Biochem. 2023;203:108043.
48. Wang F et al. A novel family of PLAC8 motif-containing/PCR genes mediates Cd tolerance and Cd accumulation in rice. Environ Sci Europe, 2019. 31(1).
49. Song WY, et al. A novel family of cys-rich membrane proteins mediates cadmium resistance in Arabidopsis. Plant Physiol. 2004;135(2):1027–39.
50. Nuruzzaman M, Sharoni AM, Kikuchi S. Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. Front Microbiol. 2013;4:248.
51. Zheng C et al. Transcription Factors Involved in Plant Stress and Growth and Development: NAC. Agronomy, 2025. 15(4).
52. Parrotta L, et al. Target or barrier? The cell wall of early- and later-diverging plants vs cadmium toxicity: differences in the response mechanisms. Front Plant Sci. 2015;6:133.

Figures

Image not available with this version

Figure 1

Growth inhibition (GGI) of the parental lines under Cd concentration gradient (30, 50, 70, and 100 μmol/L).

GGI were calculated per biological replicate using CK biomass (Bc) and Cd-treated biomass (Bt). Bars represent mean ± SEM. Asterisks indicate significant differences between accessions at each concentration by two-tailed Student's *t*-test: $p < 0.05$ (*), $p < 0.01$ **(**)**, $p < 0.001$ (***).

Image not available with this version

Figure 2

Whole-plant Cd concentration in parental lines under CK and Cd stress (70 μmol/L) after 7 d. Bars show mean ± SEM. Asterisks indicate significant differences (*P < 0.01*)

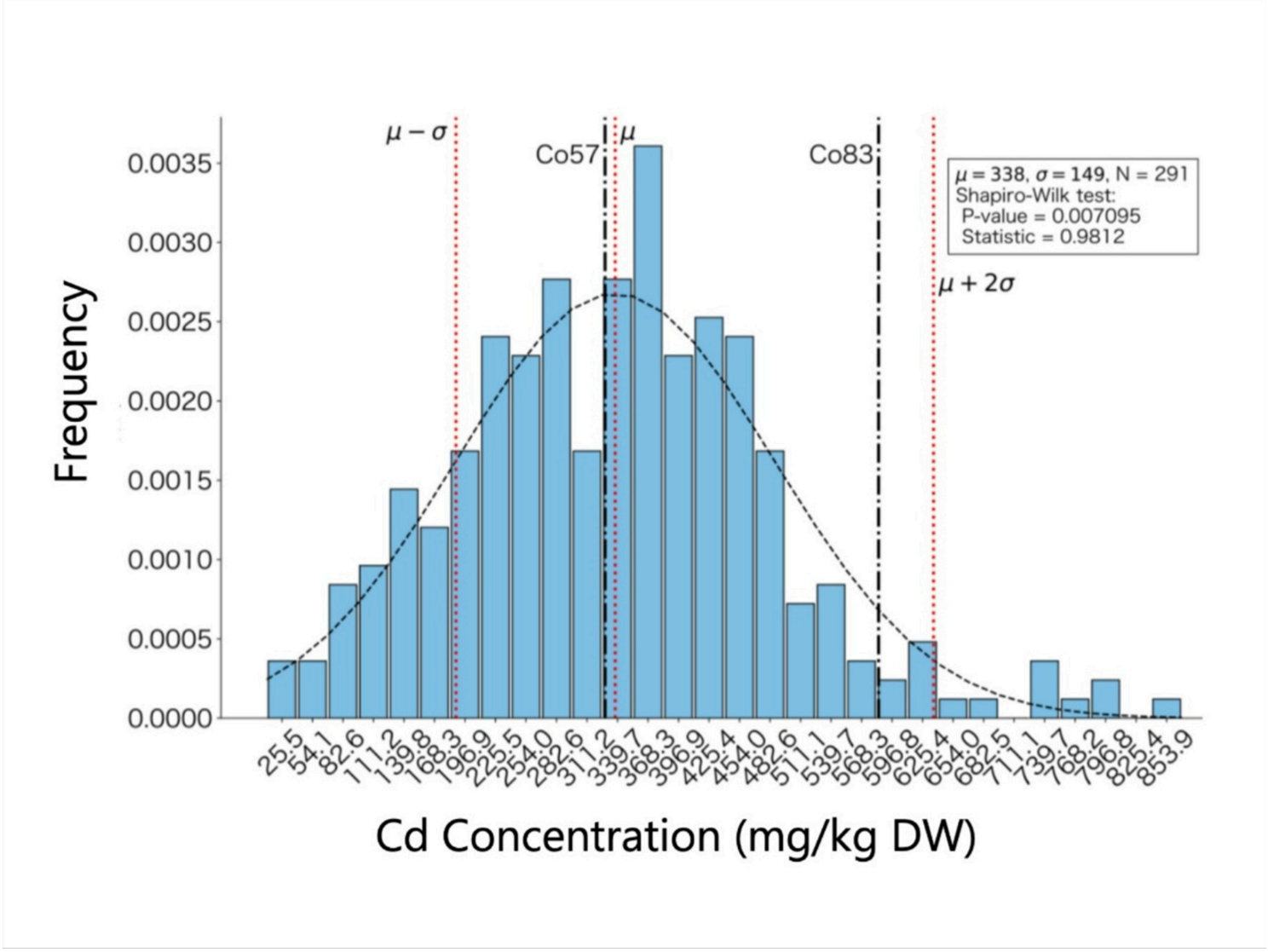

Figure 3

Frequency distribution of whole-plant Cd concentration in the F_2 population under 70 μmol/L Cd stress

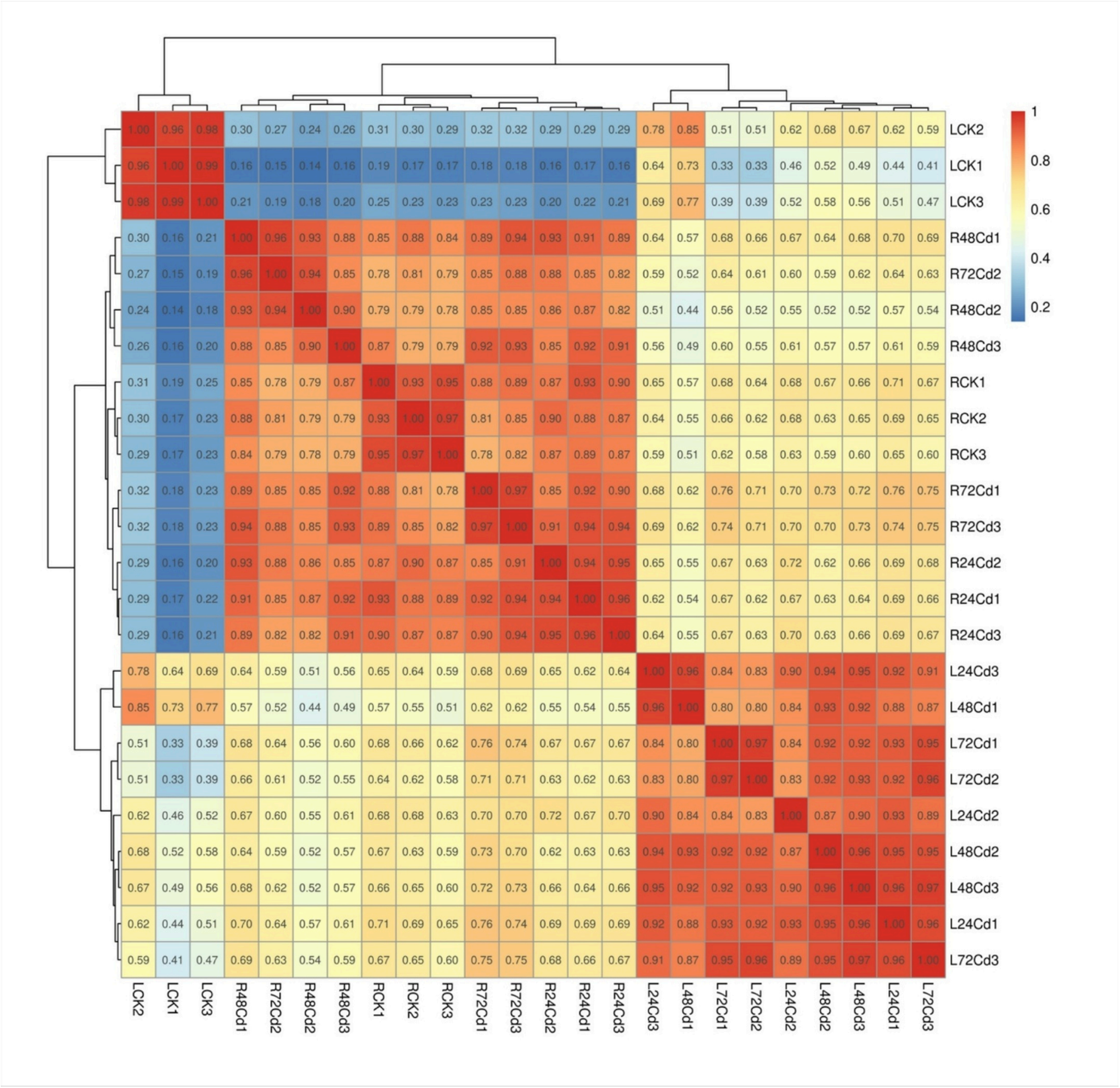

Figure 4

Pearson correlation heatmap of RNA-seq samples across tissues and time points under CK and Cd treatments

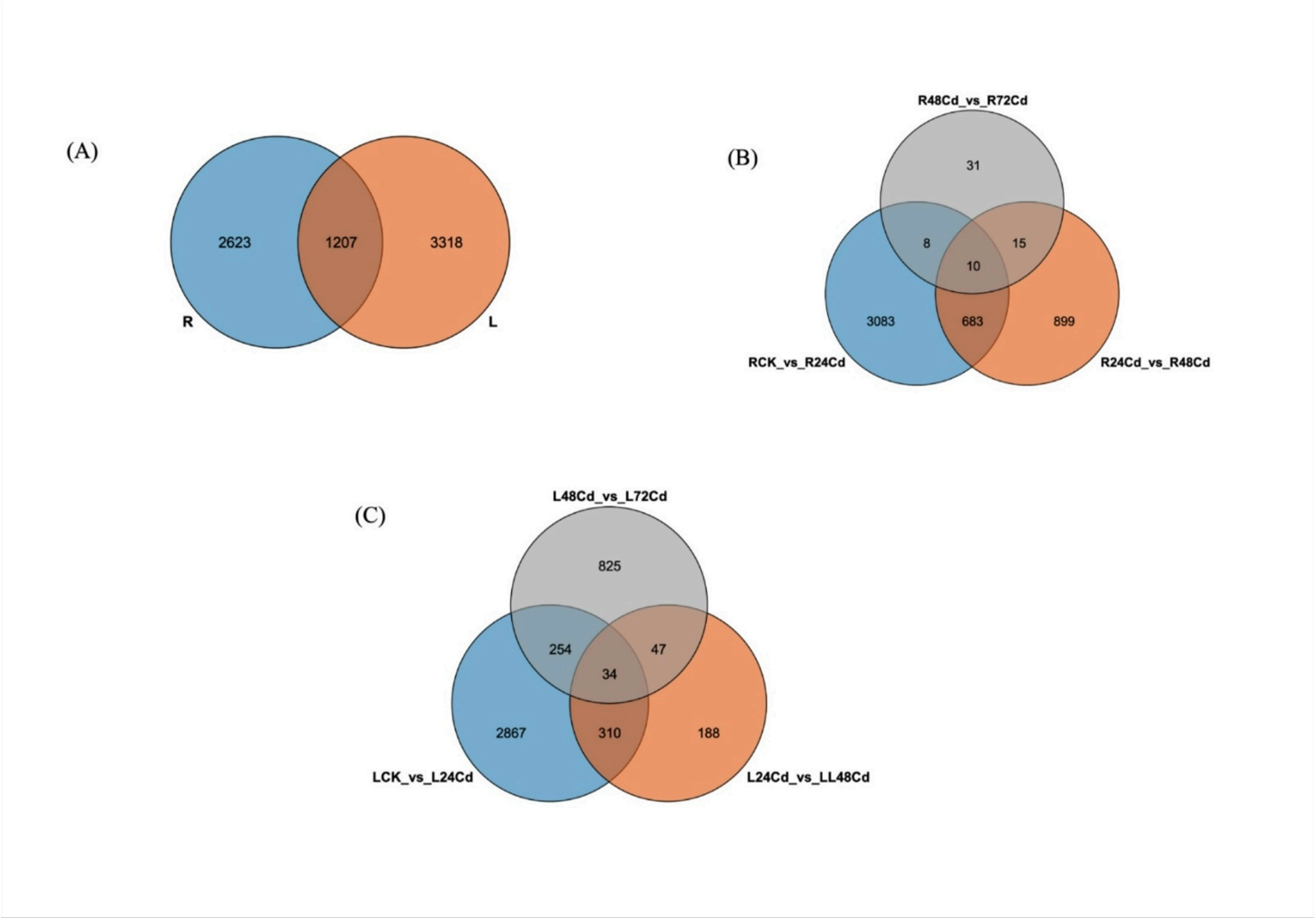

Figure 5

Overlap of Cd-responsive DEGs across tissues and time points

(A) Shared and tissue-specific DEGs between root (R) and aboveground tissues (L), (B) Time-course overlap of DEGs in roots (R), and (C) Time-course overlap of DEGs in aboveground tissues (L)

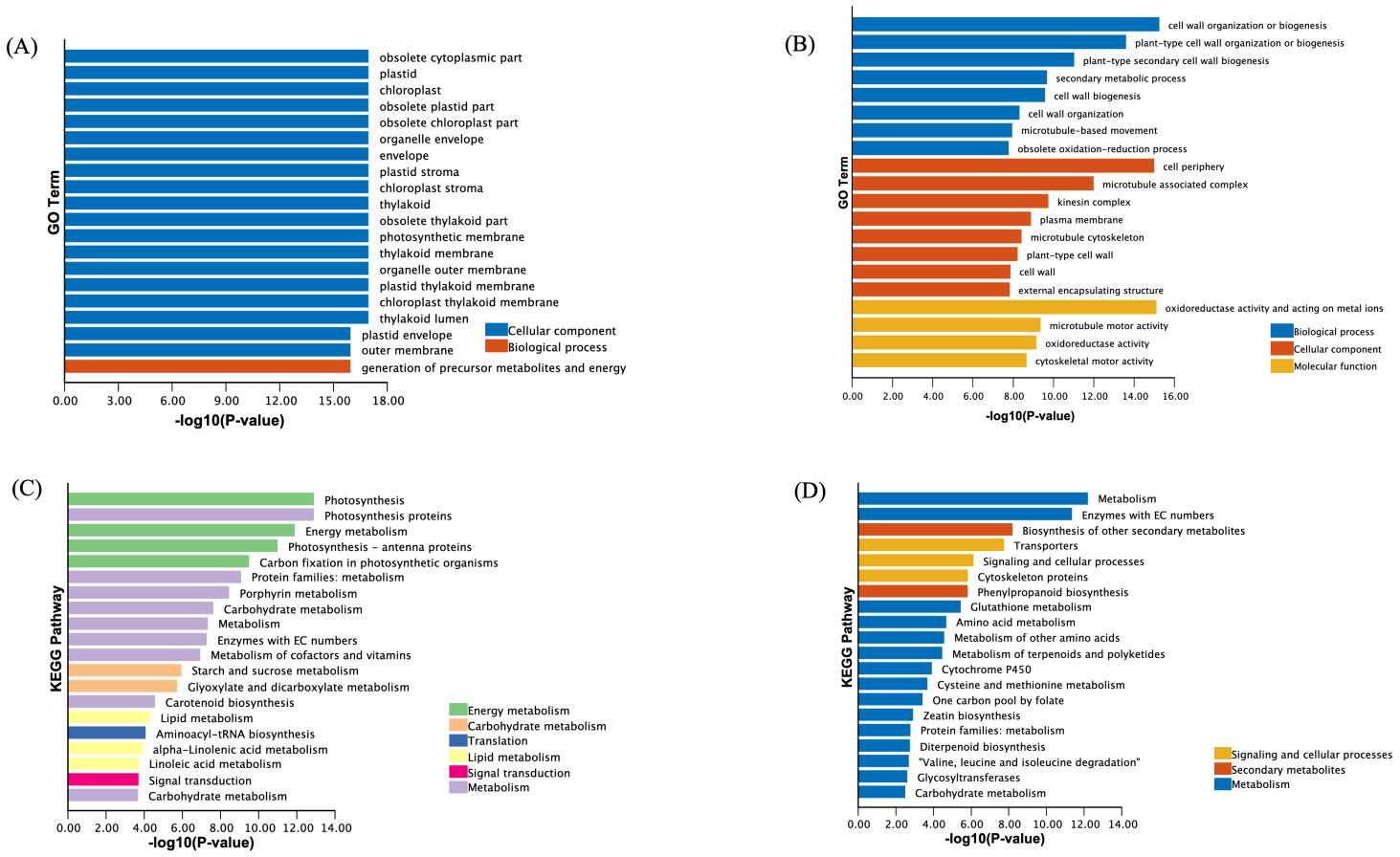

Figure 6

GO and KEGG enrichment of Cd-responsive DEGs in aboveground tissues (L) and roots (R) of the tolerant parent.

(A) Top 20 enriched GO terms in aboveground tissues (L). (B) Top 20 enriched GO terms in roots (R). (C) Top 20 enriched KEGG pathways in aboveground tissues (L). (D) Top 20 enriched KEGG pathways in roots (R). DEGs were identified by DESeq2 (padj ≤ 0.05, $|\log_2 FC| \geq 1$) at 24, 48, and 72 h after Cd exposure, and DEGs from each tissue were combined as a union set for enrichment analysis. Bars represent $-\log_{10}$ (P-value).

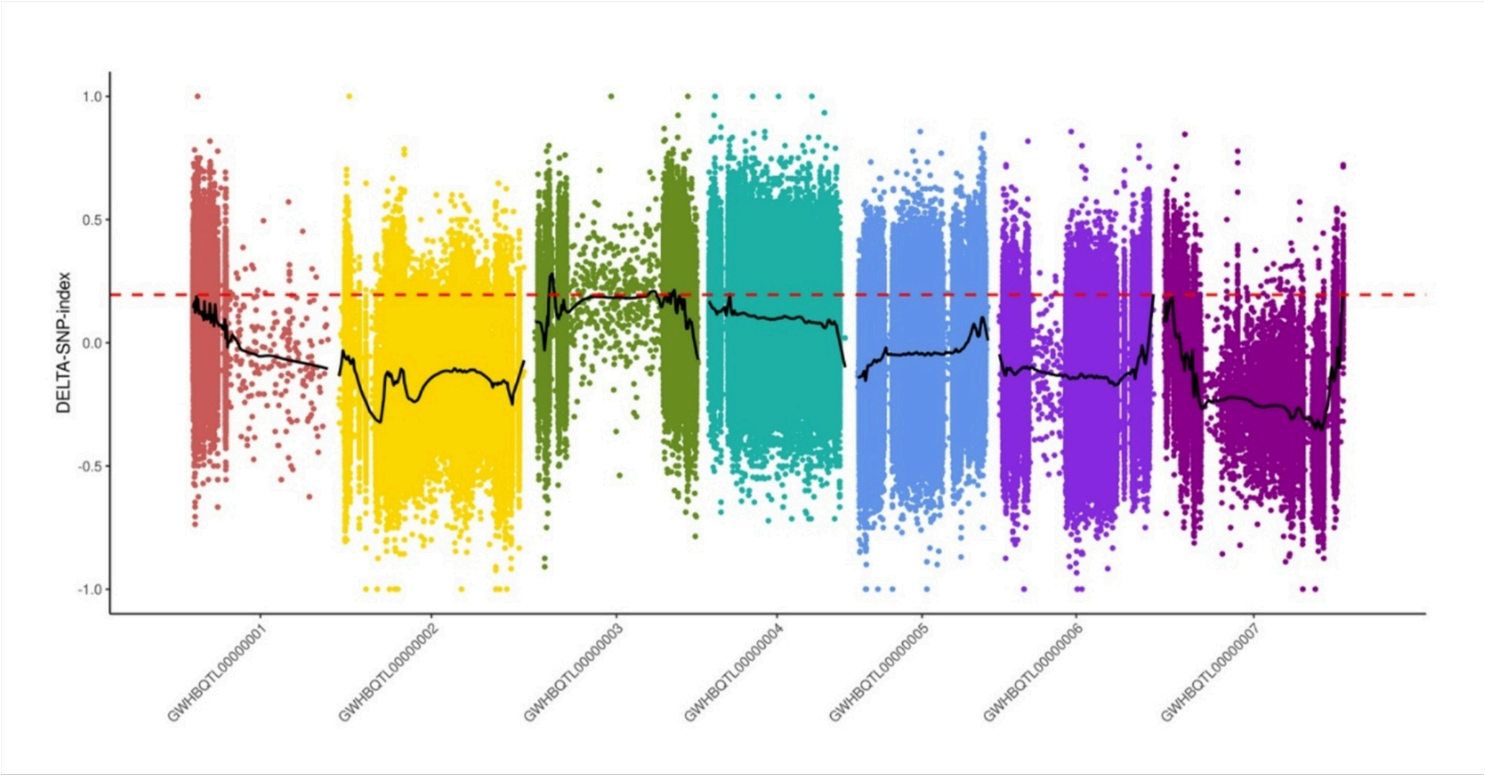

Figure 7

Genome-wide ΔSNP-index plot for the F_2 bulks across seven chromosomes. Points represent SNP-level, ΔSNP-index values and the black curve indicates the LOESS-smoothed ΔSNP-index. The red dashed line marks the significance threshold (99th percentile; ΔSNP-index = 0.02). Regions exceeding the threshold were considered candidate intervals.

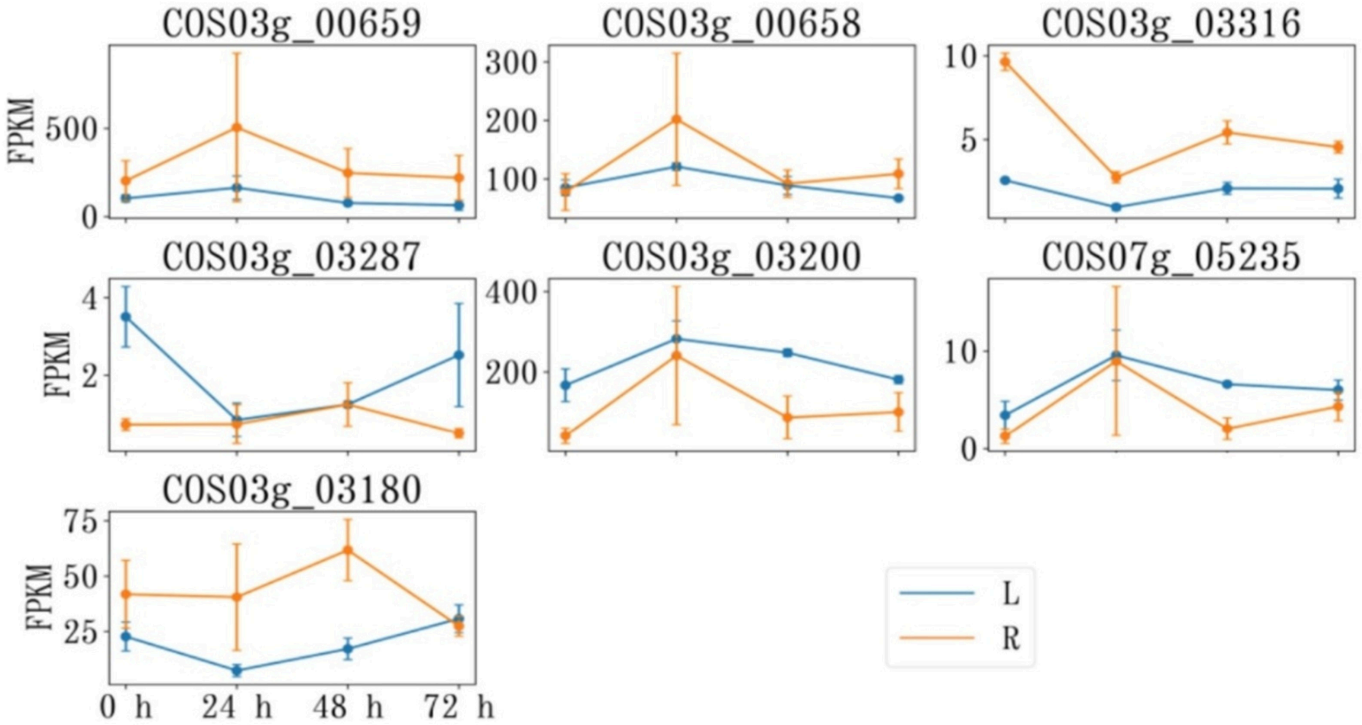

Figure 8

Time-course RNA-seq expression profiles (FPKM) of prioritized candidate genes in aboveground tissues (L) and roots (R) under Cd stress

Expression levels (FPKM) are shown for Cd-treated samples at 24, 48, and 72 h, with 0 h CK samples used as baseline controls. Points represent mean values across biological replicates and error bars indicate ±SE.

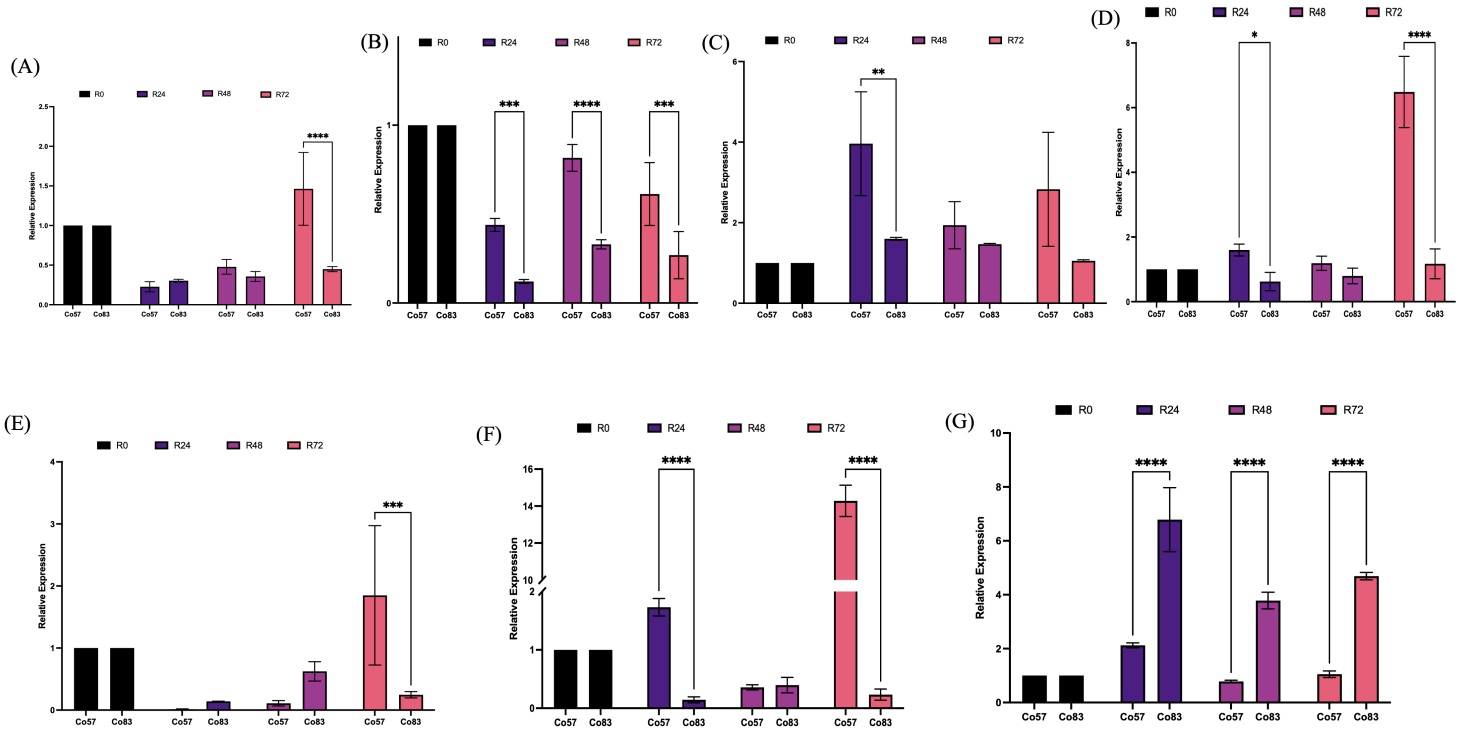

Figure 9

qRT-PCR validation of candidate genes under Cd stress

A-G represent the seven candidate genes, COS03g_03287, COS03g_03316, COS03g_03200, COS07g_05235, COS03g_03180, COS03g_00658, and COS03g_00659 respectively. Values are mean ± SE. Significant differences were assessed by t-test: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementarymaterialBMC.docx