Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Identification and expression profiling of the HMA gene family in lentil and its potential role in cadmium stress

Fawad Ali
Jeju National University
Obaid Ullah Shah
Shanghai Jiao Tong University
Arif Ali
Hainan University
Muhammad Azhar Nadeem
Department of Biotechnology
Faheem Shehzad Baloch
Department of Biotechnology
Manh Cuong Do
Duy Tan University
Yong Suk Chung
schung@hknu.ac.kr
Hankyong National University

Research Article

Keywords: Legume crop, Genome-wide analysis, Heavy metal stress, Synteny analysis, qRT-PCR

Posted Date: October 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11203382/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Identification and expression profiling of the *HMA* gene family in lentil and its potential role in cadmium stress

Fawad Ali [1], Obaid Ullah Shah [2], Arif Ali [3], Muhammad Azhar Nadeem [4,5], Faheem Shehzad Baloch [4,6], Manh Cuong Do [7,8], Yong Suk Chung [9]*

[1] Department of Plant Resources and Environment, Jeju National University, Jeju 63243, Republic of Korea

[2] School of Agriculture and Biology, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai, 200240, China

[3] School of Breeding and Multiplication (Sanya Institute of Breeding and Multiplication), School of Tropical Agriculture and Forestry, Hainan University, Sanya 572025, China

[4] Department of Biotechnology, Faculty of Science, Mersin University, Mersin 33343, Türkiye

[5] Department of Biotechnology, Institute of Biochemistry, Sh. Rashidov Samarkand State University, Samarkand 140104, Uzbekistan

[6] Sh. Rashidov Samarkand State University, Institute of Biochemistry, Department of Genetics, Samarkand, Uzbekistan

[7] Institute of Theoretical and Applied Research, Duy Tan University, Hanoi 100000, Viet Nam

[8] College of Medicine and Pharmacy, Duy Tan University, Da Nang 550000, Vietnam

[9] Department of Plant Life and Environment, Hankyong National University, 17579, Republic of Korea

Correspondence:

* schung@hknu.ac.kr

Abstract

Background

Lentil is an important food legume, but its growth and productivity can be severely constrained by exposure to toxic heavy metals. Heavy metal ATPase (HMA) proteins mediate metal ion transport, detoxification, and homeostasis, but their gene family remains unexplored in lentil.

Results

Here, we identified 13 *HMA* genes in the lentil genome. Phylogenetic analysis with *Arabidopsis* and rice HMAs divided them into a Zn/Cd/Co/Pb-type clade (*LcuHMA3* and *LcuHMA11*) and a Cu/Ag-type clade (the other 11 genes), spanning five of the six phylogenetic groups (II-VI). The genes were unevenly distributed across five chromosomes (Chr2, Chr3, Chr4, Chr5, and Chr7), with two on an unplaced scaffold. Of 50 duplicated gene pairs, three were tandem and 47 segmental, indicating that segmental duplication predominated, and 49 pairs were under purifying selection (Ka/Ks < 1). Synteny analysis revealed that lentil *HMA* genes share the greatest number of conserved genomic relationships with those in *Glycine max* among the species examined. qRT-PCR of nine *LcuHMA* genes after 2, 12, and 24 h of exposure to 3 mM $CdCl_2$ showed that *LcuHMA3*, *LcuHMA4*, and *LcuHMA5* were up-regulated at all time points, *LcuHMA11* at 2 and 24 h, and *LcuHMA7* and *LcuHMA10* were down-regulated. These results show that *LcuHMA* genes respond transcriptionally to cadmium and identify *LcuHMA3*, *LcuHMA4*, *LcuHMA5*, and *LcuHMA11* as priority candidates for functional validation.

Conclusions

Our analysis offers insight into the evolution of the lentil *LcuHMA* family and identifies candidates for future functional validation.

Keywords: Legume crop, Genome-wide analysis, Heavy metal stress, Synteny analysis, qRT-PCR

Introduction

Lentil (*Lens culinaris* Medik.) is a self-pollinating, cool-season legume ranking fifth among global pulse crops [1]. It is nutritionally valuable, rich in protein, carbohydrates, fiber, vitamins, minerals, and phytochemicals, with a low glycemic index and fat content (<2.5%) that support glycemic and insulin regulation [2,3,4]. Amid climate variability and rising pulse demand, lentil is vital to food and nutritional security [5,6], though its productivity remains constrained by heavy metal stress, particularly cadmium (Cd) exposure [7].

Cd contamination of agricultural soils arises from both natural and human sources. These include the weathering of parent material, atmospheric deposition, phosphate fertilizers, mining and smelting, sewage sludge application, and irrigation with contaminated water [8]. Once in the soil, Cd can stay bioavailable for a long time, depending on soil properties such as pH, organic matter, carbonate content, and texture. This creates a lasting risk to crop production and food safety [9]. Unlike metals that occur in several oxidation states with different mobilities, Cd exists in soil mainly as the divalent ion Cd^{2+}. This form is highly soluble, mobile, and available to plant roots, which makes Cd one of the most phytotoxic heavy metals even at low concentrations [10].

Roots are the main entry point for Cd from the soil. Because plants have no Cd-specific uptake transporter, Cd^{2+} enters root cells through broad-specificity systems that normally acquire

Fe, Mn, and Zn, including the NRAMP and ZIP/IRT transporters [11]. Ca-permeable cation channels may also contribute under some conditions [12]. Cd then accumulates in the roots, where it damages cells, causes oxidative stress, impairs nutrient and water uptake, inhibits photosynthesis, induces chlorosis, and reduces growth and yield [13]. Heavy metal ATPases (HMAs), a subfamily of P-type ATPases, use energy from ATP hydrolysis to move metal cations such as Cu^{2+}, Co^{2+}, Zn^{2+}, Pb^{2+}, and Cd^{2+} across membranes [14,15]. The HMA domain controls the spatial and temporal precision of this transport through its interactions with cellular enzymes and cofactors [16].

HMA proteins are typically defined by six to eight transmembrane domains and two functionally essential motifs—a histidine-proline (HP) motif and a cysteine-proline-cysteine/serine-proline-cysteine (CPC/SPC) motif—required for metal binding and transport [17]. Based on substrate specificity, they are classified into Cu/Ag-ATPases and Zn/Co/Cd/Pb-ATPases [18,19,20]. While *HMA* gene families have been molecularly characterized in species such as *Arabidopsis thaliana*, *Zea mays*, *Oryza sativa*, and *Triticum aestivum* [21,22,23], their functional roles vary notably across plant taxa.

Functional studies across species highlight the diverse roles of *HMA* genes: *AtHMA1* detoxifies excess chloroplastic Zn(II) [24]; *OsHMA2*, *AtHMA2*, and *AtHMA4* mediate root-to-shoot Zn^{2+}/Cd^{2+} translocation [25,26]; and *AtHMA3* overexpression boosts metal tolerance via vacuolar Cd, Zn, Pb, and Co sequestration [27]. In rice, *OsHMA1* and *OsHMA2* handle tonoplast Zn transport, *OsHMA3* manages Cd transport, *OsHMA4* directs Cu into root vacuoles, and *OsHMA5* translocates Cu^{2+} to shoots [21,28,29,30,31,32,33]. Other species show similar diversity: *SpHMA3* confers Cd hyper-tolerance in *Sedum plumbizincicola* [34], *CsHMA3*/*CsHMA4* transport

Cd^{2+}, Pb^{2+}, and Zn^{2+} toward stems [35], and *HvHMA1* exports chloroplastic metals while mobilizing Zn and Cu during germination and grain filling [36].

Despite extensive HMA characterization across plant species, lentil remains unexamined. This study fills that gap through genome-wide identification and multi-omics analysis of the *LcuHMA* gene family, covering phylogenetics, gene structure, conserved motifs, chromosomal mapping, and expression under Cr stress. The identified genes showed uneven distribution across lentil chromosomes, providing a foundation for understanding *LcuHMA* roles in metal homeostasis, Cr tolerance, and phytoremediation.

Material and methods

HMA sequence collection and protein characterization

A lentil reference genome was sourced from Phytozome (https://phytozome-next.jgi.doe.gov/; accessed 1 June 2026). To assemble a set of previously described HMA proteins, we consulted the NCBI Protein database (https://www.ncbi.nlm.nih.gov/protein) and Phytozome (https://phytozome-next.jgi.doe.gov/) on 1 June 2026 [37], together with published HMA sequences for *Arabidopsis thaliana* [38], *Oryza sativa*, and *Glycine max* [39]. We evaluated the domain composition of the HMA proteins with InterPro (https://www.ebi.ac.uk/interpro/; accessed 15 June 2026) [40]. Their predicted molecular masses, theoretical isoelectric points, and grand average of hydropathicity (GRAVY) values were then obtained through ExPASy ProtParam (https://web.expasy.org/protparam/; accessed 1 July 2026) [41].

Domain characterization of LcuHMA proteins

Each of the 13 LcuHMA protein sequences was examined for conserved domains using NCBI CD-Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi; accessed 15 July 2026), with

Pfam version 34.0 selected as the search database. The domain annotations were subsequently visualized in TBTools to compare HMA domain architecture across the proteins.

Phylogenetic analysis of the *LcuHMA* gene family

The complete amino-acid sequences of HMA family members from *L. culinaris*, *O. sativa*, and *A. thaliana* were obtained from the Phytozome v13 database on 20 July 2026 [42]. The 30 HMA protein sequences were aligned in MEGA 11 [43]. A maximum-likelihood phylogeny was then inferred, with branch support assessed through 1,000 bootstrap replicates, and the tree was displayed in iTOL [44].

Distribution and analysis of conserved motif patterns

Conserved motifs in the protein sequences of the 13 LcuHMA members were analyzed with MEME Suite (https://meme-suite.org/meme/tools/meme; accessed 25 July 2026). Default parameters were retained, except that the maximum number of motifs was set to 10 [45].

Chromosomal localization of *LcuHMA* genes

Genomic sequences and annotation records for the 13 lentil *HMA* genes were downloaded from Phytozome v13 (https://phytozome-next.jgi.doe.gov/; accessed 1 August 2026) [42]. Chromosome lengths were determined using the "FASTA stats" feature within TBTools [46], which was subsequently used to generate a chromosome ideogram (http://visualization.ritchielab.org/phenograms/plot; accessed 1 August 2026). The resulting image displays the mapped positions of all *LcuHMA* genes across the respective chromosomes.

Analysis of evolutionary relationships and gene duplication in *LcuHMA*

Duplication events within the *LcuHMA* gene family were examined using MCScanX (https://github.com/wyp1125/MCScanX; accessed 10 August 2026) and visualized with Circos

[47]. For each gene pair, we estimated the nonsynonymous (Ka) and synonymous (Ks) substitution rates [48]. Divergence time was then calculated as T = Ks / (2R), using R = 1.5×10^{-8} substitutions per synonymous site per year, as described by Shah et al. (2024) [49]. We used MCScanX to assess synteny among *G. max*, *A. thaliana*, and *O. sativa* and displayed the resulting relationships with TBtools.

LcuHMA genes expression profiling

We obtained expression values for the 13 *LcuHMA* genes from the lentil seed-development dataset available through PulseDB (https://www.pulsedb.org/Analysis/3741283) [50,51]. We compared their expression across embryo, seed-coat, and whole-seed samples collected at successive developmental stages to assess spatial and temporal patterns.

Promoter *cis*-acting element analysis of *LcuHMA* genes

We used TBtools to retrieve the 2,000-bp region upstream of each *LcuHMA* start codon from the lentil genome sequence [46]. These regions were then screened for *cis*-acting elements using PlantCARE [52].

Protein functional-association network

A predicted functional-association network of LcuHMA proteins was generated using STRING. Lentil proteins were mapped to the closest available reference orthologs, and medium-confidence associations were retained without additional interactors. Nodes represent LcuHMA proteins, while colored edges indicate STRING evidence channels. The network was interpreted as orthology-informed functional associations, not experimentally validated direct protein–protein interactions.

GO enrichment analysis

GO biological-process enrichment of LcuHMA proteins was performed in STRING using mapped reference orthologs. Enriched terms were tested against the STRING reference background using Fisher's exact test, followed by Benjamini–Hochberg FDR correction. Terms with FDR-adjusted $P<0.05$ were considered significant. Bubble plots show enrichment signal, gene count, and FDR.

Phylogenetic co-occurrence analysis

Phylogenetic co-occurrence was assessed using the STRING co-occurrence evidence channel. This analysis compares ortholog presence–absence patterns across taxa to identify proteins with similar evolutionary distributions. Columns represent LcuHMA proteins or mapped orthologs, and colored cells indicate occurrence patterns across bacterial, eukaryotic, and archaeal groups. Results indicate evolutionary co-distribution and possible functional conservation, not direct physical interactions.

Plant materials and growth conditions

The lentil check cultivar "晋扁豆 2 号" was used as the plant material and grown under hydroponic conditions. Seeds of the lentil genotype were provided by Hainan University, China during the first author postdoctoral research. Growth chamber conditions were set to 22°C during the day and 18°C at night (±2°C), with relative humidity maintained at 45%. For the qRT-PCR experiment, seedlings were cultivated in a full Hoagland nutrient solution containing 2.5 mM $Ca(NO_3)_2$, 2 mM KCl, 1 mM $MgSO_4$, 0.5 mM $NH_4H_2PO_4$, 0.5 mM $CaCl_2$, 25 µM H_3BO_3, 5 µM $MnSO_4$, 0.8 µM $ZnSO_4$, 0.3 µM $CuSO_4$, 0.1 µM $(NH_4)_6Mo_7O_{24}$, and 100 µM Na_2EDTA-Fe (Shah et al., 2024) [49].

Cd stress treatment

The Cd concentrations used in this study were determined through a preliminary experiment in which seedlings were exposed to 2.5, 3, and 3.5 mM of Cd. Based on these results, 3 mM concentration was selected to induce stress. Seven days after germination, uniform seedlings were

selected and assigned to one of two treatments: control (0 mM Cd) or Cd stress (3 mM Cd). Tissue samples were collected immediately before Cd exposure (0 h) and at 2, 12, and 24 h after treatment began. Each sample was frozen in liquid nitrogen and stored at -80 °C until analysis. The experiment used a completely randomized design with three biological replicates per treatment, each comprising five seedlings.

RNA extraction and quantitative real-time PCR analysis

Total RNA was isolated from plant samples with the RNAprep Pure Plant Kit (TIANGEN, Beijing, China), and its concentration was measured using a NanoDrop 2000C spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Before cDNA synthesis, 1 μg of RNA per sample was treated to remove genomic DNA, followed by reverse transcription with the QuantiTect Reverse Transcription Kit (Qiagen, Shanghai, China). Quantitative real-time PCR (qRT-PCR) was performed on a LightCycler 480 system (Roche) using SYBR Green Master Mix (Novogene, Shanghai, China). *LcuACTIN* (NCBI accession MH167390), selected from a previous study, served as the reference gene for normalizing target-gene expression. Three biological replicates were analyzed, with three technical PCR replicates for each biological sample. Relative expression was determined by the $2^{-\Delta\Delta CT}$ method [53].

Results

Genome-wide identification and characterization of lentil *HMA* genes

A genome-wide search of the lentil genome yielded 13 putative *HMA* family members, named *LcuHMA1–LcuHMA13*. Their physicochemical properties are presented in Table 1, while chromosomal positions and predicted subcellular localizations are listed in Supplementary Table S1. Protein lengths ranged from 159 amino acids (*LcuHMA1*) to 1,093 amino acids (*LcuHMA11*),

with molecular weights spanning 17.38–118.57 kDa. GRAVY scores ranged from −0.247 (*LcuHMA1*) to 0.346 (*LcuHMA7*); nine proteins had positive scores and four had negative scores. Predicted isoelectric points (pI) varied from 5.12 (*LcuHMA2*) to 10.29 (*LcuHMA12*). Most LcuHMA proteins were predicted to localize to the plasma membrane, whereas others were assigned to the extracellular, chloroplast, or nucleus (Table S1).

Table 1 Physicochemical properties of the predicted *LcuHMA* genes in lentil

Transcript ID	Gene name	AA	Weight (KDs)	PI	Instability index	Aliphatic index	GRAVY
Lcu.2RBY.2g001670	*LcuHMA1*	159	17.38385	8.6	44.45	85.22	-0.247
Lcu.2RBY.2g019550	*LcuHMA2*	991	106.7002	5.12	33.03	105.81	0.283
Lcu.2RBY.2g020730	*LcuHMA3*	782	85.03666	5.58	31.41	109.44	0.207
Lcu.2RBY.3g064660	*LcuHMA4*	884	94.06844	6.12	34.95	106.24	0.208
Lcu.2RBY.4g035460	*LcuHMA5*	954	103.9057	5.58	36.12	106.26	0.204
Lcu.2RBY.4g060650	*LcuHMA6*	941	100.23856	8.86	37.66	101.14	0.14
Lcu.2RBY.4g063350	*LcuHMA7*	993	106.24513	5.19	32.36	109.06	0.346
Lcu.2RBY.5g002970	*LcuHMA8*	987	106.88428	6.68	32.45	104.1	0.207
Lcu.2RBY.5g002980	*LcuHMA9*	1002	107.88969	5.61	32.06	105.81	0.235
Lcu.2RBY.7g002240	*LcuHMA10*	956	102.76228	5.9	34.35	107.26	0.27
Lcu.2RBY.7g074660	*LcuHMA11*	1093	118.5653	6.54	36	90.01	-0.197
Lcu.2RBY.L005370	*LcuHMA12*	213	23.22005	10.29	62.18	86.48	-0.109
Lcu.2RBY.L005380	*LcuHMA13*	215	23.18968	9.85	62.61	86.6	-0.209

Phylogenetic relationships

The phylogenetic analysis included 30 HMA protein sequences: 13 from lentil, eight from *Arabidopsis thaliana*, and nine from *Oryza sativa*. The total HMA genes were clustered into two major clades and six groups (groups I–VI). Group I clustered two members (*OsHMA1* and *AtHMA1*). Group II clustered seven members (*OsHMA3*, *LcuHMA3*, *OsHMA2*, *AtHMA4*, *AtHMA2*, *AtHMA3*, *LcuHMA11*). Group III clustered seven members (*OsHMA4*, *LcuHMA10*, *LcuHMA5*, *AtHMA5*, *OsHMA5*, *LcuHMA9*, *LcuHMA8*). Group IV clustered five members (*OsHMA9*, *OsHMA6*, *LcuHMA7*, *AtHMA7*, *LcuHMA2*). Group V clustered six members (*OsHMA7*, *LcuHMA1*, *AtHMA6*, *LcuHMA12*, *LcuHMA13*, *LcuHMA6*). Group VI clustered three members (*OsHMA8*, *AtHMA8*, *LcuHMA4*). Furthermore, groups I and II formed the Zn/Cd/Co/Pb-type clade, while groups III, IV, V, and VI formed the Cu/Ag-type clade (Fig. 1).

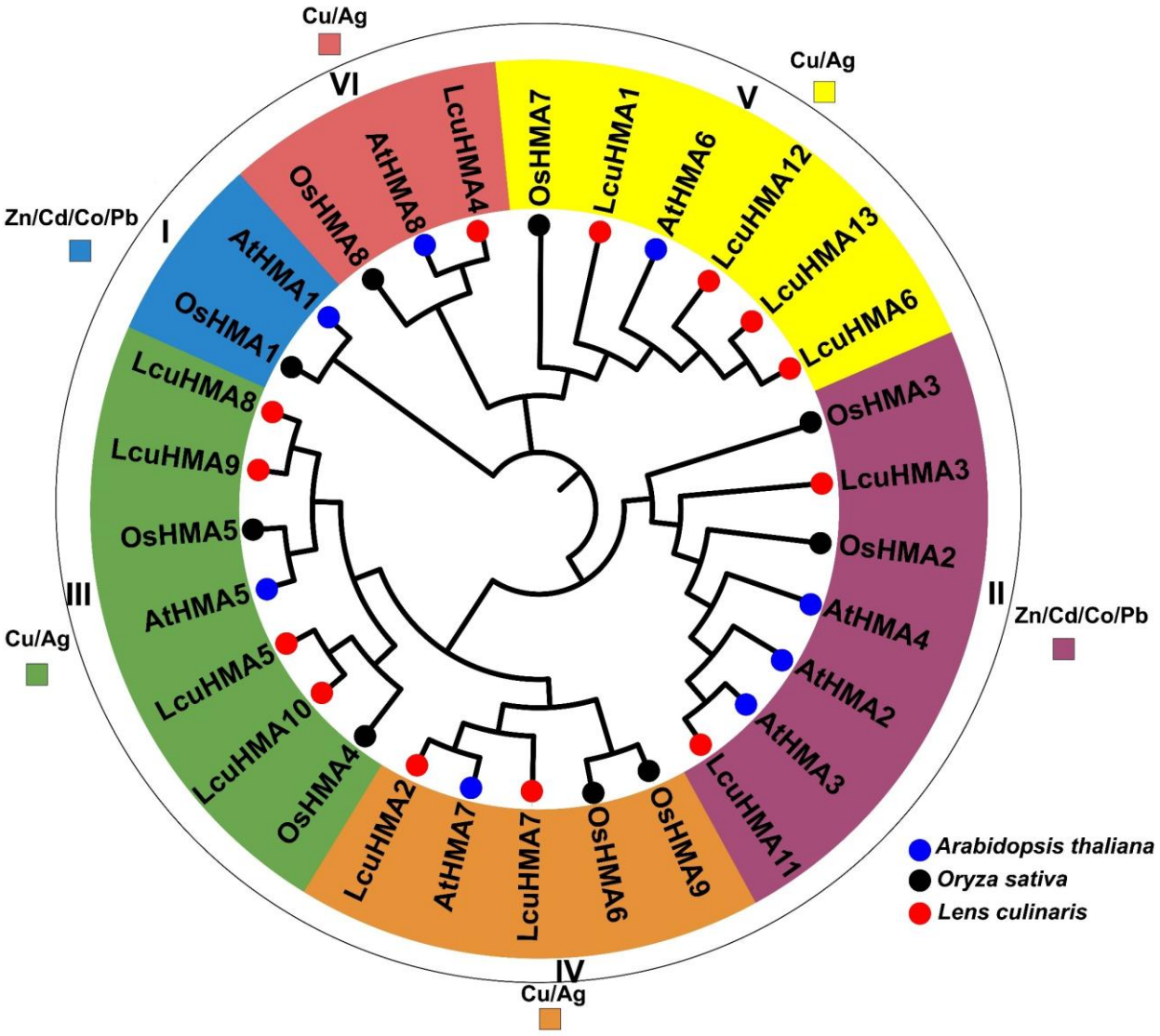


Fig. 1 Phylogenetic relationships among *HMA* genes from lentil, *O. sativa*, and *A. thaliana*. The sequences were assigned to six color-coded groups: I (blue), II (purple), III (green), IV (orange), V (yellow), and VI (red). Groups I and II belong to the Zn/Cd/Co/Pb-type clade, while groups III, IV, V, and VI belong to the Cu/Ag-type clade

Conserved motifs, domain, and gene structural analysis

To further characterize the *LcuHMA* gene family, conserved protein motifs, domain, and exon-intron organization were analyzed for all identified members (Fig. 2). MEME analysis identified 10 conserved motifs (Motifs 1–10) across the LcuHMA proteins, with motif composition and arrangement closely mirroring the phylogenetic groupings established earlier. Motif 4 was the most universally conserved element, present in all 13 LcuHMA proteins, whereas *LcuHMA1*, *LcuHMA12*, and *LcuHMA13* contained only this single motif, consistent with their comparatively short and structurally simple sequences. In contrast, most other members—including *LcuHMA2*, *LcuHMA4*, *LcuHMA5*, *LcuHMA6*, *LcuHMA7*, *LcuHMA8*, *LcuHMA9*, and *LcuHMA10*—shared a highly conserved motif arrangement comprising Motifs 4, 5, 8, 2, 7, 1, 10, 9, 6, and 3 in a consistent order, while *LcuHMA3* and *LcuHMA11* lacked Motifs 5 and 8. All motif sequences and lengths are provided in Supplementary Table S2.

Domain architecture analysis was largely consistent with the motif distribution patterns. Most LcuHMA proteins possessed a canonical tripartite domain organization typical of heavy-metal-transporting P-type ATPases, comprising an N-terminal HMA (heavy metal-associated) superfamily domain, a central E1-E2_ATPase superfamily domain, and a C-terminal Hydrolase superfamily domain. This full domain complement was observed in *LcuHMA2*, *LcuHMA4*, *LcuHMA5*, *LcuHMA6*, *LcuHMA7*, *LcuHMA8*, *LcuHMA9*, and *LcuHMA10*, consistent with their conserved motif arrangements described above. *LcuHMA3* and *LcuHMA11* similarly retained the HMA superfamily and E1-E2_ATPase superfamily domains along with the Hydrolase superfamily domain, though with a comparatively shorter overall domain span. In marked contrast, *LcuHMA1*, *LcuHMA12*, and *LcuHMA13* lacked both the E1-E2_ATPase and Hydrolase superfamily domains, retaining only a truncated HMA superfamily domain, consistent with their reduced motif content and simplified protein structures. This close correspondence between domain composition, motif

arrangement, and phylogenetic clustering further supports the classification of the *LcuHMA* family as heavy-metal-associated P-type ATPases with varying degrees of domain retention (Fig. 2).

This pattern of conservation was mirrored at the gene structure level. Exon-intron analysis revealed considerable structural diversity within the family, with exon number ranging from 2 to 10. *LcuHMA6* contained numerous exons (up to 10) interspersed with long intronic regions exceeding 20 kb. In contrast, *LcuHMA12* and *LcuHMA13* had the simplest structures, each with only two exons. Genes such as *LcuHMA2*, *LcuHMA5*, *LcuHMA7*, and *LcuHMA10* showed intermediate and relatively similar exon-intron patterns, in line with their close clustering in the phylogenetic tree, while most other members displayed a moderate number of exons together with untranslated regions (UTRs) at both ends. Notably, closely related paralogous pairs such as *LcuHMA8*/*LcuHMA9* and *LcuHMA12*/*LcuHMA13* displayed nearly identical exon-intron organization, further supporting their close evolutionary relatedness (Fig. 2).

Taken together, these results show that members within the same phylogenetic group generally share conserved motif compositions and exon-intron structures, whereas more distantly related genes (e.g., *LcuHMA1*, *LcuHMA12*, and *LcuHMA13*) display simplified gene architectures lacking several of the conserved motifs found in other family members. This concordance between structural features and phylogenetic relationships suggests that structural divergence in both motif composition and gene architecture has accompanied the evolutionary diversification of the *LcuHMA* gene family in lentil.

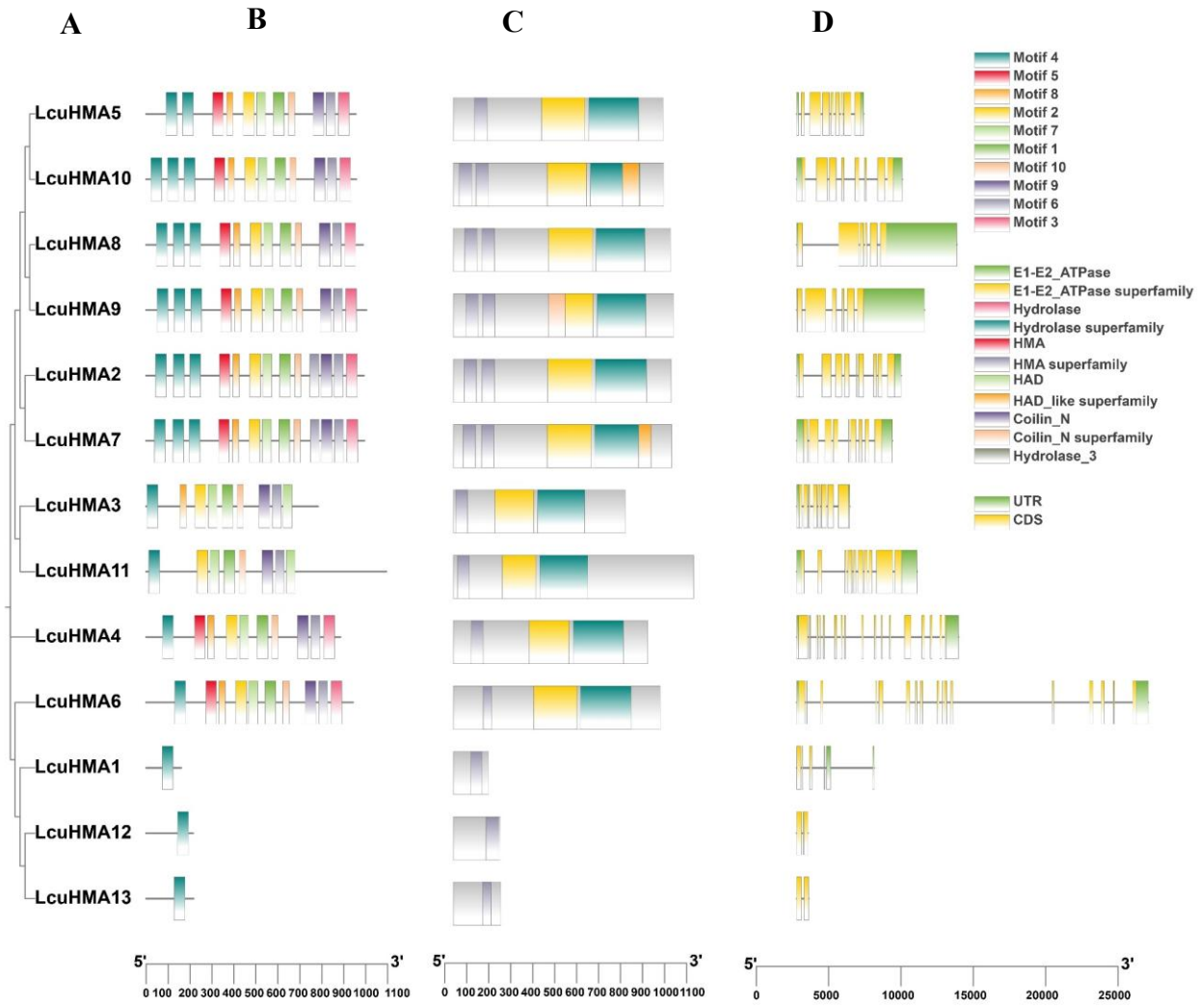


Fig. 2 Phylogenetic relationships, conserved motif distribution, domain architecture, and gene structure of the *LcuHMA* gene family. **(A)** Phylogenetic tree of the 13 LcuHMA proteins constructed using the maximum likelihood method with 1000 bootstrap replicates. **(B)** Distribution of conserved motifs (Motifs 1–10) identified using the MEME tool and visualized with TBtools; each motif is represented by a distinctly colored box, with the color key shown at right. **(C)** Predicted conserved domains of LcuHMA proteins, including the HMA, E1-E2_ATPase, and Hydrolase superfamily domains and color-coded as indicated in the legend. **(D)** Exon–intron organization of the corresponding *LcuHMA* genes, with coding sequences (CDS) and untranslated

regions (UTRs) shown in yellow and green, respectively, and introns represented as connecting lines. Scale bars below each panel indicate relative protein/motif length (B, C) and genomic length in base pairs (D)

Chromosomal distribution of *LcuHMA* genes

Using genomic coordinates from the lentil genome assembly available through Phytozome (accessed 1 July 2026), we assigned 11 *LcuHMA* genes to five of the seven lentil chromosomes; the remaining two were located on an unplaced scaffold. The chromosomal genes were distributed unevenly. Chromosomes 2 and 4 each contained three *LcuHMA* genes, chromosome 3 contained one, and none were detected on chromosomes 1 or 6 (Fig. 3).

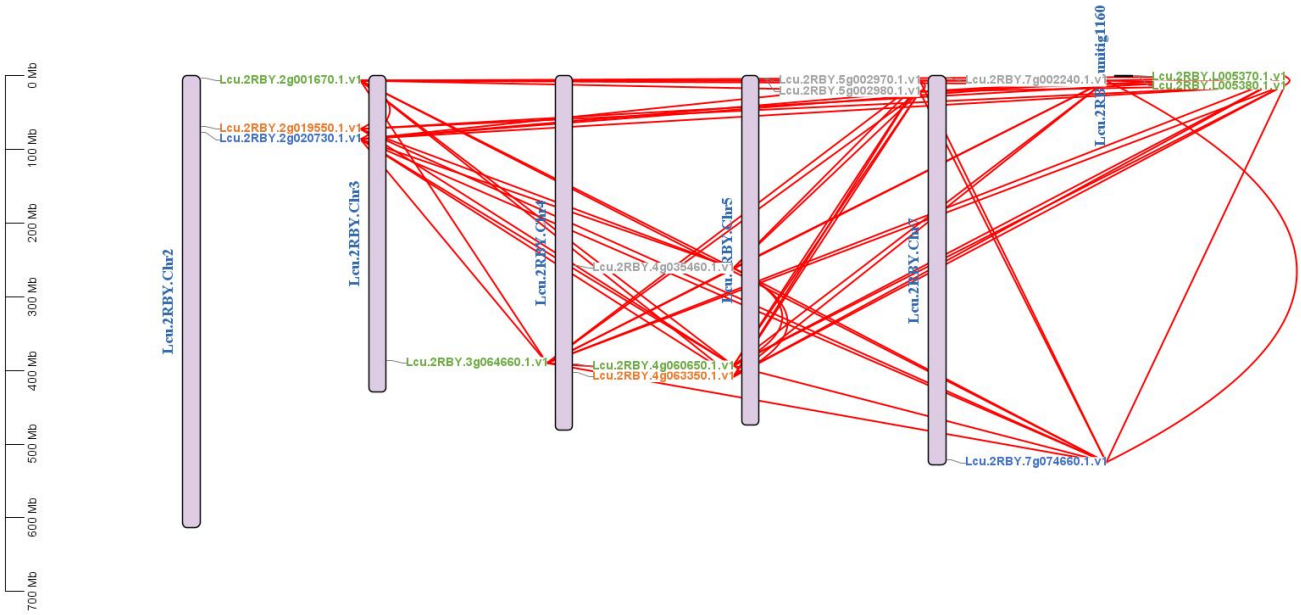


Fig. 3 Genomic positions of the 13 *LcuHMA* genes in the lentil assembly. Eleven genes are shown on five chromosomes, and two are shown on an unplaced scaffold. Labels identify individual genes at their mapped positions

Duplication patterns of *LcuHMA* genes

We assessed duplication within the *LcuHMA* family by comparing gene sequences and examining syntenic relationships (Fig. 4). This analysis revealed a total of 50 duplicated gene pairs among the *LcuHMA* genes (Table S3). Sequences located within a 100 kb region of a duplicated genomic segment were considered tandemly duplicated gene pairs. Based on these criteria, 3 gene pairs — *LcuHMA1–LcuHMA2*, *LcuHMA5–LcuHMA6*, and *LcuHMA8–LcuHMA9* — were classified as tandem duplicates. The remaining 47 gene pairs were identified as segmental duplicates, distributed across chromosomes 1 through 7 and additional unplaced scaffolds, as illustrated by the connecting lines in the circos plot (Fig. 4; Table S3).

We used the ratio of nonsynonymous to synonymous substitution rates (Ka/Ks) to assess patterns of selection among duplicated gene pairs. Values below 1 are consistent with purifying (negative) selection, values near 1 with approximately neutral evolution, and values above 1 may indicate diversifying (positive) selection. Among the 50 *LcuHMA* duplicated gene pairs, Ka/Ks ratios ranged from 0.139 to 0.929 for the majority of pairs, and 49 of the 50 pairs were found to be evolving under negative (purifying) selection (Table S3). This finding suggests that purifying selection has played a dominant role in maintaining the structural and functional conservation of *LcuHMA* genes throughout their evolutionary history. Notably, the gene pair *LcuHMA12–LcuHMA13* showed a Ka/Ks ratio of 1.451, indicating positive (diversifying) selection and possibly reflecting a recent functionally significant divergence event. Divergence time (T) estimates for the duplicated gene pairs averaged approximately 150.98 MYA, ranging from 4.59

MYA to 231.73 MYA (Table S3). These results indicate that *LcuHMA* gene duplications occurred over a broad evolutionary timescale, with most pairs retained under strong purifying selection while at least one pair may have acquired a novel or diversified function. Comparing conserved gene order across chromosomal regions provided additional context for the evolutionary history of the *LcuHMA* family.

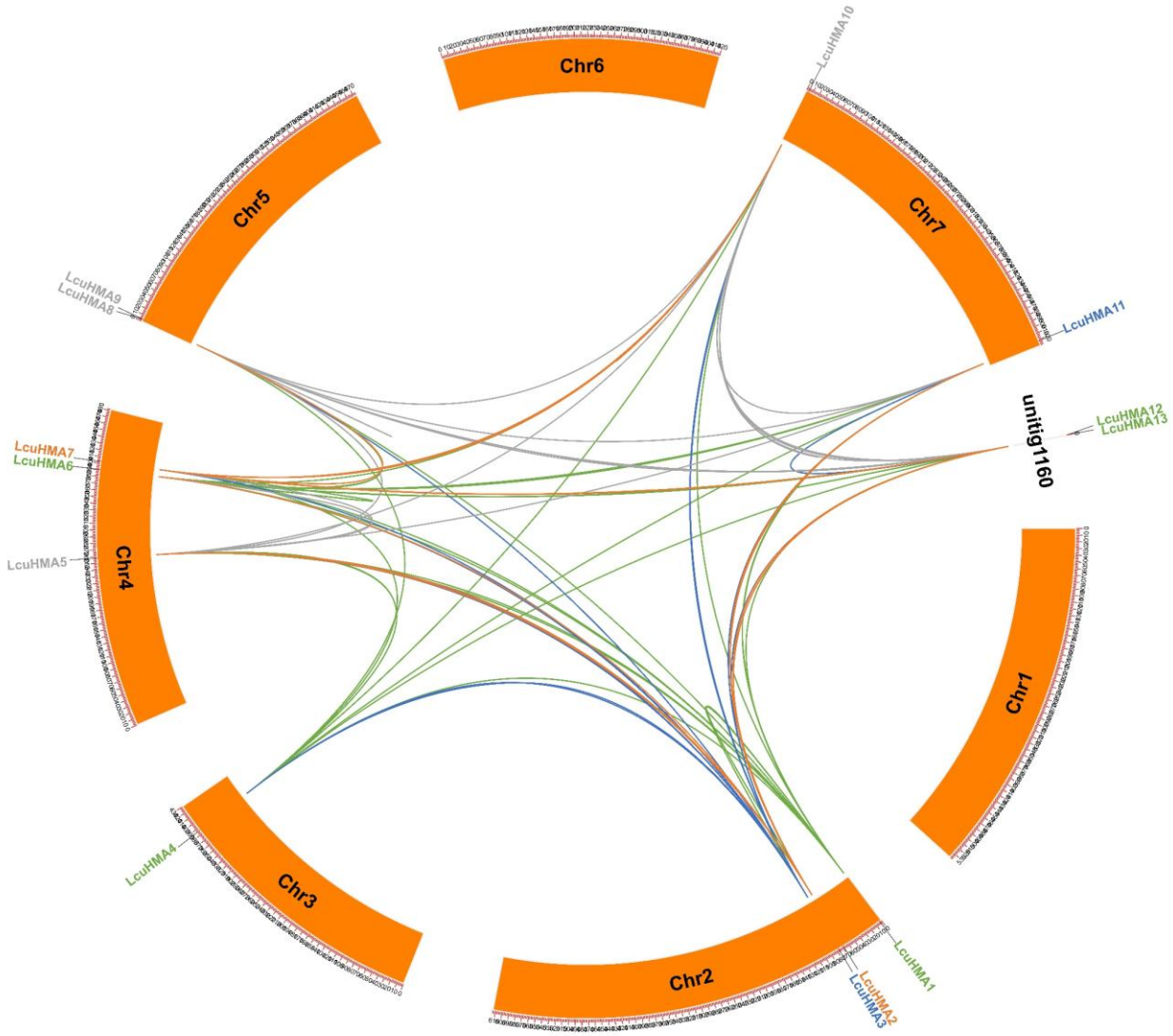


Fig. 4 Circos representation of duplicated *LcuHMA* gene pairs in the lentil genome. Links connect genes classified as duplicated pairs, and colors distinguish the HMA subgroups

We compared the genomic regions containing lentil *HMA* genes with those of *G. max*, *A. thaliana*, and *O. sativa* to identify conserved syntenic relationships. Chromosomes of *L. culinaris*, *A. thaliana*, *G. max*, and *O. sativa* are shown in orange, green, gray, and pink, respectively (Fig. 5). The analysis identified the greatest number of HMA-associated syntenic links between lentil and *G. max* (Fig. 5; Table S4).

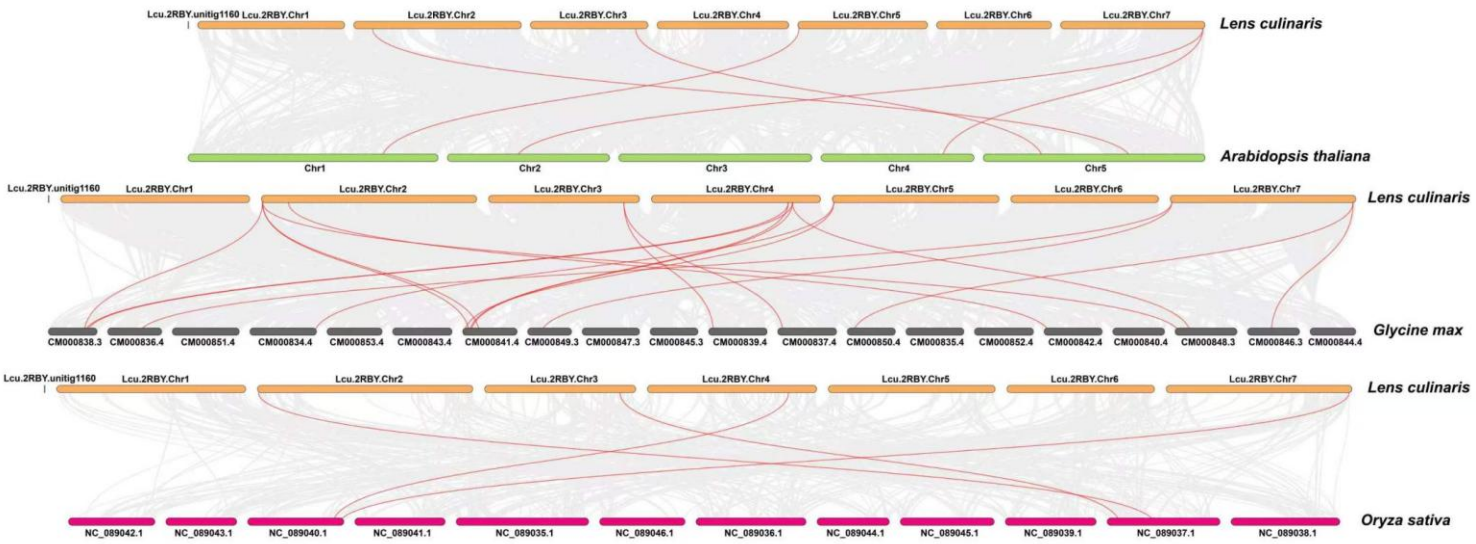


Fig. 5 Syntenic relationships between lentil *HMA* genes and their corresponding genomic regions in *G. max*, *A. thaliana*, and *O. sativa*. Red lines indicate *HMA* gene pairs linked by the synteny analysis

Cis-acting elements in *LcuHMA* upstream regions

We used PlantCARE to screen the 2,000-bp sequence upstream of each *LcuHMA* start codon for predicted *cis*-acting elements. Elements were detected in all 13 regions (Fig. 6). The predicted motifs included stress- and hormone-responsive elements, alongside other regulatory elements. The stress-responsive elements identified included the MYB binding site (MBS and MBS-D, involved in drought inducibility), the low-temperature-responsive element (LTR), the anaerobic-responsive element (ARE), and the defense- and stress-responsiveness element (DSR), along with additional regulatory motifs such as CDM-LR, MSER, ESE, and CCRE. Similarly, we identified several hormone-responsive elements, including the abscisic acid-responsive element (ABRE), the auxin-responsive element (AuxRE), the gibberellin-responsive elements (GARE and CARE), and the jasmonic acid- and salicylic acid-responsive elements (JARE and SARE, respectively). In addition to these, core promoter elements (TATA-box and CAAT-box) and light-responsive/regulatory motifs (LRE, WRE, ATBP-1, LRM, PBS, 60K PBS, MYBHv1, and ZMR) were also detected across the promoter sequences (Fig. 2; Table S5). These results suggest that *LcuHMA* genes are regulated by a diverse array of stress- and hormone-responsive elements, consistent with their proposed roles in heavy metal homeostasis and stress adaptation.

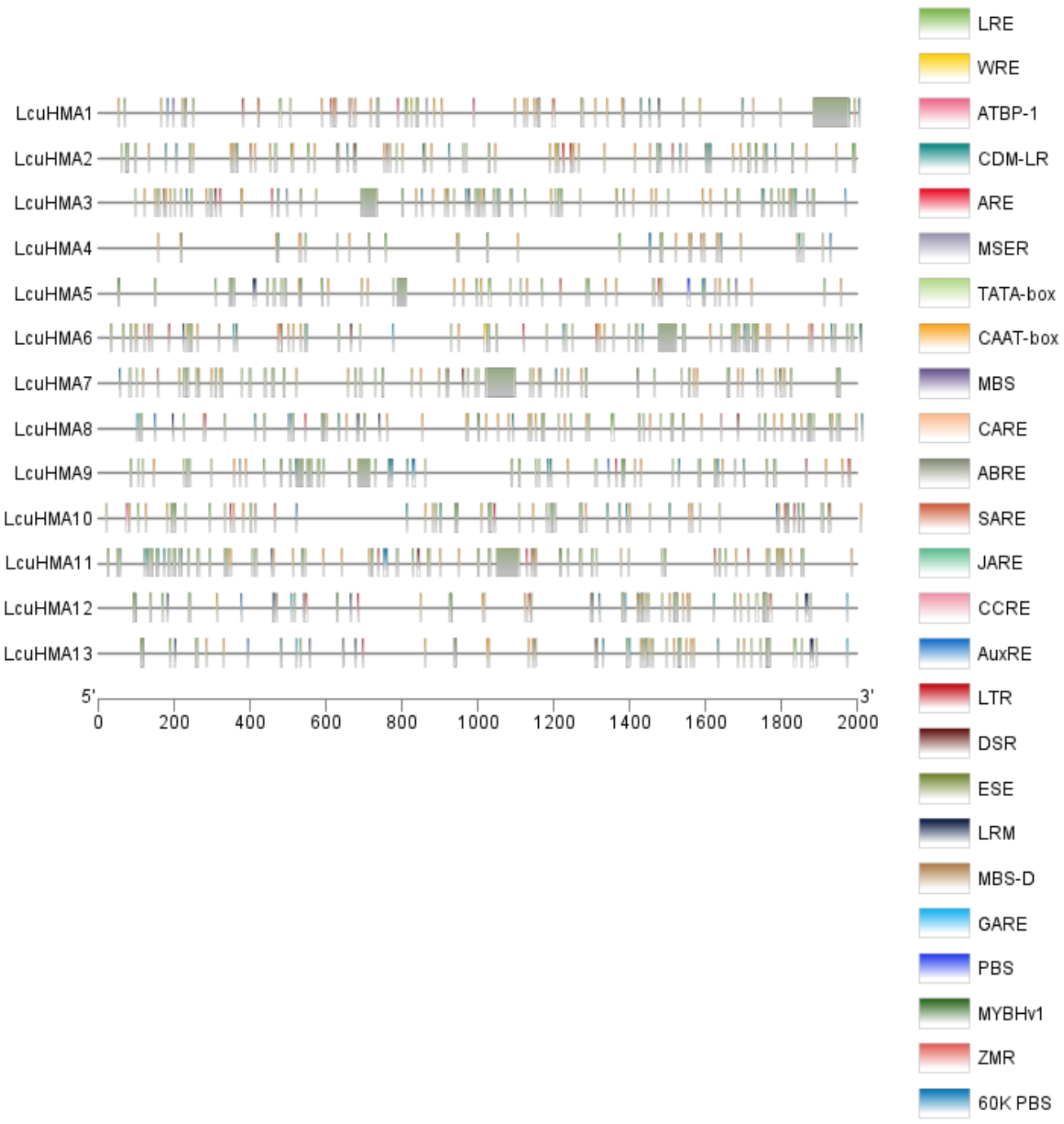


Fig. 6 Predicted *cis*-acting elements within the 2,000-bp regions upstream of the start codons of 13 *LcuHMA* genes. The plot shows the positions and categories of elements associated with hormone, stress, light, and core promoter regulation

STRING analysis

STRING analysis was performed for the 13 LcuHMA proteins, but only seven were represented in the network (Fig. 7). Six of these formed a connected predicted functional-association network: *LcuHMA1*, *LcuHMA4*, *LcuHMA7*, *LcuHMA9*, *LcuHMA10*, and *LcuHMA13*. *LcuHMA9* and *LcuHMA10* showed relatively high connectivity. In contrast, *LcuHMA11* was isolated, indicating no STRING association above the selected confidence threshold. These links represent predicted, partly orthology-supported functional associations rather than confirmed physical interactions.

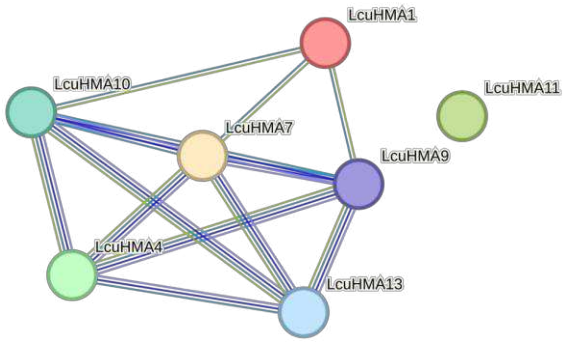


Fig. 7 Predicted functional-association network of LcuHMA proteins in lentil. Nodes represent LcuHMA proteins, while edges represent STRING-derived functional associations. The isolated node had no association above the selected confidence threshold

GO enrichment indicated significant association of LcuHMA proteins with inorganic cation transmembrane transport, copper ion homeostasis, metal ion homeostasis, and photosynthetic electron transport chain. These results support predicted roles of LcuHMA proteins in metal transport, copper homeostasis, and chloroplast-associated functions (Fig. 8).

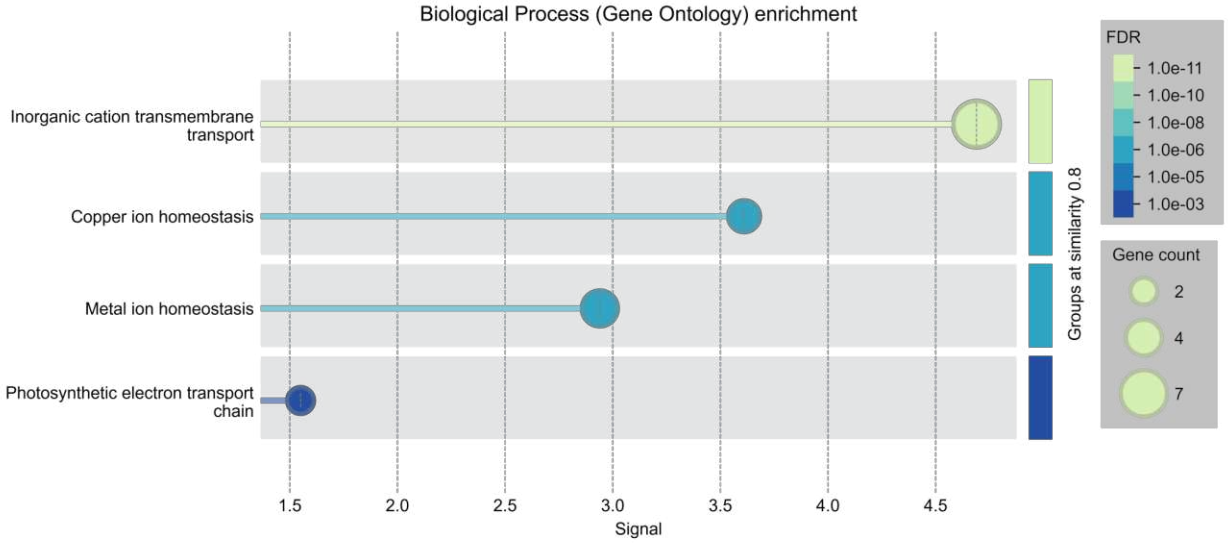


Fig. 8 GO biological-process enrichment of LcuHMA proteins. Bubble size indicates gene number, the x-axis shows enrichment signal, and bubble color represents FDR-adjusted significance

Phylogenetic co-occurrence analysis showed that LcuHMA proteins or their mapped orthologs are broadly conserved across bacterial, eukaryotic, and archaeal groups, with stronger conservation in Viridiplantae. Similar occurrence profiles among several proteins suggest evolutionary conservation of metal-transport functions; however, co-occurrence does not demonstrate common origin or direct protein interaction (Fig. 9).

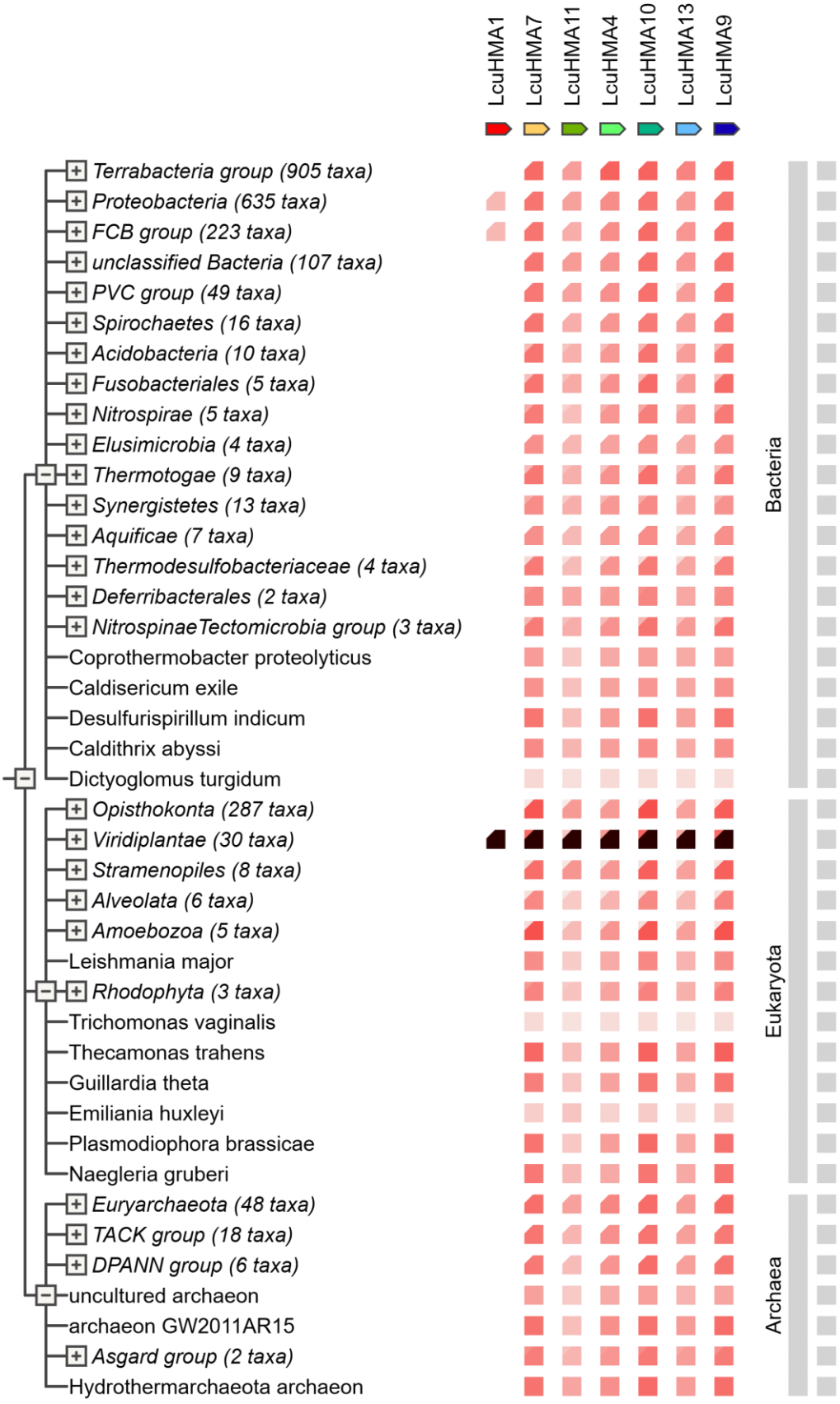

LcuHMA1
LcuHMA7
LcuHMA11
LcuHMA4
LcuHMA10
LcuHMA13
LcuHMA9
Terrabacteria group (905 taxa)
Proteobacteria (635 taxa)
FCB group (223 taxa)
unclassified Bacteria (107 taxa)
PVC group (49 taxa)
Spirochaetes (16 taxa)
Acidobacteria (10 taxa)
Fusobacteriales (5 taxa)
Nitrospirae (5 taxa)
Elusimicrobia (4 taxa)
Thermotogae (9 taxa)
Synergistetes (13 taxa)
Aquificae (7 taxa)
Thermodesulfobacteriaceae (4 taxa)
Deferribacterales (2 taxa)
NitrospinaeTectomicrobia group (3 taxa)
Coprothermobacter proteolyticus
Caldisericum exile
Desulfurispirillum indicum
Caldithrix abyssi
Dictyoglomus turgidum
Opisthokonta (287 taxa)
Viridiplantae (30 taxa)
Stramenopiles (8 taxa)
Alveolata (6 taxa)
Amoebozoa (5 taxa)
Leishmania major
Rhodophyta (3 taxa)
Trichomonas vaginalis
Thecamonas trahens
Guillardia theta
Emiliania huxleyi
Plasmodiophora brassicae
Naegleria gruberi
Euryarchaeota (48 taxa)
TACK group (18 taxa)
DPANN group (6 taxa)
uncultured archaeon
archaeon GW2011AR15
Asgard group (2 taxa)
Hydrothermarchaeota archaeon
Bacteria
Eukaryota
Archaea

Fig. 9 Phylogenetic co-occurrence profile of LcuHMA proteins. Colored cells show the evolutionary occurrence patterns of LcuHMA proteins or mapped orthologs across major taxonomic groups

LcuHMA genes expression analysis in lentil

HMA genes are essential for various plant processes, including the transport and homeostasis of heavy metals and micronutrients such as zinc, Cd, copper, chromium, and lead, as well as broader roles in metal detoxification, ion transport, growth and development, and stress tolerance. In this study, 13 *LcuHMA* gene family members were identified in lentil. Their expression patterns were investigated in the embryo (E1-E8), seed coat (S1-S8), and whole seed (W1-W8) across developmental stages, and in the control sample (UO). RNA-seq analysis showed that *LcuHMA* genes have distinct spatial and temporal expression patterns. *LcuHMA6* and *LcuHMA7* were highly expressed in nearly all tissues and stages, with *LcuHMA7* peaking sharply at seed coat stages S5-S8 and *LcuHMA6* peaking at the late embryo stage E8, suggesting a possible role in metal loading or remobilization during seed maturation. *LcuHMA1*, *LcuHMA2*, and *LcuHMA4* were moderately to highly expressed in all tissues and stages, indicating broad, constitutive-like expression (Fig. 10).

In contrast, *LcuHMA3* showed very low expression in most embryo and seed coat stages, with a modest increase only at late embryo stages (E7-E8) and late whole-seed stages (W7-W8), suggesting a more restricted, stage-specific function. *LcuHMA5* was variable: expression was low at E2-E8 and S5-S8, but higher at E1, S1-S4, W1-W4, and highest in UO. *LcuHMA11* was moderately expressed in early embryo stages, rose at E8, and increased sharply at seed coat stages

S5-S8, peaking at S7 and S8, before falling to moderate levels in whole-seed stages, pointing to a potential role in metal accumulation during seed coat development. *LcuHMA10* showed relatively uniform, moderate expression across all tissues and stages. *LcuHMA13* was expressed at low but detectable levels throughout. *LcuHMA8*, *LcuHMA9*, and *LcuHMA12* showed no detectable expression in any tissue or stage examined, suggesting these members may be silenced, pseudogenized, or expressed only under specific stress or metal-exposure conditions not captured in this dataset (Fig. 10).

The dynamic and tissue-specific expression patterns of the 13 *LcuHMA* genes, as revealed by the RNA-seq data, suggest that individual family members may play distinct, stage-specific roles in heavy metal transport and homeostasis during lentil embryo, seed coat, and whole-seed development, which could have implications for biofortification and metal tolerance breeding strategies in lentil.

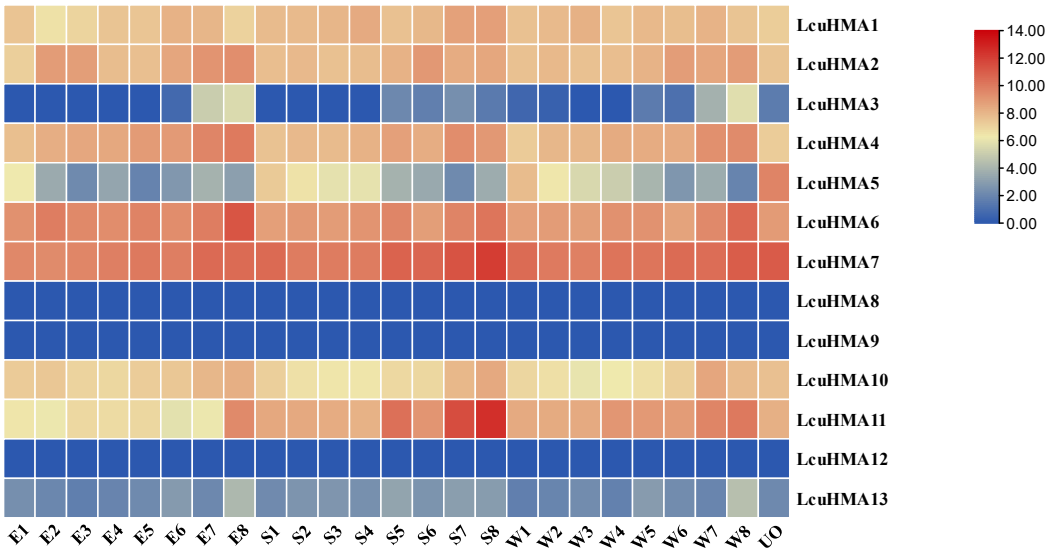


Fig. 10 Heatmap showing expression of 13 *LcuHMA* genes across samples. Rows represent genes, and columns represent eight developmental stages each for embryo (E1–E8), seed coat (S1–S8), and whole seed (W1–W8), plus the untreated control (UO). Blue indicates lower expression, while yellow to dark red indicates higher expression

LcuHMA gene expression analysis under Cd stress

Nine *LcuHMA* genes (*LcuHMA*1, 2, 3, 4, 5, 6, 7, 10, and 11) were selected based on their RNA-seq expression profiles, and their expression was analyzed by qRT-PCR in lentil seedlings exposed to 3 mM Cd for 2, 12, and 24 h, with untreated seedlings as the control (CK). Expression patterns differed among genes and time points, and all genes except *LcuHMA1*, which changed only slightly, responded clearly to Cd stress (Fig. 11). *LcuHMA3*, *LcuHMA4*, and *LcuHMA5* were up-regulated at all three time points. *LcuHMA3* increased progressively, reaching about 3.4-fold at 24 h, whereas *LcuHMA4* peaked early (about 2.5-fold at 2 h) and *LcuHMA5* peaked at 12 h (about 2.6-fold). *LcuHMA11* was strongly induced at 2 h (about 2.4-fold) and 24 h (about 3.0-fold) but returned to the control level at 12 h. In contrast, *LcuHMA7* and *LcuHMA10* were down-regulated at all time points. *LcuHMA7* declined steadily to about 0.4-fold at 24 h, while *LcuHMA10* reached its lowest level at 12 h (about 0.3-fold) and partially recovered at 24 h. *LcuHMA2* and *LcuHMA6* showed a biphasic pattern, with down-regulation at 2 and 12 h followed by up-regulation at 24 h. *LcuHMA1* was slightly reduced at 2 h and 24 h. Overall, the diverse and time-dependent expression patterns of the *LcuHMA* genes under Cd stress suggest that they have distinct roles in Cd uptake, transport, and detoxification in lentil. The primers used in this study are listed in Supplementary Table S6.

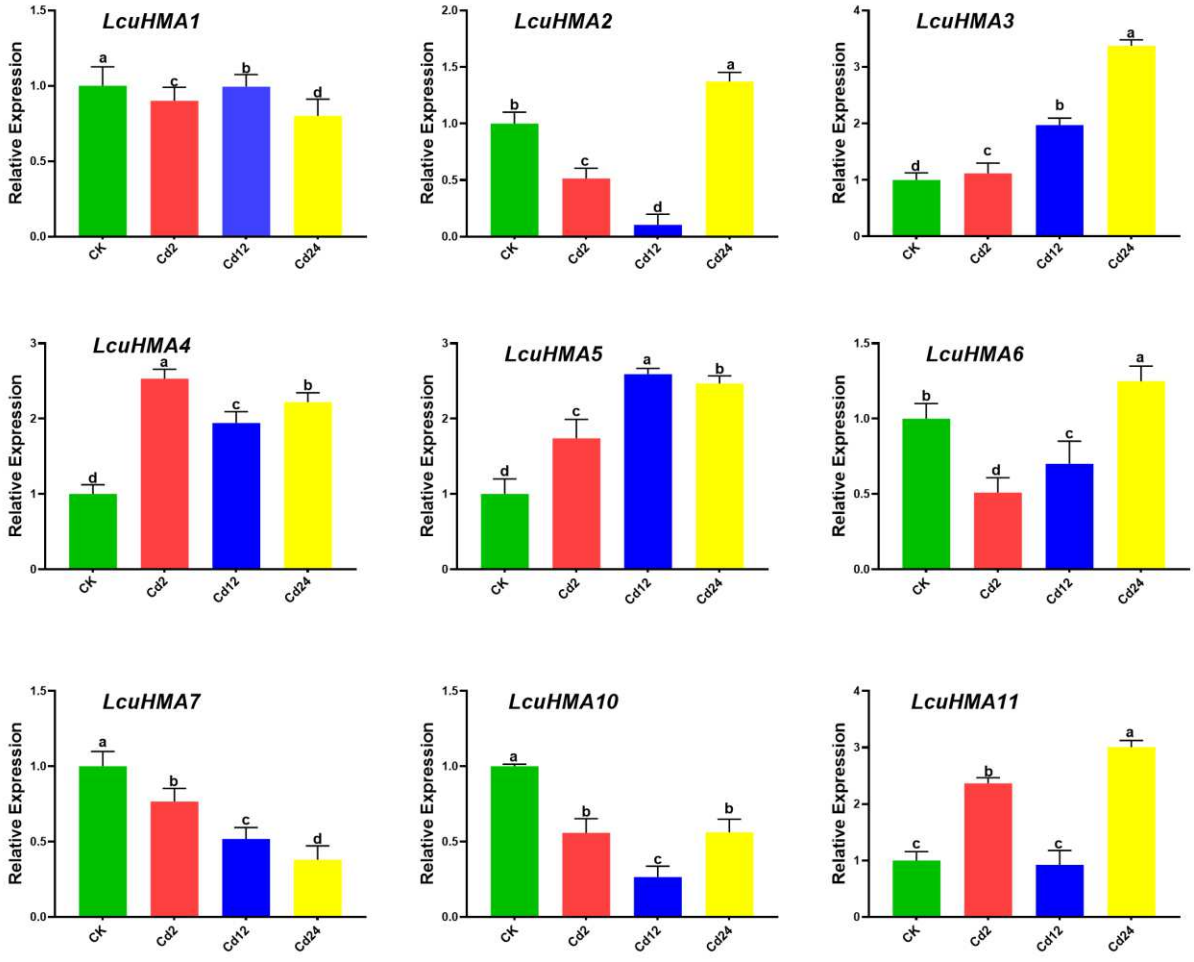


Fig. 11 Relative expression of nine *LcuHMA* genes in lentil seedlings following cadmium (Cd) exposure, measured by qRT-PCR. The x-axis shows the untreated control (CK, green) and samples collected after 2 h (Cd2, red), 12 h (Cd12, blue), and 24 h (Cd24, yellow) of Cd treatment. The y-axis shows relative transcript abundance. Bars show the mean of three biological replicates, each assayed in three technical replicates. Error bars represent standard deviations. Different lowercase letters indicate significant differences among treatments within each gene ($P \leq 0.05$)

Discussion

Bioinformatics tools have been used to characterize *HMA* gene families in several plant species, including *Arabidopsis thaliana*, *Oryza sativa*, *Glycine max*, and *Citrullus lanatus*. However, the *HMA* gene family in lentil has not yet been comprehensively characterized [54]. *HMA* genes participate in the transport and cellular distribution of metal ions, and some members contribute to the sequestration of potentially toxic metals [55,56]. Studies in *Arabidopsis* and *Oryza sativa* have reported changes in *HMA* gene expression during development and in response to Cd, Zn, or Cu exposure [57]. These findings provide a rationale for identifying lentil *HMA* genes and examining their expression under Cd stress, while their specific transport activities and contributions to tolerance remain to be tested.

The present study discovered 13 *LcuHMA* genes, designated as *LcuHMA1* to *LcuHMA13*, within the lentil genome database. The 13 identified *LcuHMA* genes were unevenly distributed across five of the seven lentil chromosomes, with two additional genes located on an unplaced genomic scaffold (unitig1160). No *LcuHMA* genes were detected on Chr1 or Chr6. Most genes (9 of 13) clustered within the proximal region of their respective chromosomes (approximately 0–100 Mb), while a smaller number were located in more distal regions, including *LcuHMA* genes on Chr3 (~380 Mb), Chr4 (~400–410 Mb), and Chr7 (~550 Mb) (Fig. 3). This clustered, uneven distribution pattern is broadly consistent with the non-random, unevenly clustered chromosomal distribution reported for *HMA* gene families in other plant genomes, including *Triticum aestivum* [58], *Arachis hypogaea* [59], *Areca catechu* [60], and *Stevia rebaudiana* [61].

The gene structure and conserved motif analysis of the *LcuHMA* gene sequences revealed distinct clade-specific patterns of exon-intron organization and conserved motif composition (Fig. 2). Genes clustering together in the phylogenetic tree shared highly similar motif arrangements and comparable exon-intron structures. For example, *LcuHMA5* and *LcuHMA10* (sister genes

within Group III) and *LcuHMA2* and *LcuHMA7* (Group IV), all of which belong to the Cu/Ag-transporting clade, shared highly similar motif arrangements (Motifs 4, 5, 8, 2, 7, 1, 10, 9, 6, and 3) and comparable exon-intron structures. In contrast, more divergent members of Group V, such as *LcuHMA1*, *LcuHMA12*, and *LcuHMA13*, retained only Motif 4 and displayed markedly simplified gene structures with fewer exons. This variation in motif composition and exon-intron structure across groups is consistent with functional divergence or subfunctionalization commonly observed in expanded gene families, and may warrant further functional characterization to determine whether *LcuHMA* members have acquired roles distinct from *HMA* genes previously characterized in crops such as *G. max* and *S. undatus* L. [19,62].

Both tandem and segmental duplication have contributed to the evolutionary diversification of *HMA* gene families [59]. In this study, the identification of 50 duplicated gene pairs among the 13 *LcuHMA* genes, comprising 3 tandem duplication events and 47 segmental duplication events, indicates that segmental duplication has been the predominant driver of *LcuHMA* gene family expansion in lentil, with tandem duplication also contributing to a lesser extent. Similarly, in peanut (*Arachis hypogaea*), one tandem duplication event was identified among 21 *AhHMA* genes, alongside ten whole-genome duplication-derived pairs [59], indicating that segmental duplication, rather than tandem duplication, is the dominant driver of *HMA* family expansion across most legume and crop species characterized to date. Accumulated evidence, including the findings of this study, supports the conclusion that duplication events — particularly segmental duplication — have been essential drivers of *HMA* gene family expansion across plant species.

For the phylogenetic analysis, we compared the protein sequences of 13 lentil LcuHMA members with 17 HMA proteins from *A. thaliana* and *O. sativa*. The *LcuHMA* genes were

categorized into two clades: Zn/Co/Cd/Pb-type and Cu/Ag-type, with the latter further divided into six groups: group-I to group-VI (Fig. 1). The Cu/Ag-type clade contained the most genes (11), with groups III and V each containing four genes, whereas the Zn/Cd/Co/Pb-type clade contained only two genes, both in group II, and group I contained no *LcuHMA* genes. A similar imbalance in *HMA* subfamily membership has been reported for *G. max* and *T. aestivum* [19,58]. This pattern is consistent with the unequal distribution observed among the lentil HMA groups. Analysis of protein sequences and differential expression profiles in lentil can identify candidate genes potentially involved in metal transport, growth and development, and abiotic stress responses. The regions upstream of *LcuHMA* coding sequences contain predicted *cis*-acting elements associated with light, stress, hormone responses, and development. Together, these findings suggest that *LcuHMA* genes may participate in metal-related stress responses, although their regulatory and transport functions require experimental validation. Similar predicted elements have been reported in *HMA* gene upstream regions in *Triticum aestivum* and *Sedum plumbizincicola* [58,33].

The predicted functional-association network and GO enrichment analyses indicate that LcuHMA proteins are mainly associated with inorganic cation transport, metal-ion homeostasis, and copper regulation [63]. The enrichment of photosynthetic electron transport likely reflects PAA1/HMA6- and PAA2/HMA8-like chloroplast copper ATPases, which may support copper delivery to photosynthetic components [64]. The connected network suggests functional similarity among several LcuHMA members, whereas the isolated LcuHMA11 may represent a divergent protein or one with limited available interaction evidence. The broad phylogenetic occurrence of LcuHMA orthologs, particularly in Viridiplantae, supports the evolutionary conservation of metal-transporting P-type ATPases [65]. However, these results are based on STRING and orthology-

supported predictions and require validation through expression analysis, localization prediction, and functional assays under heavy metal stress.

Expression profiling is a useful way to prioritize candidate genes for functional characterization (Fig. 10). The nine *LcuHMA* genes selected for qRT-PCR represent five of the six phylogenetic groups of the family (Fig. 1). *LcuHMA3* and *LcuHMA11* belong to the Zn/Cd/Co/Pb-type Group II, while the seven Cu/Ag-type genes fall into Group III (*LcuHMA5*, *LcuHMA10*), Group IV (*LcuHMA2*, *LcuHMA7*), Group V (*LcuHMA1*, *LcuHMA6*), and Group VI (*LcuHMA4*). Group I contain no *LcuHMA* genes. All nine genes responded to Cd stress, although *LcuHMA1* changed only slightly, and the direction, magnitude, and timing of the response differed markedly (Fig. 11). *LcuHMA3* (Group II), *LcuHMA4* (Group VI), and *LcuHMA5* (Group III) were up-regulated at all time points. *LcuHMA11* (Group II) was strongly induced at 2 h and 24 h (about 2.4- and 3.0-fold) but returned to the control level at 12 h.

Members of the same phylogenetic group often showed contrasting expression patterns under Cd stress (Fig. 11). In Group III, *LcuHMA5* was induced whereas *LcuHMA10* was repressed. In Group IV, *LcuHMA7* was repressed at all time points, whereas *LcuHMA2* was repressed early and rose above the control at 24 h. In Group V, *LcuHMA6* showed a biphasic response, whereas *LcuHMA1* was nearly unchanged. In Group II, both genes were induced but differed in timing: *LcuHMA3* increased progressively, whereas *LcuHMA11* returned to the control level at 12 h. Group VI was represented only by *LcuHMA4*. Similar divergence among closely related *HMA* genes has been reported in other species. In *Populus trichocarpa*, the *PtHMA5* paralogs differed in when their transcripts peaked after Cu exposure, and *PtHMA1* did not show the expression pattern of its *Arabidopsis* counterpart *AtHMA1* [66]. In soybean, *GmHMA19* and *GmHMA5* were highly expressed in leaves, unlike their counterpart *HMA1* in *Arabidopsis* and *Oryza sativa*, which

suggests functional divergence [19]. Such differences fit the known specialization of HMA proteins even within the Cu-type clade. In *Arabidopsis*, *HMA5* acts in root Cu detoxification, *HMA6* and *HMA8* are located in chloroplasts, and *HMA7* is located in the Golgi [21,67]. These patterns suggest that *LcuHMA* genes within the same group have diverged in regulation and probably in function. Whether this divergence arose after gene duplication, for example through subfunctionalization, should be tested by comparing the duplication history and promoter *cis*-elements of the divergent pairs.

Overall, these findings identify *LcuHMA* genes for further investigation of Cd responses in lentil. Their sequence features and expression patterns provide a basis for testing potential roles in metal uptake, transport, and sequestration, but those functions remain unconfirmed. Functional studies could determine whether any of these genes are useful targets for breeding or genome editing aimed at improving metal tolerance or limiting Cd accumulation in edible tissues.

Conclusions

We identified 13 putative *LcuHMA* genes in the lentil genome and characterized their phylogenetic relationships, gene structures, conserved motifs, and duplication patterns. Expression analysis showed that the family responds differently to Cd exposure. Among the nine genes examined by qRT-PCR, *LcuHMA3*, *LcuHMA4*, and *LcuHMA5* were up-regulated at every sampled post-treatment time point, whereas *LcuHMA11* was induced at 2 and 24 h. These four genes are priorities for functional studies to determine whether they contribute to Cd transport or tolerance. The findings establish a starting point for investigating HMA-mediated Cd responses in lentil; their value for breeding remains to be demonstrated.

Data availability

Datasets used in the current study are available from the corresponding author on reasonable request.

Contributions

FA, OUS, AA conceived and designed research, and conduct experiment; MAN analyzed the data; FA wrote the manuscript; FSB, MCD provided valuable suggestions on the manuscript; YSC, FSB, MCD revised the manuscript; All authors read and approved the manuscript.

Funding

The authors declare that no external funding was received for this study.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

[1] Ali F, Wang Z, Hu X, Li L, Zhao Y, Baloch FS, et al. GWAS-driven discovery and KASP validation of flowering time and seed weight loci in lentil (*Lens culinaris* Medik.). Int J Biol Macromol. 2026a;154304. https://doi.org/10.1016/j.ijbiomac.2026.154304.

[2] Liu S, Yin H, Pickard M, Ai Y. Influence of infrared heating on the functional properties of processed lentil flours: A study focusing on tempering period and seed size. Int Food Res. 2020;136:109568. https://doi.org/10.1016/j.foodres.2020.109568.

[3] Ren Y, Setia R, Warkentin TD, Ai Y. Functionality and starch digestibility of wrinkled and round pea flours of two different particle sizes. Food Chem. 2021;336:127711. https://doi.org/10.1016/j.foodchem.2020.127711.

[4] Liu S, Ren Y, Yin H, Nickerson M, Pickard M, Ai Y. Improvement of the nutritional quality of lentil flours by infrared heating of seeds varying in size. Food Chem. 2022;396:133649. https://doi.org/10.1016/j.foodchem.2022.133649.

[5] Jimenez-Lopez JC, Singh KB, Clemente A, Nelson MN, Ochatt S, Smith PM. Legumes for global food security. Front Plant Sci. 2020;11:926. https://doi.org/10.3389/fpls.2020.00926.

[6] Ali F, Zhao Y, Zeb S, Shah OU, Arif MAR, Ali A, et al. Screening lentil genotypes for chromium resilience: A study of growth characteristics, oxidative stress responses, and expression analysis. Environ Technol Innov. 2025;39:104318. https://doi.org/10.1016/j.eti.2025.104318.

[7] Hossain MS, Abdelrahman M, Tran CD, Nguyen KH, Chu HD, Watanabe Y, et al. Modulation of osmoprotection and antioxidant defense by exogenously applied acetate enhances cadmium stress tolerance in lentil seedlings. Environ Pollut. 2022;308:119687. https://doi.org/10.1016/j.envpol.2022.119687

[8] Samet H. Physiological Responses, Cadmium Partitioning, and Mineral Nutrient Disruption in *Brassicaceae* Crops Exposed to Cadmium Stress. Plants. 2026; 15:1019. https://doi.org/10.3390/plants15071019

[9] Kabata-Pendias A. Trace Elements in Soils and Plants. 4th ed. CRC Press; Boca Raton, FL, USA: 2011.

[10] Zulfiqar U, Jiang W, Xiukang W, Hussain S, Ahmad M, Maqsood MF, et al. Cadmium Phytotoxicity, Tolerance, and Advanced Remediation Approaches in Agricultural Soils: A Comprehensive Review. Front Plant Sci. 2022;13:773815. https://doi.org/10.3389/fpls.2022.773815.

[11] Tao J. Lu L. Advances in genes-encoding transporters for cadmium uptake, translocation, and accumulation in plants. Toxics. 2022;10:411. **https://doi.org/10.3390/toxics10080411**.

[12] Zhang X, Xue W, Zhang C, Wang C, Huang Y, Wang Y, et al. Cadmium pollution leads to selectivity loss of glutamate receptor channels for permeation of Ca^{2+}/Mn^{2+}/Fe^{2+}/Zn^{2+} over Cd^{2+} in rice plant. J Hazard Mater. 2023;452:131342. 10.1016/j.jhazmat.2023.131342.

[13] Zhang QH, Chen YQ, Li ZB, Tan XT, Xin GR, He CT. Defense guard: strategies of plants in the fight against cadmium stress. Adv Biotechnol. 2024;2:44. https://doi.org/10.1007/s44307-024-00052-6.

[14] Imran QM, Falak N, Hussain A, Mun BG, Sharma A, Lee SU, et al. Nitric oxide responsive heavy metal-associated gene *AtHMAD1* contributes to development and disease resistance in Arabidopsis thaliana. Front Plant Sci. 2016;7:1712. https://doi.org/10.3389/fpls.2016.01712.

[15] Zhang C, Yang Q, Zhang X, Zhang X, Yu T, Wu Y, et al. Genome-wide identification of the *HMA* gene family and expression analysis under Cd stress in barley. Plants. 2021;10:1849. **https://doi.org/10.3390/plants10091849**.

[16] He G, Qin L, Tian W, Meng L, He T, Zhao D. Heavy metal transporters-associated proteins in Solanum tuberosum: genome-wide identification, comprehensive gene feature, evolution and expression analysis. Genes. 2020;11:1269. **https://doi.org/10.3390/genes11111269**.

[17] Wu Y, Li X, Chen D, Han X, Li B, Yang Y, et al. Comparative expression analysis of heavy metal ATPase subfamily genes between Cd-tolerant and Cd-sensitive turnip landraces. Plant Divers. 2019;41:275-283. https://doi.org/10.1016/j.pld.2019.02.001.

[18] Smith AT, Smith KP, Rosenzweig AC, Diversity of the metal-transporting P1B-type ATPases. J Biol Inorg Chem. 2014;19:947-960. https://doi.org/10.1007/s00775-014-1129-2.

[19] Fang XL, Wang L, Deng XJ, Wang P, Ma QB, Nian H, et al. Genome-wide characterization of soybean P1B-ATPases gene family provides functional implications in cadmium responses. BMC Genom. 2016;17:376. https://doi.org/10.1186/s12864-016-2730-2.

[20] Batool TS, Aslam R, Gul A, Paracha RZ, Ilyas M, De Abreu K, et al. Genome-wide analysis of heavy metal ATPases (HMAs) in Poaceae species and their potential role against copper stress in *Triticum aestivum*. Sci Rep. 2023;13:1-13. https://doi.org/10.1038/s41598-023-32023-7.

[21] Williams LE, Mills RF. P1B-ATPases: an ancient family of transition metal pumps with diverse functions in plants. Trends Plant Sci. 2005;10:491-502. https://doi.org/10.1016/j.tplants.2005.08.008.

[22] Zhiguo E, Tingting L, Chen C, Lei W. Genome-wide survey and expression analysis of P 1B -ATPases in Rice, Maize and Sorghum. Rice Sci. 2018;25:208-17. https://doi.org/10.1016/j.rsci.2018.06.004.

[23] Zahra S, Shaheen T, Qasim M, Mahmood-Ur-Rahman, Hussain M, Zulfiqar S, et al. Genome-wide survey of *HMA* gene family and its characterization in wheat (*Triticum aestivum*) PeerJ. 2023;11:1-20. https://doi.org/10.7717/peerj.14920.

[24] Kim YY, Choi H, Segami S, Cho HT, Martinoia E, Maeshima M, et al. *AtHMA1* contributes to the detoxification of excess Zn (II) in Arabidopsis. Plant J. 2009;58:737-753. https://doi.org/10.1111/j.1365-313X.2009.03818.x.

[25] Satoh-Nagasawa N, Mori M, Nakazawa N, Kawamoto T, Nagato Y, Sakurai K, et al. Mutations in rice (*Oryza sativa*) heavy metal ATPase 2 (OsHMA2) restrict the translocation of zinc and cadmium. Plant Cell Physiol. 2012;53:213-224. https://doi.org/10.1093/pcp/pcr166.

[26] Fan W, Liu C, Cao B, Qin M, Long D, Xiang Z, et al. Genome-wide identification and characterization of four gene families putatively involved in cadmium uptake, translocation and sequestration in mulberry. Front Plant Sci. 2018;9:879. https://doi.org/10.3389/fpls.2018.00879.

[27] Morel M, Crouzet J, Gravot A, Auroy P, Leonhardt N, Vavasseur A, et al. *AtHMA3*, a P1B-ATPase allowing Cd/Zn/co/Pb vacuolar storage in Arabidopsis. Plant Physiol. 2009;149:894-904. https://doi.org/10.1104/pp.108.130294.

[28] Suzuki M, Bashir K, Inoue H, Takahashi M, Nakanishi H, Nishizawa NK, et al. Accumulation of starch in Zn-deficient rice. Rice. 2012;5:9. https://doi.org/10.1186/1939-8433-5-9.

[29] Ueno D, Yamaji N, Kono I, Huang, CF, Ando T, Yano M, et al. Gene limiting cadmium accumulation in rice. Proc Natl Acad Sci. 2010; 107:16500-16505. https://doi.org/10.1073/pnas.1005396107.

[30] Miyadate H, Adachi S, Hiraizumi A, Tezuka K, Nakazawa N, Kawamoto T, et al. OsHMA3, a P1B-type of ATPase affects root-to-shoot cadmium translocation in rice by mediating efflux into vacuoles. New Phytol. 2011;189:190-199. https://doi.org/10.1111/j.1469-8137.2010.03459.x.

[31] Deng F, Yamaji N, Xia J, Ma JF. A member of the heavy metal P-type ATPase *OsHMA5* is involved in xylem loading of copper in rice. Plant physiol. 2013;163:1353-1362. https://doi.org/10.1104/pp.113.226225.

[32] Yan J, Wang P, Wang P, Yang M, Lian X, Tang Z, et al. A loss-of-function allele of *OsHMA3* associated with high cadmium accumulation in shoots and grain of Japonica rice cultivars. Plant Cell Environ. 2016; 39:1941-1954. https://doi.org/10.1111/pce.12747.

[33] Huang Q, Qiu W, Yu M, Li S, Lu Z, Zhu Y, et al. Genome-wide characterization of *Sedum plumbizincicola HMA* gene family provides functional implications in cadmium response. Plants. 2022;11:215. https://doi.org/10.3390/plants11020215.

[34] Liu H, Zhao H, Wu L, Liu A, Zhao FJ, Xu W. Heavy metal ATPase 3 (*HMA3*) confers cadmium hypertolerance on the cadmium/zinc hyperaccumulator *Sedum plumbizincicola*. New Phytol. 2017;215:687-698. https://doi.org/10.1111/nph.14622.

[35] Feng S, Shen Y, Xu H, Dong J, Chen K, Xiang Y, et al. RNA-Seq identification of Cd responsive transporters provides insights into the association of oxidation resistance and Cd accumulation in Cucumis sativus L. Antioxidants. 2021;10:1973. https://doi.org/10.3390/antiox10121973.

[36] Mikkelsen MD, Pedas P, Schiller M, Vincze E, Mills RF, Borg S, et al. Barley *HvHMA1* is a heavy metal pump involved in mobilizing organellar Zn and Cu and plays a role in metal loading into grains. PLoS On. 2012;7:e49027. https://doi.org/10.1371/journal.pone.0049027.

[37] Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: A comparative platform for green plant genomics. Nucleic Acids Res. 2011;40:D1178–86. https://doi.org/10.1093/nar/gkr944.

[38] Baxter I, Tchieu J, Sussman MR, Boutry M, Palmgren MG, Gribskov M, et al. Genomic comparison of P-type ATPase ion pumps in Arabidopsis and rice. Plant Physiol. 2003;132:618-628. https://doi.org/10.1104/pp.103.021923.

[39] Zhang J, Zhang M, Shohag MJ, Tian S, Song H, Feng Y, et al. Enhanced expression of SaHMA3 plays critical roles in Cd hyperaccumulation and hypertolerance in Cd hyperaccumulator Sedum alfredii Hance. Planta. 2016;243:577-89. https://doi.org/10.1007/s00425-015-2429-7.

[40] Blum M, Andreeva A, Florentino LC, Chuguransky SR, Grego T, Hobbs E, et al. InterPro: the protein sequence classification resource in 2025. Nucleic Acids Res. 2025;53: D444-D456. https://doi.org/10.1093/nar/gkae1082.

[41] Azimi R, Ozgul M, Kenney MC, Kuppermann BD. Bioinformatic analysis of small humanin like peptides using AlfaFold-2 and Expasy ProtParam. Invest Ophthalmol Vis Sci. 2024;65:1320. https://iovs.arvojournals.org/article.aspx?articleid=2796857.

[42] Chen C, Li F, Xie F, Chen J, Hua Q, Chen J, et al. Pitaya genome and multiomics database (PGMD): a comprehensive and integrative resource of *Selenicereus undatus*. Genes. 2022;13:745. https://doi.org/10.3390/genes13050745.

[43] Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol. Biol. Evol. 2021; 38:3022-3027. https://doi.org/10.1093/molbev/msab120.

[44] Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47:W256-W259. https://doi.org/10.1093/nar/gkz239.

[45] Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009; 37:W202-W208. https://doi.org/10.1093/nar/gkp335.

[46] Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol. Plant. 2020; 13:1194-1202. https://doi.org/10.1016/j.molp.2020.06.009.

[47] Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19:1639-1645. https://doi.org/10.1101/gr.092759.109.

[48] Ali F, Zhao Y, Shah OU, Nadeem MA, Baloch FS, Madenova A, et al. Identification and expression profiling of the *GRAS* gene family reveals candidate genes for enhancing abiotic stress tolerance in lentil. BMC Plant Biol. 2026b;26:245. https://doi.org/10.1186/s12870-025-08012-5.

[49] Shah OU, Peng J, Zhou L, Khan WU, Shanshan Z, Zhuyu P, et al. Comparative omics-based characterization, phylogeny and melatonin-mediated expression analyses of *GDSL* genes in pitaya (*Selenicereus undatus* L.) against multifactorial abiotic stresses. Physiol Mol Biol Plants. 2024; 30:1493-1515. https://doi.org/10.1007/s12298-024-01506-w.

[50] Main DS, Jung S, Coyne CJ, McGee RJ. Pulse crop database resources. 2017.

[51] Yu B, Gao P, Song J, Yang H, Qin L, Yu X, et al. Spatiotemporal transcriptomics and metabolic profiling provide insights into gene regulatory networks during lentil seed development. Plant J. 2023;115:253-74. https://doi.org/10.1111/tpj.16205.

[52] Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002; 30:325-327. https://doi.org/10.1093/nar/30.1.325.

[53] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. Methods. 2001;25:402-408. https://doi.org/10.1006/meth.2001.1262.

[54] Flores-Iga G, Lopez-Ortiz C, Gracia-Rodriguez C, Almeida A, Nimmakayala P, Reddy UK, et al. A genome-wide identification and comparative analysis of the heavy-metal-associated gene family in cucurbitaceae species and their role in Cucurbita pepo under arsenic stress. Genes. 2023; 14:1877. https://doi.org/10.3390/genes14101877.

[55] Rodrigues WFC, Pires LO, Oliveira LR, Ishida JK. Comprehensive characterization of HMA transporters in common bean: Tissue-specific expression and response to metal stress. Crop Sci. 2025; 65:e70188. https://doi.org/10.1002/csc2.70188.

[56] Bhat BA, Rather MA, Bilal T, Nazir R, Qadir RU, Mir RA. Plant hyperaccumulators: a state-of-the-art review on mechanism of heavy metal transport and sequestration. Front Plant Sci. 2025; 16:1631378. https://doi.org/10.3389/fpls.2025.1631378.

[57] Cong W, Miao Y, Xu L, Zhang Y, Yuan C, Wang J, et al. Transgenerational memory of gene expression changes induced by heavy metal stress in rice (*Oryza sativa* L.). BMC Plant Biol. 2019;19:282. https://doi.org/10.1186/s12870-019-1887-7.

[58] Shao L, Guo X, Wang Y, Lei T, Wan Y, Ma L, et al. Genome-wide survey of the HMA gene family in wheat (*Triticum aestivum* L.) and its potential role in cadmium stress. BMC Genom. 2025; 26:619. https://doi.org/10.1186/s12864-025-11746-z.

[59] Li J, Zhang Z, Shi G. Genome-wide identification and expression profiling of heavy metal ATPase (*HMA*) genes in peanut: Potential roles in heavy metal transport. Int J Mol Sci. 2024;25:613. https://doi.org/10.3390/ijms25010613.

[60] Khan NM, Ali A, Wan Y, Zhou G. Genome-wide identification of heavy-metal ATPases genes in *Areca catechu*: investigating their functionality under heavy metal exposure. BMC Plant Biol. 2024;24:484. https://doi.org/10.1186/s12870-024-05201-6.

[61] Xu D, Yu H, Xu J, Zhao Y, Teng J, Zhang G, et al. Genome-wide identification and expression analysis of the Heavy-Metal ATPase (*HMA*) gene family reveal their correlation with copper and cadmium stress adaptation in *Stevia rebaudiana*. Cells. 2026;15:1646. https://doi.org/10.3390/cells15181646.

[62] Zaman QU, Khan LU, Hussain MA, Ali A, Hui L, Khokhar AA, et al. Characterizing the *HMA* gene family in dragon fruit (*Selenicereus undatus* L.) and revealing their response to multifactorial stress combinations and melatonin-mediated tolerance. S Afr J Bot. 2023;163:145-156. https://doi.org/10.1016/j.sajb.2023.10.039.

[63] Szklarczyk D, Nastou K, Koutrouli M, Kirsch R, Mehryary F, Hachilif R, et al. The STRING database in 2025: Protein networks with directionality of regulation. Nucleic Acids Res. 2025;53:D730-D737. https://doi.org/10.1093/nar/gkae1113.

[64] Abdel-Ghany SE, Müller-Moulé P, Niyogi KK, Pilon M, Shikanai T. Two P-type ATPases are required for copper delivery in *Arabidopsis thaliana* chloroplasts. Plant Cell. 2005; 17:1233-1251. https://doi.org/10.1105/tpc.104.030452.

[65] Hanikenne M, Baurain D. Origin and evolution of metal P-type ATPases in Plantae (Archaeplastida). Front Plant Sci. 2014;4:544. https://doi.org/10.3389/fpls.2013.00544.

[66] Li D, Xu X, Hu X, Liu Q, Wang Z, Zhang H, et al. Genome-wide analysis and heavy metal-induced expression profiling of the *HMA* gene family in *Populus trichocarpa*. Front Plant Sci. 2015;6:1149. https://doi.org/10.3389/fpls.2015.01149.

[67] Kobayashi Y, Kuroda K, Kimura K, Southron-Francis JL, Furuzawa A, Kimura K, et al. Amino acid polymorphisms in strictly conserved domains of a P-type ATPase HMA5 are involved in the mechanism of copper tolerance variation in Arabidopsis. Plant Physiol. 2008;148:969-980. https://doi.org/10.1104/pp.108.119933.

Supplementary Files

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

- TableS1.docx
- TableS4.xlsx
- TableS3.docx
- TableS2.docx
- TableS6.docx
- TableS5.xlsx