

Research Article



Genome-wide Screening of Zinc-Responsive Genes in *Arabidopsis thaliana* for Enhanced Phytoremediation Potential

Mohammed Salih AL-Janaby¹, Yousry Sadoon Rasheed², Mohammed Torki Hammood¹, Mustafa M. Fahad^{3*}

¹College of Science, Department of Biotechnology, University of Anbar, Ramadi 31001, Iraq

²College of Science, Department of Biology, University of Anbar, Ramadi 31001, Iraq

³Center of Desert Studies, University of Anbar, Ramadi 31001, Iraq

ARTICLE INFO

Article history:

Received October 23, 2024

Received in revised form January 10, 2025

Accepted January 30, 2025

KEYWORDS:

Arabidopsis thaliana,
Genomic Screening,
Heavy Metal Detoxification,
Phytoremediation,
Zinc Uptake

ABSTRACT

Zinc contamination presents considerable threats to human health, agricultural practices, and ecological systems. *Arabidopsis thaliana* has demonstrated potential for use in phytoremediation, which is an environmentally friendly strategy aimed at addressing this type of pollution. This study identified genes associated with phytoremediation in *A. thaliana*. This identification was achieved through comparative bioinformatics analysis involving 27 plant species. A critical literature review was conducted. This review aimed to collect genes involved in metal uptake and transport across various hyperaccumulator plant species. A database in BLAST format was generated to enable the study of main homologous protein families, including ATP-binding Cassette transporters, acyl-CoA-binding proteins, metallothioneins, and phytochelatin synthases. *In silico* screening for zinc-binding was performed by means of a web application based on the deep-learning framework called GNINA. The results highlighted the fact that within the phytochelatin synthases-especially PCS1 and PCS2, the protein PCS2 has the highest affinity to zinc. Heavy metal ATPases involved in the vital process of metal transport and homeostasis include HMA1. Zinc transporters responsible for zinc uptake and sequestration include cation exchanger 1CAX1 and IRT2. The information thus provided explains in detail the complex molecular machinery regulating metal detoxification in *A. thaliana* and provides insights for developing transgenic plants with enhanced phytoremediation capabilities. Future research should target these genes to create plants with improved metal tolerance and accumulation, advancing phytoremediation technologies.



Copyright (c) 2025@ author(s).

1. Introduction

Soil, water, and air degradation due to heavy metals introduces environmental pollution; it remains a problem of high priority in many parts of the world, creating serious hazards to human health (Rasheed *et al.* 2024), agricultural productivity, and ecological balance. Among these pollutants, heavy metals such as cadmium (Cd), chromium (Cr), lead (Pb), arsenic (As), copper (Cu), and zinc (Zn) occur naturally in soils through geological and biological processes (Mitra *et al.* 2022).

However, human activities such as industrial emissions, agricultural runoff, and waste disposal have significantly increased the concentration of these metals, leading to widespread contamination (Sofa *et al.* 2012). The wastewater released from the textile industries contains Zn, Cd, Pb, and Cu-heavy metals, which are usually present in the colorants used by the industry. This results in contamination of the environment with great risks to human health, such as carcinogenicity (Hussain *et al.* 2020), neurological damage, or even organ failure when these pollutants enter the water system. Further, they also pollute water by decreasing dissolved oxygen, hence water's ability to support photosynthesis in water plants (Singh *et al.* 2024).

* Corresponding Author

E-mail Address: mustafa.moh@uoanbar.edu.iq

Although Zn is an essential micronutrient required for a wide range of biological processes, its heavy use, contaminating the soil from industrial processes and discharge of wastewater, is toxic (Monserrath *et al.* 2024). Besides, high Zn is not fertile for the soil and the growth of plants and poses a lethal health hazard in humans ranging from nausea, stomach ache, diarrhea, and fever (García-García *et al.* 2020; Hussain *et al.* 2022; Agni, n.d.). As such, the natural and anthropogenic releases of Zn are capable of impacting the global availability and uptake of other essential nutrients as it elicits multiple toxic effects across ecosystems.

Because of this challenge, phytoremediation-the use of plants to remove or render pollutants from soil innocuous-has emerged as a promising and environmentally benign technology. Although many types of bacterial and enzymatic processes have been utilized for the degradation of textile dyes, these processes are usually highly costly and complex in their application. In this context, plants have emerged as a sustainable, eco-friendly, and low-cost solution with the natural potential for pollutant uptake and subsequent detoxification of these heavy metals from soil and water (the phytoremediation) (Faizan *et al.* 2024; Hayashi *et al.* 2020). Plant mechanisms that allow the uptake, transport, and detoxification of Zn in these contaminated environments include a variety of metalloproteins and enzymes (zinc transporters and detoxifying proteins). In *Arabidopsis thaliana*, ZIP (Zinc/Iron-regulated transporter-like Protein) and HMA (Heavy Metal ATPase) are the types of enzymes that have been reported to supply Zn uptake and detoxification.

For this purpose, *A. thaliana* presents as a model organism widely used in plant biology for its genetic properties, e.g., self-fertilization, short life span, and heavy metal toxicity tolerance (Agni, n.d.; Chaffai & Koyama 2011). In addition, *A. thaliana* offers insights into phytoremediation processes, including metal uptake, transport, and detoxification, for example, of Zn (Wei *et al.* 2021; Amin *et al.* 2023). Although general phytoremediation mechanisms have been studied, a detailed understanding of the specific genes involved in Zn detoxification in *A. thaliana* remains limited. Many of these transporter proteins in *A. thaliana* are highly conserved during evolution and are crucial in metal soil uptake, accumulation, and translocation towards aerial tissues to allow detoxification and sequestration from the environment (Andolfo *et al.* 2015; Verma *et al.* 2017; Wang *et al.* 2021). Knowing these proteins also makes it possible to genetically modify many plant

species to produce plants with higher phytoremediation capabilities.

Although previous research has underscored the significance of the identified phytoremediation trait on contamination by heavy metals, the molecular processes governing Zn detoxification and transport in *A. thaliana* remain elusive (Wang *et al.* 2021; Escudero *et al.* 2022). Additionally, understanding such genes is a critical missing link because the identification and characterization of such genes may contribute greatly to phytoremediation improvement. Such information could potentially be used to fill this knowledge gap and improve the biosynthesis of plants to obtain higher species for the absorption of heavy metals, including zinc, from affected soils. Such enhancements may be useful for backing environmentally friendly remediation procedures and boosting the qualities of the soil to favor crop production for sustainable food production (Kumar *et al.* 2023). To achieve this objective, this work will follow a functional genomics approach to screen the *A. thaliana* genome and identify core genes involved in Zn uptake, transport, and detoxification so as to provide a basis for future phytoremediation research to improve the efficiency of soil remediation processes.

2. Materials and Methods

2.1. Identifying Phytoremediation Genes Across Plant Species

An extensive literature review was conducted to identify genes involved in phytoremediation processes across various plant species (*Arabidopsis halleri*, *Alyssum lesbiacum*, *Arabidopsis thaliana*, *Arundo donax*, *Azolla pinnata*, Barley (*Hordeum vulgare*), *Brassica napus*, *Cicer arietinum*, *Morus notabilis*, *Nicotiana tabacum*, *Populus tomentosa*, *Populus trichocarpa*, *Prosopis juliflora*, *Pteris vittata*, *Psychotria gabriellae*, *Sedum alfredii*, *Sesbania rostrata*, *Suaeda salsa*, Sweet potato (*Ipomoea batatas*), *Thlaspi caerulescens*, *Triticum aestivum* (wheat), *Vicia sativa* and *Zea mays*), specifically focusing on metal uptake and transport. The review only used two databases, namely Google Scholar and PubMed. Using a mix of subject terms with free-text keywords was the formulated search strategy; keywords used in the study include; “phytoremediation”, “metal absorption”, “transport”, “sequestration”, “detoxification”, “heavy metals genes”, “phytoanalysis”. The genes identified through this review are listed in Supplementary File 1.

2.2. Shortlisting of Phytoremediation-related Proteins in *A. Thaliana*

Genome annotation data for *A. thaliana* was obtained in GTF (Gene Transfer Format) from the National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/>) (Sayers *et al.* 2022). Symbols of phytoremediation-related genes from various hyperaccumulator plants (given above), identified through a comprehensive literature search, were compared against the *A. thaliana* genome annotation file to identify common genes potentially involved in phytoremediation functions. Subsequently, protein sequences for these genes were retrieved from UniProt (<https://www.uniprot.org/>), and their respective domain sequences were obtained from InterPro (<https://www.ebi.ac.uk/interpro/>) for both *A. thaliana* and other phytoremediation-capable plant species (Lussi *et al.* 2023). Moreover, the entire *A. thaliana* proteome was also retrieved from UniProt to identify putative phytoremediation-related proteins that are not fully annotated in the annotation GTF file.

2.3. BLAST Sequence Analysis and Homologous Gene Identification

To identify homologous proteins with potential phytoremediation functions, a custom BLAST (Basic Local Alignment Search Tool) database was constructed using the proteome of *A. thaliana* (Nishimura *et al.* 2024). From other species and *A. thaliana*, the sequences of the homologous phytoremediation genes to known and reported genes were used to probe the custom database; their corresponding domain sequences are also listed in Supplementary File 1. These proteins were then analyzed using BLAST to find homologous proteins that were used in earlier steps to identify proteins with potential phytoremediation functions. The command line tool "blastp" was employed with the following parameters: maximum 5 alignments, e-value 1e-30. The results were summarized in tabular format, and the hit IDs were also stored in text format in Supplementary File 1. To perform deeper analysis, using the 'blastdbcmd' command-line tool from the BLAST database generated earlier, sequences corresponding to the hit IDs were obtained.

2.4. Protein Structure Retrieval and Modeling

Protein structures for the shortlisted homologous proteins were retrieved from the Protein Data Bank (PDB) (<https://www.rcsb.org/>) management using the National

Resource for Biomolecular Structure Determination appointed at www.rcsb.org, which offers access to a database of experimentally inferred biomolecular structures (Burley *et al.* 2023). These structures were manually optimized using the University of California, San Francisco Chimera software, which is used for molecular visualization and manipulation (Shadidizaji *et al.* 2024). For regions lacking models, MODELLER (<https://salilab.org/modeller/>), and the best model obtained from the ones generated was chosen from the DOPE score (Webb & Sali 2017). For structures not available in the PDB, AlphaFold (<https://alphafold.ebi.ac.uk/>) and were used to derive the structures (Jumper *et al.* 2021). The structure of the ligand zinc (CID: Preliminary data on the chemical structure of compound Zi-Shen-Hong-Tang (CAS reg. No. 23994) was obtained in the PubChem data base (<http://pubchem.ncbi.nlm.nih.gov/>). To achieve this, the standard SMILES code of Zn was translated to PDB format using the chemical informatics tool called the RDKit, which is implemented in Python.

2.5. Virtual Screening Against Zn

Virtual screening was conducted using GNINA, a molecular docking program that incorporates convolutional neural networks to score and optimize ligand interactions, to assess the binding potential of the selected protein structures with zinc (Zn) (<https://github.com/gnina/gnina>) (McNutt *et al.* 2021). The protein structures of potential phytoremediation-related proteins were first minimized using the OpenMM toolkit (<https://openmm.org/>), which optimizes protein structures for more accurate docking results. These minimized protein structures were then docked against the Zn ligand. A blind molecular docking strategy was utilized, encompassing the entire protein structure as the grid box. The docking region was specified in the GNINA tool using the "autobox_ligand" parameter, which was configured based on the protein's PDB file. The grid box coordinates facilitated the molecular docking process, allowing the GNINA tool to assess the interaction between the proteins and Zn. The top complexes were selected based on the highest binding affinity. PyMOL (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC) was used for visualization of the docked complexes (Yuan *et al.* 2017).

3. Results

3.1. Identification and Categorization of Key Genes for Metal Management in Plants

A thorough literature review was conducted to identify genes essential for metal uptake, transport, sequestration, and detoxification across diverse plant species. The review highlighted 97 genes involved

in phytoremediation processes in 27 different plants. These genes fall into various families and play crucial roles in metal management. The identified gene families include ATP-binding cassette transporters, acyl-CoA-binding proteins, and metallothioneins, among others. The function of each gene and its associated plant species are detailed in Table 1.

Table 1. Gene families and respective plant species for identified genes

Family	Genes	Organisms
ATP-binding cassette transporters	<i>ABCC1, ABCG36, ABCG43</i>	<i>Populus trichocarpa, Oryza sativa</i>
Acyl-CoA-binding proteins	<i>ACBP1, ACBP4</i>	<i>A. thaliana</i>
Acyl-CoA reductases	<i>ACR2, ACR3</i>	<i>A. thaliana, Pteris vittata</i>
Auxin-related	<i>AUX1</i>	<i>Oryza sativa</i>
Calmodulin-like proteins	<i>CAL1, CAL2</i>	<i>Oryza sativa</i>
Cation exchangers	<i>CAX1</i>	<i>A. halleri</i>
Chromatin-binding proteins	<i>CBP4</i>	<i>Nicotiana tabacum</i>
Cyclic nucleotide-gated ion channel	<i>CNGCI</i>	<i>A. thaliana</i>
Iron-regulated transporters	<i>IREG1, IRT1, IRT2, IRTI</i>	<i>Psychotria gabriellae, Oryza sativa</i>
Expansins	<i>EXPA12</i>	<i>Populus tomentosa</i>
Fatty acid CoA ligase	<i>FC1</i>	<i>A. thaliana</i>
Homeobox proteins	<i>HB1, HB4</i>	<i>Nicotiana tabacum, Oryza sativa</i>
Heavy metal-associated domain-containing proteins	<i>HMA1, HMA2, HMA3, HMA4, HMA5</i>	<i>Sedum alfredii, A. halleri, Populus tomentosa</i>
Hydroxymethyltransferase	<i>HMT</i>	<i>Prosopis juliflora</i>
Heat shock transcription factor	<i>HSFA4A</i>	<i>Oryza sativa</i>
Phytoene synthase	<i>PSE1</i>	<i>A. thaliana</i>
Laccases	<i>LCD</i>	<i>Oryza sativa</i>
Low-affinity cation transporter	<i>LCT1</i>	<i>Oryza sativa</i>
Chlorophyll a-b binding protein	<i>LHCB2</i>	<i>Sedum alfredii</i>
Mannan endo-1,4-beta-mannosidase	<i>MAN3</i>	<i>A. thaliana</i>
Mercuric transport protein	<i>MERC-SYP121</i>	<i>A. thaliana</i>
Metallothioneins	<i>MT1, MTIE, MT2, MT2B, MT2C</i>	<i>Prosopis juliflora, Oryza sativa, Suaeda salsa, Iris lactea</i>
Zinc transporter proteins	<i>ZIP3, ZIP6, ZIP1, ZIP5, ZIP7, ZIP9</i>	<i>A. halleri, A. thaliana, Oryza sativa</i>
Metal tolerance protein	<i>MTP1</i>	<i>Oryza sativa</i>
Mannitol-specific enzyme IIA component	<i>MUSI</i>	<i>Sweet potato</i>
MYB transcription factor	<i>MYB4</i>	<i>Oryza sativa</i>
Nicotianamine synthases	<i>NAS2, NAS3, NAS4</i>	<i>A. halleri</i>
NAD(P)H-quinone oxidoreductase	<i>NHAD</i>	<i>Salvinia minima</i>
Nitrate reductase [NADH]	<i>NR3C1</i>	<i>Nicotiana tabacum</i>
Natural resistance-associated macrophage proteins	<i>NRAMP1, NRAMP2, NRAMP3, NRAMP4, NRAMP5, NRAMP6</i>	<i>Oryza sativa, A. thaliana, Hordeum vulgare, Sedum alfredii</i>
Purple acid phosphatase	<i>PCR1</i>	<i>Oryza sativa</i>
Phytochelatin synthases	<i>PCS1, PCS15, PCS2, PCS3, PCS5, PCSI</i>	<i>Arundo donax, Oryza sativa, Morus notabilis</i>
Pleiotropic drug resistance protein	<i>PDR8</i>	<i>A. thaliana</i>
Phosphate transporters	<i>PHT1, PHT7</i>	<i>A. thaliana</i>
tRNA methyltransferases	<i>CARMT, RICMT</i>	<i>Cicer arietinum, Oryza sativa</i>
Stress-associated protein	<i>SAP</i>	<i>Lobularia maritima</i>
Ubiquitin carboxyl-terminal hydrolases	<i>UBP15, UBP16, UBP19</i>	<i>Zea mays</i>
Vacuolar H(+)-pyrophosphatase	<i>VPI</i>	<i>Triticum aestivum</i>
Xanthine dehydrogenase/oxidase	<i>XCD1</i>	<i>A. thaliana</i>
Yellow Stripe-like ATP transporter genes	<i>YSL3, YSL5, YSL7</i>	<i>Thlaspi caerulescens</i>
Zinc transporters	<i>ZIP1, ZIP5, ZIP7, ZIP9</i>	<i>A. thaliana, Oryza sativa, A. halleri</i>
Tryptophan-rich sensory protein	<i>ZTP1</i>	<i>Thlaspi caerulescens</i>

3.2. Sequence Retrieval, Alignment, and Identification of Phytoremediation-Related Proteins in *A. thaliana*

Phytoremediation-related genes from different hyperaccumulator plants were compared with the *A. thaliana* genome. This comparison identified 325 similar phytoremediation genes in *A. thaliana*. Notably, the products of these genes exhibited multiple domains implicated in phytoremediation processes, such as A20-type, ABC transmembrane (type-1, type-1-1, type-1-2, type-2-1, type-2-2), ABC transporter (transporter-1, transporter-2), ACB, AN1-type, Cyclic nucleotide-binding, Expansin-like CBD, Expansin-like EG45, FAD-binding PCMH-type, HMA, HMA1, HMA2, HMA3, HTH myb-type 1, HTH myb-type 2, LRAT, MYND-type, Peptidase C83, Rhodanese, Sodium/calcium exchanger membrane region, t-SNARE coiled-coil homology, and USP. Additionally, these domain names were used to find proteins in the *A. thaliana* genome that are linked to gene products based on the domain names. This process led to the identification of 37 more homologous genes. Protein sequences for these 325 and 37 genes were collected from UniProt.

Moreover, protein sequences for the phytoremediation-related genes from other plant species (mentioned above) were retrieved, including a total of 41 proteins and 71 additional sequences corresponding to their respective domains; however, it should be noted that only those sequences were retrieved which were available on UniProt for the given species. Finally, these sequences were compiled and queried against a custom-built *A. thaliana* BLAST database to identify potential phytoremediation-related proteins that are not fully annotated in the genome annotation file. This search produced 75 hits, which were used to retrieve the protein sequences for further analysis. Furthermore, after removing redundantly obtained proteins from different sources (BLAST, GTF, domains, and literature review), the unique list of proteins was further utilized, as given in Supplementary File 1.

3.3. Virtual Screening and Domain Analysis of Phytoremediation Proteins

Protein structures for the shortlisted homologous proteins were obtained if they were available through experimentally determined methods (such as NMR or X-ray crystallography) or computational models (such as AlphaFold). Structures were accessible for only 93 proteins.

Protein structures of phytoremediation-related genes were screened for their binding affinity to Zn. Among the top complexes, the proteins PCS2, PCS1, EXPA12, ZIP10, CCOAOMT1, DCL1, HMA1, CAX1, USPL1, and IRT2 demonstrated high binding affinities. Specifically, PCS2 exhibited the highest binding affinity with a score of -1.77, while the other proteins showed affinities ranging from -1.72 to -1.47. Selection of the top proteins was based on the least binding energy.

The functional characteristics of these proteins were further examined through domain analysis using the InterPro database. The analysis revealed that ZIP10 and IRT2 did not have any identified domains. In contrast, other top-scoring proteins displayed various functional domains. For example, PCS1 and PCS2 both contain phytochelatin synthase domains crucial for metal detoxification. EXPA12 has domains related to expansin activity and cellulose binding, while CCOAOMT1 features a SAM-dependent O-methyltransferase domain. DCL1 showed multiple domains, including ribonuclease III and PAZ domains, which are important for RNA processing and regulation. HMA1, CAX1, and USPL1 also possess specific domains relevant to metal transport and stress responses. Table 2 summarizes the binding affinities of these proteins for zinc, showing the range and specific scores for each protein. Table 3 details the domains identified in the top-scoring proteins, providing insight into their functional roles and potential contributions to phytoremediation. Zn interactions with the top proteins are given in Figure 1.

The PCS1 protein demonstrated binding with zinc at residues 398A (THR), 409A (THR), 397A (SER), and 399A (GLN) within the Phytochelatin Synthase C-terminal domain (231-465). The HOG1

Table 2. Binding affinities of top phytoremediation genes against Zn

UniProt IDs	Zinc (Zn)	
	Genes	Affinity
Q9ZWB7	<i>PCS2</i>	-1.77
Q9LDJ3	<i>EXPA12</i>	-1.72
Q8W245	<i>ZIP10</i>	-1.67
O49499	<i>CCOAOMT1</i>	-1.56
Q9SP32	<i>DCL1</i>	-1.53
Q9M3H5	<i>HMA1</i>	-1.51
Q9S7Z3	<i>PCS1</i>	-1.5
Q39253	<i>CAX1</i>	-1.48
Q9X199	<i>USPL1</i>	-1.48
O81850	<i>IRT2</i>	-1.47

Table 3. Domain analysis of top-scoring proteins involved in phytoremediation

Genes	UniProt ID	Region	Domain name
<i>PCS1</i>	Q9S7Z3	1-221	Phytochelatin synthase, N-terminal catalytic domain
		231-465	Phytochelatin synthase, C-terminal
<i>PCS2</i>	Q9ZWB7	1-220	Phytochelatin synthase, N-terminal catalytic domain
		230-433	Phytochelatin synthase, C-terminal
<i>EXPA12</i>	Q9LDJ3	45-156	Expansin/pollen allergen, DPBB domain
		157-246	Expansin, cellulose-binding-like domain
		62-146	RlpA-like protein, double-psi beta-barrel domain
<i>ZIP10</i>	Q8W245	NULL	No Domain
<i>CCOAOMT1</i>	O49499	34-258	SAM-dependent O-methyltransferase class I-type profile
		1357-1538	Ribonuclease III domain
		1559-1730	
		648-812	Helicase, C-terminal domain-like
		1180-1342	PAZ domain
<i>DCL1</i>	Q9SP32	840-935	Dicer dimerisation domain
		250-415	Helicase/UvrB, N-terminal
		245-448	Helicase superfamily 1/2, ATP-binding domain
		1733-1796	Double-stranded RNA-binding domain
		1831-1906	
<i>HMA1</i>	Q9M3H5	433-733	P-type ATPase, haloacid dehalogenase domain
<i>CAX1</i>	Q39253	96-257	Sodium/calcium exchanger membrane region
		288-431	
<i>USPL1</i>	Q9XI99	56-280	BURP domain
<i>IRT2</i>	O81850	NULL	No Domain

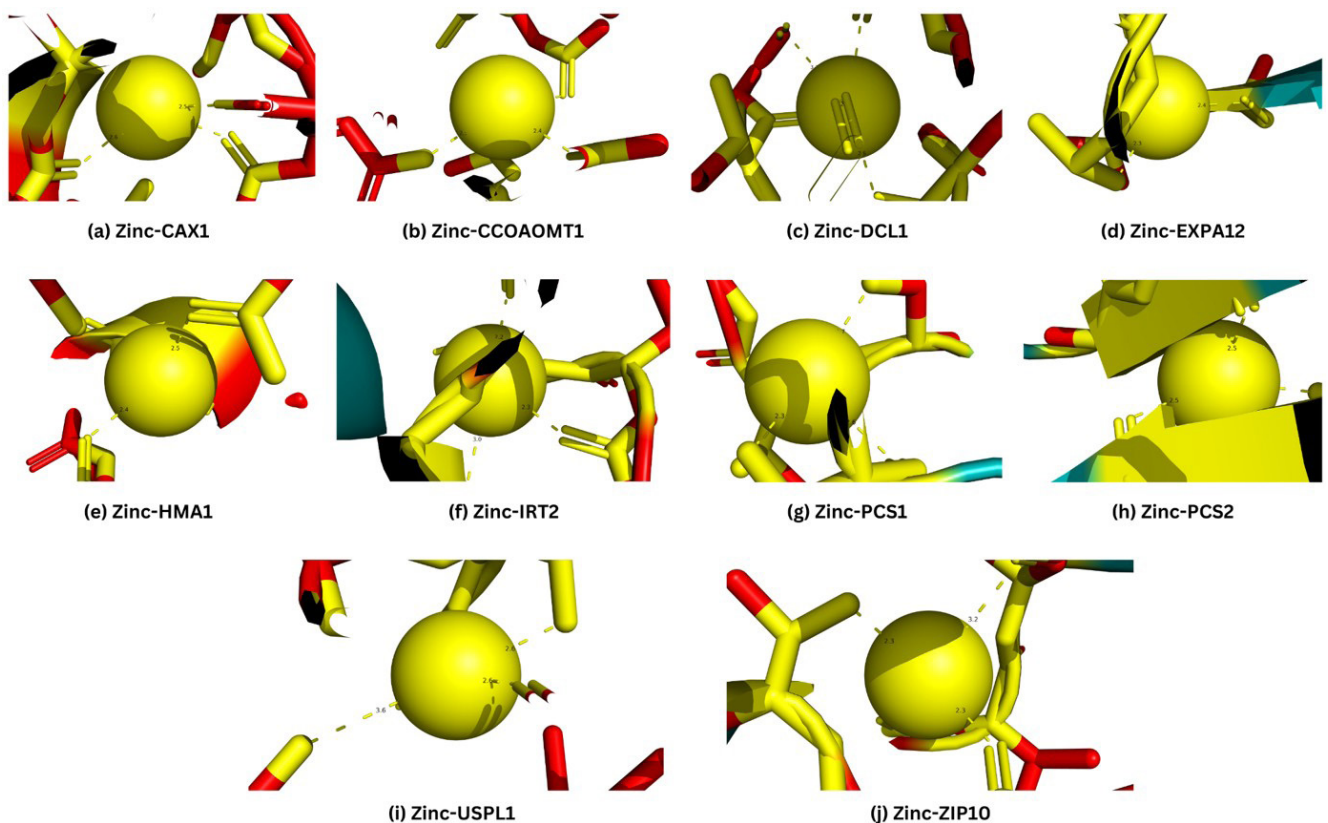


Figure 1. The zoomed view illustrates the protein structures of top phytoremediation genes screened against Zn

protein interacted with zinc at residues 161A (HIS), 210A (ARG), and 149A (TYR), occurring within the Phytochelatin Synthase N-terminal catalytic domain (230-433). Additionally, the EXPA12 protein showed interactions at residues 63A (ALA) and 104A (THR) within the Expansin/Pollen Allergen DPBB domain (506-653). The residue 104A (THR) was also involved in interactions within the RlpA-like protein, double-psi beta-barrel domain (62-146).

The CCOAOMT1 protein bound to zinc at residues 157A (ASP), 178A (LYS), 181A (TYR), 201A (ASP), and 202A (ASN) within the SAM-dependent O-methyltransferase class I-type profile domain (34-258). The DCL1 protein exhibited interactions at residues 1275A (ILE), 1276A (LYS), and 1309A (LEU) within the PAZ domain (1180-1342). The USPL1 protein interacted with residues 260A (GLU),

263A (SER), and 264A (ALA) within the BURP domain (56-280).

Similarly, the CAX1 protein showed interactions at residues 405A (THR), 407A (GLN), 408A (ASP), 410A (SER), and 415A (LYS) within the Sodium/Calcium Exchanger Membrane Region domain (288-431). The HMA protein bound to zinc at residues 729A (LYS) and 732A (GLN) within the P-type ATPase, haloacid dehalogenase domain (433-733), with additional interactions at residues 782A (ARG) and 785A (ASN) occurring outside this domain. Finally, the IRT2 and ZIP10 proteins interacted with zinc at residues 37A (CYS), 39A (ASN), 222A (SER), 42A (CYS), and 236A (THR), respectively, but no domains for these proteins have been reported in the existing literature. The protein-ligand interactions in 2D format are illustrated in Figure 2.

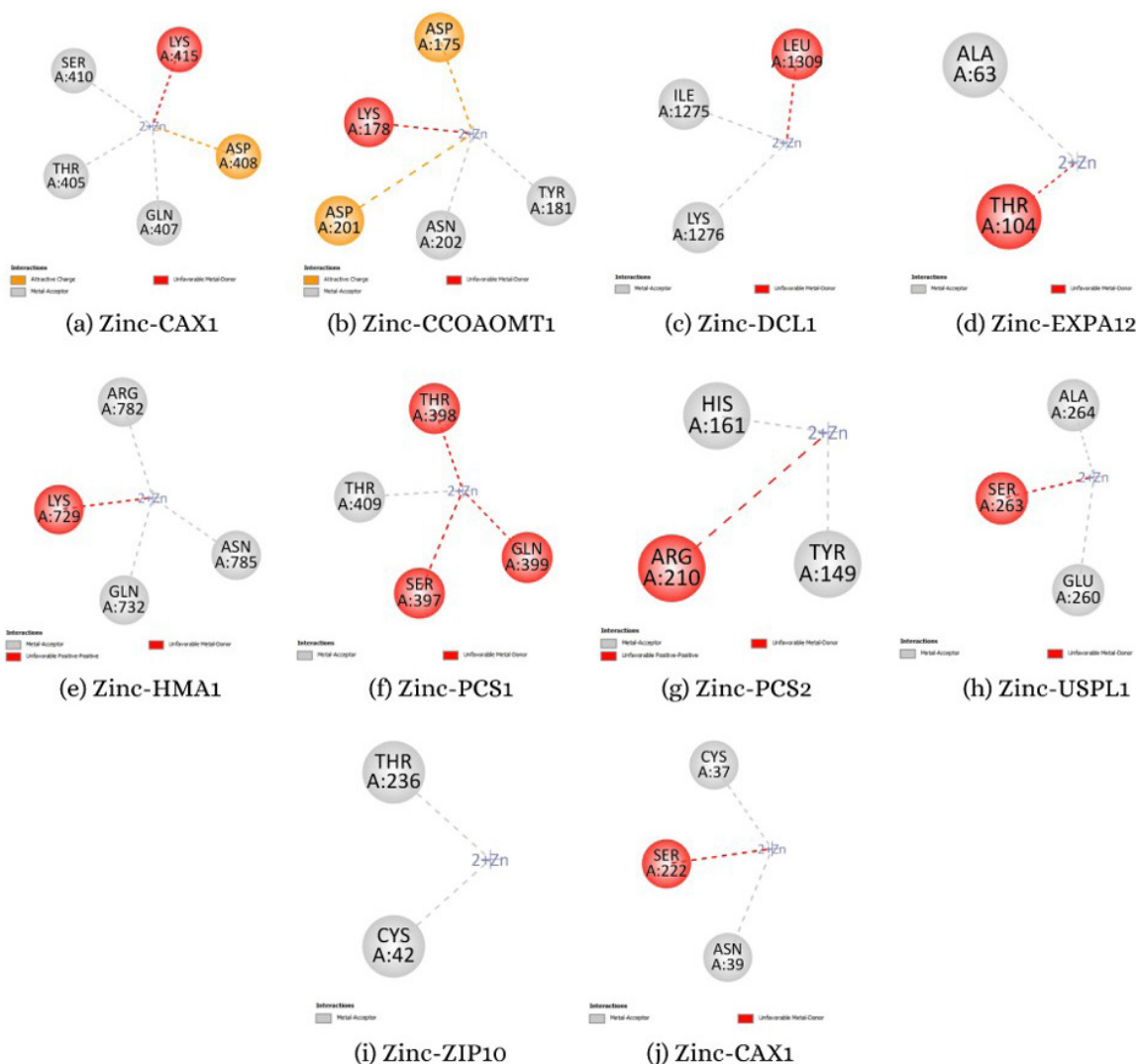


Figure 2. 2D illustration of interactions of top-scoring proteins with Zn

4. Discussion

The study provides a comprehensive overview of the molecular mechanisms underlying phytoremediation processes utilized by *A. thaliana*. By analyzing phytoremediation-related genes across 27 plant species, we have identified key gene families and elucidated their roles in metal uptake, transport, sequestration, and detoxification. Gene families such as ABC transporters, acyl-CoA-binding proteins, metallothioneins, and others play a pivotal role in phytoremediation (Andolfo *et al.* 2015; Yan *et al.* 2020; Amin *et al.* 2023). Notably, PCSs and HMAs stand out for their essential roles in metal regulation.

PCSs are vital for the synthesis of phytochelatin, sulfur-rich peptides that bind heavy metals, enabling their detoxification and sequestration within plant vacuoles. The high binding affinities observed for PCS2 (-1.77) and PCS1 (-1.5) highlight the importance of these enzymes in managing metal stress. The findings indicate that PCS proteins effectively chelate zinc ions, which become toxic at elevated concentrations, thereby playing a pivotal role in metal detoxification. Supporting evidence from other studies confirms that phytochelatin is crucial not only for detoxifying toxic metals like cadmium but also for maintaining the homeostasis of essential metals such as zinc under normal physiological conditions. Studies have shown that PCs are involved in the homeostasis of Zn via modulation of cellular availability of zinc, and mutations in PC synthase disrupt metal tolerance, further indicating their critical role in both detoxification and essential metal regulation. This sort of dual function of PCs in managing metal toxicity and homeostasis has been observed in various organisms, including plants and yeast (Tennstedt *et al.* 2009; García-García *et al.* 2020; Łuczowski *et al.* 2024).

HMAs like HMA1 play a very important role in the transportation of metals and their homeostasis. They belong to P1B-type ATPases that actively transport metal ions across cellular membranes. The binding affinity of zinc to HMA1, as evidenced by the score of -1.51, indicates its role in the selective transport of zinc and other metals, balancing essential metal uptake with the detoxification of excess metals. This active transport mechanism is of great importance in detoxification to prevent toxic accumulation in tissues for the homeostasis of metals in plants. Further, other studies have unraveled the complementary role of different HMA genes responsible for the transport and detoxification of metals such as zinc,

cadmium, and copper in plants. *A. thaliana* studies on. The roles of HMA2 and HMA3 in zinc and cadmium translocation and sequestration have been documented in *A. thaliana* and rice. In *Areca catechu*, genes such as AcHMA1 and AcHMA3 have been recognized as key players in the response to cadmium (Cd) and zinc (Zn) stress (Takahashi *et al.* 2012; Ullah *et al.* 2022; Khan *et al.* 2024).

Proteins such as CAX1 and IRT2 are integral to the transport and regulation of zinc, participating in various processes related to the uptake and movement of zinc ions. CAX1, functioning as a sodium/calcium exchanger, may contribute to the sequestration of zinc ions into vacuoles, while IRT2 is likely involved in zinc uptake from the soil. Additionally, studies have identified CAX1 and CAX2 as critical components of transport systems for both calcium and heavy metal ions, with CAX2 recognized as a high-capacity transporter for H⁺/heavy metal cations. Meanwhile, ZAT10 has been implicated in cadmium uptake and detoxification by inhibiting IRT1 and positively regulating IRT2, though its role in zinc transport remains unreported (Montes-Rocha *et al.* 2024).

The interplay of gene families such as PCSs, HMAs, CAX1, and IRT2 highlights the complex molecular mechanisms regulating metal uptake, transport, and detoxification in *A. thaliana*. These proteins play essential roles in binding and storing metals like zinc and cadmium, potentially offering valuable insights for advancing plant-based strategies to improve environmental remediation techniques. Also, studies on other plants like rice and *Areca catechu* support the idea that these transporters are important for dealing with heavy metal stress. Future research should look into how these genes work together to create transgenic plants that can clean up metal contamination better.

In conclusion, by combining all the above beneficial effects, this study presented the overall first systematic analysis of the underlying molecular mechanisms relevant to Zn phytoremediation in *A. thaliana*. The characterization of these key gene families that are plastic to metal chelation, transporters, and detoxifiers like PCS1 and PCS2 (the protein PCS2 presented the high-affinity to zinc) will provide the basis for improving the metal tolerance and root bioaccumulation ability of plants. An example of heavy metal ATPases that play an important role in the control of metal transport and homeostasis is HMA1. Psyrdoes cite zinc transporters that are implicated in the uptake and sequestration of zinc, such as cation exchanger1 CAX1 and IRT2. Validation of the methodology in field conditions and optimization of

gene functions are further steps toward the best strategy for phytoremediation.

References

- Agni, L. Unraveling the metabolites in *Arabidopsis halleri* ssp. gemmifera root exudate which facilitate zinc and cadmium leaching from contaminated soil [Dissertation]. Sendai, Japan: Tohoku University.
- Amin, H., Ahmed Arain, B., Jahangir, T.M., Abbasi, A.R., Abbasi, M.S., Amin, F., 2023. Comparative zinc tolerance and phytoremediation potential of four biofuel plant species. *International Journal of Phytoremediation*. 25, 1014–1028. <https://doi.org/10.1080/15226514.2022.2125496>
- Andolfo, G., Ruocco, M., Di Donato, A., Frusciante, L., Lorito, M., Scala, F., Ercolano, M.R., 2015. Genetic variability and evolutionary diversification of membrane ABC transporters in plants. *BMC Plant Biology*. 15, 1–15. <https://doi.org/10.1186/s12870-014-0323-2>
- Burley, S.K., Bhikadiya, C., Bi, C., Bittrich, S., Chao, H., Chen, L., Craig, P.A., Crichlow, G.V., Dalenberg, K., Duarte, J.M., 2023. RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Research*. 51, 488–508. <https://doi.org/10.1093/nar/gkac1077>
- Chaffai, R., Koyama, H., 2011. Heavy metal tolerance in *Arabidopsis thaliana*. In *Advances in Botanical Research*. 60, 1–49. <https://doi.org/10.1016/B978-0-12-385851-1.00001-9>
- Escudero, V., Ferreira Sánchez, D., Abreu, I., Sopena-Torres, S., Makarovskiy-Saavedra, N., Bernal, M., Krämer, U., Grolimund, D., González-Guerrero, M., Jordá, L., 2022. *Arabidopsis thaliana* Zn²⁺-efflux ATPases HMA2 and HMA4 are required for resistance to the necrotrophic fungus *Plectosphaerella cucumerina* BMM. *Journal of Experimental Botany*. 73, 339–350. <https://doi.org/10.1093/jxb/erab400>
- Faizan, M., Alam, P., Hussain, A., Karabulut, F., Tonny, S.H., Cheng, S.H., Yusuf, M., Adil, M.F., Sehar, S., Alomrani, S.O., 2024. Phytochelatin: a key regulator against heavy metal toxicity in plants. *Plant Stress*. 11, 100355. <https://doi.org/10.1016/j.stress.2024.100355>
- García-García, J.D., Sánchez-Thomas, R., Saavedra, E., Fernández-Velasco, D.A., Romero-Romero, S., Casanova-Figueroa, K.I., Mendoza-Cózatl, D.G., Moreno-Sánchez, R., 2020. Mapping the metal-catalytic site of a zinc-activated phytochelatin synthase. *Algal Research*. 47, 101890. <https://doi.org/10.1016/j.algal.2020.101890>
- Hayashi, S., Tanikawa, H., Kuramata, M., Abe, T., Ishikawa, S., 2020. Domain exchange between *Oryza sativa* phytochelatin synthases reveals a region that determines responsiveness to arsenic and heavy metals. *Biochemical and Biophysical Research Communications*. 523, 548–553. <https://doi.org/10.1016/j.bbrc.2019.12.093>
- Hussain, M.S., Hamid, N.M., Al-Ani, M.Q., Muden, B.K., Hamad-Allah, M.J., Sweed, O.H., Sabbar, S.S., 2020. Hematological changes in blood of smokers of cigarettes and nargyle. *Indian Journal of Forensic Medicine & Toxicology*. 14, 1771–1776. <https://doi.org/10.37506/ijfimt.v14i2.3193>
- Hussain, S., Khan, M., Sheikh, T.M.M., Mumtaz, M.Z., Chohan, T.A., Shamim, S., Liu, Y., 2022. Zinc essentiality, toxicity, and its bacterial bioremediation: a comprehensive insight. *Frontiers in Microbiology*. 13, 900740. <https://doi.org/10.3389/fmicb.2022.900740>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*. 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Khan, N.M., Ali, A., Wan, Y., Zhou, G., 2024. Genome-wide identification of heavy-metal ATPases genes in *Areca catechu*: investigating their functionality under heavy metal exposure. *BMC Plant Biology*. 24, 484. <https://doi.org/10.1186/s12870-024-05201-6>
- Kumar, K., Shinde, A., Aeron, V., Verma, A., Arif, N.S., 2023. Genetic engineering of plants for phytoremediation: advances and challenges. *Journal of Plant Biochemistry and Biotechnology*. 32, 12–30. <https://doi.org/10.1007/s13562-022-00776-3>
- Łuczowski, M., Leszczyńska, W., Wąty, J., Clemens, S., Krężel, A., 2024. Phytochelatin bind Zn(II) with micro-to picomolar affinities without the formation of binuclear complexes, exhibiting zinc buffering and muffling rather than storing functions. *Inorganic Chemistry*. 63, 10915–10931. <https://doi.org/10.1021/acs.inorgchem.4c01707>
- Lussi, Y.C., Magrane, M., Martin, M.J., Orchard, S., Consortium, U., 2023. Searching and navigating UniProt databases. *Current Protocols*. 3, e700. <https://doi.org/10.1002/cpz1.700>
- McNutt, A.T., Francoeur, P., Aggarwal, R., Masuda, T., Meli, R., Ragoza, M., Sunseri, J., Koes, D.R., 2021. GNINA 1.0: molecular docking with deep learning. *Journal of Cheminformatics*. 13, 43. <https://doi.org/10.1186/s13321-021-00522-2>
- Mitra, S., Chakraborty, A.J., Tareq, A.M., Emran, T. Bin, Nainu, F., Khusro, A., Idris, A.M., Khandaker, M.U., Osman, H., Alhumaydhi, F.A., 2022. Impact of heavy metals on the environment and human health: Novel therapeutic insights to counter the toxicity. *Journal of King Saud University-Science*. 34, 101865. <https://doi.org/10.1016/j.jksus.2022.101865>
- Monserrath, A., Sadoon Rasheed, Y., Kaur, M., Kaur, H., Salam Hussein, E., Al-khamiseh, B.M., Elawady, A., Zainul, R., Qiang, L., 2024. Biomarker detection using boron carbide layer: a first-principles study. *Fullerenes, Nanotubes and Carbon Nanostructures*. 32, 1–8. <https://doi.org/10.1080/1536383X.2024.2333838>
- Montes-Rocha, J.A., Alonso-Castro, A.J., Carranza-Álvarez, C., 2024. Mechanism of heavy metal-induced stress and tolerance. *Aquatic Contamination: Tolerance and Bioremediation*. 61-76. <https://doi.org/10.1002/9781119989318.ch4>
- Nishimura, Y., Yamada, K., Okazaki, Y., Ogata, H., 2024. DiGAlign: versatile and interactive visualization of sequence alignment for comparative genomics. *Microbes and Environments*. 39, ME23061. <https://doi.org/10.1264/jsme2.ME23061>
- Rasheed, Y.S., Al-Janaby, M.S., Ahmed, M.M., Houmady, H.Y., 2024. Effect of lead on some clinical and immunological variables for gasoline station workers and cigarette smokers in Ramadi, Iraq. *AIP Conference Proceedings*. 3097, 020005. <https://doi.org/10.1063/5.0209549>

- Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, D.C., Connor, R., Funk, K., Kelly, C., Kim, S., 2022. Database resources of the national center for biotechnology information. *Nucleic Acids Research*. 50, 20–26. <https://doi.org/10.1093/nar/gkab1112>
- Shadidizaji, A., Cinisli, K.T., Warda, M., Cicek, B., Hacimuftoglu, A., 2024. Virtual insights into the quercetin-Melampsora lini-derived effector AvrM14 interaction: an *in silico* exploration of plant defense mechanisms. *Physiological and Molecular Plant Pathology*. 129, 102200. <https://doi.org/10.1016/j.pmp.2023.102200>
- Singh, P., Pani, A., Mujumdar, A.S., Shirkole, S.S., 2024. The role of artificial intelligence in drying and biomass valorization in the field of phytoremediation of contaminated soils. *Drying Technology*. 42, 955–966. <https://doi.org/10.1080/07373937.2024.2345123>
- Sofa, A., Scopa, A., Remans, T., Vangronsveld, J., Cuypers, A., 2012. Biochemical and functional responses of *Arabidopsis thaliana* exposed to cadmium, copper and zinc. In: Anjum N, Ahmad I, Pereira M, Duarte A, Umar S, Khan N. (Eds.). *The Plant Family Brassicaceae. Environmental Pollution, Vol 21*. Dordrecht: Springer. pp. 239–263. https://doi.org/10.1007/978-94-007-3913-0_9
- Takahashi, R., Bashir, K., Ishimaru, Y., Nishizawa, N.K., Nakanishi, H., 2012. The role of heavy-metal ATPases, HMAs, in zinc and cadmium transport in rice. *Plant Signaling & Behavior*. 7, 1605–1607. <https://doi.org/10.4161/psb.22454>
- Tennstedt, P., Peisker, D., Böttcher, C., Trampczynska, A., Clemens, S., 2009. Phytochelatin synthesis is essential for the detoxification of excess zinc and contributes significantly to the accumulation of zinc. *Plant Physiology*. 149, 938–948. <https://doi.org/10.1104/pp.108.127472>
- Ullah, I., Mateen, A., Ahmad, M.A., Munir, I., Iqbal, A., Alghamdi, K.M.S., Al-Solami, H.M., Siddiqui, M.F., 2022. Heavy metal ATPase genes (HMAs) expression induced by endophytic bacteria, “AI001, and AI002” mediate cadmium translocation and phytoremediation. *Environmental Pollution*, 293, 118508. <https://doi.org/10.1016/j.envpol.2021.118508>
- Verma, A.K., Diwan, D., Raut, S., Dobriyal, N., Brown, R.E., Gowda, V., Hines, J.K., Sahi, C., 2017. Evolutionary conservation and emerging functional diversity of the cytosolic Hsp70: J protein chaperone network of *Arabidopsis thaliana*. *G3: Genes. Genomes. Genetics*. 7, 1941–1954. <https://doi.org/10.1534/g3.117.042291>
- Wang, L., Chen, K., Zhou, M., 2021. Structure and function of an *Arabidopsis thaliana* sulfate transporter. *Nature Communications*. 12, 4455. <https://doi.org/10.1038/s41467-021-24778-2>
- Webb, B., Sali, A., 2017. Protein structure modeling with MODELLER. In: Kaufmann M, Klinger C, Savelsbergh A, (Eds.). *Functional Genomics. Methods in Molecular Biology, Vol 1654*. New York: Humana Press. pp. 39–54. https://doi.org/10.1007/978-1-4939-7231-9_4
- Wei, Z., Gu, H., Van Le, Q., Peng, W., Lam, S.S., Yang, Y., Li, C., Sonne, C., 2021. Perspectives on phytoremediation of zinc pollution in air, water and soil. *Sustainable Chemistry and Pharmacy*. 24, 100550. <https://doi.org/10.1016/j.scp.2021.100550>
- Yan, A., Wang, Y., Tan, S.N., Mohd Yusof, M.L., Ghosh, S., Chen, Z., 2020. Phytoremediation: a promising approach for revegetation of heavy metal-polluted land. *Frontiers in Plant Science*. 11, 359. <https://doi.org/10.3389/fpls.2020.00359>
- Yuan, S., Chan, H.C.S., Hu, Z., 2017. Using PyMOL as a platform for computational drug design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 7, e1298. <https://doi.org/10.1002/wcms.1298>