https://doi.org/10.56093/ijas.v90i9.106596

Indian Journal of Agricultural Sciences **90** (9): 1685–9, September 2020/Article

In-silico characterization of cadmium stress response-associated Abc1-like protein and its homologues in maize (*Zea mays*)

ALLA SINGH[1]*, CHIKKAPPA KARJAGI[2], RAMESH KUMAR[3], D P CHAUDHARY[4]
and SUJAY RAKSHIT[5]

ICAR-Indian Institute of Maize Research, PAU campus, Ludhiana, Punjab 141 004, India

Received: 18 November 2019; Accepted: 24 December 2019

ABSTRACT

Cadmium is a pollutant of concern and accumulates in soil from unchecked industrial activities like smelting, incinerations and combustion. Plants interact via different mechanisms with heavy metal pollutants in soil. Understanding of the plant-soil-metal interactions and the biomolecules involved in this interaction is important to make informed choices for setting breeding goals. In *Arabidopsis thaliana*, *AtOSA1* (*A. thaliana* oxidative stress-related Abc1-like protein) is involved in response to cadmium stress. The homologous sequences have been deciphered in maize (*Zea mays* L.). The structures of the maize proteins have been determined. It is observed that multiple proteins sense cadmium and can respond to its presence. Concerted action of all the proteins may be essential to realize full response against cadmium presence in soil and mitigation of plant growth inhibition. Better understanding of the response to cadmium stress would pave way for designing effective breeding strategies for pyramiding beneficial alleles in elite germplasm.

Key words: Abcl-like proteins, Maize, Metal stress response, Phytoremediation

Among several abiotic stresses, contamination of agricultural land with heavy metals is being considered as one of the emerging problem (Raymond and Felix, 2011). Anthropogenic developmental activities like urbanization coupled with industrialization have led to contamination of agriculture lands with heavy metals mainly through release of industrial waste to rivers followed by canal irrigation of agricultural land. The term heavy metals refer to all those metal or metalloid elements of periodic table with atomic density of 5 g cm^{-3} or more, or 5 times or more greater than water. The agriculture area located in and around dense cities and industrial hubs are the one which have affected the most. The impact of heavy metal contamination ranges from agriculture soil used for crop production to the food being produced from that land. Heavy metals not only affects plant growth and development but also the soil microflora, an integral part of crop production ecosystem. Certain heavy metals like cadmium, arsenic, lead, mercury etc., which are classified as non-essential to metabolic and other biological functions are found to affect adversely several essential biochemical, physiological, and metabolic processes of plant (Srivastava *et al.* 2017).

In this context, maintaining soil quality and reclamation of polluted agricultural land are important goals for agricultural sustainability to ensure food and nutritional security. In general heavy metal contaminated soils threaten plant, animal and human health, thus affects overall quality of environment. Cadmium, the most common heavy metal is released in environment by various unregulated industrial activities like smelting, incinerations, fuel combustion, improper discard of batteries, etc. (Dotaniya 2015). Crops can negatively affected by presence of high concentration of heavy metals. While limited concentrations of copper, zinc, manganese, etc. can aid physiological processes in plants, cadmium and lead are particularly harmful, as they have no biological function in plants (Jasinski *et al.* 2008). Various strategies to mitigate the toxicity caused by cadmium ion are known. Foliar application of ascorbic acid reduced cadmium toxicity in maize (*Zea mays* L.) (Zhang *et al.* 2019). Alternatively endophytes have also been employed to reduce the cytotoxic effects of cadmium in maize. For example, plant root-associated fungi like dark septate endophyte (DSE), *Exophiala pisciphila* was found to reduce cadmium phytotoxicity in maize at physiological, cytological and molecular levels under increased soil Cd stress condition (He *et al.* 2017). The DSE colonization of maize roots also induced regulation of shoot and root specific genes related to metal uptake in plant. Thus it paved the way to envisage another possible strategy, *i.e.* identification, characterization and possible mobilization

Corresponding author e-mail: allasingh.panesar@gmail.com

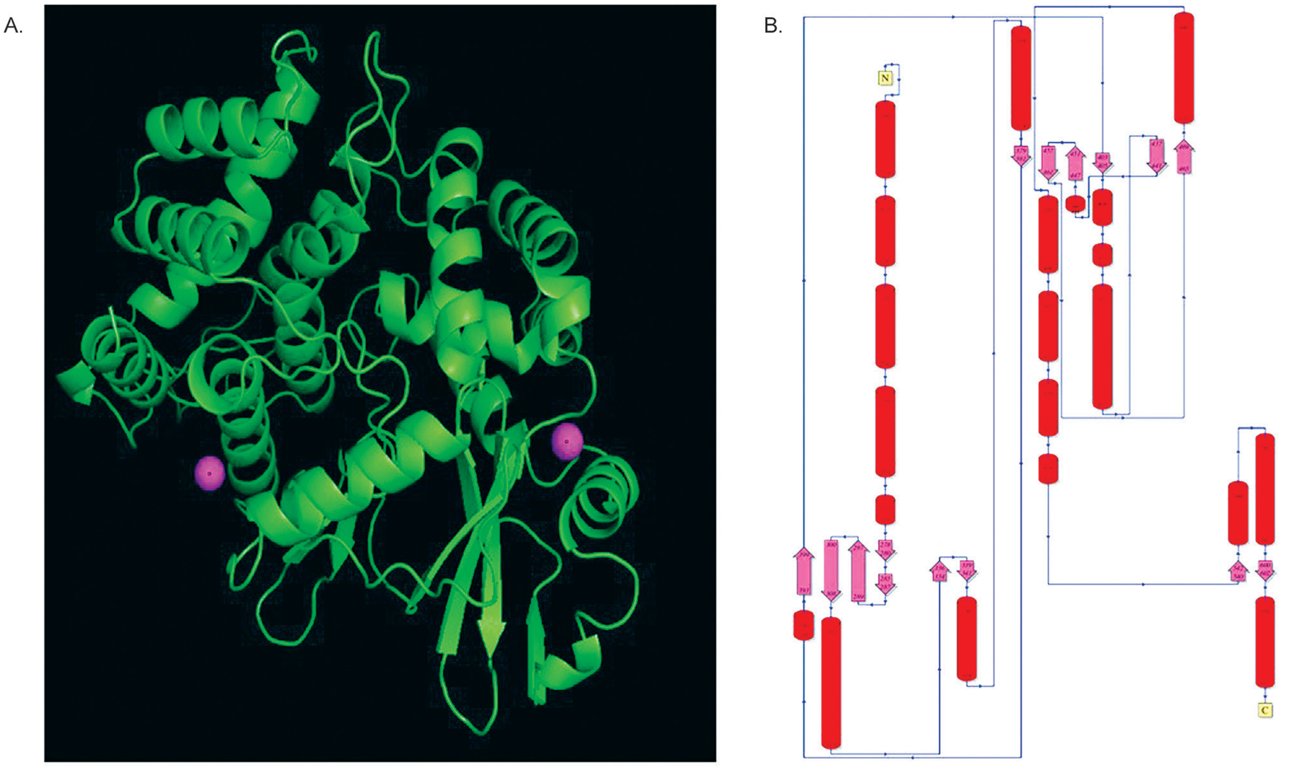

Fig 1 Structure of *Zea mays ATATH13* protein. A. 3-dimensional orientation of the molecule. The docked Cadmium ions are shown in magenta. B. Secondary structure representation of the protein. Helices and sheets are indicated by red rods and pink arrows, respectively. N- and C- termini are marked.

of cadmium stress response related genes through breeding to impart tolerance to Cd stress in maize. However, the process requires elucidation of gene/protein network that signals plant about the presence of cadmium and drives downstream signal relay for onset of protective measures.

Previous studies in *A. thaliana* using cDNA microarrays and quantitative PCR has identified *AtOSA1* (*A. thaliana* oxidative stress-related Abc1-like protein), inducible gene in the presence of cadmium, which protects against Cd phytotoxicity by balancing oxidative stress (Jasinski *et al.* 2008). Abc1-like proteins are present in prokaryotic and eukaryotic organelles like mitochondria. Abc1-like proteins are known to regulate vitamin E levels in *Arabidopsis* (Martinis *et al.* 2013). In rice, they have been implicated in regulation of grain size and weight (Li *et al.* 2015). Transcriptomic analysis has revealed auxin signalling pathways to be implicated in response of maize roots to cadmium (Yue *et al.* 2016). Given the versatile role of Abc1-like proteins and a comprehensive metabolic change brought by cadmium presence, it is imperative to elucidate the cadmium stress responsive genes and their homologous sequences. The identified genes can be introgressed into desired lines to screen against cadmium stress. In the present study, effort has been made to identify and characterize the homologues of *AtOSA1* in maize. The identified genes with putative role in imparting tolerance to Cd stress in maize will go a long-way in systematic understanding and deployment in active maize breeding programmes.

MATERIALS AND METHODS

Sequence retrieval: All the sequences were retrieved from National Center for Biotechnology Information (NCBI). Basic local alignment was done using BLASTP module. *Zea mays* taxid 4577 was used for alignments.

Generation of homology model of Zea mays ATATH13: The homology model of *Zea mays* ATATH13 was generated using SWISS-MODEL program (Benkert *et al.* 2011; Waterhouse *et al.* 2018). The secondary structure of the model was computed using PDBsum program (Laskowski 2001; Laskowski *et al.* 2018).

Multiple sequence alignment: The multiple homologous sequences were aligned using Expresso module for structural alignments in T-Coffee web server (Notredame *et al.* 2000).

Elucidation of cadmium binding on Abc1-like proteins: The interaction of proteins with cadmium was modelled through fragment transformation algorithm employed in Metal-Ion Binding Site Prediction and Docking Server (Lin *et al.* 2016).

Determination of molecular patterns of Abc1-like proteins: The protein sequence features of the Abc1 proteins were determined using Motif Scan program of Swiss Institute of Bioinformatics. *In-silico* subcellular localization analysis was done using BUSCA program (Savojardo *et al.* 2018).

RESULTS AND DISCUSSION

A. thaliana OSA1 protein was used to search maize protein database for the presence of similar sequences. Basic

Table 1 Score of top two binding interactions of cadmium with Abc1 proteins [based on scores assigned by Metal-Ion Binding Site Prediction and Docking Server (Lin *et al.* 2016).

Protein	Cadmium binding score (Top 2 Interactions)
ATATH13	1.765, 1.760
PWZ21020.1	1.388, 1.249
PWZ22977.1	1.368, 1.330
XP_020393919.1	1.372, 1.256
XP_008643622.1	1.369, 1.332

local alignment search of *A. thaliana* OSA1 protein with *Zea mays* (taxid: 4577) using BLAST tool, showed *Z. mays* ATATH13 (ZmATAH13) protein as sharing 85.51% identity with AtOSA1. Hence, ZmATATH13 was taken forward for further study. The homology model of *Z. mays* ATATH13 was determined (Fig 1A). It showed a 20.98% sequence identity with mitochondrial aarF domain containing kinase 3 (ADCK3) protein. The structure has a Global Model Quality Estimation (GMQE) and QMEAN score of 0.34 and -3.98, respectively.

Fragment transformation algorithm was used to find the cadmium binding sites in ZmATATH13. In this analysis done using Metal-Ion Binding Site Prediction and Docking Server, query protein is compared with templates without any data training. The method has been reported to predict cadmium-binding sites with an accuracy of 92.9% (Lin *et al.* 2016). It was found that cadmium bonds to two regions with high affinity (Table 1). The regions include glutamic acid residues at positions 266, 267 and 427, 431 (Fig 1A). Secondary structure of ZmATATH13 model shows that it contains 21 α-helices and 13 β-sheets (Fig 1B). The protein is majorly helical in nature.

In order to investigate if other proteins with similar structural fold as ZmATATH13 are present in maize, the ATATH13 sequence was further used to search the maize genome (taxid: 4577) for the presence of similar sequences using BLASTP module. Many similar sequences were obtained. Following sequences with their accession number and sequence identity were taken for further analysis: uncharacterized protein LOC100281496 isoform X1 (XP_008646322.1) - 99.34%; hypothetical protein Zm0014a_29828 (PWZ22977.1) - 93.72%; BC1 complex kinase 7 (XP_020393919.1) - 42.67%; hypothetical protein Zm0014a_012517 (PWZ21020.1) - 42.65%. The sequences were aligned with *Z. mays* ATATH 13 (Fig 2).

Multiple sequence alignment of ZmATATH13 with its homologues was done. It was found that ZmATATH13 aligns well with the homologues, except for relatively lower sequence identity in the N-terminal region. The central and C-terminal regions align fairly well. Structural alignment indicated the possibility that the homologues may also bind cadmium. In order to confirm this, the homology models were prepared using SWISS-MODEL program (Fig 3).

The best homology models were analyzed for presence of cadmium binding sites using fragment transformation method, as done for ZmATATH13 (Lin *et al.* 2016). The affinities of top two binding sites were noted (Table 1).

It was observed that all the four homologues of ZmATATH13 bind cadmium and therefore, can potentially play a role in cadmium stress response. *In silico* subcellular localization analysis showed that ATATH13, PWZ22977.1 and XP_008643622.1 localize in the outer membrane of chloroplast, while PWZ21020.1 and XP_020393919.1 localize in mitochondria. Hence, in spite of differences in the N-terminal region, the proteins are capable of binding cadmium and contain transit peptides for localization in either of the two energy organelles: chloroplast or mitochondria.

```
BAD AVG GOOD

ZmATATH13       MSVASA--------------AAAASLIASTSLSVPDRIRLHPRPALP--------------PLPCF
PWZ21020.1      MAMLVPDLSFSASISSVKLPRYGASKSETFFVKKPPTFRAEARPAESGKYGTNNGRATKMVPTTEF
PWZ22977.1      MSVASA--------------AAAASLIASTSLSVPDRIRLHPRPALP--------------PLPCF
XP_020393919.1  MAMLVPDLSFSASISSVKLPRYGASKSETFFVKKPPTFRAEARPAESGKYGTNNGRATKMVPTTEF
XP_008643622.1  MSVASA--------------AAAASLIASTSLSVPDRIRLHPRPALP--------------PLPCF

cons            *::  .                .**   :  :. *  :* ..*** .              * . *

ZmATATH13       RRRSRVCLVRA-VLEDRPSPPAEEDAKCYGLNGNGSGLGYDDAAVQAYLGSNGNGSASGNGGAVKQ
PWZ21020.1      RRTSGG-LVRSGTVNGSPNSAVNGST-KVLI--NGSANAIDNSSTKAVI--NGTPKVAVNGTSLVK
PWZ22977.1      RRRSRVCLVRA-VLEDRPSPPAEEDAKRYGLNGNGSGLGYDDAAVQAYLGSNGNGSASGNGGAVKQ
XP_020393919.1  RRTSGG-LVRSGTVNGSPNSAVNGST-KVLI--NGSANAIDNSSTKAVI--NGTPKVAVNGTSLVK
XP_008643622.1  RRRSRVCLVRA-ILEDRPSPPAEEDAKRYGLNGNGSGLGYDDAAVQAYLGSNGNGSASGNGGAVKQ

cons            ** *   ***:  ::. *....: .:    :  ***. . *:::.:* :  **. ..: ** :: :

ZmATATH13       NPEVAGSSAAIVPAPVPPAEGERRRKERVEEIGREDAWFKQSSAEVSVAPGGRWNRFKTYSTIQRT
PWZ21020.1      SSKTS-----LV--------------KTQKQVRPKDDPFQEELKVLPSDEGFSWAK-DNYNSVQRS
PWZ22977.1      NPE------------------------------------------VSVAPGGRWNRFKTYSTIQRT
XP_020393919.1  SSKTS-----LV--------------KTQKQVRPKDDPFQEELKVLPSDEGFSWAK-DNYNSVQRS
XP_008643622.1  NPEVAGSSAAIVPAPVPPAEGERRRKERVEEIGREDAWFKQSSAEVSVAPGGRWNRFKTYSTIQRT

cons            ..:                                          :.   *  * : ..*.::**:

ZmATATH13       LEIWGFVFTFIFKAWLNNQKFTYR---GGMTEEKKIMRRKVLAKWLKESLLRLGPTFIKIGQQFST
PWZ21020.1      IDIWSFVLSLRVRVLFDNAKWAYP---GGFSEEKQKVRRRKTASWLREQVLQLGPTFIKLGQLSST
PWZ22977.1      LEIWGFVFTFIFKAWLNNQKFTYR---GGMTEEKKIMRRKVLAKWLKESLLRLGPTFIKIGQQFST
XP_020393919.1  IDIWSFVLSLRVRVLFDNAKWAYP---GGFSEEKQKVRRRKTASWLREQVLQLGPTFIKLGQLSST
XP_008643622.1  LEIWGFVFTFIFKAWLNNQKFTYRASAGGMTEEKKIMRRKVLAKWLKESLLRLGPTFIKIGQQFST

cons            ::**.**::: .:. ::* *::*    **::***: :**:  *.**:*.:*:*******:**  **
```

Fig 2 Multiple sequence alignment of ZmATATH13 and its homologues. Sequence alignment up to first 232 residues in the N-terminal are shown. The colour index for quality of alignments is shown on top left. Except for sequence in the N-terminal region, the sequence in the central and C-terminal regions is fairly well-aligned in the homologues.

In order to elucidate the putative role of differences

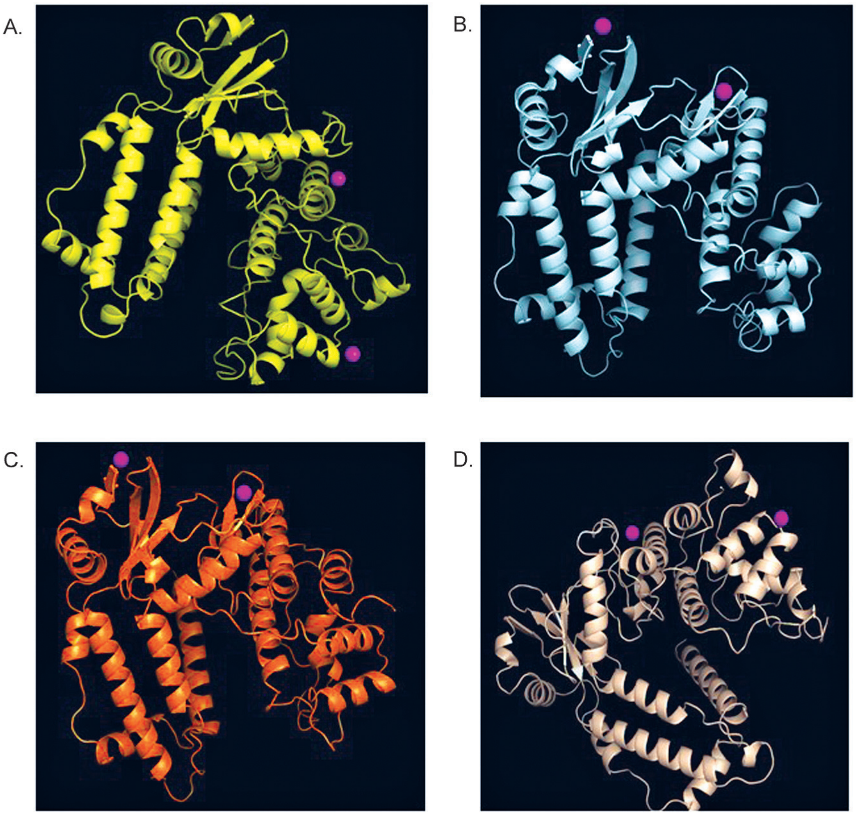

Fig 3 Structures of the four homologues of ZmATAH13. A. PWZ21020.1, B. PWZ22977.1, C. XP_020393919.1, D. XP_008643622.1. The cadmium ions are shown in magenta.

Table 2 Variation in the number of phosphorylation sites in the different protein sequences [based on subcellular localization analysis done using BUSCA program (Savojardo *et al.* 2018)].

Protein	Asparagine glycosylation	cAMP phosphorylation	Casein Kinase 2 phosphorylation	Myristylation	Protein Kinase C phosphorylation	Tyrosine phosphorylation
ATAH13	02	01	08	11	05	--
PWZ21020.1	04	02	03	07	11	02
PWZ22977.1	02	01	07	10	04	--
XP_020393919.1	05	04	04	12	12	02
XP_00864322.1	02	01	08	11	04	--

in the N-terminal sequence, Motif Scan program (Swiss Institute of Bioinformatics) was used to scan motifs present in the amino acid sequences of ATATH13 and its homologues. Analysis of the protein sequence patterns showed that all the proteins have kinase and ATP-binding activities. Analysis of the protein sequence also revealed that the different Abc1-like proteins contain variable number of phosphorylation sites (Table 2). The protein different with respect to Asparagine glycosylation, phosphorylation by cAMP, caseing kinase 2, protein kinase C and tyrosine, and also in myristylation. Many, but not all phosphorylation sites are present in the N-terminal region. Thus, it appears that all the proteins respond to cadmium stress, but relay it differently in the cell, either in mitochondria or chloroplasts. Combined action of all cadmium stress responsive Abc1 proteins may lead to gross changes in cell transcriptome, proteome and metabolome to offer protection against the heavy metal, especially in the energy organelles and prepare the cell for mitigation of the stress.

The present study also characterized certain hitherto uncharacterized or hypothetical proteins and show that they bind cadmium, albeit with little less efficiency than ZmATATH13. The presence of phosphorylation- and ATP-binding sites, along with protein kinase activity indicates that the homologous Abc1 proteins are competent to bind cadmium and influence downstream signalling. The differential cadmium binding affinity (Table 1) and presence of variable number of phosphorylation sites (Table 2) indicates that the above Abc1 proteins may act via different mechanisms to respond to cadmium stress. Determining the network and key molecules of the cadmium response system is necessary to understand the genetics of cadmium stress tolerance and devise efficient breeding programs for integrating different components of the response network together in elite germplasm. Genetic dissection of the plant response to cadmium stress and the estimation of role played in it by genetic regions coding for Abc1-like protein and its homologues will result in elucidation of effective loci for gene introgression for developing cadmium-tolerant maize. Alternatively, the responsible proteins can also be employed for genetic improvement using transgenic approaches.

REFERENCES

Benkert P, Biasini M and Schwede T. 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**: 343-50.

He Y, Yang Z, Li M, Jiang M, Zhan F, Zu Y, Li T and Zhao Z. 2017. Effects of a dark septate endophyte (DSE) on growth, cadmium content, and physiology in maize under cadmium stress. *Environ Sci Pollut Res Int* **24**: 18494-18504.

Jasinski M, Sudre D, Schansker G, Schellenberg M, Constant S, Martinoia E and Bovet L. 2008. AtOSA1, a member of the

Abc1-like family, as a new factor in cadmium and oxidative stress response. *Plant Physiol* **147**: 719-31.

Laskowski R A. 2001. PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res* **29**: 221-2.

Laskowski R A, Jablonska J, Pravda L, Varekova R S and Thornton J M. 2018. PDBsum: Structural summaries of PDB entries. *Protein Sci* **27**: 129-134.

Li T, Jiang J, Zhang S, Shu H, Wang Y, Lai J, Du J, Yang C. 2015. OsAGSW1, an ABC1-like kinase gene, is involved in the regulation of grain size and weight in rice. *J Exp Bot* **66**: 5691-701.

Lin Y F, Cheng C W, Shih C S, Hwang J K, Yu C S and Lu C H. 2016. MIB: Metal ion-binding site prediction and docking server. *J Chem Inf Model* **56**: 2287-2291.

Dotaniya M and Saha J. 2016. Heavy metal polluted soils in India: status and counter measures. *JNKVV Res J*. **49**: 320-337.

Martinis J, Glauser G, Valimareanu S and Kessler F. 2013. A chloroplast ABC1-like kinase regulates vitamin E metabolism in Arabidopsis. *Plant Physiol* **162**: 652-62.

Notredame C, Higgins D G and Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**: 205-17.

Raymond A W and Felix E O. 2011. Heavy metals in contaminated soils: A review of sources, chemistry, risks and best available strategies for remediation. *ISRN Ecology* 2011: 20.

Savojardo C, Martelli P L, Fariselli P, Profiti G and Casadio R. 2018. BUSCA: an integrative web server to predict subcellular localization of proteins. *Nucleic Acids Res* **46**: W459-W466.

Srivastava V, Sarkar A, Singh S, Singh P, de Araujo A S F and Singh R P. 2017. Agroecological responses of heavy metal pollution with special emphasis on soil health and plant performances. *Frontiers in Environmental Science* **5**. DOI=10.3389/fenvs.2017.00064

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R., Heer F T, de Beer T A P , Rempfer C, Bordoli L, Lepore R and Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**: W296-W303.

Yue R, Lu C, Qi J, Han X, Yan S, Guo S, Liu L, Fu X, Chen N, Yin H, Chi H and Tie S. 2016. Transcriptome analysis of cadmium-treated roots in maize (*Zea mays* L.). *Front Plant Sci* **7**: 1298.

Zhang K, Wang G, Bao M, Wang L and Xie X. 2019. Exogenous application of ascorbic acid mitigates cadmium toxicity and uptake in maize (*Zea mays* L.). *Environ Sci Pollut Res Int* **26**: 19261-19271.