**The Journal of Animal & Plant Sciences, 29(2): 2019, Page: 539-548**
**ISSN: 1018-7081**

# GENOMIC COMPARISON AMONG THREE *ARABIDOPSIS* SPECIES REVEALED HEAVY METAL RESPONSIVE GENES

S. M. Jazayeri[1*], L. C. García Cruzatty[2] and R. Villamar-Torres[3,4,5,6]

[1]Departamento de Biología, Facultad de Ciencias, Universidad Nacional de Colombia, Bogota, Colombia; [2]Facultad de Agronomía, Universidad Técnica de Manabí, Campus La Teodomira, Portoviejo, Manabí 130105, Ecuador; [3]Programa de Maestría en Agricultura, mención Protección Vegetal, Centro de Posgrado de la Universidad Técnica de Babahoyo, km 2,5 vía Babahoyo – Montalvo; [4]Facultad de Ciencias Agrarias, Universidad Técnica Estatal de Quevedo, Los Ríos, Ecuador; [5]Université de Montpellier, Montpellier, France; [6]Facultad de Ingeniería Agronómica. Universidad Técnica de Manabí. Campus Experimental "La Teodomira", km 13 ½ vía Santa Ana, EC131301, Santa Ana - Ecuador

*Corresponding author's Email: smjazayeri@unal.edu.co

## ABSTRACT

*Arabidopsis* is the ever most well-known plant model. Publishing genome of *A. thaliana* has restructured biological sciences towards more knowledge on multi and interdisciplinary researches based on genomic data. Eleven approved species of *Arabidopsis* were classified into 4 groups. Despite of its importance in biology, only three *Arabidopsis* species genomes are available. Genomes, proteins and gene models in three *Arabidopsis* species including *A. halleri, A. lyrata* and *A. thaliana* were compared. Genomic comparison disclosed 80-90 % of similarity for proteins and gene models among three species. The clustered protein sets showed 79% of common orthologue proteins among three species. Transcript pairwise comparisons among three species showed a similarity range of 82-92%. Twenty-six species-specific clustered proteins related to heavy metal response are highlighted in *A. halleri*. They function in transport, phytochelat, protein degradation, cell wall formation, detoxification and signaling.

**Key words:** Arabidopsis, genes, genomic comparison, heavy metal, tolerance.

## INTRODUCTION

*Arabidopsis* is the reference model in plants The genome of *A. thaliana* published in 2000 (Arabidopsis Genome Initiative, 2000) has revolutionized the genomic era and provided a very useful source and reference for publishing genomes of other organisms from prokaryotes to eukaryotes. Homology among genomes and genes have answered many biological questions based on orthologues and revealed several life secrets but depending on the under study organism, its speciation divergence and convergence, gene annotation and function (Jones *et al.,* 2008).

There are 11 *Arabidopsis* species (http://www.theplantlist.org/browse/A/Brassicaceae/Arabidopsis/) those possess different chromosome numbers and can be listed into 4 groups (Table 1). The chromosome number show how they have diverged during speciation process (Levin and Wilsont, 1976). Despite of its importance in genus level for genetic and genomic studies in plant kingdom, the genome of only three *Arabidopsis* species have been published; *A. thaliana* (Arabidopsis Genome Initiative, 2000), *A. lyrata* (Hu *et al.,* 2011) and *A. halleri* (Briskine *et al.,* 2017). However, nowadays it is unavoidable to publish genome of other species to understand biological concepts, evolutionary mechanisms, and speciation events. Genomic data in plants provide necessary information for plant selection and are of help for understanding which biological mechanisms exist and revealing gene function and annotation (Jazayeri and Villamar Torres, 2017). The data generated by *Arabidopsis* model studies reveal that similar processes exist in other plants because of gene homology and orthology (Bauer *et al.,* 2004). However, such similarities and differences among living organisms make biology interesting and informative (Jones *et al.,* 2008).

As shown by chromosome number and 4 groups those were classified into two somewhere else (Koch and Matschinger, 2007), *A. thaliana* belonging to the Group1is the only species with 2n=10 separated from the rest of the genus (Table 1). The Group 2 comprises the species possessing 2n=16 including *A. croatica, A. halleri, A. lyrata, A. cebennensis, A. pedemontana,* as well *A. neglecta, A. umezawana* that are closely related species pairs (Bomblies and Weigel, 2010). *A. arenosa*, with 2n=26 and *A. umezawana* with 2n=32 belong to the Group3 and Group4 respectively suggesting that they are totally separated from the other species by genome duplication/deletion events (del Pozo and Ramirez-Parra, 2015). However, as the genome of all species is not published yet, it remains unclear how they have separated from their ancestor or diverged from each other. Four major groups based on genome-wide polymorphism have been reported corresponding to the widely distributed species *A. thaliana*, *A. halleri* , *A. lyrata* and *A. arenosa*, and three minor groups, corresponding to the geographically limited *A. croatica, A. cebennensis* and *A. pedemontana* (Novikova *et al.,* 2016).

**Table 1. Eleven confirmed *Arabidopsis* species by The Plant List in alphabetic order.**

| Name | Chromosome number 2n | Group |
|---|---|---|
| *A. arenosa* (L.) Lawalrée | 32 | 4 |
| *A. cebennensis* (DC.) O'Kane and Al-Shehbaz | 16 | 2 |
| *A. croatica* (Schott) O'Kane and Al-Shehbaz | 16 | 2 |
| *A. halleri* (L.) O'Kane and Al-Shehbaz | 16 | 2 |
| *A. lyrata* (L.) O'Kane and Al-Shehbaz | 16 | 2 |
| *A. neglecta* (Schult.) O'Kane and Al-Shehbaz | 16 | 2 |
| *A. pedemontana* (Boiss.) O'Kane and Al-Shehbaz | 16 | 2 |
| *A. petrogena* (A.Kern.) V.I.Dorof. | 16 | 2 |
| *A. suecica* (Fr.) Norrl. | 26 | 3 |
| *A. thaliana* (L.) Heynh. | 10 | 1 |
| *A. umezawana* Kadota | 16 | 2 |

For three published-genome species, *A. thaliana* is considered as inbreeder for closely out-crossing species *A. halleri* and *A. lyrata* whose divergence from the *A. thaliana* is estimated to have occurred around 5 million years ago (MYA) (Ramos-Onsins *et al.,* 2004). However, reported somewhere else, *A. thaliana* and *A. lyrata* shared an ancestor 10 MYA (Hollister *et al.,* 2011). *A. halleri* and *A. lyrata* have around 1.5-fold larger genome than *A. thaliana*. Although they share more than 80% sequence identity in their genomes, there are some differences among these three species (Hu *et al.,* 2011). *A. lyrata* is mainly a allogamous species whereas *A. thaliana* is an autogamous (Hollister *et al.,* 2011). One species-specific ecological trait belongs to *A. halleri* that is able to habit in the zones highly enriched in zinc (Zn) and cadmium (Cd) accumulating heavy metals (HM) in its aerial parts while *A. lyrata* and *A. thaliana* are both non-accumulators and sensitive to Zn and Cd (Roux *et al.,* 2011).

Despite these extremely different physiological traits, they belongs to the sister clade phylogenetically (Talke *et al.,* 2006). Their natural distribution is also different. *A. halleri* is mostly seen in continental Europe, and partially in Eastern Eurasia for one of its subspecies (*A. halleri* ssp. *gemmifera*) (Roux *et al.,* 2011). In comparison, *A. lyrata* is distributed in Western and Central Europe but with a circumboreal distribution. *A. thaliana* occurs in a worldwide manner and geographically is native of Europe and Central Asia (Al-Shehbaz and O'Kane, 2002). A study on 8 linkage groups of *A. halleri* revealed QTLs of rich regions of Zn tolerance (Willems *et al.,* 2007). *A. halleri* has some capacities by which it can tolerate excess of HMs and grow in the zones rich in HM. *A. halleri* roots are able to export metals like Cd, Zn toward its leaves (Cosio *et al.,* 2004) . This feature comes from HM ATPase HMA4 gene that has an increased copy number in *A. halleri* more than *A. thaliana* and *A. lyrata* (Bomblies and Weigel, 2010). *Brassicaceae* family show more taxa with hyperaccumulation of HMs such as Cd, Zn and Ni distributed in the species of *Alyssum, Thlaspi caerulescens* and *A. halleri* (Talke *et al.,* 2006). In one study metallothionein 2b and 3, APX and MDAR4 in the ascorbate-glutathione pathway and certain metal transporters in P(1B)-type ATPase, ZIP, Nramp, and CDF families were overexpressed in *A. halleri* than in *A. thaliana* ( Chiang *et al.,* 2006).

Previously *A. halleri* showed gene divergent of the 'HM–associated domain' category (Krämer, 2010; Verbruggen *et al.,* 2009). HMs in ionic form are highly reactive and toxic specially while being accumulated in plants as a critical agricultural problem (Olaniran *et al.,* 2013). There are two groups of HM; essential mineral elements like Cu, Zn, Ni and Mo required in low concentrations for adequate growth and development and non-essential metals with toxic character like Cd and Pb (Arif *et al.,* 2016). However, an excess of both groups is toxic. For example Cd can harm plants by DNA mutation, modifying protein side chains and demolishing phospholipid (Oono *et al.,* 2014).

The present study is aimed to *in silico* compare three genome-published *Arabidopsis* species including *A. halleri, A. lyrata* and *A. thaliana*. We discuss the similarities and differences among the three species in genomic level. We highlight the proteins related to HM hyperaccumulation in *A. halleri* as species-specific.

## MATERIALS AND METHODS

The reference protein, CDS (coding DNA sequences) and assembled genome files were downloaded from Phytozome (Goodstein *et al.,* 2012) version 12 (https://phytozome.jgi.doe.gov/). According to Phytozome, these files were for *A. thaliana* Assembly-version TAIR9, Annotation-version TAIR10; for *A. lyrata* Assembly-version v1, Annotation-versionv2.1 and for *A. halleri* Assembly-version v1 and Annotation-version v1.1. Ortho Venn was used to compare protein datasets among three species. Ortho Venn is a web-based program that can compare up to six protein datasets by clustering the input proteins of datasets and generate Venn diagram, Swiss Prot (Uni Prot) data (Poux *et al.,*

2016) and GO enrichment for the compared datasets (Wang *et al., 2015*). The default parameters were employed. The SwissProt hits and protein names were compared again manually for the clusters. Finally, the non-similar hits were studied to use for comparison of their function among three species and to extract the species-specific proteins of *A. halleri*. Blast+ V2.7.0 (Camacho *et al.,* 2009) was used to compare the gene models of three species using their CDS. The criteria for filtering were pident >80%, and Evalue < 1e-05. The blast-based pairwise comparisons were performed in two reciprocal directions by changing query and subject. GC content was calculated by the version 5.0.0 of the geecee program (Rice *et al.,* 2000) in Galaxy (Blankenberg *et al.,* 2007) for the genome and CDS of three species (https://main.g2.bx.psu.edu/).

## RESULTS AND DISCUSSION

**Protein clusters and singletons:** The proteins of three species formed 24816 clusters, 23500 orthologous clusters (at least contains two species), 1316 species-specific cluster sand 16371 singletons (those were not clustered and are unique to the species) (Figure 1 and Table-2). Although the genome of *A. thaliana* is smaller than two other species but its protein number is bigger likely due to better assembly and availability of data for *A. thaliana*. *A. halleri* shows less singletons, clusters, and unique proteins maybe because of its lower number of proteins compared to the other two. The common clusters were 17921 containing 62966 genes among all three species while pairwise common proteins between two species were 12769 distributed in 4733 clusters (Figure 1 and Table 3). Three species showed 66% common proteins. The more pairwise common proteins were seen between *A. thaliana* and *A. lyrata* with 22-23% for each species. The least belongs to *A. thaliana* while compared with *A. halleri* with 6.58%. These results are not in accordance with the chromosome number groups as it is expected that *A. halleri* and *A. lyrata* belongs to the same group and *A. thaliana* is diverged differently from them in another single species group. This might be due to the lower proteins reported for *A. halleri* and better annotation and assembly for *A. thaliana* and *A. lyrata*. *A. lyrata* showed the most species-specific proteins including clustered and singletons. The analyses about singletons for three species will be reported in further studies.

**Table 2. The number of clusters, proteins and singletons of three species.**

| Species | Proteins | Clusters | Singletons | Species clusters | Species specific clustered genes |
|---|---|---|---|---|---|
| *A. thaliana* | 35386 | 22822 | 5766 | 445 | 1086 |
| *A. lyrata* | 33132 | 23153 | 6044 | 685 | 1768 |
| *A. halleri* | 26911 | 20262 | 4561 | 186 | 469 |
| Total | 95429 | 66237 | 16371 | 1316 | 3323 |

**Table 3. The number of proteins for each comparison among three species. Species-specific proteins are sum of clustered proteins and singletons.**

| Comparison | Protein number | Percent of total species protein (%) |
|---|---|---|
| *A. thaliana* specific | 5766+1086 | 19.36 |
| *A. lyrata* specific | 6044+1768 | 23.57 |
| *A. halleri* specific | 4561+469 | 18.69 |
| *A. thaliana* _ *A. lyrata* | 7850 | *A. thaliana* 22.18<br>*A. lyrata* 23.69 |
| *A. thaliana* _ *A. halleri* | 2328 | *A. thaliana* 6.58<br>*A. halleri* 8.65 |
| *A. halleri* _ *A. lyrata* | 2591 | *A. halleri* 9.63<br>*A. lyrata* 7.82 |
| *A. halleri* _ *A. lyrata* _ *A. thaliana* | 62966 | 66 |

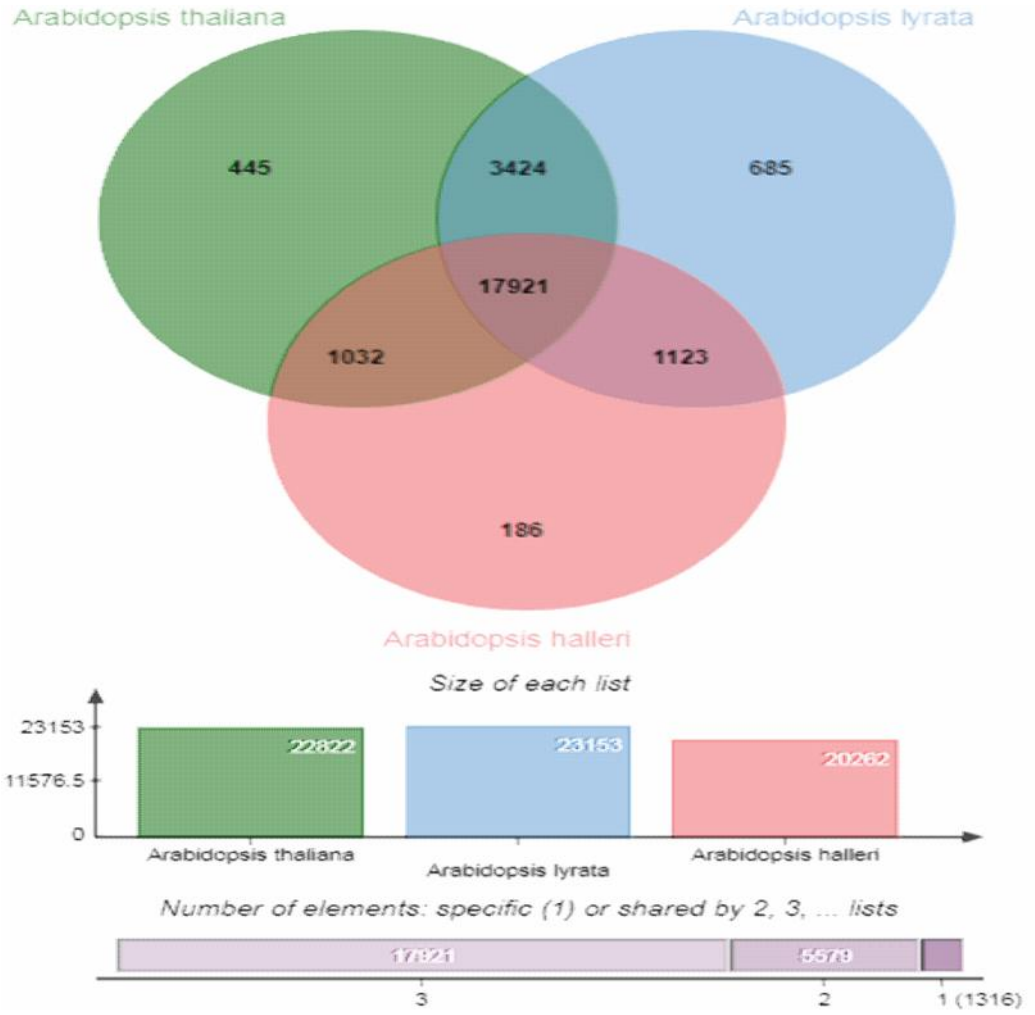


**Figure 1. Venn diagram for three *Arabidopsis* species based on the clusters of proteins. Below of Venn diagram the cluster number for each species is depicted. The three cluster groups and their number are shown as 3 shared by all three species, 2 shared between two species and 1 as specific cluster number.**

**Gene comparison:** The blast-based pairwise comparisons of CDS showed that there is at least more than 81% similarity among all three *Arabidopsis* species in general. Interestingly, the most similarity exist between *A. halleri* (query) and *A. lyrata* (subject) and the least between *A. lyrata* (query) and *A. halleri* (subject) (**Error! Not a valid bookmark self-reference.**). *A. halleri* showed more similarity with the two other species when it is query but when it is subject the least similarities seen. These results might be due to this fact that blast is based on the number of the genes in the database and there are less genes for *A. halleri*. The blast outputs when *A. halleri* is subject are similar to the outputs generated by Ortho Venn. These results are in accordance with the results obtained for protein clusters.

**Table 4. Transcript (CDS) pairwise comparisons of three *Arabidopsis* species.**

| Query<br>Subject | *A. halleri*<br>26911 | *A. lyrata*<br>33132 | *A. thaliana*<br>35386 |
|---|---|---|---|
| ***A. halleri*** | - | 27121 (81.8%) | 30228 (85.4%) |
| ***A. lyrata*** | 24854 (92%) | - | 32483 (91.8%) |
| ***A. thaliana*** | 24246 (90.1%) | 28756 (86.8%) | - |

GC content or G + C domains can help to understand structural, organizational and biological functions of genomes (Gao and Zhang, 2006) and comprehend ecology and evolution of particular taxa.

Because of its influence on CpG islands, DNA methylation, DNA stability, exon length and transcription, it plays a vital role in gene and genome regulation (Singh *et al.,* 2016; Tirado-Magallanes *et al.,* 2017). It determines the physical properties of the genome. It provides a vision on how genome/ chromosome size can be among different species, genera, towards kingdom (Li and Du, 2014). GC content has been reported 36% for genome of *A. thaliana* (Meister, 2005) and its ESTs (Shangguan *et al.,* 2013) that is in accordance with our results taking into account that we have used the most recently assembled genome. However, *A. lyrata* and *A. halleri* showed very close GC content that is more than that of *A. thaliana*. This suggests that they might be closer in terms of GC/AT domains (**Error! Not a valid bookmark self-reference.**). GC content of transcriptome show very close similarities among three species suggesting similar transcription pattern among them. Our results of GC content for transcripts are similar to the GC levels of coding sequences of *Oryza sativa* (43.6%), *Glycine max* (45%) and *A. thaliana* (44%) (Severin *et al.,* 2010; Smarda *et al.,* 2012). Interestingly, CDS size is not in coherent with genome size and the biggest one belongs to *A. thaliana* with the smallest genome. This might be due to high copy number of genes in *A. thaliana* although its genome size and chromosome number is less than the other two species. This needs further studies to reveal how they have diverged from their common ancestor and which evolutionary events have occurred

**Table 5. GC content and number of base pairs for 3 *Arabidopsis* species.**

| Species | GC content (genome) | Genome size (bp) | GC content (CDS) | CDS size (bp) |
|---|---|---|---|---|
| *A. halleri* | 37.05 | 127615339 | 43.74 | 31383222 |
| *A. lyrata* | 37.74 | 206667935 | 44.62 | 38615903 |
| *A. thaliana* | 36 | 119667750 (119146348) | 44.40 | 43546761 |

**Genes related to resistance or response to HMs:** Interestingly, among the different genes, those are species-specific for *A. halleri*, there are the 26 genes from 10 clusters related to HM (like Cd, Al, As and Zn) resistance and response. They belong to genes with different functions like transporter, different detoxifying enzymes, signaling and regulation (**Error! Not a valid bookmark self-reference.**). These genes have been reported to be involved in HM response and tolerance in plants (DalCorso *et al.,* 2010). Most of these genes reported by Chen *et al.,* are differentially expressed in *Fagopyrum esculentum* Moench functioning in resistance and response to aluminum (Al) (Chen *et al.,* 2017). Meantime, they are related to other HM responses in plants as well.

**Table 6. The species-specific genes of *A. halleri* related to HM response.**

| Protein | SwissProt hit | HM | Function | Reference |
|---|---|---|---|---|
| Araha.10286s0004.2.p<br>Araha.10286s0004.1.p | oligopeptide transporter | Al, Cd | Transport, phytochelatin | (Patel 2007; Chen *et al.,* 2017) |
| Araha.3113s0004.2.p<br>Araha.3113s0004.1.p | putative purine permease | Al | Transport | (Chen *et al.,* 2017) |
| Araha.11872s0001.5.p<br>Araha.11872s0001.1.p<br>Araha.11872s0001.2.p<br>Araha.11872s0001.3.p<br>Araha.11872s0001.4.p<br>Araha.11872s0001.6.p | Aldehyde dehydrogenase family 2-member C4 Q56YU0 | Al, Cd | Detoxification | (Chen *et al.,* 2017) |
| Araha.7027s0001.1.p<br>Araha.7027s0001.2.p | Cellulose synthase-like protein O23386 | Al, Cd | Cell wall | (Song *et al.,* 2013; Chen *et al.,* 2017) |
| Araha.32714s0001.4.p<br>Araha.32714s0001.3.p | Importin subunit alpha-1 Q96321 | Al, Cu, Si | Transport | (Chen *et al.,*, 2017; Jang *et al.,* 2018; Wang *et al.,*, 2017) |
| Araha.23960s0001.1.p<br>Araha.23960s0001.2.p | Cysteine synthase 1 P47998 | Cd, Zn | Detoxification | (Kawashima *et al.,* 2004; Kim *et al.,* 2011; Konlechner *et al.,* 2013) |
| Araha.0018s0018.1.p | Plasmodesmata callose-binding | Al | Transport | (Chen *et al.,* 2017) |

| | | | | |
|---|---|---|---|---|
| Araha.0018s0018.2.p<br>Araha.0018s0018.3.p | protein Q9M2K6 | | | |
| Araha.7955s0008.4.p<br>Araha.7955s0008.2.p<br>Araha.7955s0008.3.p | Probable BOI-related E3 ubiquitin-protein ligase 3 Q9LDD1 | Al | Degradation, | (Chen *et al.,* 2017) |
| Araha.51292s0001.1.p<br>Araha.51292s0001.2.p | SNF1-related protein kinase catalytic subunit alpha KIN11 P92958 | Al, Cd | Signaling, ROS | (Kulik *et al.,* 2012; Chen *et al.,* 2017) |
| Araha.8372s0003.1.p<br>Araha.8372s0003.2.p | Translocase Q9SLF3 | Al, Cd | Transport | (Chen *et al.,* 2017) |

The genes with GO terms related to Zn and Cd compared. The findings show that *A. halleri* possess the enzyme "Cysteine synthase 1" under GO term response to Cd as species-specific term (Kawashima *et al.,* 2004; Kim *et al.,* 2011). In plants resistant to Cd stress cysteine is synthesized in order to regulate glutathione GSH biosynthesis. Then GSH a non-protein thiol acts as an antioxidant to detoxification of Cd (Gill and Tuteja, 2011). On the other hand, sulfur (S) and Cd show a direct relation as uptake and assimilation of S can determine Cd resistance in crops via cysteine synthesis targeting GSH. *A. halleri* possesses two proteins of Cysteine synthase.

Oligopeptide transport (OPT) gene family of integral membrane proteins is involved in diverse functions like long-distance sulfur distribution, nitrogen mobilization, metal homeostasis, and HM sequestration through GSH transport, metal-chelates, and peptides (Lubkowitz, 2011). In addition, phytochelatins/metal chelats are synthesized in plants resistant to Cd as a part of detoxification (Gill and Tuteja, 2011). It is proved that overexpression of OPT causes Cd hypersensitivity and accumulation in *A. thaliana* roots (Mendoza-Cózatl *et al.,* 2014). *A. halleri* has two species-specific members of OPT those can make it tolerant to HM and may facilitate Cd hyperaccumulation in *A. halleri*.

Oxidation of toxic aldehydes to carboxylic acids is of functions of ALDHs those belong to a family of NAD(P)+-dependent enzymes. They are highly expressed in response to drought, salinity, water excess, oxidative stress and HMs (Jimenez-Lopez *et al.,* 2016). ALDH2 has been reported that function in Al stress (Singh *et al.,* 2013). ALDH can function via decreasing accumulation of lipid peroxidation-derived reactive aldehydes as a detoxification pathway to increase tolerance in plants (Sunkar *et al.,* 2003). We found 6 ALDHs in *A. halleri* those may function similarly in response to HM.

*Sedum alfredii* Hance hyperaccumulating ecotype has shown constitutively upregulation of cellulose synthase gene while being exposed to Cd (Gao *et al., 2013*). In *Elsholtzia splendens*, a hyperaccumulator of Cu, the composition and distribution of root cell wall and cellulose was altered in excess of Cu that can improve cellulose, hemicellulose and pectin contents during HM exposure (Le Gall *et al.,* 2015). In a low-Cd mutant of rice, a correlation between cellulose synthase subunit 9 and Cd accumulation in roots, leaves, grains and altered vascular structure have been reported. This mutation of CESA9 (E101K) affects Cd transport efficiency targeting low Cd accumulation by altering cell-wall properties in the conducting tissues (Song *et al.,* 2013). However, cell wall and xylem alteration are involved in HM tolerance mechanisms by preventing HM translocation or deposition (Vollenweider *et al.,* 2006). Two copies of cellulose synthase are found in *A. halleri*.

Importin is involved in nuclear import and export processes transporting other proteins intro nucleus by binding to nuclear localization signal (NLS) motifs (Tran *et al.,* 2014). It has two subunits; α and β. Importin α binds to NLS and β binds to molecules transporting them to the nucleus. A C2-domain-like protein with NLS motif is highly expressed in response to different HMs in barley (Ouelhadj *et al.,* 2006). In response to Cu and Si, importin αis induced and can act with NLS (Wang *et al.,* 2017; Jang *et al.,* 2018). In *A. halleri* with these two proteins, it may play a similar role in binding to such motif to tolerate HM excess.

When plants are exposed to HM (Al, Cd, Cu etc.), they accumulate callose at plasmodesmata causing alteration in their permeability (De Storme and Geelen, 2014). This then regulate heavy metal transport via plasmodesmata (O'Lexy, 2017). In order to activate such permeability, plasmodesmata callose-binding protein plays a role by interacting with other transporters like CTL1 (Gao *et al.,* 2017) in vesicle trafficking. This protein can manage metal transport and function in metallic ion channels (Gao *et al.,* 2017). They may facilitate channels from which HM can import and export to root cells toward leaves. Three clustered proteins of plasmodesmata callose-binding protein suggest the existence of similar mechanism for *A. halleri*.

Ubiquitination pathway is one of the ways whose participation in plant abiotic stress tolerance has been proved (Dametto *et al.,* 2015) . Protein degradation by the ubiquitin/proteasome 26S system plays an important role in HM stress response via removing damaged proteins (Dametto *et al.,* 2015). BOI-related E3 ubiquitin-protein ligase 3 as an E3 ubiquitin-protein ligase functioning in ubiquitination has been reported that is involved in the regulation of pathogen and abiotic stress responses by facilitating degradation of

MYB108/BOI (Luo *et al.,* 2010). One OSHIR1 (E3 ubiquitin-protein ligase) is involved in As and Cd uptake (Hasan *et al.,* 2017) by interacting with OsTIP4;1 a tonoplast intrinsic protein in the plasma membrane. This degradation and regulatory function may occur in *A. halleri* by three discovered proteins.

SNF1-related protein kinase catalytic subunit alpha KIN11 belongs to the SNF1-related protein kinase family that is involved in many biological processes like signaling cascade, senescence, transcriptional and post transcriptional regulatory, phosphorylation of enzymes involved in various pathways (Crozet *et al.,* 2014; Wang, 2017). SnRK KIN11 control convergent reprogramming to darkness, sugar level and stress (Kulik *et al.,* 2011). One of this family SnRK2 functions in Cd resistance by regulating ROS accumulation in plant root (Kulik *et al.,* 2012). By employing these two disclosed proteins, *A. halleri* might follow the similar pattern.

Translocase is one protein located in inner and outer membrane of mitochondrion. It mediates transport of proteins from nuclear coding genes in mitochondrion. Its expression has increased in ROS scavenging in response to Cd (Yu *et al.,* 2017) and Al (Chen *et al.,* 2017). However, its exact role is not clear yet.

Phytochelatins, phytohormones like ABA, BR, SA and transporters are crosstalk elements in diverse plant stresses that are also involved in HM stress (Núñez Vázquez *et al.,* 2013). HM responsive genes share similar function as they employ almost identical mechanisms to respond to other abiotic stresses like osmotic, oxidative, drought, salinity, as well as biotic stress (Chinnusamy *et al.,* 2004; Zhu, 2016). As there exist diverse stresses at the same time surrounding plants, these crosstalk and common responsive genes can collaborate to mitigate the impacts of stresses on plants (Chinnusamy *et al.,* 2004). It can be worthy then to study different stresses at the same time to understand how different stresses (HM and osmotic stress for example) can share alike responses and genes (Jazayeri, 2015).

**Conclusions:** The findings disclosed that in species level there are around 10-20% difference among three *Arabidopsis* species taking into account the criteria for similarity. Speciation might count on such differences that we can see among close species of a genus. Interestingly, *A. halleri* has a genetically physiological speciation character different from two other species; i.e. resistance to HMs. The findings imply that *A. halleri* has acquired the built-in mechanism to tolerate and detoxify HMs during speciation process and evolutionary events. *A. halleri* is naturally hyperaccumulator of HMs like Zn and Cd and hypertolerant to HMs while two other species are not tolerant. However, the function in resistance process in *Arabidopsis* needs further expressional and functional analyses to disclose the role of the presented genes in HM detoxification and accumulation. HM stress as an abiotic stress can share common mechanisms causing tolerance in plants. Signaling network, functional enzymes and regulatory system are involved in several stress response in plants. We hypothesize that oxidative, osmotic, drought and salinity stress responses in plants can make networks of coexpressing genes by which plants control diverse environmental factors and ecological impacts. Further studies are required to disclose common responses among abiotic stress with HMs.

## REFERENCES

Al-Shehbaz, I., and S. O'Kane (2002). Taxonomy and Phylogeny of *Arabidopsis* (Brassicaceae). Page *in* C. Somerville and E. Meyerowitz, editors. The Arabidopsis Book. American Society of Plant Biologists. USA.

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature. 408(6814):796–815.

Arif, N., V. Yadav, S. Singh, S. Singh, P. Ahmad, R. K. Mishra, S. Sharma, D. K. Tripathi, N. K. Dubey, and D. K. Chauhan (2016). Influence of high and low levels of plant-beneficial heavy metal ions on plant growth and development. Frontiers in Environmental Science. 4:69.

Bauer, P., T. Thiel, M. Klatte, Z. Bereczky, T. Brumbarova, R. Hell, and I. Grosse (2004). Analysis of sequence, map position, and gene expression reveals conserved essential genes for iron uptake in Arabidopsis and tomato. Plant Physiology. 136(4):4169–83.

Blankenberg, D., J. Taylor, I. Schenck, J. He, Y. Zhang, M. Ghent, N. Veeraraghavan, I. Albert, W. Miller, K. D. Makova, R. C. Hardison, and A. Nekrutenko (2007). A framework for collaborative analysis of ENCODE data: making large-scale analyses biologist-friendly. Genome Research. 17(6):960–4.

Bomblies, K., and D. Weigel (2010). Arabidopsis and relatives as models for the study of genetic and genomic incompatibilities. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. 365(1547):1815–23.

Briskine, R. V., T. Paape, R. Shimizu-Inatsugi, T. Nishiyama, S. Akama, J. Sese, and K. K. Shimizu (2017). Genome assembly and annotation of *Arabidopsis halleri* , a model for heavy metal hyperaccumulation and evolutionary ecology. Molecular Ecology Resources. 17(5):1025–1036.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden (2009). BLAST+: architecture and applications.

BMC Bioinformatics. 10:421.

Chen, W. W., J. M. Xu, J. F. Jin, H. Q. Lou, W. Fan, and J. L. Yang (2017). Genome-wide transcriptome analysis reveals conserved and distinct molecular mechanisms of Al resistance in Buckwheat (*Fagopyrum esculentum* Moench) leaves. International J. Molecular Sciences. 18(9).

Chiang, H.-C., J.-C. Lo, and K.-C. Yeh (2006). Genes associated with heavy metal tolerance and accumulation in Zn/Cd hyperaccumulator *Arabidopsis halleri*: a genomic survey with cDNA microarray. Environmental Science & Technology. 40(21):6792–6798.

Chinnusamy, V., K. Schumaker, and J.-K. Zhu (2004). Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. J. Experimental Botany. 55(395):225–236.

Cosio, C., E. Martinoia, and C. Keller (2004). Hyperaccumulation of cadmium and zinc in *Thlaspi caerulescens* and *Arabidopsis halleri* at the leaf cellular level. Plant Physiology. 134(2):716–25.

Crozet, P., L. Margalha, A. Confraria, A. Rodrigues, C. Martinho, M. Adamo, C. A. Elias, and E. Baena-González (2014). Mechanisms of regulation of SNF1/AMPK/SnRK1 protein kinases. Frontiers in Plant Science. 5:190.

DalCorso, G., S. Farinati, and A. Furini (2010). Regulatory networks of cadmium stress in plants. Plant Signaling & Behavior. 5(6):663–667.

Dametto, A., G. Buffon, É. A. Dos Reis Blasi, and R. A. Sperotto (2015). Ubiquitination pathway as a target to develop abiotic stress tolerance in rice. Plant Signaling & Behavior. 10(9):e1057369.

Le Gall, H., F. Philippe, J.-M. Domon, F. Gillet, J. Pelloux, and C. Rayon (2015). Cell wall metabolism in response to abiotic stress. Plants (Basel, Switzerland). 4(1):112–66.

Gao, F., and C.-T. Zhang (2006). GC-Profile: a web-based tool for visualizing and analyzing the variation of GC content in genomic sequences. Nucleic Acids Research. 34(Web Server issue):W686-91.

Gao, J., L. Sun, X. Yang, and J.-X. Liu (2013). Transcriptomic analysis of cadmium stress response in the heavy metal hyperaccumulator *Sedum alfredii* Hance. PLoS ONE. 8(6):e64643.

Gao, Y.-Q., J.-G. Chen, Z.-R. Chen, D. An, Q.-Y. Lv, M.-L. Han, Y.-L. Wang, D. E. Salt, and D.-Y. Chao (2017). A new vesicle trafficking regulator CTL1 plays a crucial role in ion homeostasis. PLoS Biology. 15(12):e2002978.

Gill, S. S., and N. Tuteja (2011). Cadmium stress tolerance in crop plants: Probing the role of sulfur. Plant Signaling & Behavior. 6(2):215.

Goodstein, D. M., S. Shu, R. Howson, R. Neupane, R. D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam, and D. S. Rokhsar (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research. 40 (Database issue):D1178-86.

Hasan, M. K., Y. Cheng, M. K. Kanwar, X.-Y. Chu, G. J. Ahammed, and Z.-Y. Qi (2017). Responses of plant proteins to heavy metal stress—A review. Frontiers in Plant Science. 8(1492).

Hollister, J. D., L. M. Smith, Y.-L. Guo, F. Ott, D. Weigel, and B. S. Gaut (2011). Transposable elements and small RNAs contribute to gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata*. Proceedings of the National Academy of Sciences of the United States of America. 108(6):2322–7.

Hu, T. T., P. Pattyn, E. G. Bakker, J. Cao, J.-F. Cheng, R. M. Clark, N. Fahlgren, J. A. Fawcett, J. Grimwood, H. Gundlach, G. Haberer, J. D. Hollister, S. Ossowski, R. P. Ottilar, A. A. Salamov, K. Schneeberger, M. Spannagl, X. Wang, L. Yang, M. E. Nasrallah, J. Bergelson, J. C. Carrington, B. S. Gaut, J. Schmutz, K. F. X. Mayer, Y. Van de Peer, I. V Grigoriev, M. Nordborg, D. Weigel, and Y.-L. Guo (2011). The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. Nature Genetics. 43(5):476–481.

Jang, S.-W., Y. Kim, A. L. Khan, C.-I. Na, and I.-J. Lee (2018). Exogenous short-term silicon application regulates macro-nutrients, endogenous phytohormones, and protein expression in *Oryza sativa* L. BMC Plant Biology. 18(1):4.

Jazayeri, S. M. (2015). Characterization of genes related to oil palm (Elaeis guineensis Jacq.) drought stress responses. PhD thesis. National University of Colombia. Colombia.

Jazayeri, S. M., and R. Villamar Torres (2017). Genomic and transcriptomic approaches toward plant selection. J. Science and Research: Revista Ciencia e Investigación. 2(8):54.

Jimenez-Lopez, J. C., F. J. Lopez-Valverde, P. Robles-Bolivar, E. Lima-Cabello, E. W. Gachomo, and S. O. Kotchoni (2016). Genome-wide identification and functional classification of tomato (*Solanum lycopersicum*) Aldehyde Dehydrogenase (ALDH) gene superfamily. PLoS ONE. 11(10).

Jones, A. M., J. Chory, J. L. Dangl, M. Estelle, S. E. Jacobsen, E. M. Meyerowitz, M. Nordborg, and D. Weigel (2008). The impact of Arabidopsis on human health: diversifying our portfolio. Cell. 133(6):939–43.

Kawashima, C. G., M. Noji, M. Nakamura, Y. Ogra, K. T. Suzuki, and K. Saito (2004). Heavy metal tolerance of transgenic tobacco plants over-expressing cysteine synthase. Biotechnology Letters. 26(2):153–157.

Kim, Y.-N., J.-S. Kim, S.-G. Seo, Y. Lee, S.-W. Baek, I.-S. Kim, H.-S. Yoon, K.-R. Kim, S.-H. Kim, and K.-H. Kim (2011). Cadmium resistance in tobacco plants expressing the MuSI gene. Plant Biotechnology Reports. 5(4):323–329.

Koch, M. A., and M. Matschinger (2007). Evolution and genetic differentiation among relatives of *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America. 104(15):6272–7.

Konlechner, C., M. Türktaş, I. Langer, M. Vaculík, W. W. Wenzel, M. Puschenreiter, and M.-T. Hauser (2013). Expression of zinc and cadmium responsive genes in leaves of willow (*Salix caprea* L.) genotypes with different accumulation characteristics. Environmental Pollution. 178:121–127.

Krämer, U. (2010). Metal Hyperaccumulation in Plants. Annual Review of Plant Biology. 61(1):517–534.

Kulik, A., A. Anielska-Mazur, M. Bucholc, E. Koen, K. Szymanska, A. Zmienko, E. Krzywinska, I. Wawer, F. McLoughlin, D. Ruszkowski, M. Figlerowicz, C. Testerink, A. Sklodowska, D. Wendehenne, and G. Dobrowolska (2012). SNF1-related protein kinases type 2 are involved in plant responses to cadmium stress. Plant Physiology. 160(2):868–83.

Kulik, A., I. Wawer, E. Krzywińska, M. Bucholc, and G. Dobrowolska (2011). SnRK2 protein kinases--key regulators of plant response to abiotic stresses. Omics : a J. Integrative Biology. 15(12):859–72.

Levin, D. A., and A. C. Wilsont (1976). Rates of evolution in seed plants: Net increase in diversity of chromosome numbers and species numbers through time (polyploidy/aneuploidy/population structure/speciation/fossil record). Genetics. 73(6):2086–2090.

Li, X.-Q., and D. Du (2014). Variation, evolution, and correlation analysis of C+G content and genome or chromosome size in different kingdoms and phyla. PloS ONE. 9(2):e88339.

Lubkowitz, M. (2011). The oligopeptide transporters: A small gene family with a diverse group of substrates and functions? Molecular Plant. 4(3):407–415.

Luo, H., K. Laluk, Z. Lai, P. Veronese, F. Song, and T. Mengiste (2010). The Arabidopsis Botrytis Susceptible1 Interactor defines a subclass of RING E3 ligases that regulate pathogen and stress responses. Plant Physiology. 154(4):1766–82.

Meister, A. (2005). Calculation of binding length of base-specific DNA dyes by comparison of sequence and flow cytometric data. Application to *Oryza sativa* and *Arabidopsis thaliana*. J. Theoretical Biology. 232(1):93–97.

Mendoza-Cózatl, D. G., Q. Xie, G. Z. Akmakjian, T. O. Jobe, A. Patel, M. G. Stacey, L. Song, D. W. Demoin, S. S. Jurisson, G. Stacey, and J. I. Schroeder (2014). OPT3 is a component of the iron-signaling network between leaves and roots and misregulation of OPT3 leads to an over-accumulation of cadmium in seeds. Molecular Plant. 7(9):1455–1469.

Novikova, P. Y., N. Hohmann, V. Nizhynska, T. Tsuchimatsu, J. Ali, G. Muir, A. Guggisberg, T. Paape, K. Schmid, O. M. Fedorenko, S. Holm, T. Säll, C. Schlötterer, K. Marhold, A. Widmer, J. Sese, K. K. Shimizu, D. Weigel, U. Krämer, M. A. Koch, and M. Nordborg (2016). Sequencing of the genus *Arabidopsis* identifies a complex history of nonbifurcating speciation and abundant trans-specific polymorphism. Nature Genetics. 48(9):1077–1082.

Núñez Vázquez, M., Y. R. Guerrero, L. M. González, and W. Torres De La Noval (2013). Brassinosteroids and plant responses to heavy metal stress. An overview. Open J. Metal. 3:34–41.

O'Lexy, R. (2017). Exposure to heavy metal stress regulates intercellular signaling via callose deposition and breakdown. PhD thesis. AAI10636796. University of Pennsylvania. USA

Olaniran, A. O., A. Balgobind, and B. Pillay (2013). Bioavailability of heavy metals in soil: impact on microbial biodegradation of organic compounds and possible improvement strategies. International J. Molecular Sciences. 14(5):10197–228.

Oono, Y., T. Yazawa, Y. Kawahara, H. Kanamori, F. Kobayashi, H. Sasaki, S. Mori, J. Wu, H. Handa, T. Itoh, and T. Matsumoto (2014). Genome-wide transcriptome analysis reveals that cadmium stress signaling controls the expression of genes in drought stress signal pathways in rice. PLoS ONE. 9(5):e96946.

Ouelhadj, A., P. Kuschk, and K. Humbeck (2006). Heavy metal stress and leaf senescence induce the barley gene HvC2d1 encoding a calcium-dependent novel C2 domain-like protein. New Phytologist. 170(2):261–273.

Patel, A. A. (2007). Role of the Arabidopsis peptide transporter AtOPT6 in heavy metal detoxification and plant-pathogen interaction. PhD thesis. University of Missouri-Columbia. USA.

Poux, S., C. N. Arighi, M. Magrane, A. Bateman, C.-H. Wei, Z. Lu, E. Boutet, H. Bye-A-Jee, M. L. Famiglietti, and B. Roechert (2016). On expert curation and sustainability: UniProtKB/Swiss-Prot as a case study. bioRxiv.:094011.

del Pozo, J. C., and E. Ramirez-Parra (2015). Whole genome duplications in plants: An overview from *Arabidopsis*. J. Experimental Botany. 66(22):6991–7003.

Ramos-Onsins, S. E., B. E. Stranger, T. Mitchell-Olds, and M. Aguadé (2004). Multilocus analysis of variation and speciation in the closely related species *Arabidopsis halleri* and *A. lyrata*. Genetics. 166(1):373–88.

Rice, P., I. Longden, and A. Bleasby (2000). EMBOSS: the European Molecular Biology Open Software Suite. Trends in Genetics : TIG. 16(6):276–7.

Roux, C., V. Castric, M. Pauwels, S. I. Wright, P. Saumitou-Laprade, and X. Vekemans (2011). Does speciation between *Arabidopsis halleri* and *Arabidopsis lyrata* coincide with major changes in a molecular target of adaptation? PLoS ONE. 6(11):e26872.

Severin, A. J., J. L. Woody, Y.-T. Bolon, B. Joseph, B. W. Diers, A. D. Farmer, G. J. Muehlbauer, R. T. Nelson, D. Grant, J. E. Specht, M. A. Graham, S. B. Cannon, G. D. May, C. P. Vance, and R. C. Shoemaker (2010). RNA-Seq atlas of *Glycine max*: a guide to the soybean transcriptome. BMC Plant Biology. 10:160.

Shangguan, L., J. Han, E. Kayesh, X. Sun, C. Zhang, T. Pervaiz, X. Wen, and J. Fang (2013). Evaluation of genome sequencing quality in selected plant species using expressed sequence tags. PloS ONE. 8(7):e69890.

Singh, R., R. Ming, and Q. Yu (2016). Comparative analysis of GC content variations in plant genomes. Tropical Plant Biology. 9:10.1007/s12042-016-9165-4.

Singh, S., C. Brocker, V. Koppaka, Y. Chen, B. C. Jackson, A. Matsumoto, D. C. Thompson, and V. Vasiliou (2013). Aldehyde dehydrogenases in cellular responses to oxidative/electrophilic stress. Free Radical Biology & Medicine. 56:89–101.

Smarda, P., P. Bureš, J. Smerda, and L. Horová (2012). Measurements of genomic GC content in plant genomes with flow cytometry: a test for reliability. The New phytologist. 193(2):513–21.

Song, X.-Q., L.-F. Liu, Y.-J. Jiang, B.-C. Zhang, Y.-P. Gao, X.-L. Liu, Q.-S. Lin, H.-Q. Ling, and Y.-H. Zhou (2013). Disruption of secondary wall cellulose biosynthesis alters cadmium translocation and tolerance in rice plants. Molecular Plant. 6(3):768–780.

De Storme, N., and D. Geelen (2014). Callose homeostasis at plasmodesmata: molecular regulators and developmental relevance. Frontiers in Plant Science. 5:138.

Sunkar, R., D. Bartels, and H.-H. Kirch (2003). Overexpression of a stress-inducible aldehyde dehydrogenase gene from *Arabidopsis thaliana* in transgenic plants improves stress tolerance. The Plant J.: for Cell and Molecular Biology. 35(4):452–64.

Talke, I. N., M. Hanikenne, and U. Krämer (2006). Zinc-dependent global transcriptional control, transcriptional deregulation, and higher gene copy number for genes in metal homeostasis of the hyperaccumulator *Arabidopsis halleri*. Plant Physiology. 142(1):148–167.

Tirado-Magallanes, R., K. Rebbani, R. Lim, S. Pradhan, and T. Benoukraf (2017). Whole genome DNA methylation: beyond genes silencing. Oncotarget. 8(3):5629–5637.

Tran, E. J., M. C. King, and A. H. Corbett (2014). Macromolecular transport between the nucleus and the cytoplasm: Advances in mechanism and emerging links to disease. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 1843(11):2784–2795.

Verbruggen, N., C. Hermans, and H. Schat (2009). Molecular mechanisms of metal hyperaccumulation in plants. New Phytologist. 181(4):759–776.

Vollenweider, P., C. Cosio, M. S. Günthardt-Goerg, and C. Keller (2006). Localization and effects of cadmium in leaves of a cadmium-tolerant willow (*Salix viminalis* L.). Environmental and Experimental Botany. 58(1–3):25–40.

Wang, C., J. Wang, X. Wang, Y. Xia, C. Chen, Z. Shen, and Y. Chen (2017). Proteomic analysis on roots of Oenothera glazioviana under copper-stress conditions. Scientific Reports. 7(1):10589.

Wang, Y. (2017). The Role of SNF1-Related Protein Kinase 1 (SnRK1) in regulating intermediary metabolism in *Arabidopsis thaliana*. PhD thesis. University of Guelph. Canada.

Wang, Y., D. Coleman-Derr, G. Chen, and Y. Q. Gu (2015). OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Research. 43(W1):W78-84.

Willems, G., D. B. Dräger, M. Courbot, C. Godé, N. Verbruggen, and P. Saumitou-Laprade (2007). The genetic basis of zinc tolerance in the metallophyte *Arabidopsis halleri* ssp. *halleri* (Brassicaceae): An analysis of quantitative trait loci. Genetics. 176(1):659.

Yu, R., Y. Tang, C. Liu, X. Du, C. Miao, and G. Shi (2017). Comparative transcriptomic analysis reveals the roles of ROS scavenging genes in response to cadmium in two pak choi cultivars. Scientific Reports. 7(1):9217.

Zhu, J.-K. (2016). Abiotic stress signaling and responses in plants. Cell. 167(2):313–324.