Oryza Vol. 55 No. 2, 2018 (248-259)
DOI 10.5958/2249-5266.2018.00031.0

Computational analysis of genes encoding for molecular determinants of arsenic tolerance in rice (*Oryza sativa* L.) to engineer low arsenic content varieties

Chinmaya Kumar Das[1]*, D Bastia[2], SC Swain[3] and SS Mahapatra[1]

[1]Regional Research & Technology Transfer Station, Ranital, OUAT, Bhubaneswar, Odisha, India

[2]College of Agriculture, Orissa University of Agriculture & Technology, Bhubaneswar, Odisha, India

[3]Agro-Polytechinique Centre, Ranital, OUAT, Bhubaneswar, Odisha, India

**Corresponding author e-mail: iitk.chinmaya@gmail.com*

Received : 04 February 2018 Accepted : 29 May 2018 Published : 27 June 2018

ABSTRACT

Rice (Oryza sativa L.) stands at the centre stage of the dietary landscape as one of the principal staple food crop in the globe. Unfortunately, it is known to be a potential accumulator of toxic metalloid arsenic, a non-threshold, class 1 human carcinogen, in its grains. Therefore, it is imperative to device strategies to minimize the level of arsenic toxicity in rice grains. In this investigation, we have analyzed the transcriptional regulatory architecture of genes involved in arsenic metabolism in rice by analyzing their gene coexpression modules, their spatiotemporal expression profiles in different tissues during developmental phases of rice and scanning of their upstream promoter regions for potential transcription factor binding sites (TBFS) using computational tools & genomic/ transcriptomic data available in a set of genomic databases (RiceFREND, RiceXPro, PlantPAN2.0, RAP-DB, OryzaBase, MSU Rice Genome Annotation Project Database & Resources and NCBI) and manually written R-Scripts to understand the molecular basis of arsenic tolerance for designing varietal interventions with minimum arsenic toxicity. In the analysis of gene coexpression modules of candidate genes of arsenic metabolism, MYB or MYB-related TF encoding genes were enriched in the constructed gene-coexpression network of one candidate gene OsHAC1;1 which encodes for an arsenate reductase enzyme in rice. Further, potential TBFS of MYB were found on the upstream promoter region of OsHAC1; 1 gene. Upregulated and downregulated gene expression pattern of both OsMYB48 (A P-type R2R3 MYB transcription factor) and OsHAC1;1 in root tissues and most of the reproductive structures (inflorescences, anthers, pistils, ovaries, embryos and endosperms), respectively were found in the spatiotemporal analysis of gene expression . Therefore, contrasting differential gene expression pattern of both OsMYB48 and OsHAC1;1 genes in root and reproductive organs has a potential significance in restricting toxic arsenites in the cellular vacuoles of root tissues with canonical ABC transporter (OsABCC1) at vegetative and reproductive stages and thereby limiting the translocation of toxic As into the grains. Hence, functional validation of OsMYB48 mediated transcriptional modulation of OsHAC1; 1 gene in the root and reproductive tissues would help in gaining insights into the basis of arsenic tolerance in rice and designing genetic strategies for evolving low grain-arsenic content varieties of paddy using genetic engineering and plant breeding tools.

Key words: *Rice, arsenic tolerance, gene-expression*

INTRODUCTION

Rice (*Oryza sativa* L.) is one of the widely cultivated cereal crops that feed nearly 3 billion people (Chen et al., 2017) in the globe. It is a vital and an integral component of the food security in the world. Unfortunately, it is known to be a hyper-accumulator of toxic metalloid arsenic (As) (Williams et al., 2005; Sohn, 2014). Paddy cultivated in arsenic contaminated soils can accumulate up to 2 mg Kg^{-1} As in grains and up to 92 mg Kg^{-1} in straw (Williams et al., 2005). Consumption of rice grains polluted with toxic arsenic

imposes significant and serious threat to human health. Long-term exposure to As is associated with diseases such as hyper-pigmentation, keratosis, and skin & internal cancers (Argos et al., 2010). As rice is a globally consumed crop, it is imperative to device strategies to minimize the level of arsenic in rice grains to prevent As-associated health risks in human population. Arsenic has been detected both in inorganic form [as arsenite (AsIII) and arsenate (AsV)] and organic form [as dimethylarsinic acid (DMAV)] in rice grains (Williams et al., 2005). Therefore, solution space to minimize the level of arsenic in rice grains lies in designing genetic interventions to reduce root to grain movement of As. In this direction, many genes involved in arsenic metabolism have been identified in different genetic model organisms including rice. In the genome of rice, Os02g0745100 (*LSI1*), Os03g0107300 (*LSI2*), Os10g0444700 (*OsPT8*), Os04g0620000 (*OsABCC1*), Os07g0258400 (*OsNRAMP1*), Os02g0102300 (*OsHAC1;1*), Os01g0368900 (*OsGRX4*), Os01g0955700 (*OsCLT1*) and Os03g0195800 (*OsSultr1;1*) are the important genes involved in the arsenic metabolism (Chen et al., 2017) whose genomic atlas and gene anatomy are depicted in Fig. 1A and Fig. 1B.

Arsenic metabolism in simple terms involves four broad steps including cellular trafficking, translocation, reduction and sequestration of arsenic molecules from the soil which is summarized in Fig. 1C. Arsenate molecules from the soil enter into the root through a phosphate transporter named as *OsPT8* (*OsPht1;8*) (Wu et al., 2011; Wang et al., 2016). *OsPT8* is one of the 13 putative high affinity Pi transporters in rice (Paszkowski et al., 2002) and is expressed in both root and shoot tissues of rice (Paszkowski et al., 2002; Jia et al., 2011). Overexpression of *OsPT8* has been reported to significantly enhance the uptake and translocation of Pi and arsenate (Wu et al., 2011). Similarly, trafficking of arsenites from the soil to the xylem cells of roots occur by the concerted action of nodulin 26-like intrinsic membrane proteins (NIPs) type aquaporin channel family of transporters which

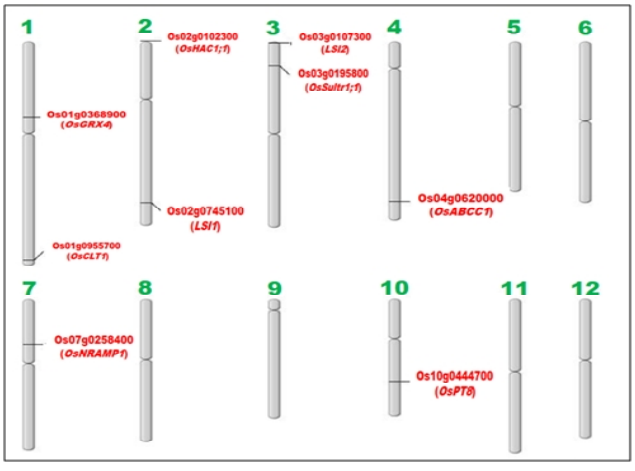

Fig. 1A. Genomic atlas of genes encoding for molecular determinants of arsenic metabolism in rice: The chromosomal location of each of the candidate genes involved in arsenic metabolism in rice are depicted in this figure.

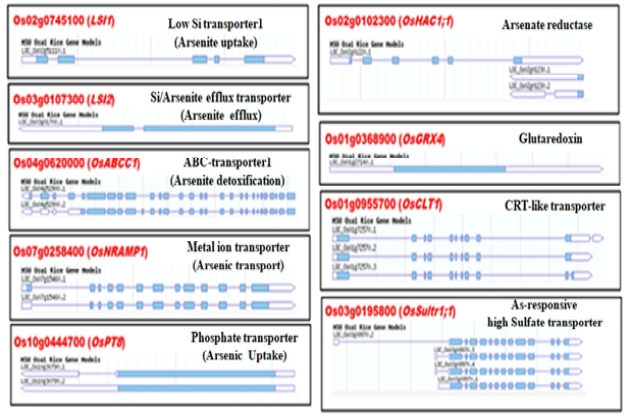

Fig. 1B. Genic architecture of candidate genes involved in arsenic tolerance in rice: Anatomy of candidate genes with their exons and introns are depicted in this figure.

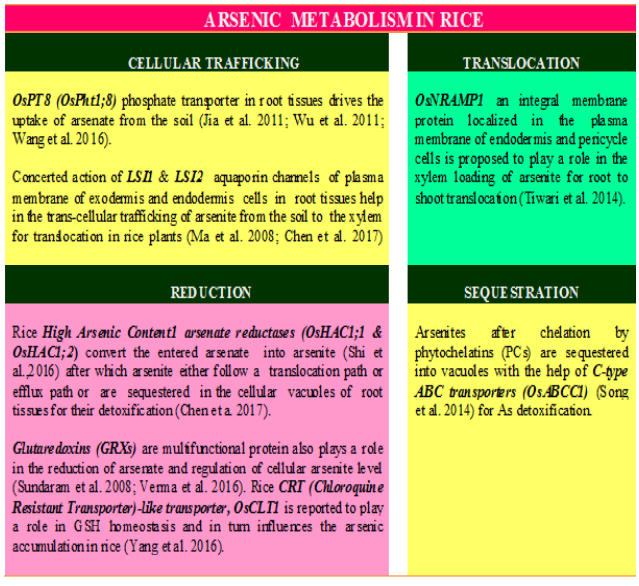

Fig. 1C. Major steps of arsenic metabolism in rice: Important steps of arsenic metabolism including uptake, celluar trafficking , reduction, translocation, sequestration and detoxification are presented in this figure.

are permeable to both silicon and arsenite molecules (Ma et al., 2008). Presence of an influx type silicon transporter LSI1 on distal side and an efflux type silicon transporter LSI2 on the proximal side of both exodermal and endodermal cells of roots permit the effective trans-cellular transport of both Si and arsenite from the paddy soil towards xylem for their translocation through their simultaneous pumping action (*LSI1* helps in influx of Si and arsenite into the cell while *LSI2* helps in their efflux from the cell) in rice plants (Ma et al., 2008; Chen et al., 2017). Arsenate (AsV) is reduced to arsenite (*AsIII*) in rice with the help of High Arsenic Content1 arsenate reductases (*OsHAC1;1* & *OsHAC1;2*) (Shi et al., 2016) in rice. Over-expression of *OsHAC1; 1* and *OsHAC1;2* has been reported to decrease the accumulation of As in rice by increasing the AsIII efflux into the external medium (Chain et al., 2017). In addition to arsenate reductase, glutaredoxin play an important role in reduction of arsenate (AsV) thereby regulating the cellular arsenite (AsIII) levels (Chen et al., 2017; Sundaram et al., 2008). Arsenite can be detoxified by chelating them with glutathione (GSH) or phytochelatins (PC) (Pal and Rai, 2010). Recently, the role of rice CRT (Chloroquine Resistant Transporter)-like transporter, *OsCLT1* has been demonstrated to play a role in GSH homeostasis and in turn influences the arsenic accumulation in rice (Yang et al., 2016). AsIII after chelation by PCs are sequestered into vacuoles with the help of C-type ABC transporters (ABCC) (Song et al., 2014) for As detoxification.

The knowledge of genes involved in the modulation of arsenic metabolism is increasing day by day based on studies in different genetic models including rice. However, transcription factor mediated regulation of genes involved in arsenic metabolism is not clearly known. Recently, the role of an R2R3 MYB transcription factor, OsARM1, has been reported in the regulation of arsenic uptake and root to shoot translocation in rice (Wang et al., 2017). Further, authors have demonstrated that OsARM1 binds to the conserved MYB-binding site in the genomic regions of promoters of key arsenic transporters encoding genes including *OsLSI1, OsLSI2* and *OsLSI6* (Wang et al. 2017). Similarly, WRKY6 transcription factor is reported to regulate the expression of arsenate/ phosphate transporters in *Arabidopsis thaliana* (Castrillo et al., 2013). Similarly, the role of tonoplast-localized ABC transporter in the arsenic detoxification by controlling the transport of arsenic accumulation in grain though the sequestration of As (III)-PC complex in to the vacuoles of the phloem companion cells at the nodes (Song et al., 2014) is reported. However, transcriptional regulation of *OsHAC1; 1* in rice is not clear.

Therefore, in this investigation, an attempt has been made to analyze the HyperTree of associated networks of coexpressed genes constructed for each of the genes involved in arsenic metabolism (*OsPT8, LsSi1, LSi2, OsHAC1;1, OsNRAMP1, OsABCC1* and *OsGrx4*) and in depth analysis has been made to understand the transcriptional regulation of *OsHAC1;1* gene which is involved in the reduction of arsenate into arsenite for detoxification in arsenic metabolism, which will provide insights for designing genetic interventions for engineering low-arsenic-content rice varieties.

MATERIALS & METHODS

Retrieval of genomic sequence and RAP-DB gene ID of candidate genes and their visualization in the genomic landscape of rice

Genomic DNA sequences of putative genes reported in literature encoding for proteins involved in arsenic metabolism (Chen et al., 2017) namely Os02g0745100 (*LSI1*), Os03g0107300 (*LSI2*), Os10g0444700 (*OsPT8*), Os04g0620000 (*OsABCC1*), Os07g0258400 (*OsNRAMP1*), Os02g0102300(*OsHAC1;1*), Os01g0368900 (*OsGRX4*), Os01g0955700 (*OsCLT1*) and Os03g0195800(*OsSultr1;1*) in rice were retrieved from the Rice Annotation Project Database (RAP-DB) (http://rapdb.dna.affrc.go.jp/) (Sakai et al., 2013). The details of each gene were obtained by seeding each gene name in the search box of RAP-DB and out of which genomic sequence, RAP-DB gene ID of each of the searched genes were fished out. Each of the RAP-DB gene IDs were utilized for visualization of each candidate genes in their respective chromosomes using chromosome map tools of Oryzabase (http://viewer.shigen.info/oryzavw/maptool/MapTool.do) in the genomic landscape of rice.

Depiction of the genic architecture of candidate genes

The RAP-DB ID (Os ID) of each candidate genes was transformed in to MSU ID (LOC_Os ID) using ID converter tool of Rice Annotation Project Database (RAP-DB) (http://rapdb.dna.affrc.go.jp/) (Sakai et al., 2013). Gene features including exons and introns, and splice variants of the selected genes were visualized by seeding MSU ID in the search box of rice locus identifier search option of MSU Rice Genome Annotation Project Database and Resource (http://rice.plantbiology.msu.edu/index.shtml ; Kawahara et al., 2013).

Construction of HyperTree of co-expressing genes for each candidate genes

A set of coexpressed genes for each of the candidate genes were identified by seeding their RAP-DB ID as guide gene in the coexpression search box for "Single Guide Gene" present in the "Rice Functionally Related gene Expression Network Database (RiceFREND) (http://ricefrend.dna.affrc.go.jp/)" (Sato et al., 2012). Further, HyperTree of coexpressed genes for each candidate guide genes with hierarchy of 3 and mutual rank of 7 were computed to visualize the interaction between them using HyperTree option available in the gene coexpression database RiceFREND.

Scanning of upstream promoter region of candidate genes for potential transcription factor binding sites (TBFS)

In order to identify TBFS in the promoter region, 1000bp upstream DNA sequence from the start codon of each the selected candidate genes involved in the arsenic metabolism in rice were retrieved from the nucleotide database of NCBI (https://www.ncbi.nlm.nih.gov/) (Coordinators, 2016). Cis-regulatory elements present in the upstream regions of the selected genes for binding of transcription factors were identified by seeding retrieved upstream DNA sequence of each gene in fasta format in the "promoter analysis search box" of PlantPAN2.0 (http://plantpan2.itps.ncku.edu.tw/index.html) which is an on line plant promoter analysis navigator (Chow et al., 2015).

Mining of transcriptomic data of candidate genes from the microarray database

Relative gene expression profile in different tissues at different developmental timings of rice for each of the selected candidate gene and identified regulatory MYB or MYB related TFs encoding genes in the HyperTree were retrieved by seeding their RAP-DB IDs in the "Global Gene Expression Profile" search box of "RiceXPro (http://ricexpro.dna.affrc.go.jp/)" microarray database (Sato et al., 2010).We retrieved the normalized global expression profile of cDNA accessions namely AK067143, AK069842, AK101092, AK101170, CI007435, AK103557, AK107858, AU075970 and AF493792, for *OsHAC1;1, LSI1, LSI2, OsPT8, OsABCC1, OsNRAMP1, OsGRX4, OSCLT1* and *OsSultr1;1*, respectively, candidate genes of the arsenic metabolism pathway in rice. Similarly, we retrieved the normalized expression profiles of four cDNA accessions (AY39858, AJ237661, AK066180 and AK061823) for OsMYB48 (Os01g0975300), two cDNA accessions namely AK099223 and AK111634 for Os05g0114700 and one cDNA accession each namely AK111960, AK111626, AK112054 and CI437851 for *OsMYB58/63*(Os04g0594100), Os11g0700500, Os01g0635200 and Os05g0579600, respectively, for MYB or MYB-related TF encoding genes found in the HyperTree of *OsHAC1; 1* gene.

Visualization of transcriptomic data in the heatmap for spatiotemporal analysis of gene expression

All the retrieved transcriptomic data from RiceXPro database were processed in to a .csv file format which was used for construction of a heatmap through a manually written R script using "agridat R package" (Wright, 2013) in an Ubuntu-Linux platform. Heatmap was used for spatiotemporal analysis of gene expression pattern in different tissues during the different developmental phases of rice plant.

RESULTS AND DISCUSSION

MYB or MYB-related TF encoding genes were enriched in the constructed gene-coexpression network of arsenate reductase encoding gene (*OsHAC1;1*)

HyperTree of gene coexpression network of each of the selected genes in arsenic metabolism were constructed with the help of RiceFREND (Sato et al., 2012) by using each of them as guide gene to gain knowledge of their transcriptional regulatory modules.

Table 1. List of transcription factor encoding genes co-expressed with guide genes.

S.N	Name of the guide gene	Co-expressed transcription factor encoding genes	Remarks
01	OsPT8	WRKY transcription factor 28-like (WRKY5) (Os01g0584900)	WRKY Transcription factor77
		DNA-binding protein WRKY2-like (WRKY16) (Os01g0626400)	WRKY Transcription factor11
		Conserved hypothetical protein (Os01g0289600)	WRKY Transcription factor
		Similar to Scarecrow-like 9 (Fragment) (Os11g0705200)	GRAS Transcription Factor
		Homeodomain-like containing protein (Os06g0670300)	G2-Like Transcription factor
		Homeodomain-like containing protein (Os09g0379600)	HB-Transcription factor)
		No apical meristem (NAM) protein domain containing protein (Os08g0535800)	NAM Transcription factor
		No apical meristem (NAM) protein domain containing protein (Os09g0509100)	NAM Transcription factor
02	OsHAC1;1	Similar to Typical P-type R2R3 MYB protein (Fragment) (Os01g0975300)	MYB Transcription Factor48(OsMYB48)
		Similar to Typical P-type R2R3 MYB protein (Fragment) (Os04g0594100)	MYB transcription factor 58/63, (OsMYB58/63)
		Similar to Snapdragon MYB protein 305 homolog (Os11g0700500)	MYB Transcription Factor
		MYB, DNA-binding domain containing protein (Os05g0114700)	MYB-related Transcription Factor
		Homeodomain-like containing protein (Os01g0635200)	Myb-related Transcription Factor
		Homeodomain-like containing protein (Os05g0579600)	Myb-related Transcription Factor
		Similar to Phosphate starvation regulator protein (Regulatory protein of P- starvation acclimation response Psr1) (Os05g0488600)	G2-like Transcription Factor
		Similar to Ethylene response factor 2 (Os07g0617000)	AP2-EREBP Transcription Factor
03	OsABCC1	Zinc finger, B-box domain containing protein (Os04g0540200)	Orphan Transcription factor
		Zinc finger, B-box domain containing protein (Os02g0646200)	Orphan Transcription factor
		Similar to Two-component response regulator ARR11 (Receiver-like protein 3) (Os01g0971800)	G2-like Transcription factor
		Similar to Gt-2(Os03g0113500)	Trihelix Transcription factor
04	OsNRAMP1	Basic helix-loop-helix dimerisation region bHLH domain containing protein (Os01g0952800)	Orphan Transcription factor
05	OsGRX4	Zinc finger, CCCH-type domain containing protein (Os01g0917400)	C3H Transcription factor
		Similar to Transfactor-like protein (Os12g0586300)	G2-like Transcription factor
		Zinc finger, B-box domain containing protein(Os04g0540200)	Orphan Transcription factor

Analysis of HyperTree of each of the guide genes revealed a number of transcription factor encoding genes in the network of some of the guide genes which are summarized in Table 1. In the HyperTree of *OsPT8, OsHAC1;1, OsABCC1, OsGRX* and *OsNRMPA1* guide genes, eight (3 WRKY TFs, 1 GRAS TF, 1 G2-like TF, 1 HB TF and 2 NAM TFs), eight (6 MYB or MYB-related TFs, 1 G2-like TF, 1 AP2-EREBP TF), four (2 Orphan TF, 1 G2-like TF, 1 Trihelix TF), three (1 C3H TF, 1 G2-like TF and 1 Orphan TF) transcription factor encoding genes, respectively, were found. Analysis of different TFs found in the HyperTree of each guide genes revealed the preponderance of a number of MYB or MYB-related transcription factor encoding genes in the gene coexpression network of *OsHAC1;1* (Fig. 2B and Fig. 2C).

There were three MYB [2 Typical P-type R2R3 MYB protein (Fragment) representing OsMYB48 (Os01g0975300) and OsMYB58/63 (Os04g0594100)] and 1 Snapdragon MYB protein 305 homolog (Os11g0700500)] and three MYB-related [2 Homeodomain-like containing protein (Os01g0635200, Os05g0579600) and 1 MYB, DNA-binding domain containing protein (Os05g0114700)] genes in the HyperTree of *OsHAC1;1* (Table 1, Fig. 2B and Fig. 2C).

Cis-elements for binding of MYB or MYB-related TFs were found in the 1000bp upstream genomic region from the translational start site

of *OsHAC1; 1* gene

As many MYB or MYB-related genes including Typical P-type R2R3 MYB protein encoding genes were found in the HyperTree of *OsHAC1; 1* compared to other guide genes, we investigated whether there is any connection between these transcription factors and *OsHAC1;1*. To know the relationship between MYB TFs and *OsHAC1; 1*, we scanned 1000bp the upstream genomic promoter region of *OsHAC1; 1* gene. Scanning revealed the presence of cis-elements for binding of MYB or MYB-related TFs in the upstream promoter region of *OsHAC1; 1* (Fig. 2A).

Spatiotemporal analysis of gene expression profiles revealed the role of P-type R2R3 *OsMYB48* & *OsMYB58/63* and other MYB-related genes in the modulation of gene expression of *OsHAC1;1* and other genes (*OsABCC1*, *OsPT8*, *LSI1*, *OsCLT1* and *OsSultra1;1*) involved in arsenic metabolism in root tissues of rice

As we found a preponderance of MYB or MYB-related genes in the gene coexpression module of *OsHAC1;1* and presence of cis-elements for binding of MYB or MYB-related TFs in the upstream promoter region of *OsHAC1;1*, we asked whether MYB TFs are modulating the gene expression pattern of *OsHAC1;1* at transcriptional level. To address this question, we mined gene expression data for MYB or MYB-related genes found in the HyperTree of *OsHAC1; 1* and other selected genes linked to arsenic metabolism in different tissues of rice during its different developmental stages from the "Rice Expression Profile Database (RiceXPro)" (Sato et al., 2010) and did a spatiotemporal analysis of gene expression pattern by constructing a heatmap using "agridat package of R" (Wright, 2013) in a Ubuntu platform (Fig. 3A).

Expression profiling analysis revealed that all the *MYB or MYB-related genes except OsMYB58/63*, and *OsHAC1;1* were recorded a contrasting degree of expression pattern in vegetative organs (Leaf blades, leaf sheaths, roots and stems) and reproductive organs (inflorescences, anthers, pistils, lemmas, paleas, ovaries, embryos and endosperms) of rice. They were up-regulated in vegetative parts while down-regulated in reproductive structures during different developmental stages of rice. At the same time, other genes involved in arsenic metabolism including *OsABCC1, OsNRAMP1, LSI1, LSI2, OsPT8, OsGRX4, OsCLT1* and *OsSultr1;1* recorded mixed expression patterns in vegetative and reproductive

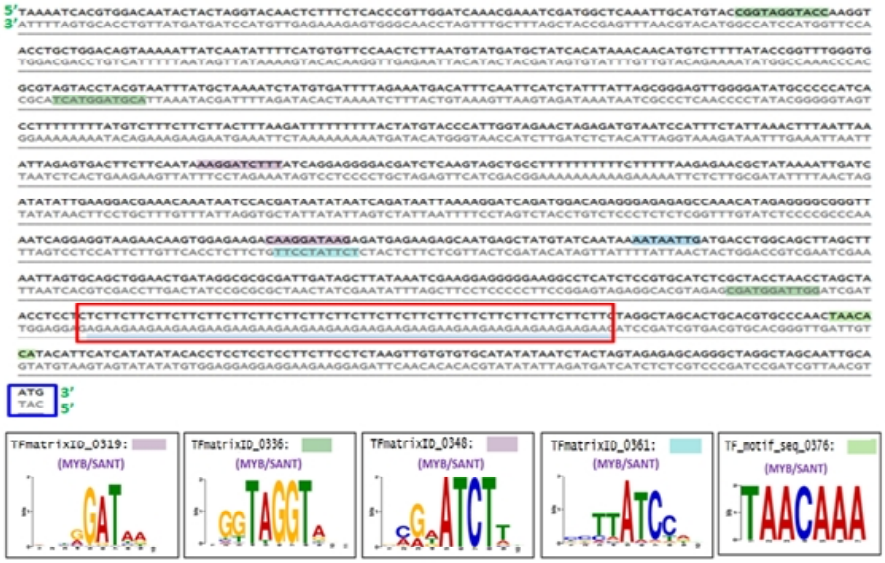

Fig. 2A. Visualization of MYB/SANT transcription factor binding sites (TFBS) in the upstream region of the OsHAC1;1 gene: In this figure, translational start site (ATG codon) is highlighted with a blue box while tadem repeats are marked with a red box. Four different consensus DNA motifs for binding of MYB/SANT are highlighted in different colors in the 1000bp upstream region of OsHAC1;1 gene .

organs. Among all these genes, *OsHAC1;1, LSI1, OsGRX4*, and *OsPT8* recorded an elevated expression pattern in leaf blades during vegetative, reproductive & ripening stage and in leaf sheaths during vegetative and reproductive stages, while LSI2 and *OsSultr1;1* were down-regulated at all the stages (Fig. 3A). OsABCC1 witnessed an increased expression in leaf blade and leaf sheaths at all stages except for the vegetative stage of leaf sheath. Further, *OsNRAMP1* gene recorded mixed gene expression pattern throughout the life cycle of rice. Interestingly, in root tissues, *LSI1, LSI2, OsPT8, OsABCC1, OsCLT1, OsSUltra1;1, OsHAC1;1* and all the MYB/MYB-related genes except M6 (Os05g0579600) exhibited an increased expression during vegetative and reproductive stages. From this evidence, we reasoned whether MYB/MYB-related TFs modulate the expression pattern of these genes. Scanning of the upstream 1000bp genomic region from the translational start site of these genes revealed a number of MYB/SANT binding sites in *LSI1, OsPT8, OsABCC1, OsCLT1, OsSultra1;1* and *OsHAC1;1* genes implicated in the arsenic metabolism. All these evidences supported that MYB/MYB-related TFs may be involved in the modulation of the gene expression pattern of all these genes in root tissues. Then, we asked which MYB/MYB-related genes are involved in the modulation of gene expression of aforesaid genes in root tissues. Initially, we identified 6 MYB/MYB-related genes in the HyperTree of the *OsHAC1;1* gene. Out of all these candidate MYB genes, OsMYB48, OsMYB58/63 has similarity to the typical P-Type R2R3 MYB proteins.

Therefore, we analyzed the expression pattern of all the splice variants (M1a, M1b, M1c & M1d) of M1 (OsMYB48) gene and M2 (*OsMYB58/63*) gene (Fig. 3A and Fig. 3B). We found an elevated expression for all the splice variants of M1 (OsMYB48) gene and M2 (OsMYB58/63) gene in root tissues at vegetative and reproductive stages of rice plant. Apart from these typical MYB protein-encoding genes, splice variants of M3 (MYB gene, Os05g0114700) gene, M4 (Snapdragon MYB-protein 305 homolog) and M5 (Os01g0635200) were recorded to have increased expression in root tissues at vegetative and reproductive stages (Fig. 3A and Fig. 3B).

Arsenic toxicity is a serious global issue as its exposure is associated with serious health problems including cancer. As paddy is highly vulnerable to hyper-accumulate toxic arsenic in its grains, it is imperative to design genetic interventions to minimize the arsenic levels in paddy grains to combat human health issues as rice is a lifeline for fulfilling the calorie-needs of billions of people in the globe. In this investigation, an attempt has been made to understand the molecular basis of arsenic tolerance at transcriptomic level to identify transcriptional regulatory player in paddy using computational biology tools and analyzing publicly available genomic/transcriptomic resources and microarray data.

Initially, we constructed HyperTree of co-expressing genes for all the candidate guide genes involved in arsenic metabolism (*OsLSI1, OsLSI2, OsCLT1, OsSultr1;1, OsPT8, OsHAC1;1, OsABCC1, OsNRAMP1* and *OsGRX4*) and identified a number of TF-encoding genes in the HyperTree of few guide genes (*OsPT8, OsHAC1;1, OsABCC1, OsNRAMP1* and *OsGRX4*). Interestingly, in the HyperTree of guide gene *OsHAC1;1*, a number of MYB or MYB-related TF-encoding genes were enriched (Fig. 2B and Fig. 2C). Further, analysis of the

Fig. 2B. HyperTree of OsHAC1;1 gene : In this figure, different associated coexpressed genes of OsHAC1;1 are depicted in the form of a HyperTree using RiceFREND. Transcription factor encoding genes are shown as orange square boxes while guide gene OsHAC1;1 is highlighted in the centre of the circular HyperTree with a small yellow circle.

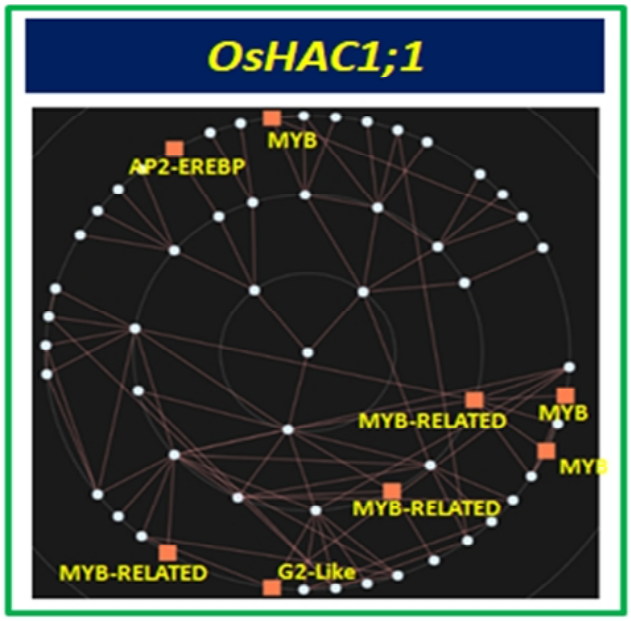

Fig. 2C. Visualization of transcription factor encoding genes of OsHAC1;1 HyperTree: MYB or MYB-related enriched transcription factor encoding genes and other TF encoding genes (G2-like and AP2EREBP) are depicted in this figure.

upstream 1000bp genomic region from the translational start site of *OsHAC1;1* revealed the presence of binding sites for MYB transcription factors (Fig. 2A). Therefore, coexpression of MYB or MYB-related TFs and *OsHAC1;1*, and presence of MYB binding sites in the promoter region of *OsHAC1;1* gene indicated the hint for transcriptional modulation of *OsHAC1;1* gene by MYB transcription factors.

Further, we did a spatiotemporal analysis of gene expression patterns of *OsHAC1;1* and *MYB or MYB-related* genes along with all other genes (*OsABCC1, OsNRAMP1, LSI1, LSI2, OsPT8, OsGrx4, OsCLT1 OsSultra1*) involved in arsenic metabolism in different tissues of rice at different developmental phases to understand the basis of transcriptional regulation of *OsHAC1;1* gene by MYB transcription factors using the microarray data obtained from the "Rice Expression Profile Database (RiceXPro)" (Sato et al., 2010) by constructing a heatmap. In the heatmap, OsHAC1;1 gene along with all the MYB or MYB-related genes except OsMYB58/63 gene, recorded a contrasting degree of expression patterns in vegetative organs (up-regulation in leaf blades, leaf sheaths, roots and stems) and reproductive organs (down-regulation in inflorescences, anthers, pistils, lemmas, paleas, ovaries, embryos and endosperms) at different developmental timings in paddy while *OsABCC1, OsNRAMP1, LSI1, LSI2, OsPT8, OsGrx4, OsCLT1, OsSultra1* genes recorded mixed expression patterns in vegetative and reproductive organs. Further, an elevated gene expression pattern of *OsHAC1;1* along with typical R2R3 P-type MYB (OsMYB48 and OsMYB58/63) and MYB-related (Os05g0114700, Os01g0635200 and Snapdragon MYB-protein 305 homolog) genes and other genes (*LSI1, LSI2, OsPT8, OsABCC1, OsCLT1, OsSUltra1;1*) of arsenic metabolism in root tissues during both vegetative and reproductive stages were noticed in the heatmap. Further, MYB/SANT binding sites were also found in the promoter region of *LSI1, LSI2, OsPT8, OsABCC1, OsCLT1, OsSUltra1;1*) genes. Recently, the role of an R2R3 MYB transcription factor, OsARM1, has been reported in the regulation of arsenic uptake and root to shoot translocation in rice, where authors have demonstrated that OsARM1 binds to the conserved MYB-binding site in the genomic regions of promoters of key arsenic transporters encoding genes including *OsLSI1, OsLSI2 and OsLSI6* (Wang et al., 2017). All these findings supports and points towards the vital role of MYB transcription factors in the transcriptional modulation of arsenic metabolism especially in the modulation of *OsHAC1;1*.

OsHAC1;1 encodes for arsenate reductase enzyme which play a vital role in reduction of arsenates into arsenites in rice. All these evidences, suggest that MYB transcription factors (OsMYB48 and OsMYB58/63) may be playing a role in the transcriptional regulation of *OsHAC1;1* gene in the root tissues of rice by binding to its promoter region. Therefore, significantly, higher degree of expression of *OsHAC1;1* in the root tissue compared to other tissues might be helping in the reduction of accumulated arsenates via OsPT8 transporter from the soil in to arsenites which in turn are sequestrated in the vacuoles of root tissue with the help of ATP-binding cassette (ABC) transporters (OsABCC1)-phytochelatin complex.

This may be feasible as we noticed an elevated expression of OsPT8 in root tissues during vegetative stage which may help in pumping arsenates from the root zone of the soil followed by their reduction into arsenite by arsenate reductase encoded by *OsHAC1;1* and sequestration and detoxification by tonoplast-bound

OsABCC1 transporter which were found to be highly upregulated in the root tissues during vegetative and reproductive stages. Reduction of arsenates to arsenites in the root tissue by arsenate reductase help either by transporting arsenites to the external environment or by the sequestration of arsenites in the cellular vacuoles of root tissues which help in decreasing the accumulation of toxic arsenic in the grains. Earlier, the role of OsABCC1 transporters in arsenic detoxification and reduction in grain has been demonstrated by Song et al. (2014). Further, OsABCC1 is shown to be highly expressed in the flag leaf blade tissue and root tissue during flowering and grain filling stage, respectively (Song et al., 2014). Recently, role of *OsMYB48* gene in drought and salinity tolerance is reported and this gene is differentially expressed in different tissue (Xiong et al. 2014). One splice variant of this gene MYB TF48-1 has been reported to exhibit higher degree of expression in root tissues during seedling stage and in both root and leaf tissues during reproductive stage, compared to reproductive organs such as panicle (Xiong et al., 2014). Similarly, elevated expression of both *LSI1* and *LSI2* genes in the root tissues during vegetative stage as witnessed in the heatmap may play a role in

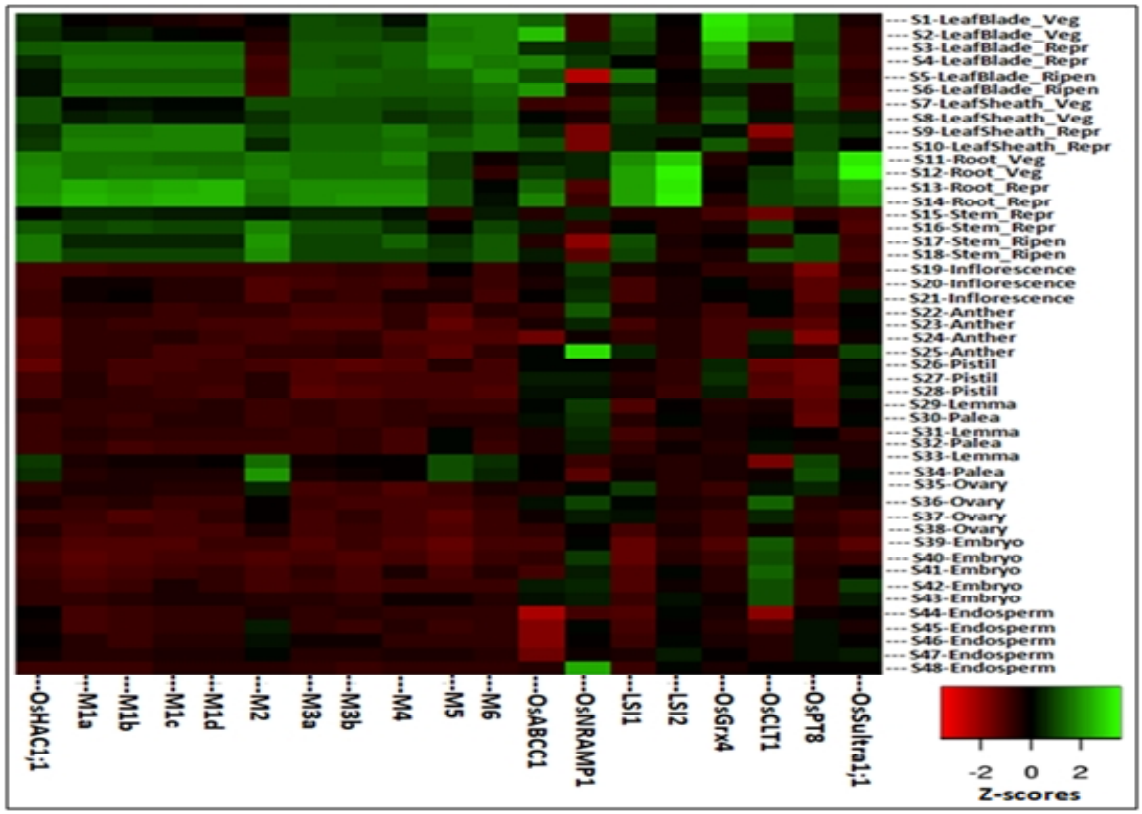

Fig. 3A. Heatmap visualization spatiotemporal pattern of gene expressions of MYB or MYB-related TFs and genes involved in arsenic metabolism at different developmental stages of rice plant: Expression patterns of different MYB or MYB-related TFs found in the hyper tree of OsHAC1;1 and different genes of arsenic metabolism have been depicted in the heatmap in Figure-3A. S1 to S48 refers different developmental stages (example: S1:Leafblade_Veg……S48: Endosperm mentioned in figure). M1a (Acc. AY398581), M1b (Acc. AJ237661) , M1c (Acc. AK066180) & M1d (Acc. AK061823) are different splice variants of a typical P-type R2R3 MYB gene (Os01g0975300, OsMYB48). M2 (Acc. AK111960): A typical P-type R2R3 MYB gene (Os04g0594100, OsMYB58/63). M3a (Acc. AK099223) & M3b (Acc. AK111634) are the splice variants of a MYB gene (Os05g0114700). M4 (Acc. AK111626): A MYB-related gene (Os11g0700500) similar to Snapdragon MYB protein 305 homolog. M5 (Acc. AK112054) (Os01g0635200) & M6 (Acc. CI437851) (Os05g0579600) are MYB-related TFs similar to homeodomain-like containing proteins. In the heatmap, green color indicates upregulation while red color indicates downregulation of gene expression based on their z-scores. Acc. : refers to the accession numbers for cDNAs of genes.

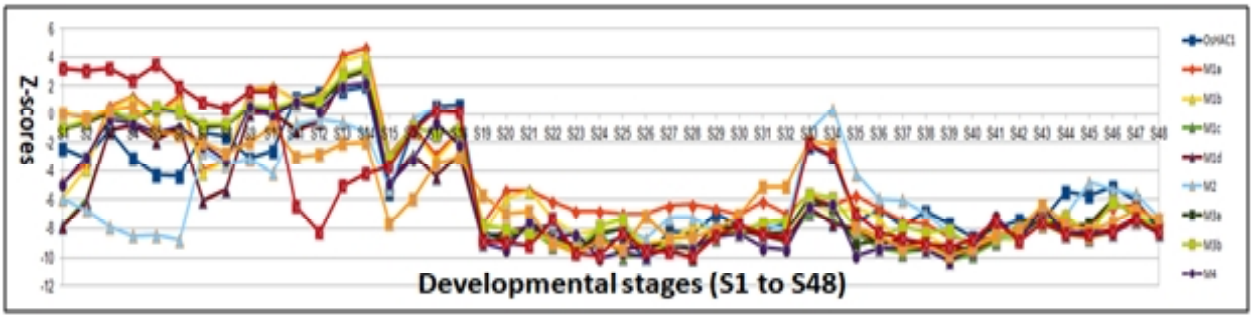

Fig. 3B. Comparison of expression profiles of MYB or MYB-related genes and OsHAC1;1 : In Figure-3B, expression pattern of OsHAC1;1 and MYB or MYB-related genes are depicted in the form of a line diagram at different developmental stages (S1 to S48 as depicted in Figure-3A, example: S1:Leafblade_Veg……S48: Endosperm).The contrasting gene expression pattern of all the MYB or MYB-related genes found in the HyperTree of OsHAC1;1 in vegetative and reproductive organs at different developmental stages have been shown in the heatmap and is highlighted in Figure-3B. All the MYB or MYB related genes except M2(OsMYB58/63) along with OsHAC1;1 were up-regulated leaf blades, leaf sheaths, roots and stems at different developmental stages, while they were drastically down-regulated in most of the reproductive organs including inflorescences, anthers, pistils, lemmas, paleas, ovaries, embryos and endosperms. M1a, M1b, M1c & M1d (OsMYB 48 cDNA Acc), M2 (OsMYB58/63), M3a & M3b (cDNA acc. Of a MYB gene), M4, cDNA acc. of a MYB-related gene (Refer Figure-3A, for details). Acc. : refers to the accession numbers for cDNAs of genes

pumping of toxic arsenites from root zone of soil to the vacuolar membrane transporter OsABCC1 for their sequestration and detoxification (Fig. 3A).

All these findings are consistent with our hypothesis that differential spatiotemporal expression of MYB (OsMYB48 & OsMYB58/63) transcription factors and genes involved in arsenic metabolism are contributing to the sequestration of toxic arsenites in the cellular vacuoles of root tissues by the canonical molecular machineries including OsABCC1 transporters for detoxification and thereby minimizing toxic arsenic levels in grains.

CONCLUSIONS

In this investigation, we have discovered the role of a P-type R2R3 MYB transcription factor OsMYB48 in the contrasting differential modulation of OsHAC1; 1 gene in root tissues and most of the reproductive tissues which is supported by the upregulated expression of *OsMYB48* and *OsHAC1; 1* during the vegetative phase & reproductive phase of rice in root tissues and a drastic downregulation of their expression pattern in inflorescences, anthers, pistils, ovaries, embryos and endosperms. In addition, many MYB or MYB related genes including OsMYB48 were enriched in the HyperTree of *OsHAC1;1* guide gene and MYB transcription factor binding sites were found in the genomic promoter region of *OsHAC1;1*. Therefore, OsMYB48 might be playing a role in the contrasting differential modulation of expression pattern of *OsHAC1; 1* gene in the root and reproductive tissues (inflorescences, anthers, pistils, ovaries, embryos and endosperms) during different growth stages of rice. This finding has an potential implication in the reduction of pumped arsenate by OsPT8 into arsenite by the *OsHAC1;1* and sequestration of toxic arsenites by the OsABCC1 transporters for detoxification in the root tissues and thereby minimizing toxic arsenic levels in grains .

Many transcription factors play a vital role in stress response processes (Das et al., 2017; Das, 2013). MYB transcription factors are reported to play a vital role in abiotic stress response processes. Recently, role of OsMYB48 gene in drought & salinity tolerance is reported and this gene is differentially expresses in different tissue (Xiong et al., 2014). In this investigation we have discovered the role of OsMYB48 in the transcriptional modulation of *OsHAC1;1* gene which encodes for arsenate reductase involved in the reduction of arsenates to arsenites using computational biology tools and transcriptomic analysis. Therefore, functional validation of OsMYB48 mediated transcriptional modulation of *OsHAC1;1* gene in the root and reproductive tissues would help in gaining insights into the basis of arsenic tolerance in rice and designing genetic strategies for evolving low grain-

arsenic content varieties of paddy using genetic engineering and plant breeding tools.

REFERENCES

Argos M, Kalra T, Rathouz PJ, Chen Y, Pierce B, Parvez F, Islam T, Ahmed A, Rakibuz-Zaman M, Hasan R and Sarwar G (2010). Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): a prospective cohort study. The Lancet. 376(9737): 252-8

Castrillo G, Sánchez-Bermejo E, de Lorenzo L, Crevillén P, Fraile-Escanciano A, Mohan TC,Mouriz A, Catarecha P, Sobrino-Plata J, Olsson S and del Puerto YL (2013). WRKY6 transcription factor restricts arsenate uptake and transposon activation in Arabidopsis. The Plant Cell 25(8): 2944-57

Chen Y, Han YH, Cao Y, Zhu YG, Rathinasabapathi B and Ma LQ (2017). Arsenic transport in rice and biological solutions to reduce arsenic risk from rice. Frontiers in Plant Science 8: 268

Chow CN, Zheng HQ, Wu NY, Chien CH, Huang HD, Lee TY, Chiang-Hsieh YF, Hou PF, Yang TY and Chang WC (2015). PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. Nucleic acids research 44(D1): D1154-60

Coordinators NR (2016). Database resources of the national center for biotechnology information. Nucleic acids research. 44 (Database issue): D7

Das CK (2013). Genetic and biochemical approaches to explore the basis of lifespan regulation and stress response using small molecule based tools. PhD Thesis, IIT Kanpur

Das CK, Mohapatra SS and Mohanty MR (2017). Computational analysis of the molecular architecture and cis-elements of Heat-shock-protein Hsp70 encoding gene in different ecotypes of rice (*Oryza sativa* L.), Proceedings of International Conference On Organic Farming for Sustainable Agriculture , 2nd-3rd June, 2017, OUAT, Bhubaneswar, Abstract No. OFSA-2017/CD-39, pp.77

Jia H, Ren H, Gu M, Zhao J, Sun S, Zhang X, Chen J, Wu P and Xu G (2011). The phosphate transporter gene OsPht1; 8 is involved in phosphate homeostasis in rice. Plant Physiology.156(3): 1164-75

Kawahara Y, De la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu J, Zhou S and Childs KL (2013). Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. Rice 6(1):4

Ma JF, Yamaji N, Mitani N, Xu XY, Su YH, McGrath SP and Zhao FJ (2008). Transporters of arsenite in rice and their role in arsenic accumulation in rice grain. Proceedings of the National Academy of Sciences.105(29): 9931-5

Pal R and Rai JP (2010). Phytochelatins: peptides involved in heavy metal detoxification. Applied biochemistry and biotechnology.160(3): 945-63

Paszkowski U, Kroken S, Roux C and Briggs SP (2002). Rice phosphate transporters include an evolutionarily divergent gene specifically activated in arbuscular mycorrhizal symbiosis. Proceedings of the National Academy of Sciences. 99(20):13324-9

Sakai H, Lee SS, Tanaka T, Numa H, Kim J, Kawahara Y, Wakimoto H, Yang CC, Iwamoto M, Abe T and Yamada Y (2013). Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. Plant and Cell Physiology. 54(2):e6

Sato Y, Antonio BA, Namiki N, Takehisa H, Minami H, Kamatsuki K, Sugimoto K, Shimizu Y, Hirochika H and Nagamura Y (2010). RiceXPro: a platform for monitoring gene expression in japonica rice grown under natural field conditions. Nucleic acids research. 39(suppl_1):D1141-8

Sato Y, Namiki N, Takehisa H, Kamatsuki K, Minami H, Ikawa H, Ohyanagi H, Sugimoto K, Itoh JI, Antonio BA, Nagamura Y (2012). RiceFREND: a platform for retrieving coexpressed gene networks in rice. Nucleic acids research. 41(D1): D1214-1221

Shi S, Wang T, Chen Z, Tang Z, Wu Z, Salt DE, Chao DY, Zhao F (2016). *OsHAC1; 1* and *OsHAC1; 2* function as arsenate reductases and regulate arsenic accumulation. Plant physiology.pp-01332. doi: 10.1104/pp.16.01332

Sohn E (2014). The toxic side of rice. Nature. 514(7524): S62-3

Song WY, Yamaki T, Yamaji N, Ko D, Jung KH, Fujii-Kashino M, An G, Martinoia E, Lee Y and Ma JF (2014). A rice ABC transporter, OsABCC1, reduces arsenic accumulation in the grain. Proceedings of the National Academy of Sciences.111(44): 15699-15704. doi: 10.1073/pnas. 1414968111

Sundaram S, Rathinasabapathi B, Ma LQ and Rosen BP

(2008). An arsenate-activated glutaredoxin from the arsenic hyperaccumulator fern Pteris vittata L. regulates intracellular arsenite. Journal of Biological Chemistry. 283(10): 6095-101. doi: 10.1074/jbc.M704149200

Tiwari M, Sharma D, Dwivedi S, Singh M, Tripathi RD, Trivedi PK (2014). Expression in Arabidopsis and cellular localization reveal involvement of rice NRAMP, OsNRAMP1, in arsenic transport and tolerance. Plant, Cell & Environment 37(1): 140-52

Verma PK, Verma S, Pande V, Mallick S, Deo Tripathi R, Dhankher OP and Chakrabarty D (2016). Overexpression of rice glutaredoxin OsGrx_C7 and OsGrx_C2. 1 reduces intracellular arsenic accumulation and increases tolerance in Arabidopsis thaliana. Frontiers in plant science. 1(7): 740

Wang FZ, Chen MX, Yu LJ, Xie L, Yuan LB, Qi H, Xiao M, Guo W, Chen Z, Yi K and Zhang J (2017). OsARM1, an R2R3 MYB Transcription Factor, is involved in Regulation of the Response to Arsenic Stress in Rice. Frontiers in plant science.8:1868

Wang P, Zhang W, Mao C, Xu G and Zhao FJ (2016). The role of OsPT8 in arsenate uptake and varietal difference in arsenate tolerance in rice. Journal of experimental botany. 67(21):6051-9. doi:10.1093/jxb/erw362

Williams PN, Price AH, Raab A, Hossain SA, Feldmann J and Meharg AA (2005). Variation in arsenic speciation and concentration in paddy rice related to dietary exposure. Environmental Science & Technology. 1; 39(15): 5531-40

Wright K (2013). agridat: Agricultural datasets. R package version 2013;1

Wu Z, Ren H, McGrath SP, Wu P and Zhao FJ (2011). Investigating the contribution of the phosphate transport pathway to arsenic accumulation in rice. Plant Physiology. 157(1): 498-508

Xiong H, Li J, Liu P, Duan J, Zhao Y, Guo X, Li Y, Zhang H, Ali J and Li Z (2014). Overexpression of OsMYB48-1, a novel MYB-related transcription factor, enhances drought and salinity tolerance in rice. PLoS One. 9(3):e92913

Yang J, Gao MX, Hu H, Ding XM, Lin HW, Wang L, Xu JM, Mao CZ, Zhao FJ and Wu ZC (2016). OsCLT1, a CRTlike transporter 1, is required for glutathione homeostasis and arsenic tolerance in rice. New Phytologist.211 (2):658-70. doi: 10.1111/nph.13908