

Characterizing genetic factors that determine arsenic tolerance in plants

A Thesis
Submitted to the Faculty
in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

in

Biological Sciences

by Todd Warczak

Guarini School of Graduate and Advanced Studies
Dartmouth College
Hanover, New Hampshire

September 2020

Examining Committee:

Mary Lou Guerinot
(chair) Mary Lou Guerinot, Ph.D.

Thomas Jack
Thomas Jack, Ph.D.

Erik Griffin

Erik Griffin, Ph.D.

Om Parkash Dhankher
Om Parkash Dhankher, Ph.D.

F. Jon Kull, Ph.D.

Dean of the Guarini School of Graduate and Advanced Studies

Abstract

Arsenic is a non-essential, toxic metalloid that poses serious risks towards crop yields and human health. Among the crops highly consumed by humans, rice grain contains the highest levels of the metalloid, potentially exposing billions of people to the negative effects of arsenic toxicity. Although much work has been done to understand arsenic detoxification in plants, many questions remain regarding how plants tolerate the various forms of arsenic and the genetic factors regulating the transport of arsenic from the soil into different plant tissues.

In this thesis, we combined fluorescently activated cell sorting (FACS) with RNA-sequencing to generate high resolution, genome-wide expression datasets for three plant root cell-types when plants are exposed to As^{III} stress. This work provides the first cell-type specific expression map for every gene thought to be involved in As^{III} acquisition, efflux, sequestration, etc. for root epidermal, cortical, and endodermal cells.

Additionally, we identify how natural variation in the expression of a nodulin 26-like intrinsic protein primarily determines arsenic tolerance in *Arabidopsis*. We utilize a population of 527 *Arabidopsis thaliana* Multiparent Advanced Generation Inter-Cross (MAGIC) recombinant inbred lines (RILs) to identify *NIP1;1* expression as the major quantitative trait locus (QTL) which influences As^{III} tolerance.

This thesis improves the accuracy of current arsenic detoxification models by showing many genes within these models have cell-type specific roles in the roots. This work also elevates the importance of *NIP1;1* in determining As^{III} tolerance and offers insights into how strategically controlling *NIP1;1* expression could help design crops that

do not accumulate arsenic or conversely, crops with enhanced plant arsenic accumulation for phytoremediation.

Lastly, this thesis provides opportunities for further discovery of novel genes involved in arsenic tolerance, either through the cell-type specific expression or by identifying the genes responsible for the minor QTLs found in our GWAS.

Acknowledgements

Foremost, I would like to express my gratitude to my advisor Mary Lou Guerinot for supporting my thesis work, for the rare opportunity you gave me to study arsenic tolerance in plants as a Superfund Trainee, and for your patience with me during my years at Dartmouth. I am deeply indebted to the other Guerinot lab scientists. Ali Sivitz and Heng-Hsuan Chu were particularly helpful in my early years, while I heavily relied on Suna Kim, Garo Akmakjian, Zhong Tang, Karina Lopes, Anne Lichtner, and Nabila Riaz in the later years. Thank you all so much.

I would also like to thank the brilliant scientists, administrators, and organizers in the Dartmouth SRP for their support, friendship, and expertise. It was an honor to work alongside such a diverse group of talented people.

A special thank you to Heng-Hsuan Chu and Sarah Jennewein for the hundreds of hours spent under the hood with me, plating and screening MAGIC lines and mutants, and to Gary Ward at DHMC for being my cell-sorting guru. To my thesis committee Tom Jack, Erik Griffin, and Om Parkash Dhankher, thank you for your necessary critique and encouragement. And thank you to Granola softball and Dartmoose hockey for the bruises, dehydration, and trash talk. I need that stuff.

Lastly, my deepest appreciation to the small group of supporters firmly in my corner. To Melisa Fuentes for teaching me to walk again while surviving the Trump years. And to my babies, JM and Auden. I'll love you, dears, I'll love you till China and Africa meet, and the river jumps over the mountain and the salmon sing in the street. This is for you.

Table of Contents

Abstract	ii
Acknowledgements	iv
List of Illustrations	vii
List of Tables	viii
Chapter 1: Introduction	1
Arsenic Toxicity in Humans	2
Natural Sources	3
Anthropogenic Sources	4
Arsenic in the Food Chain	7
Arsenic Toxicity in Plants	8
Arsenic Uptake from Soil	12
Phytochelatin Mediated Sequestration	16
Root to Shoot Arsenic Transport	19
Shoot to Seed Arsenic Transport	21
Plant Genetic Engineering Strategies to Decrease the Arsenic Content of Plants	21
Conclusion	31
Chapter 2: Cell-type specific expression profiles in <i>Arabidopsis thaliana</i> roots in response to As ^{III}	34
Abstract	35
Introduction	37
Results	41
Discussion	59
Materials and Methods	67

Supplemental Data	70
Chapter 3: Natural variation in <i>NIP1;1</i> expression determines As ^{III} tolerance in <i>A. thaliana</i> accessions	87
Abstract	88
Introduction	89
Discussion.....	102
Materials and methods.....	108
Supplemental Figures	112
Chapter 4: Conclusions	119
Cell-Type Specific Expression in Response to As^{III}	121
Gene Ontology Analysis.....	123
Arsenic Responsive Genes of Interest.....	124
MAGIC Population Identifies Multiple QTL for As^{III} Tolerance	125
<i>NIP1;1</i> Expression Responsible for Major QTL	126
Appendix I: MATE family of proteins in plants.....	128
Appendix II: <i>A. thaliana</i> mutants screened for As ^{III} sensitivity	137
Appendix III: Understanding arsenic dynamics in agronomic systems to predict and prevent uptake by crop plants	145
Appendix IV: Opportunities and Challenges for Dietary Arsenic Intervention	158
References.....	165

List of Illustrations

Figure 1.1: Generalized Diagram of Arsenic Transport and Metabolism in Plants	11
Figure 2.1: Unsupervised Clustering & Exploratory Data Analysis	45
Figure 2.2: Differentially Expressed Genes by Cell-Type	48
Figure 2.3: GO-Term Enrichment	52
Figure 2.4: As ^{III} Induced Expression by Cell-Type	55
Supplemental Figure 2.1: MA Plots for Differentially Expressed Genes	71
Supplemental Figure 2.2: GFP-Reporter Lines	72
Supplemental Figure 2.3: GFP-marker lines used for FACS	74
Supplemental Figure 2.4: Complete GO-Term Enrichment (with redundant terms)	78
Supplemental Figure 2.5: ICP-MS by Cell-Type	79
Supplemental Figure 2.6: As ^{III} Induced Expression by Cell-Type	81
Supplemental Figure 2.7: <i>A. thaliana</i> GO Term Limitations	82
Figure 3.1: MAGIC Line Root Scoring for Arsenic Tolerance	93
Figure 3.2: GWAS for Arsenic Tolerance Found in MAGIC Population	95
Figure 3.3: MAGIC Line Haplotypes at Major and Minor QTL	99
Figure 3.4: MAGIC founder <i>NIP1;1</i> expression	106
Supplemental Figure 3.1: MAGIC Founder Lines; ordered by As ^{III} sensitivity rank	113
Supplemental Figure 3.2: Variants of Interest	115
Supplemental Figure 3.3: GWAS with Condensed SNP Library Showing Association with As ^{III} Tolerance	116
Supplemental Figure 3.4: Root Scores of Minor & Major QTL Haplotype Combinations	117
Supplemental Figure 3.5: Counts of MAGIC line Founders Per Genome	118
Figure A1.1: MATE protein phylogeny in selected plants	131
Figure A1.2: As ^{III} Induced Expression by Cell-Type (MATEs)	133
Figure A1.3: MATE yeast spot assay on As ^{III}	135
Figure A2.1: Single <i>abc</i> mutant lines show elevated As ^{III} tolerance	142
Figure A2.2: Quadruple <i>pip</i> mutant line shows elevated As ^{III} tolerance	144

List of Tables

Supplemental Table 2.1: RNA Library Counts & Alignment.....	83
Supplemental Table 2.2: Gene Expression (RPKM) Counts from Up-Regulated DEGs by Cell-Type and Condition.....	84
Supplemental Table 2.3: Gene Expression (RPKM) Counts from Down-Regulated DEGs by Cell-Type and Condition.....	85
Supplemental Table 2.4: DEG Counts Per Cell-Type for GO Enrichment Analysis	86
Table A2.1: Single mutant As ^{III} phenotyping	141

Chapter 1: Introduction

Arsenic Toxicity in Humans

Arsenic (As) is a non-essential, naturally occurring, toxic metalloid found in soil and water. Arsenic poisoning resulting from ingesting concentrated arsenic is rare, but can be lethal and will produce immediate symptoms of vomiting, abdominal pain, and diarrhea in humans (World Health Organization, 2018). The WHO estimates more than 200 million people worldwide are exposed to water with potentially unsafe levels of arsenic. Environmental exposure to arsenic through plant-based foods also affects millions of people globally. Chronic arsenic exposure often leads to various diseases including skin, lung, and bladder cancers or other adverse health effects such as cardiovascular diseases, type II diabetes, neurotoxicity, and developmental issues (Clemens and Ma, 2016; Cubadda et al., 2017; World Health Organization, 2018).

Once ingested, detoxification of inorganic arsenic in animals relies heavily on enzymes that catalyze arsenic methylation. Excreted through urine, monomethylated (MMA) and dimethylated (DMA) arsenic clear faster than inorganic arsenic (Buchet et al., 1981; Marafante et al., 1987). Most of this methylation is accomplished by arsenic (+3 oxidation state) methyltransferase (AS3MT), a conserved protein produced by organisms ranging from sea squirts to humans (Drobna et al., 2005; Thomas et al., 2007; De Chaudhuri et al., 2008; Thomas et al., 2010; Yokohira et al., 2010). Growing evidence suggests gut microbiota also contribute to arsenic detoxification in humans by generating methylated and thiolated arsenic species through different metabolic pathways (Van de Wiele et al., 2010; Alava et al., 2012; Yin et al., 2015; Yu et al., 2016).

Natural Sources

Soil

The global average for soil arsenic concentration is approximately 5 ppm, but ranges by region (Koljonen et al., 1989). Arsenic in European soils range from 2.5 - 410 ppm with a median of 6 ppm (Lado et al., 2008). Data from the U.S. Geological Survey showed median arsenic was 5.2 ppm in U.S. soils, ranging from < 0.6 ppm up to 830 ppm in one Nevada location (Smith et al., 2014). The inorganic arsenic species arsenite (As^{III} ; H_3AsO_3 and H_2AsO_3^-) and to a lesser degree, arsenate (As^{V} ; H_2AsO_4^- and HAsO_4^{2-}), are widely regarded as more toxic to living organisms than any organic arsenicals like monomethylarsonic acid (MMA; CH_5AsO_3) or dimethylarsenic acid (DMA; $\text{C}_2\text{H}_7\text{AsO}_2$). The oxidation state (+3 or +5) of inorganic arsenic fluctuates based on O_2 levels, environmental factors, and the presence of certain microbial flora. O_2 rich soils contain arsenic mostly as As^{V} bound to iron (Fe) oxyhydroxides molecules (Duan et al., 2017; Cui and Jing, 2019). Over time, an environment lacking O_2 will reduce As^{V} to As^{III} , but anoxic soils also promote growth of anaerobic bacteria and archaea which respire using As^{V} as an electron acceptor in place of oxygen, a process that greatly accelerates As^{V} to As^{III} reduction (Islam et al., 2004; Qiao et al., 2018). Microbially mediated arsenic reduction, which also separates arsenic from the ferric oxyhydroxides, is viewed as the dominant mechanism for mobilization of arsenic in groundwater (Nickson et al., 2000; Smedley and Kinniburgh, 2002; Akai et al., 2004; Islam et al., 2004; Qiao et al., 2018; Cui and Jing, 2019).

Water

Millions of people around the world are heavily dependent on groundwater containing elevated level of arsenic for drinking purposes (Podgorski and Berg, 2020). Groundwater with high levels of arsenic can also be extensively used for crop irrigation during dry seasons (Meharg and Rahman, 2003; de la Fuente et al., 2010; Rosas-Castor et al., 2014; Kabir et al., 2016). For example, more than half of Bangladesh farm land is irrigated during the dry winter, with 58% of the farmland using groundwater from shallow wells often high in arsenic (Smith et al., 2000; Kinniburgh and Smedley, 2001). These irrigation practices increase the arsenic concentration of local crops, which adds another source of arsenic to Bangladeshi communities already suffering from arsenic contaminated drinking water (Kurosawa et al., 2008; Dittmar et al., 2010). Rice, wheat, maize, potato, sugar beet, carrot, taro, cabbage, tomato and other crops grown in various regions can all acquire higher than normal levels of arsenic due to irrigation with arsenic contaminated groundwater (Alam et al., 2003; Norra et al., 2005; Tripathi et al., 2007; Moyano et al., 2009; Ruiz-Huerta et al., 2017).

Anthropogenic Sources

Pesticides

Soils used for agriculture may contain higher than normal arsenic levels stemming from anthropogenic sources such as arsenic-based pesticides, herbicides, and wood preservatives, arsenic-contaminated irrigation water, or runoff from nearby mining operations (**Punshon** et al., 2017). Paris green ($\text{Cu}(\text{C}_2\text{H}_3\text{O}_2)_2 \cdot 3\text{Cu}(\text{AsO}_2)_2$), an emerald-

green powder once used to kill rats in Paris sewers, was also applied to American crops from 1867 until 1900 to control Colorado potato beetle, tobacco budworm, and mosquito populations (Denver DEH, 2004; Renshaw et al., 2006). Paris green use was largely replaced by the less toxic lead arsenate ($PbHAsO_4$) or calcium arsenate ($Ca_3(AsO_4)_2$), which were used around the world until the mid-1900s. In the U.S. alone, peak uses of arsenic for agriculture was approximately 86,440,000 lbs. for lead arsenate in 1944 and 80,195,000 lbs. of calcium arsenate in 1942 (Murphy and Aucott, 1998). By 1960, use of lead arsenate, mainly in apple and cherry orchards, was phased out after studies were reporting concerns over negative health effects in orchard workers and illnesses due to consumption of American apples (Frisbie, 1936; Nelson et al., 1973). The US EPA officially banned lead arsenate use on food crops in 1988. The organic arsenicals DMA and MMA were used as pesticides on cotton and herbicides for golf courses and along highways until 2013.

Wood Treatments

Chromated copper arsenate (CCA) was used as a wood preservative starting in the 1930s to protect the material from microbes and insects. Arsenic from CCA-treated utility poles, decks, and footbridges has been shown to leach into nearby soils (Townsend et al., 2003; Pouschat and Zagury, 2006) and the US EPA, together with the lumber industry, agreed to a voluntary manufacturer withdrawal of CCA in 2003.

Mining

Mining operations also contribute to added arsenic in nearby soils. Naturally high arsenic concentrations are often found within the structure of different oxide minerals or inorganic metal compounds (Bowell et al., 2014). Arsenopyrite is the (FeAsS) is the most abundant arsenic-containing mineral in soils, but realgar (AsS), orpiment (As₂S₃), and olivenite (Cu₂OHAsO₄) are also present. Arsenic sulfides formed deep beneath the surface are prone to leaching in surface environments after extraction, leading to increased desorption and mobility of arsenic (Craw and Bowell, 2014). Many factors determine arsenic leaching near mines, including: the presence of iron sulfide, silicate, or clay minerals; susceptibility of those minerals to weathering conditions; local pH; exposure to air and water; metabolic activity and speciation of local microbes; and climate (Craw and Bowell, 2014). Ultimately, arsenic becomes a highly concentrated byproduct of mine waste where it can leach and resorb into nearby topsoil used for farming (Lee et al., 2008; Garcia-Sanchez et al., 2010; Craw and Bowell, 2014). For example, in Trentino, Italy, soils, leaves, and grape berries were sampled from 10 vineyards close to mining areas used in the Middle Ages (but no longer in use) and compared to 28 control vineyards (Bertoldi et al., 2013). Total arsenic in soil and arsenic concentrations in vine leaves and berries were all significantly higher in samples collected near the ancient mining areas. High levels of inorganic arsenic in rice have also been traced to nearby mining operations in Hunan Province, China (Zhu et al., 2008; Williams et al., 2009; Ma et al., 2016). Craw & Bowell highlight the common occurrence of elevated arsenic concentrations in gold (Au) mine waste from sites around the world, with a focus on gold mines in New Zealand (Craw and Bowell, 2014).

Arsenic in the Food Chain

Arsenic in the soil enters the food chain through plant roots. Most arsenic absorbed into root cells remains in the roots, but a portion of arsenic is transported to other tissues, some which are consumed by humans like seeds (grain). Although every land plant can acquire arsenic this way, the dominant route for human dietary exposure to arsenic is through rice consumption (Ma et al., 2014).

Rice grains can contain 10-times more total arsenic than other crop grains, partly because flooded paddies increase the amount of As^{III} available to plant roots compared to unflooded soils (Xu et al., 2008). Rice is also more efficient at both As^{III} uptake from the soil and moving that As^{III} from roots into shoots (stems, leaves, etc.) compared to other crops (Williams et al., 2007; Su et al., 2010). This increased efficiency is a direct consequence of Si uptake and translocation.

Si is an important nutrient needed by plants for defense against many different abiotic and biotic stresses (Ma, 2004). Concentrations of Si in plants typically range from 0.1% to several percent, but in rice plants Si can account for 10% of dry weight (Epstein, 1994). Rice acquires higher levels of Si mainly because its roots actively direct Si towards the vasculature through polarly localized Si transporters in specific cell-types (Ma, 2010). These same Si transporters are also permeable to As^{III}, driving the flux of both Si and As^{III} towards the xylem (Ma et al., 2008; Zhao et al., 2010).

Wheat and barley actually transport arsenic from the shoots into developing grains at higher rates than rice yet still accumulate much lower total concentrations (Williams et

al., 2007). In rice, Williams *et al.* found median As_{shoot}/As_{soil} was nearly 50 times higher than wheat or barley and median As_{grain}/As_{soil} in rice was 13 and 20 times greater than median values for wheat and barley, respectively. Median As_{grain}/As_{shoot} for wheat and barley, however, were 4 times greater than rice. Therefore, when soil conditions are equal, the rate at which arsenic moves from the roots to shoots is a better indicator of total grain arsenic than the rate of arsenic transport from shoots to grain.

Arsenic Toxicity in Plants

Like animals, plants suffer from arsenic exposure. It is well documented that arsenic leads to the production of reactive oxygen species (ROS), including superoxide (O_2^-), hydroxyl radicals (OH), and hydrogen peroxide (H_2O_2) (Hartley-Whitaker et al., 2001; Singh et al., 2006). In addition, As^V and phosphate (Pi) are chemically similar and As^V competes with Pi to bind transport proteins on the root surface, reducing Pi acquisition (Ullrich-Berius et al., 1989). As^V also reduces cellular ability to produce ATP and maintain normal metabolism (Tawfik and Viola, 2011). ATP synthases in the mitochondrial inner membrane and plastid thylakoid membrane can utilize As^V in the reactions that normally phosphorylate ADP to produce ATP, instead forming ADP-As^V, a molecule with no metabolic use in the cell (Gresser, 1981; Moore et al., 1983). Arsenic also limits the uptake and/or transport of essential elements like selenium (Se), nickel (Ni), and zinc (Zn), although the cause of these associations are not understood (Williams et al., 2009; Duan et al., 2013). Williams *et al.* showed that as soil arsenic increases, rice shoots acquired less

Se and Zn. Ni acquisition in rice shoots was unaffected by soil arsenic, but as shoot arsenic increased, less Ni, Se, and Zn was transferred to the grain.

Stunted growth, brown spots, and scorching (browning at leaf margins and tips) are symptoms of arsenic toxicity in rice plants (Khan et al., 2010). Ultimately, rice grain yields become severely reduced as soil arsenic concentrations increase (Duxbury et al., 2003; Heikens et al., 2007). MMA and DMA induce straighthead disease, a disorder identified in fully mature rice by sterile and/or empty spikelets, distorted husks, and erect ('straight') panicles ('head') (Rahman et al., 2008; Tang et al., 2020). Healthy rice panicles are weighted down by multiple spikelets each filled with grain, causing mature panicles to sag or bend over under the weight of many fully developed grains. Recent work by Tang *et al* showed straighthead disease was induced in WT rice when 2-10 μM DMA was added to the soil at tillering to flowering stages (Tang et al., 2020). Tang *et al* also developed transgenic rice expressing a bacterial arsenite methyltransferase (*arsM*) gene, which allowed the plant to methylate As^{III} to MMA and (mostly) DMA. Seed setting rates of these transgenic *arsM* rice lines, after exposure to 2 μM As^{III} for 2 weeks, were only 20-39% of WT; typical similar to that seen in rice with straighthead diseased rice (Tang et al., 2020). Arsenic contaminated soils, especially those used to grow rice, thus pose agro-economic challenges in addition to human health and safety concerns.

The order of phytotoxicity for inorganic and organic arsenic is plant species specific and possibly dependent on soil conditions. Many studies have been conducted showing inorganic and organic arsenic toxicity varies in crops like maize, wheat, and rice, although not every arsenic species has been assayed with each plant. For example, the order of arsenic toxicity follows As^V > As^{III} > DMA for maize (Abbas and Meharg, 2008), As^{III} >

DMA > As^V for wheat (Liu et al., 2005; Yoon et al., 2015; Duncan et al., 2017), and MMA > As^{III} > As^V > DMA for rice (Marin et al., 1992). Phytotoxicity for two *Spartina* grass species was shown to be DMA = MMA > As^{III} > As^V, different from any crop species tested (Carbonell-Barrachina et al., 1998).

Work by Yoon *et al* assaying effects of As^{III}, As^V, and DMA on 10 different crop plants demonstrated variation in relative toxicity to the different arsenic compounds and that soil conditions can alter arsenic toxicity in some plants (Yoon et al., 2015). Of the plant species tested they showed sensitivity to As^{III} was mung bean > pea > cucumber > wheat > kale > barley > sorghum, similar to As^V which was mung bean > pea > cucumber > wheat, but sensitivity to DMA differed with mung bean > barley > wheat > cucumber (Yoon et al., 2015). They also showed organic matter content in soil could alter arsenic toxicity for mung bean, wheat, cucumber, and sorghum. For example, the toxic effect (seedling growth and germination) of As^{III} on wheat was higher in soil with less organic matter, while the toxic effect of DMA on wheat was higher in soil with more organic matter (Yoon et al., 2015).

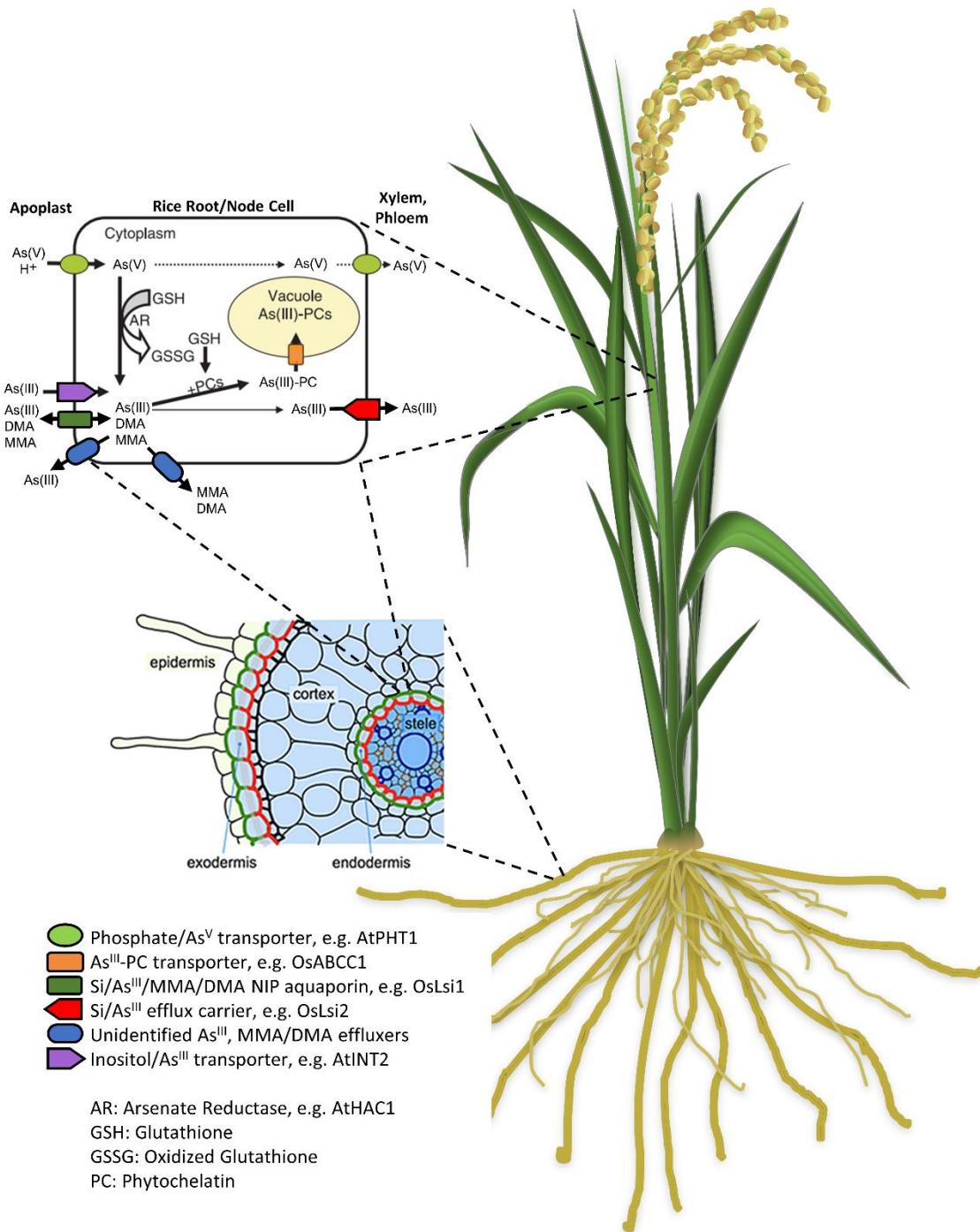


Figure 1.1: Generalized Diagram of Arsenic Transport and Metabolism in Plants

Modified from Punshon *et al.* (2017), Zhao *et al.* (2009), and Ma *et al.* (2007).

Arsenic Uptake from Soil

The amount of different arsenic species taken up by plants from the soil is most often $\text{As}^{\text{III}} > \text{As}^{\text{V}} > \text{DMA} > \text{MMA}$ (Raab et al., 2007; Finnegan and Chen, 2012). While inorganic arsenic is taken into root cells at higher rates than organic arsenic, MMA and DMA move from roots to shoots to grain more efficiently. MMA and DMA have not been shown to sequester into plant vacuoles, which ostensibly allows these molecules to move unimpeded through the plant. A study of 46 different plant species found arsenic translocation rates from roots to shoots generally reflected $\text{DMA} > \text{MMA} > \text{As}^{\text{V}}$ in most species (Raab et al., 2007). MMA and DMA contaminated soils, therefore, might result in crops with higher total arsenic concentrations in the shoots and grain than when As^{III} or As^{V} are the dominant soil species (Finnegan and Chen, 2012; Cubadda et al., 2017).

The various arsenic species enter root cells via different membrane transport proteins on the plasma membrane that allow ions and molecules to cross with varying levels of selectivity, or substrate specificity. Similarities in chemical structure between As^{V} and phosphate (PO_4^{3-}), and between As^{III} and silicic acid ($\text{H}_2\text{O}_3\text{Si}$) allow arsenic to enter using transporters for essential molecules. In flowering plants like *Arabidopsis thaliana* and rice (*Oryza sativa*), nearly all arsenic first enters root cells as As^{III} in anaerobic soils or As^{V} in aerobic soils (Clemens and Ma, 2016). The process of growing rice in continuously flooded paddies, instead of irrigating using wet/dry cycles, is a unique cultivation strategy among staple crops. A negative consequence of growing rice this way is the As^{III} concentration of the soil rises during prolonged flooding (Xu et al., 2008). Over time flooding depletes the soil of oxygen, allowing communities of anaerobic microbes to

thrive in the soil where rice roots are growing. These microbes reduce As^V-Fe complexes, releasing the arsenic as As^{III}, which is more easily taken up by Si transporters in rice roots (Islam et al., 2004; Williams et al., 2007; Xu et al., 2008; Zhu et al., 2014). MMA and DMA are also unintended substrates of the same Si proteins that transport As^{III}, with uptake specificity of OsLsi1, being As^{III} >> MMA > DMA (**Figure 1.1**) (Abedin et al., 2002; Li et al., 2009). Plants themselves cannot methylate arsenic, but many bacteria, archaea, and fungi possess the conserved *arsM* gene required for arsenic methylation (Andres and Bertin, 2016; Dunivin et al., 2019). Soil microorganisms are therefore considered fully responsible for converting inorganic arsenic to endogenous (non-pesticide, herbicide, etc.) organic forms, including MMA and DMA (Lomax et al., 2012; Jia et al., 2013; Xu et al., 2016).

A detailed understanding of arsenic uptake and transport in plants is needed if we are to reduce arsenic accumulation in food. In the following sections, we briefly summarize what is known and which genes have been identified to date.

Aerobic Soils

In aerobic soils, As^V can enter root cells via phosphate transporters. Members of the Phosphate transporter 1 family such as AtPht1;1, AtPht1;4 in *A. thaliana* and OsPT1, OsPT8 in rice are responsible for As^V uptake (**Figure 1.1**) (Catarecha et al., 2007; Seo et al., 2008; Zhao et al., 2009; Sun et al., 2012; Kamiya et al., 2013; LeBlanc et al., 2013; Wang et al., 2016; Ye et al., 2017; Tang and Zhao, 2020). Cytosolic As^V is almost entirely

reduced to As^{III} by arsenate reductase enzymes in the roots, although As^V reduction can also occur in the shoots (Chao et al., 2014; Sanchez-Bermejo et al., 2014; Shi et al., 2016).

Anaerobic Soils

In anaerobic soils, As^{III} enters root cells via passive, bi-direction channels known as aquaporins or Membrane Intrinsic Proteins (MIPs). Plant MIP subfamilies include Nodulin 26-like Intrinsic Proteins (NIPs), Plasma membrane Intrinsic Proteins (PIPs), Tonoplast Intrinsic Proteins (TIPs), Small basic Intrinsic Proteins (SIPs) and uncharacterized X-Intrinsic Proteins (XIPs). MIPs are characterized by having six transmembrane domains connected by five loops, with two loops containing alpha-helix domains with NPA (Asparagine-Proline-Alanine) motifs that form the main pore, and both the N-terminus and C-terminus are localized in the cytosol (Bezerra-Neto et al., 2019). MIPs are organized into tetramers in the cell membranes which are stabilized by hydrogen bonds and interactions among the monomer loops. The individual monomers, however, each act as individual pores for water and other substrates (Jiang et al., 2006; Bezerra-Neto et al., 2019).

NIP proteins have been shown to allow As^{III} to move in, out, and between root cells (**Figure 1.1**). As^{III} permeable NIPs in plants also transport substrates such as water, hydrogen peroxide, silicon, glycerol, boron, antimony, or aluminum (Weig and Jakob, 2000; Kamiya and Fujiwara, 2009; Sasaki et al., 2016; Sadhukhan et al., 2017; Wang et al., 2017; Tang and Zhao, 2020).

Arabidopsis plants lacking AtNIP1;1 and rice plants lacking its closest rice homolog OsNIP2;1 (Lsi1) have increased tolerance to As^{III}, likely by reducing the amount of As^{III} that enters the roots. These mutant lines have lower total inorganic arsenic in the roots, shoots, and seeds (Ma et al., 2008; Kamiya et al., 2009). Of the nine NIP aquaporins in Arabidopsis, six (AtNIP1;1, AtNIP1;2, AtNIP3;1, AtNIP5;1, AtNIP6;1, and AtNIP7;1) are capable of bidirectional As^{III} transport (**Figure 1.1**) (Bienert et al., 2008; Kamiya et al., 2009; Xu et al., 2015; Lindsay and Maathuis, 2016), but due to their low expression in root cells their As^{III} transport contribution in vivo is less clear compared to AtNIP1;1. OsNIP2;1/Lsi1 is an important Si transporter, which along with Lsi2 (an anion transporter family member unrelated to NIPs), facilitates Si transport towards the root vasculature in rice (**Figure 1.1**) (Ma et al., 2007; Ma et al., 2008). Relative to the xylem in rice roots, OsNIP2;1/Lsi1 localizes on the distal plasma membrane of exodermis and endodermis cells to influx Si while Lsi2 localizes on the proximal plasma membrane of exodermis and endodermis cells to efflux Si (Ma et al., 2006; Ma et al., 2007). While both proteins contribute to As^{III} accumulation, Lsi2 has a greater impact than OsNIP2;1/Lsi1, as field grown *lsi2* mutant rice showed lower concentrations of arsenic in husks, grain, straw, and entire shoots than WT and *lsi1* mutants (Ma et al., 2008). Homologs of Lsi2 in other crops, however, do not show polar localization (Ma, 2010). The ability of rice roots to take up and mobilize both Si and As^{III} through this unique Lsi1/Lsi2 pathway results in higher Si/As^{III} flux through the roots and is ultimately responsible for much higher Si/As^{III} concentrations in rice grain compared to barley, maize, rye, sorghum or wheat grain (Tamai and Ma, 2003; Ma et al., 2008; Ma, 2010).

Phytocelatin Mediated Sequestration

Animals have evolved an arsenic detoxification system that first utilizes an AS3MT-like enzyme to methylate inorganic arsenic, which can then be filtered through specialized tissues and excreted (Thomas et al., 2010; Yokohira et al., 2010). After humans ingest inorganic arsenic (iAs), arsenic is excreted in urine as DMA >> MMA > iAs, although results can vary depending on population, diet, etc. (Buchet et al., 1981; Lai et al., 2004; Gilbert-Diamond et al., 2011; Morton and Leese, 2011). Higher land plants do not contain an *AS3MT*-like gene to methylate inorganic arsenic or a mechanism to remove methylated arsenic. In fact, when Tang *et al* expressed a bacterial arsenite methyltransferase (*arsM*) gene in rice, the plants grown on 2 μ M As^{III} converted most As^{III} to DMA, which had little effect on vegetative tissues, but severely penalized reproductive tissue development (Tang et al., 2020). The panicles of *arsM* expressing rice either failed to emerge or produced unfilled or under-filled seeds with deformed husks, classic symptoms of straighthad disease (Tang et al., 2020).

Instead of arsenic methylation/excretion, plants sequester inorganic arsenic and other toxic molecules into the vacuole, a membrane-bound organelle present in every cell of the plant. It has been reported only 10% of total As^{III} taken up by rice plants from the soil reaches the shoots and only 3.3% ends up in the grain (Zhao et al., 2012). The low percentage of As^{III} reaching the grain is mostly attributed to phytocelatin (PC) mediated complexation of As^{III}, which is required for vacuolar sequestration in plants. Phytocelatin Synthase (PCS) enzymes synthesize PC peptides in the form of $(\gamma\text{-GluCys})_n\text{-Gly}$ ($n = 2\text{-}11$) by polymerizing glutathione (GSH, $\gamma\text{-GluCysGly}$) (Zenk, 1996; Cobbett and

Goldsbrough, 2002). This process has been widely studied in higher plants, but PCS proteins have also been identified in algae (Tsuji et al., 2002; Brautigam et al., 2011), fungi (Ha et al., 1999), invertebrate animals (Vatamaniuk et al., 2001; Brulle et al., 2008; Rigouin et al., 2013), and even bacteria (Tsuji et al., 2004). The cysteine thiol groups found within GSH and PCs have a high affinity for metal binding, especially for cadmium (Cd^{2+}) and As^{III} (Schmoger et al., 2000). PC affinity for these heavy metals in plants is $PC_2 > PC_3 >> PC_{4-5}$ for As^{III} and $PC_3 >> PC_4$ for Cd^{2+} (Schmoger et al., 2000).

In every cell As^{III} enters between the roots and grain, there is an opportunity for PCs to bind As^{III}. Transporter proteins on the vacuolar membrane (tonoplast) that recognize $PC_n\text{-As}^{III}$ can then move the chelated arsenic into the vacuole where As^{III} toxicity is mitigated. PC mediated As^{III} sequestration is the main detoxification strategy in angiosperms, which generally results in most arsenic being trapped in the roots, although species specific tissues like rice nodes can also be important locations for As^{III} sequestration (Song et al., 2010; Song et al., 2014; Zhang et al., 2018).

Once As^{III} has entered a cell it can either be complexed to glutathione/phytochelatin molecules and sequestered inside the vacuole via ABCC transporters, moved from cell to cell towards the vasculature (Tang et al., 2019), or effluxed back into the soil (**Figure 1.1**) (Zhao et al., 2010; Chao et al., 2014; Tang and Zhao, 2020). Rice OsABCC1 and Arabidopsis AtABCC1 and AtABCC2 are vacuole localized transporters of heavy metal/metalloid-PC complexes As^{III}-PC₂, Cd^{II}-PC₂, and Hg^{II}-PC₂ (**Figure 1.1**) (Song et al., 2010; Song et al., 2014). Approximately 60% to 80% of inorganic arsenic that enters rice roots is effluxed back to the external environment as As^{III} by the NIP aquaporin Lsi1 and other currently unidentified transporters (**Figure 1.1**) (Xu et al., 2007; Zhao et al., 2010).

Lsi1 accounts for 10-15% of total As^{III} efflux, suggesting different transporter proteins contribute to this process as well (Zhao et al., 2010). ABCC proteins are responsible for sequestering the remaining As^{III} in root vacuoles, preventing most As^{III} from reaching the shoots (**Figure 1.1**) (Liu et al., 2010). Recent work also suggests AtVPT1, a tonoplast-localized phosphate transporter responsible for vacuolar Pi sequestration, is indirectly involved in As^V resistance, not by transporting As^V into the vacuole, but by modulating phosphate balance between the cytoplasm and vacuole (Luan et al., 2018). When VPT1 is knocked out in Arabidopsis plants, less Pi enters vacuoles, elevating cytosolic Pi levels. The roots respond to this pseudo-Pi-sufficient state by down-regulating expression of PHT1 family genes to restrict Pi (and consequently As^V) uptake (Luan et al., 2018). Luan *et al* also showed VPT1 overexpression reduced As^V tolerance.

Hyperaccumulators

Not all plants utilize PCs for sequestering As^{III} into the vacuole. In the Chinese brake fern (*Pteris vittata*), a plant that hyperaccumulates arsenic, an As^{III} transporter, PvACR3, was identified by screening *P. vittata* genes capable of complementing the yeast *acr3* mutant (Indriolo et al., 2010). Although in yeast ScACR3 transports As^{III} out of the cell, PvACR3 was found to localize to the vacuolar membrane of *P. vittata* cells and transport As^{III} without it being complexed to GSH or PC_n. Most of the arsenic is sequestered in the vacuoles of the fronds (leaves), allowing the fern to accumulate and tolerate high levels of arsenic (Chen et al., 2013). PvACR3 orthologs have been found in several plant species, but appear lost in flowering plants (Zhang et al., 2018; Perez-Palacios

et al., 2019). The PvACR3 protein sequence most closely aligns with a bile acid:sodium symporter (BASS) family protein (At1g78560; BASS1) in the bile/arsenite/riboflavin/transporter (BART) protein superfamily (Indriolo et al., 2010). The BASS1 protein was recently shown to localize exclusively to the chloroplast inner envelope and is therefore unlikely to contribute to arsenic tolerance (Huang et al., 2018).

Root to Shoot Arsenic Transport

The process of moving As^{III} from the roots to the shoots is not well understood. Natural Resistance-Associated Macrophage Proteins (NRAMPs) might be involved as Arabidopsis plants ubiquitously expressing GFP-tagged *OsNRAMP1* using the CaMV35S promoter, accumulated more arsenic in roots and shoots with increased biomass compared to WT (Tiwari et al., 2014). The GFP-tagged protein only localized to endodermal and pericycle cells in the roots, suggesting post-transcriptional regulation of this NRAMP. The phenotype of these transgenic plants, together with the cell-type specific translation and localization, suggests NRAMPs might have a role in loading As^{III} into the xylem. It should also be noted that yeast cells expressing *OsNRAMP1* showed a dramatic increase in Cd uptake, but relatively little difference in As^{III} uptake or As^{III} tolerance (Tiwari et al., 2014). Recently Arabidopsis inositol transporters AtINT2 and AtINT4 were shown to assist in loading As^{III} into phloem cells (Duan et al., 2016), opening the door for other sugar transporters to become candidates for novel As^{III} transport (**Figure 1.1**).

In the last decade, research into how plants respond to arsenic has identified many genes important for arsenic transport and detoxification. Most of this work has been

focused on plant roots responding to As^V exposure. The greatest source of dietary arsenic that threatens human health, however, comes from rice grown in As^{III} rich environments and some proteins contributing to known As^{III} detoxification pathways have yet to be identified. For example, 60% to 80% of the As^V entering root cells can be effluxed back into the external environment as As^{III} (Xu et al., 2007; Liu et al., 2010; Zhao et al., 2010), and while bi-directional passive channels contribute, NIPs like OsLsi1 can only account for 15-20% of the total As^{III} that is effluxed (Zhao et al., 2010). Therefore, it is highly likely other transport proteins outside the NIP family of aquaporins contribute to this efflux process but remain elusive still.

WRKY6 and WRKY45 transcription factors regulate the expression of *Pht1;1* by limiting As^V uptake in Arabidopsis (Castrillo et al., 2013; Wang et al., 2014). In response to As^V, WRKY6 represses *Pht1;1* expression, which is accompanied by endocytosis of the PHT1;1 protein from the plasma membrane (Castrillo et al., 2013). WRKY45, in contrast, activates *Pht1;1* expression. Overexpression of *WRKY45* increased *PHT1;1* expression, leading to higher Pi uptake and Pi content in 17-day-old seedlings (Wang et al., 2014). The *WRKY45* overexpression plants were also more sensitive to As^V while *wrky45* knockdown RNAi lines were slightly more tolerant to As^V (Wang et al., 2014). An ER-localized protein, Phosphate Transporter Traffic Facilitator 1 (PHF1), affects the localization of PHT1;1 protein. Mutations to *PHF1* led to PHT1;1 retention in the ER, with no PHT1;1 reaching the plasma membrane, even during Pi starvation (Gonzalez et al., 2005). Mutant *phf1* plants were also highly resistant to As^V (Gonzalez et al., 2005). Apart from a recent study demonstrating an Arabidopsis calcium-dependent protein kinase, CPK31, might target and regulate NIP1;1 for As^{III} tolerance (Ji et al., 2017), little is known about

transcriptional or translational regulation of As^{III} transporters during As^{III} stress. Cell-type specific information for known arsenic genes is also lacking.

Shoot to Seed Arsenic Transport

Very little is known about how arsenic exits phloem sap and is loaded into the developing grain. Using Synchrotron X-ray fluorescence (SXRF) microtomography, Carey *et al* (2011) and Punshon *et al* (2018) were able to localize arsenic in the rice grain (Carey et al., 2011; Punshon et al., 2018). By delivering As^V, As^{III}, or DMA through cut flag leaves during grain filling, SXRF analysis showed As^V and As^{III} localized to the ovular vascular trace (OVT) in the bran layer of rice grains and very little As^V or As^{III} entered the starchy endosperm tissue (Carey et al., 2011; Punshon et al., 2018). DMA, in contrast, readily entered the starchy endosperm and in higher concentrations than As^V or As^{III} OVT concentrations (Punshon et al., 2018).

Recently, a putative peptide transporter, OsPTR7, was identified in rice as a potential DMA transporter (Tang et al., 2017; Tang and Zhao, 2020). *OsPTR7* is normally highly expressed in rice leaves, node 1, and roots. Knockout mutants of *OsPTR7* decreased root to shoot DMA transport and accumulated less DMA in the grain than WT (Tang et al., 2017).

Plant Genetic Engineering Strategies to Decrease the Arsenic Content of Plants

Developing effective, economically feasible, and sustainable methods for both remediation of arsenic from soils and generating new rice varieties with low grain arsenic content are urgently needed to mitigate human exposure to arsenic. To achieve these results, future transgenic approaches that utilize the collective knowledge of plant biochemical, molecular, and genetic processes to modulate arsenic levels in plants are desirable for their relatively low cost, environmental safety, and scalability (Kumari et al., 2018). Essentially, developing transgenic plants to manipulate arsenic localization, that can be substituted or rotated into growing cycles is more attainable than conventional physicochemical methods like chemical precipitation, electrolytic recovery, ion chelation, solvent extraction, etc. which would need to be implemented at every problematic location. While some of these non-plant-based approaches are promising (Cegłowski and Schroeder, 2015; Kumar et al., 2017; Naseri et al., 2017), they can be cost-intensive, ineffective when metal concentrations in the soil are low (Hammami et al., 2003), and difficult to enforce in the private-sector of economies.

Modulating Arsenic Uptake

Many transgenic approaches have already begun that alter arsenic accumulation and tolerance in plants, although most focus on only one step in arsenic transport or metabolism. Modulating single plant genes responsible for arsenic uptake, translocation, sequestration, or transformation have shown the most promise for altering arsenic accumulation. Knocking out arsenic transporter genes like *AtNIP1;1* in Arabidopsis or *OsLsi1* in rice, is effective at reducing arsenic seed/grain concentration (Ma et al., 2008;

Kamiya et al., 2009). However, removing these genes also reduces Si uptake and grain yields (Ma et al., 2006; Ma and Yamaji, 2006; Ma et al., 2007). Lowering arsenic accumulation without compromising plant growth, Si uptake, or grain yield can be accomplished, however, by overexpressing different NIP aquaporins in rice, *OsNIP1;1* or *OsNIP3;3* (Sun et al., 2018). Using a maize ubiquitin promoter to drive strong constitutive expression of *OsNIP1;1* or *OsNIP3;3* in all cells, the OsNIP1;1 and OsNIP3;3 transporters localized to all cells and all surfaces of the plasma membrane, unlike OsLsi1 and OsLsi2, which are polarly localized in certain root cell-types (**Figure 1.1**) (Sun et al., 2018). Sun *et al* argue these transgenic lines probably disrupt the directional radial transport of As^{III} to the stele formed by Lsi1 and Lsi2, essentially providing an additional route for As^{III} to leak out of root cells and the stele into the apoplast (Sun et al., 2018). Their data also suggest OsNIP1;1 and OsNIP3;3 are equally permeable as OsLsi1 to As^{III}, but less permeable to Si. Therefore, as Si and As^{III} enter root cells through Lsi1, As^{III} could be preferentially effluxed by OsNIP1;1 or OsNIP3;3 into the apoplast. When additional Si was added to the medium, however, these transgenic lines acquired more As^{III} in roots and shoots than WT plants (Sun et al., 2018). Normally excess Si or As^{III} in the medium cause plants to down-regulate most Si/As^{III} transporters in the roots, but the overexpression lines maintained high *NIP1;1* or *NIP3;3* expression and thus high Si and As^{III} uptake (Sun et al., 2018). Future transgenic rice involving NIP aquaporins must be capable of adapting to soils with different Si and As^{III} availability.

NIPs belong to a larger Major Intrinsic Proteins (MIP) superfamily of aquaporins (Kapilan et al., 2018; Bezerra-Neto et al., 2019) that also include Plasma membrane Intrinsic Proteins (PIPs), Tonoplast Intrinsic Proteins (TIPs), Small basic Intrinsic Proteins

(SIPs) and X Intrinsic Proteins (XIPs). PIPs fall into two subgroups in plants; PIP1 and PIP2. PIPs can form hetero-tetramers, which affects substrate permeability (Fetter et al., 2004; Yaneff et al., 2014). The genomic sequences within PIP1 and PIP2 subfamilies are very similar to each other compared to NIPs, SIPs, or TIPs subfamilies (Johanson et al., 2001; Bezerra-Neto et al., 2019). The high similarity between PIPs suggests this subfamily has expanded more recently than the other subfamilies. PIPs might therefore perform more redundant functions than NIPs, TIPs, and SIPs, which have had more time to evolve novel functions. There are 13 *A. thaliana* PIPs and 11 *O. sativa* PIPs. In addition to water transport, PIPs have been shown to facilitate CO₂ (Heckwolf et al., 2011) and H₂O₂ (Tian et al., 2016; Rhee et al., 2017). PIPs might also be involved with arsenic transport. Overexpression of PIP aquaporins *OsPIP2;4*, *OsPIP2;6*, or *OsPIP2;7* enhanced rice tolerance to As^{III} while increasing plant biomass (Mosa et al., 2012).

A TIP aquaporin gene from the arsenic hyperaccumulating plant *Pteris vittata*, *PvTIP4;1* was shown to increase arsenic accumulation when expressed in Arabidopsis, although the plants became more sensitive to arsenic (He et al., 2016).

Genomic changes involving phosphate transporter genes can alter As^V accumulation. Lower arsenic was measured in xylem sap of rice *OsPT4* null mutants (Cao et al., 2017), while overexpressing *OsPT4* in rice showed increased sensitivity to As^V and higher rates of As^V uptake (Ye et al., 2017). Overexpression of *OsPT1* or *OsPT8* in rice or *AtPHT1;1* in *A. thaliana* also increased As^V uptake (Wu et al., 2011; Kamiya et al., 2013; Su et al., 2015). The rice *aus* variety Kasalath is normally more tolerant to As^V than the *japonica* variety Nipponbare, but null mutations of *OsPT8* in each variety decreased As^V uptake and increased As^V tolerance (Wang et al., 2016). The difference in As^V

tolerance between the two varieties was also abolished in the *ospt8* Kasalath and *ospt8* Nipponbare mutants, suggesting natural mutations in this one gene generated changes to rice As^V tolerance.

Modulating Arsenic Sequestration

Quickly reducing As^V to As^{III}, chelating As^{III} with PC, and trapping that PC-As^{III} in vacuoles is a major component of arsenic detoxification in angiosperms (Zhang et al., 2018). PC-As^{III} transporters located on the vacuolar membrane greatly influence which cells retain arsenic so cells in other plant tissues can evade the toxic metalloid. In angiosperms, this process occurs mostly in the roots. Overexpression of the PC-As^{III} transporter *AtABCC1* gene in Arabidopsis plants only resulted in increased arsenic tolerance when co-overexpressed with the phytochelatin synthase *AtPCS1* gene, presumably by providing more PCs to chelate As (Song et al., 2010).

Recent work by Chen *et al.* (2017) showed expressing a vacuolar membrane As^{III} antiporter gene *PvACR3;1* from *Pteris vittata* in Arabidopsis and tobacco successfully increased arsenic retention, without PC, in root vacuoles, leading to lower arsenic translocated to shoots and more arsenic tolerant plants overall (Chen et al., 2017). Similar results were seen when the *PvACR3;1* construct was expressed in rice, again leading to higher arsenic retention in the roots and lower arsenic in the shoots and grain (Chen et al., 2019).

Deng *et al* (2018) recently developed transgenic rice plants ubiquitously expressing a bacterial γ -glutamylcysteine synthetase gene, γ -ECS, along with two different As^{III}

vacuolar transporter genes, yeast *ScYCF1* and rice *OsABCC1*, under the control of a tissue specific promoter driving expression in root cortex, node phloem, and internode cell types (Deng et al., 2018). Transgenic rice lines expressing all three genes greatly reduced root-to-shoot arsenic translocation due to elevated vacuolar sequestration of As^{III} in root cortex cells. Cortex cells are particularly important for this process because the adjacent cell layer proximal to the vasculature is the endodermis, where Lsi1 and Lsi2 efflux transporters coordinate to move Si and As^{III} into the xylem. Rice root cortex cells are therefore the final layer of root cells that effectively trap As^{III} to prevent it from reaching the shoots. Their transgenic rice was also able to reduce arsenic translocation from the internode to the grain. Ultimately these rice lines produced grain with 70% less arsenic without any agronomic trait penalties (Deng et al., 2018).

Modulating Arsenic Efflux

Plant roots can efflux arsenic back into soil. Although MIP aquaporins passively contribute to efflux through bidirectional diffusion of As^{III}, other transport protein(s) that efflux As^{III} out of the roots have yet to be identified. Angiosperms appear to have lost the ACR3-like antiporter gene that can transport unchelated As^{III} across membranes (Duan et al., 2012). Two research groups, however, were able to express the yeast plasma membrane As^{III} transporter gene *ScACR3* in plants (Ali et al., 2012; Duan et al., 2012). In yeast, ScACR3 effluxes As^{III} back into the external medium. ScACR3, when expressed in both Arabidopsis and rice, localized to the plasma membrane and increased As^{III} efflux activity

into the external medium, improving tolerance to As^{III} and As^V by reducing arsenic accumulation (Ali et al., 2012; Duan et al., 2012).

Chen et al (2013) were able to express the *P. vittata* vacuolar membrane As^{III} antiporter gene *PvACR3* (not to be confused with *PvACR3;1* discussed in the section on modulating arsenic sequestration), in Arabidopsis and found this protein localized to the plasma membrane, not the vacuolar membrane (Chen et al., 2013). *PvACR3* and *ScACR3* appear to function in the same way when expressed in Arabidopsis, effluxing As^{III} out of the cytoplasm, back into the external medium. *PvACR3* transgenic lines dramatically reduced root and shoot arsenic concentrations compared to WT when grown on 5 µM As^{III} and greatly reduced arsenic concentrations in the root, but not shoots compared to WT when grown on 150 µM As^{III} (Chen et al., 2013). When grown for 70 days in 10 ppm As^V soil *PvACR3* plants also had elevated shoot arsenic concentrations. *PvACR3* transgenic plants also accumulated more shoot biomass and had longer roots than WT in all assays.

Chen et al (2013) suggest the elevated shoot arsenic is caused by *PvACR3* in the roots effluxing As^{III} toward or into the xylem, but this is unlikely given how little arsenic reaches the shoots when the *PvACR3* transgenic plants are grown on As^{III} medium (Chen et al., 2013). It is more likely the elevated shoot arsenic occurs because *PvACR3* in roots efflux As^{III} quickly enough that the roots never experience As^{III} stress, which would explain the increased biomass and longer roots. *PvACR3* plants likely maintain normal levels of Pi/As^V transporters, arsenate reductases, etc. in the roots, causing higher levels of As^V to enter roots and less efficient As^V to As^{III} reduction than WT. Even though nearly all As^{III} is effluxed back outside the roots in *PvACR3* plants, more As^V might reach the shoots before it can be reduced to As^{III}.

Regardless of mechanistic details, expressing *ScACR3* or *PvACR3* in rice appears to be a promising strategy in reducing arsenic accumulation in rice, but only if the rice is grown in flooded paddies, where most arsenic is present as As^{III} and not As^V.

Natural Resistance-Associated Macrophage Proteins (NRAMPs) are also candidates for arsenic transport that could affect overall arsenic accumulation. Heterologous expression of rice *OsNRAMP1* in Arabidopsis was able to increases both As^{III} accumulation and tolerance, while also increasing overall plant biomass (Tiwari et al., 2014).

Modulating Arsenate Reduction

When crops, including rice, are grown in aerated, non-flooded soils As^V is the dominant arsenic species taken up by roots (Clemens and Ma, 2016). Under these conditions quickly reducing As^V to As^{III} in the roots becomes critically important so As^{III} can be either effluxed back to the soil or sequestered in vacuoles. Recent work that alters a plants As^V reduction capacity appears to be an effective strategy for modulating arsenic tolerance, accumulation, and localization within the plant. Over-expressing either *OsHAC1;1* or *OsHAC1;2*, the two arsenate reductase genes in rice, resulted in decreased arsenic accumulation in rice grains (Shi et al., 2016). In contrast, single knockout mutant lines of either *OsHAC1;1* or *OsHAC1;2* increased arsenic accumulation in roots, shoots, and grains, while decreasing As^{III} efflux back to the hydroponic medium (Shi et al., 2016). Double knockout lines of both *OsHAC1;1* and *OsHAC1;2* genes had a cumulative effect,

further increasing arsenic accumulation in roots and shoots with even lower As^{III} efflux compared to single knockouts (Shi et al., 2016).

Xu *et al.* 2017 showed knockout and over-expression lines of a different rice arsenate reductase, *OsHAC4*, also showed decreased As^V reduction, and increased arsenic accumulation in roots and shoots (Xu et al., 2017). In their report OsHAC4 only appears active in root epidermal cells, whereas OsHAC1;1 could function in multiple root cell types, in addition to stem, leaf, and node tissues (Xu et al., 2017). It therefore appears likely specific cell-types have unique roles within the plant that contribute to overall arsenic detoxification.

Modulating Thiol Metabolism

Cysteine thiol groups within GSH and PCs have a high a high affinity for As^{III} and angiosperms use these thiol containing molecules for As^{III} sequestration. Genes for transporters and enzymes involved in cysteine, glutathione, or phytochelatin metabolism will therefore be useful when developing transgenic plants with altered arsenic accumulation. Early steps for PC biosynthesis depend on cellular sulfur (S) acquisition. Mutant *Arabidopsis* plants lacking a functional *SULTR1;2* sulfate transporter gene acquire less S in roots and shoots, rendering these mutant plants more sensitive to arsenic than WT (Nishida et al., 2016). GSH serves as a major reservoir for reduced sulfur in plants (Noctor, 2006) and enzymes in the GSH biosynthesis pathway are required for normal root cell division in plant development (Vernoux et al., 2000; Foyer and Noctor, 2005).

Maintaining GSH and PC reservoirs for root development during arsenic stress should be considered when designing transgenic plants to increase GSH and PC production. To preserve this thiol balance, Guo *et al* simultaneously overexpressed a phytochelatin synthase gene, *AsPCS1*, from garlic (*Allium sativum*) and a γ -glutamylcysteine synthetase gene, *GSH1*, from yeast (*S. cerevisiae*) in *Arabidopsis thaliana* (Guo et al., 2008). Under As^{III} stress, dual-gene overexpression lines accumulated nearly 10 times more arsenic than WT while single-gene overexpression lines accumulated only 3.5 times more arsenic than WT. Dual-gene overexpression lines were also healthier, with longer roots than WT and single-gene overexpression lines.

The bacterial γ -glutamylcysteine synthetase gene, *γ -ECS* described above in the *Modulating Arsenic Sequestration* section had previously been expressed in Arabidopsis plants with a strong constitutive promoter, which moderately increased resistance to mercury (Hg) and arsenic (Li et al., 2005). In another study, transgenic Arabidopsis lines were developed to ubiquitously overexpress *γ -ECS* along with leaf specific expression of *ArsC*, a bacterial arsenate reductase gene (Dhankher et al., 2002). When grown on medium containing As^V, single *ArsC* expressing plants were hypersensitive to As^V compared to WT (Dhankher et al., 2002). *ArsC* transgenic plants quickly reduce As^V reaching the shoots to As^{III} and are unable to efficiently detoxify As^{III} in the shoots. Transgenic plants expressing both *γ -ECS* and *ArsC*, however, showed strong tolerance to As^V compared to single *γ -ECS* lines and WT, while also accumulating more than double total shoot arsenic than WT or either single transgenic lines (Dhankher et al., 2002). Using a similar *γ -ECS + ArsC* approach is an exciting strategy for developing plants capable of phytoremediation on As^V contaminated soils.

It should be noted that in the *Modulating Arsenic Sequestration* section above, all transgenic rice plants in Deng *et al.* (2018) expressing γ -ECS and *OsABCC1* accumulated more than twofold higher arsenic in their roots than WT and rice plants expressing only *OsABCC1* (Deng et al., 2018). Only after combining *OsABCC1* and γ -ECS constructs did they see more arsenic being retained in the roots, suggesting that thiol content can be a limiting factor when designing transgenic rice to increase As^{III} sequestration capacity in the roots (Deng et al., 2018).

Glutaredoxin proteins have been shown to confer arsenic tolerance. Overexpressing either rice glutaredoxin genes *OsGrx_C7* or *OsGrx_C2.1* in Arabidopsis increased tolerance to As^V and As^{III} based on higher germination rates, root growth, and whole plant growth (Verma et al., 2016). Overexpression *OsGrx_C7* and *OsGrx_C2.1* transgenic plants contain more intracellular reduced glutathione (GSH) and less intracellular oxidized glutathione (GSSG) content after 1 week As^{III} or As^V exposure (Verma et al., 2016). Glutaredoxin proteins, therefore, might play a crucial role in maintaining GSH levels and redox status of cells under arsenic stress, which would be useful for transgenic crops needing to accumulate less arsenic or tolerate higher levels of arsenic.

Conclusion

In this thesis, we discuss how plants acquire, transport, store, and tolerate arsenic so future crops can be developed to contain less arsenic in their grain and edible tissues.

This thesis also supports phytoremediation efforts to genetically engineer plants capable of accumulating high levels of arsenic from arsenic contaminated soils and groundwater.

We first identified how natural variation in the expression of a nodulin 26-like intrinsic protein explains arsenic toxicity in *Arabidopsis*. We utilized a population of 527 *Arabidopsis thaliana* Multiparent Advanced Generation Inter-Cross (MAGIC) recombinant inbred lines (RILs) to identify *NIP1;1* as the major quantitative trait locus (QTL), located on chromosome 4, that controls As^{III} tolerance. We have also identified a second, minor QTL on chromosome 3 that we are currently investigating.

Additionally, we combined fluorescently activated cell sorting (FACS) with RNA-Seq technology to generate high resolution, genome-wide expression datasets for three plant root cell-types under As^{III} stress. Before this thesis work, transcriptional response data for individual plant root cell-types responding to As^{III} did not exist. This work provides the first cell-type specific expression map for every gene thought to be involved in As^{III} acquisition, efflux, sequestration, etc. for root epidermal, cortical, and endodermal cells. We also provide the first dataset that can be used for identifying genes with novel As^{III} detoxification functions specific to individual cell-types, whose unique expression profiles have been diluted in previous arsenic stress studies that derived transcriptional data from whole roots. Several Plasma-membrane Intrinsic Proteins (PIPs), for example, are highly expressed in all cell-types, but much higher in endodermal cells. In fact, the expression of these *PIPs* follow endodermal >> cortex > epidermal, the opposite pattern of *NIP1;1*, which is epidermal >> cortex > endodermal (**Figure 2.4A**). Collectively, these PIPs might perform the same function as NIP1;1, but in distinct root cells.

Strategically developing transgenic plants to modulate multiple processes involving arsenic uptake, efflux, sequestration, etc. is a priority for future research. Currently most transgene constructs are designed to alter arsenic accumulation by over-expressing a short list of model plant genes using ubiquitously active promoters. To more strategically design the low-arsenic-crop varieties and arsenic phytoremediators of the future, this thesis promotes conducting genomic research across a broader subset of plant ecotypes to help identify the best candidate genes available in nature. This thesis also demonstrates the benefits of utilizing innovative technology to enhance the magnification and resolution of gene expression data in order to more intimately understand plant response mechanisms like arsenic detoxification.

Chapter 2: Cell-type specific expression profiles in *Arabidopsis thaliana* roots in response to As^{III}

Abstract

Arsenic is a non-essential toxic metalloid that decreases crop yields and poses serious risks to human health. Rice grain contains the highest levels of arsenic among staple foods, potentially exposing billions of people to the negative effects of arsenic toxicity. Improved understanding of arsenic uptake from the soil, translocation, grain filling, and detoxification mechanisms in plants is critical for developing rice and other crop varieties with lower arsenic concentrations in the grain as well as engineering plants capable of arsenic phytoremediation on contaminated soils. Many questions remain regarding the genetic basis of these arsenic related processes and how different cell-types in the plant contributes to each mechanism in order to tolerate and detoxify arsenic. Here, we use Fluorescently Activated Cell Sorting (FACS) of *Arabidopsis thaliana* root cells, combined with RNA-Sequencing, to generate high-resolution transcriptome profiles for individual plant root cell types in response to arsenic stress. Our analysis showed that 7543 genes were differentially expressed across root epidermal, cortex, and endodermal cell-types in response to arsenite (As^{III}) stress. Many genes, including those already known to be involved with arsenic tolerance and detoxification, responded to As^{III} in cell-type specific patterns. *NIP1;1*, a Nodulin 26-like Intrinsic Protein known to transport As^{III} and several Plasma-membrane Intrinsic Proteins (PIPs) all show cell-type specific expression, but in opposite cell layers. *NIP1;1* expression follows epidermal >> cortex > endodermal, while *PIP* expression follows endodermal >> cortex > epidermal. Collective PIP function might therefore mimic *NIP1;1* function, but in distinct cells. We also found that between the two ATP-Binding Cassette C (ABCC) proteins most critical for As^{III} sequestration,

ABCC1 and ABCC2, ABCC2 is likely responsible for most As^{III} sequestration in these root cells based on higher ABCC2 expression across each cell-type. Ultimately, this work provides the first cell-type specific expression map for every gene thought to be involved in As^{III} acquisition, efflux, sequestration, etc. for root epidermal, cortical, and endodermal cells. This work also provides opportunities for further discovery of novel genes involved in arsenic tolerance.

Introduction

Arsenic is a highly toxic metalloid element found ubiquitously in the environment. In humans, arsenic exposure is associated with higher rates of cancer, diabetes, respiratory and cardiovascular diseases and therefore elevated levels of arsenic in drinking water and food pose a significant health risk (Abdul et al., 2015; Clemens and Ma, 2016). Several different organic and inorganic forms of arsenic can be present in the soil, but in flowering plants like *Arabidopsis thaliana* and rice (*Oryza sativa*), nearly all arsenic first enters root cells as inorganic arsenite (As^{III}) in anaerobic soils or as inorganic arsenate (As^{V}) in aerobic soils (Clemens and Ma, 2016). Among staple crops, rice accumulates the most arsenic due to polarly localized As^{III} /silicon (Si) transporters that direct the flow of $\text{As}^{\text{III}}/\text{Si}$ towards the xylem in the roots and because As^{III} concentrations in flooded rice paddies are higher than in aerated soils (Ma et al., 2008; Ma, 2010). Prolonged flooding depletes the soil of oxygen allowing anaerobic microbes to thrive. These microbes reduce $\text{As}^{\text{V}}\text{-Fe}^{\text{III}}$ complexes, releasing the arsenic as As^{III} which can then be taken up by plant roots (Islam et al., 2004; Williams et al., 2007; Xu et al., 2008; Zhu et al., 2014).

In aerobic soils, As^{V} enters root cells via phosphate (Pi) transporters. Members of the Phosphate transporter 1 family such as AtPht1;1, AtPht1;4 in *A. thaliana* and OsPT1, OsPT8 in rice are responsible for As^{V} uptake (Catarecha et al., 2007; Seo et al., 2008; Zhao et al., 2009; Sun et al., 2012; LeBlanc et al., 2013; Wang et al., 2016). Cytosolic As^{V} is almost entirely reduced to As^{III} by arsenate reductase enzymes in the roots, although As^{V} reduction can also occur in the shoots (Chao et al., 2014; Sanchez-Bermejo et al., 2014; Shi et al., 2016). In anaerobic soils, As^{III} enters via nodulin 26-like intrinsic proteins

(NIPs), but these passive channels act bidirectionally, allowing As^{III} to move in, out, and between root cells. As^{III} permeable NIPs in plants also transport substrates such as water, hydrogen peroxide, silicon, glycerol, boron, antimony, or aluminum (Kamiya and Fujiwara, 2009; Sasaki et al., 2016; Sadhukhan et al., 2017; Wang et al., 2017). Arabidopsis plants lacking AtNIP1;1 and rice plants lacking its closest rice homolog OsNIP2;1 (Lsi1), an important Si transporter, have increased tolerance to As^{III}, likely by reducing the amount of As^{III} that enters the roots (Kamiya et al., 2009). These mutant lines have lower As^{III} concentrations in the roots, shoots, and seeds (Ma et al., 2008). AtNIP1;2, AtNIP3;1, AtNIP5;1, AtNIP6;1, and AtNIP7;1 are also capable of bidirectional As^{III} transport as shown by heterologous expression in yeast and oocytes (Bienert et al., 2008; Kamiya et al., 2009; Xu et al., 2015), but their ability to transport As^{III} *in planta* is less clear compared to AtNIP1;1. OsNIP2;1/Lsi1 and Lsi2, a putative anion transporter unrelated to NIPs, facilitate Si transport towards the root vasculature in rice (Ma et al., 2007; Ma et al., 2008). Relative to the vasculature, Lsi1 protein is enriched on the distal plasma membrane (PM) of exodermal and endodermal cells, whereas Lsi2 protein is enriched on the proximal PM of the same cells. Together, these proteins coordinate the flux of Si and unfortunately As^{III}, towards the xylem and ultimately to the shoots.

Once As^{III} has entered a cell, it can either be complexed to glutathione/phytochelatin (GSH/PC) molecules and sequestered inside the vacuole via ABCC transporters, moved from cell to cell towards the vasculature (Tang et al., 2019), or effluxed back into the soil (Zhao et al., 2010; Chao et al., 2014). Rice OsABCC1 and Arabidopsis AtABCC1 and AtABCC2 are vacuole-localized transporters of heavy metal/metalloid-PC complexes As^{III}-PC₂, Cd^{II}-PC₂, and Hg^{II}-PC₂ (Song et al., 2010; Song

et al., 2014). While approximately 60% to 80% of arsenic that enters the roots is effluxed back to the external environment as As^{III}, the ABCC proteins are responsible for sequestering the remaining As^{III} in root vacuoles, preventing most As^{III} from reaching the shoots (Xu et al., 2007; Liu et al., 2010; Zhao et al., 2010).

The process of moving As^{III} from the roots to the shoots is not well understood. Natural Resistance-Associated Macrophage Proteins (NRAMPs) might be involved as Arabidopsis plants ubiquitously expressing GFP-tagged *OsNRAMP1* driven by the CaMV35S promoter accumulated more arsenic in roots and shoots with increased biomass compared to WT (Tiwari et al., 2014). The GFP-tagged protein only localized to endodermal and pericycle cells in the roots, suggesting post-transcriptional regulation of this NRAMP. The phenotype of these transgenic plants, together with the cell-type specific localization, suggests NRAMPs might have a role in loading As^{III} into the xylem. It should also be noted that yeast cells expressing *OsNRAMP1* showed a dramatic increase in cadmium (Cd) uptake, but relatively little difference in As^{III} uptake and As^{III} tolerance (Tiwari et al., 2014). Recently Arabidopsis inositol transporters AtINT2 and AtINT4 were shown to assist in loading As^{III} into phloem cells (Duan et al., 2016), opening the door for other sugar transporters to become candidates for novel As^{III} transport.

In only the last decade, research into how plants respond to arsenic has identified many genes important for arsenic transport and detoxification. Much of this work has been focused on plant roots responding to As^V exposure. The greatest source of dietary arsenic that threatens human health, however, comes from rice grown in As^{III} rich environments and many proteins contributing to known As^{III} detoxification pathways have yet to be identified. For example, 60% to 80% of the As^V entering root cells can be effluxed back

into the external environment as As^{III} (Xu et al., 2007; Liu et al., 2010; Zhao et al., 2010), and while bi-directional passive channels contribute, NIPs like OsLsi1 can only account for 15-20% of the total As^{III} that is effluxed (Zhao et al., 2010). WRKY6 and WRKY45 transcription factors regulate the expression of *Pht1;1* and thus As^V uptake in Arabidopsis (Castrillo et al., 2013; Wang et al., 2014). Phosphate Transporter Traffic Facilitator 1 (PHF1), an ER-localized protein, affects the localization of Pht1;1 transporters (Gonzalez et al., 2005). Apart from a recent study demonstrating an Arabidopsis calcium-dependent protein kinase, CPK31, might target and regulate NIP1;1 for As^{III} tolerance (Ji et al., 2017), little is known about transcriptional or translational regulation of As^{III} transporters during As^{III} stress.

Cell-type specific information for how plants tolerate arsenic is also lacking. To determine how As^{III} detoxification is partitioned by cell-type and highlight the gaps in our molecular models, we combined Fluorescently Activated Cell Sorting (FACS) with RNA-sequencing to generate the first high resolution, genome-wide expression datasets for *A. thaliana* root epidermal, cortex, and endodermal cells in response to As^{III} stress. Using this dataset, we identified differentially expressed genes by individual cell-type and organized them by expression pattern. Here we show how expression of arsenic-related genes like *NIP1;1* respond to As^{III} stress in obvious cell-type specific patterns within root tissue, whereas genes like *ABCC2* respond the same in each cell-type. By analyzing gene-family expression profiles, we reinforce previous work that argues, for example, *NIP1;1* is the single most important NIP impacting arsenic tolerance. We also present the first evidence that multiple members of the PIP family may collectively perform the same function as NIP1;1, but in different cell-types. The expression profiles of other major arsenic tolerance

genes like *PCS1*, supplemented with cell-type specific inductively coupled plasma mass spectrometry (ICP-MS) data, shows how detoxification processes like As^{III} sequestration might vary by cell-type. In contrast to previous reports, we found *ABCC2* and to a lesser extent *ABCC1*, are regulated by arsenic in root cells and that *ABCC2* likely plays a larger role than *ABCC1* in these three cell-types. Additionally, we demonstrate the limitations of typical gene-ontology (GO) enrichment analysis and how cell-type expression data like ours can improve this widely used technique. Finally, this work identifies candidate genes to examine in future studies which may contribute to overall plant tolerance to arsenic.

Results

FACS of Root Cell-Types and RNA-Sequencing

Although multiple studies have explored the transcriptional response to arsenic by various plant species (Abercrombie et al., 2008; Norton et al., 2008; Norton et al., 2008; Huang et al., 2012; Yu et al., 2012; Castrillo et al., 2013; Fu et al., 2014; Shukla et al., 2018; Kumar et al., 2019), it remains unclear which specific cells in the plant are responding to arsenic stress and how different cells coordinate to perform specific roles to best help the plant detoxify arsenic. To address the lack of cell-type specific transcription data for plant As^{III} detoxification, we generated high resolution expression profiles for three *A. thaliana* root cell layers. We used previously characterized GFP-marker lines to isolate root epidermal (WER::GFP), cortex (Cortex::GFP), and endodermal (SCR::GFP) cells (Sena et al., 2004; Dinneny et al., 2008) via fluorescently activated cell sorting (FACS) (Bargmann and Birnbaum, 2010) from plants grown under control and As^{III} stress

conditions. Root tissue from 8-day old plants that had been exposed to control or 50 μM As^{III} conditions for the previous 24 hours was harvested, followed by protoplast isolation which were sorted using FACS.

We generated 24 high quality RNA libraries used in this study. Three biological replicates for each of the three sorted cell-types and entire Col-0 roots, for both conditions. The read depths for the 24 single-end RNA libraries used in this study ranged from 14.20 to 42.15 million reads per sample (**Supplemental Table S2.1**). 75.8% to 88.2% of cleaned reads mapped to the *A. thaliana* genome (Araport11) using TopHat (Trapnell et al., 2012) with only 1.1% to 3.1% of reads aligning to multiple loci (**Supplemental Table S2.1**). To assess the relative expression of transcripts across the Arabidopsis genome, we generated Reads Per Kilobase of transcript per Million mapped reads (RPKM) units from normalized read counts for each gene in every sample using Cufflinks (Trapnell et al., 2012).

To check the biological reproducibility of our samples we performed principal component analysis (PCA) on sample transcriptomes and saw strong clustering of replicates across cell-types and conditions with cortex and endodermal (SCR) samples more closely grouping to each other than to epidermal (WER) samples (**Figure 2.1A**). 14% and 79% of the variance in the data is explained by the three different cell-types and two experimental conditions (0 vs. 50 μM As^{III}), respectively. Pairwise hierarchical clustering of sample transcriptomes also showed replicates grouping in a similar pattern, with endodermal (SCR) and cortex samples more tightly correlated to each other than with epidermal (WER) samples (**Figure 2.1B**).

Dinneny et al was able to identify many genes in ‘super clusters’ that exhibit cell-type enriched expression patterns for individual cell-types under control, NaCl stress, and

Fe deprivation conditions (Dinneny et al., 2008). To confirm the capacity of our FACS assay to identify cell-type specific gene expression, we compiled a list of genes similar to the Dinneny *et al* ‘super clusters’ which exhibited epidermal, cortex, and endodermal specific expression under their control conditions and asked whether these genes behave the same in our control condition. Our analysis showed the same enriched expression patterns as the genes in the super clusters identified in (Dinneny et al., 2008) under control conditions, including the expression of *At5g14750* (WER::GFP), *At1g09750* (Cortex::GFP), and *At3g54220* (SCR::GFP), which confirmed the cell-type specificity of our FACS assay (**Figure 2.1C**). The promoters of *At5g14750*, *At1g09750*, and *At3g54220* were used to drive the expression of GFP in the WER::GFP, Cortex::GFP, and SCR::GFP lines, respectively.

We also found expression of endogenous *At5g14750* and *At3g54220* were specific to epidermal and endodermal cell-types, respectively, in both control and As^{III} conditions, but expression of *At1g09750* was only specific to cortex cells in control conditions (**Supplemental Figure 2.2F**), which partly explains the lower % of sorted cortex cells with Cortex::GFP lines under our As^{III} conditions. Endogenous *At1g09750* was not expressed under our As^{III} conditions in any sample. However, GFP was still sufficiently present in cortex cells of Cortex::GFP lines after 24 hours As^{III} exposure to perform FACS (**Supplemental Figure 2.2C-E**). Negligible GFP signal was found in Cortex::GFP lines after 2 days As^{III} exposure. Future FACS assays using Cortex::GFP lines should therefore be limited to \leq 24 hours in As^{III} conditions.

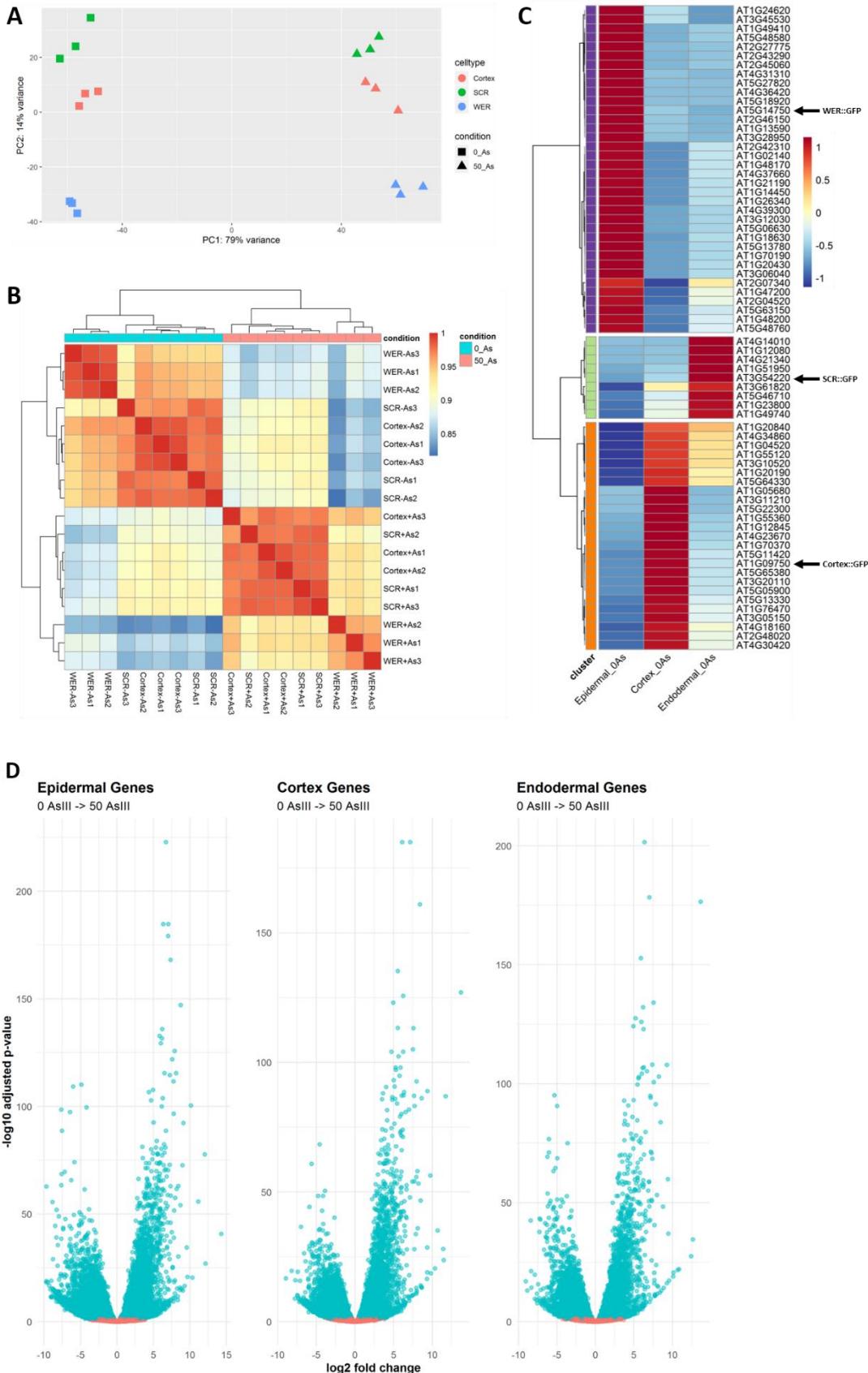


Figure 2.1: Unsupervised Clustering & Exploratory Data Analysis

(A) Principal Component Analysis (PCA) plot showing the samples clustering based on the two principal components, cell-type and condition, which explains 14% and 79% of the variation, respectively, in the RNA-seq data. (B) Hierarchical clustering heatmap showing similarities in global gene expression between both the biological replicates and the different sample groups. Perfect correlation = 1. (C) Heatmap of genes in control conditions that show cell-type specific expression in Dinneney et al. 2008 and this study. The promoters of At5g14750, At3g54220, and At1g09750 were used to drive expression of GFP in the WER::GFP, SCR::GFP, and Cortex::GFP marker lines. Color scale based on difference between RPKM of gene in cell-type and mean RPKM of gene in all samples, divided by the standard deviation $(x - \text{mean}(x)) / \text{sd}(x)$. (D) Volcano plots showing the fold changes relative to the adjusted p-values for all genes in epidermal, cortex, and endodermal cell-types. Blue dots are statistically significant differentially expressed genes $\text{padj} < 0.05$, red dots are not significant.

Cell-Type Specific Gene Expression Profiling

We performed differential expression analysis using the *DEseq2* R package (Love et al., 2014). A Generalized Linear Model (GLM) was fit to the normalized read counts with Wald tests to make pairwise comparisons of expression between the sample groups of our experimental conditions, for each cell-type. For example, epidermal DEGs were identified from contrasting gene counts in WER + 0As to WER + 50As samples. Using a significance threshold of adjusted p-value < 0.05 and an absolute value of log2 fold change ($\log FC$) > 1, we initially identified 10921 DEGs in the epidermal cell-type, 9336 DEGs in the cortex cell-type, and 8910 DEGs in the endodermal cell-type. To visualize the range of fold changes and adjusted p-values in our data we generated volcano plots for each cell-type, highlighting only significant DEGs (**Figure 2.1D**). MA plots were also generated for each cell-type to visualize the relationship between normalized read counts and $\log FC$ of every gene, highlighting the DEGs (**Supplemental Figure 2.1**).

We only considered DEGs with an $\text{abs}(\log FC) > 1.0$ in at least one cell-type and a $\text{RPKM} > 6$ in the As^{III} condition for up-regulated genes or $\text{RPKM} > 6$ in the control condition for down-regulated genes. We noticed some genes that have been implicated in arsenic tolerance have a maximum RPKM value under 10 in any one sample. For example, *ABCC1* has a maximum RPKM of 8 (epidermal cells, As^{III} condition) and *NIP5;1* has a maximum RPKM of 10 (epidermal cells, control condition). To keep these and similar arsenic tolerance genes in our analysis, we therefore set an $\text{RPKM} > 6$, which limited the number of DEGs to around 7500 for initial analysis. Using these filters, we found 7543 total DEGs; 3219 up-regulated and 3326 down-regulated epidermal DEGs, 3116 up-

regulated and 2933 down-regulated cortex DEGs, 2752 up-regulated and 3333 down-regulated endodermal DEGs (**Figure 2.2**).

We separated the 7543 DEGs into 21 groups by Venn diagram considering only if $\text{abs}(\text{logFC}) > 1$ in each cell-type (**Figure 2.2B, Supplemental Table S2.2**). These results show most genes were differentially expressed in multiple cell-types. 1859 of 3219 (58%) up-regulated genes were up-regulated in all 3 cell-types and 1676 of 3326 (50%) down-regulated genes were down-regulated in all 3 cell-types. 669, 247, and 306 up-regulated genes were specific to epidermal, cortex, and endodermal cells, respectively. 885, 314, and 841 down-regulated genes were specific to epidermal, cortex, and endodermal cells, respectively.

To find genes “differentially expressed” in multiple cell-types which meet our logFC and RPKM thresholds, yet still respond “specifically” (expression responds to As^{III} substantially more in one cell-type) by cell-type, we also used the “ward.D2” hierarchical clustering method within the *pheatmap* R package to group the 7543 DEGs into 8 clusters based on the RPKM values across all samples (**Figure 2.2A, Supplemental Table S2.3**). With some exceptions, clusters 1-4 contain genes up-regulated in one or more cell-types, while clusters 5-8 contain genes down-regulated in one or more cell-types. Cluster 1 contains genes predominantly up-regulated in epidermal cells. Cluster 2 is less specific and contains genes up-regulated in two or three different cell-types. Cluster 3 contains genes up-regulated in multiple cell-types, but highest in cortex cells. Cluster 4 contains genes up-regulated in cortex and endodermal cells, but not epidermal cells. Cluster 5 contains genes predominantly down regulated in epidermal cells. Cluster 6 contains genes predominantly down regulated in endodermal cells. Cluster 7 contains genes

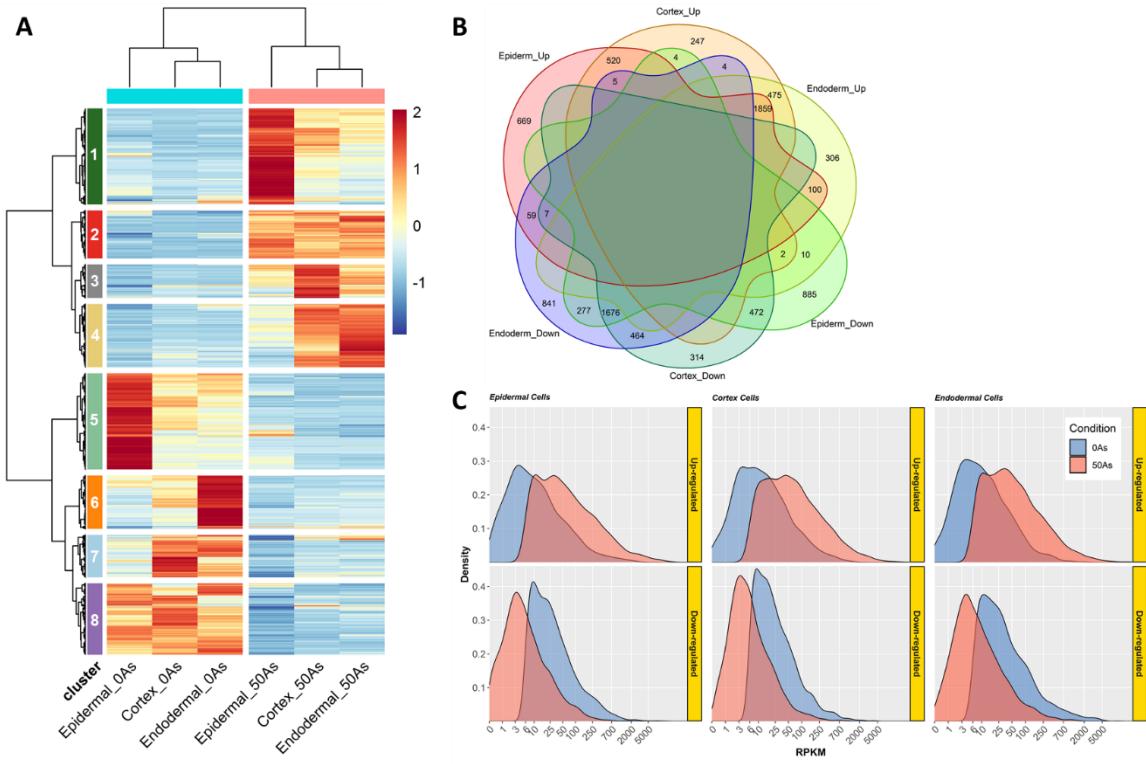


Figure 2.2: Differentially Expressed Genes by Cell-Type

(A) Hierarchical clustering of 7543 significant differentially expressed genes with $\text{abs}(\log\text{FC}) > 1$ and $\text{RPKM} > 6$ in at least one cell-type of any condition. 8 clusters can be seen showing relative DEG expression in each cell-type and condition. Color scale the difference between RPKM of gene in cell-type and mean RPKM of gene in all samples, divided by the standard deviation $(x - \text{mean}(x)) / \text{sd}(x)$. (B) Venn diagram of same 7543 DEGs differentially expressed in unique or combinations of cell-types. (C) Density plots of the up (bottom row) and down (top row) DEG RPKM values in epidermal, cortex, or endodermal cells in either control (blue) or As^{III} (red) conditions.

predominantly down regulated in cortex and endodermal cells, but not epidermal cells. Cluster 8 is less specific and contains genes down-regulated in two or three different cell-types.

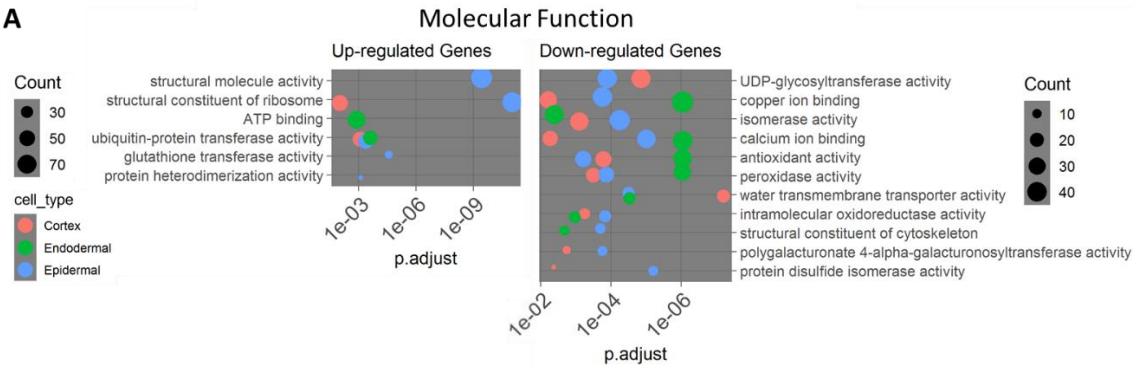
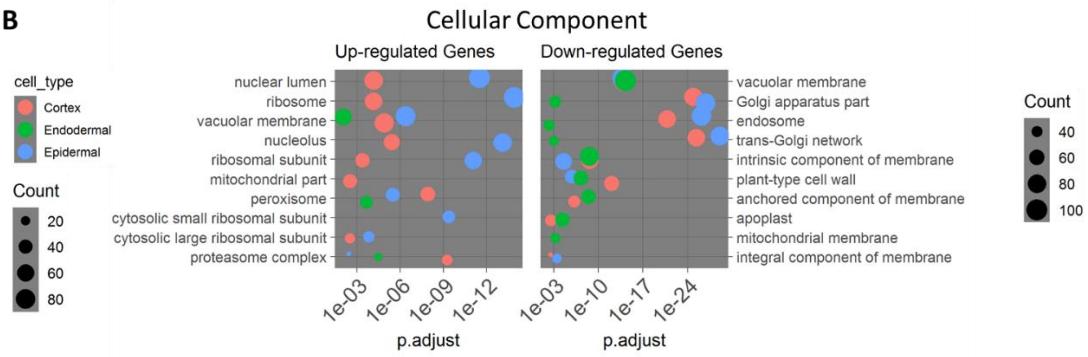
To better understand how As^{III} affects overall DEG expression in our assay and to visualize the underlying probability distribution of DEG RPKM values, we produced RPKM density plots for each cell-type and condition (**Figure 2.2C**). Overall, the densities of RPKM values for up-regulated DEGs were similar across cell-types and the densities of RPKM values for down-regulated DEGs were similar across cell-types. For up-regulated DEGs (top plots) in all cell-types under control conditions (blue), RPKM values have a normal distribution concentrated at a peak near 3-10 RPKM and rarely exceed 250 (**Figure 2.2C, Supplemental Table S2.4**). Under As^{III} conditions (red), RPKM values for these up-regulated DEGs show a binomial, right-skewed distribution, concentrated at peaks near 10 and 40 RPKM, with maximum values near 5000 RPKM and approximately 50% of up-regulated DEGs having > 50 RPKM in all cell-types. RPKM values for down regulated DEGs (bottom plots) in all cell-types show a slightly right skewed distribution concentrated at a peak near 10 RPKM in control conditions (blue) and rarely exceed 700 (**Figure 2.2C, Supplemental Table S2.5**). Under As^{III} conditions (red), RPKM values for these down regulated DEGs show a normal distribution concentrated near the peak at 0-6 RPKM, with maximum values around 700 RPKM and > 90% of down-regulated DEGs having < 50 RPKM in all cell-types.

Gene-Ontology Enrichment Analysis

To capture the most physiologically relevant genes for our gene-ontology (GO) analysis, we filtered the DEGs using $\text{abs}(\log\text{FC}) > 1.3$ and either a $\text{RPKM} > 8$ in the As^{III} condition for up-regulated genes or $\text{RPKM} > 8$ in control condition for down-regulated genes (**Supplementary Table S2.6**). These parameters resulted in 2000-2500 DEGs per cell-type. GO analysis does not narrow the focus to individual genes, so setting more stringent filters to reduce the number of DEGs allows us to focus on the GO terms of the genes most affected by As^{III} . Each cell-type responds to As^{III} uniquely by altering the expression of different genes annotated with Biological Processes (BP), Cellular Components (CC), and Molecular Functions (MF) GO terms. To visualize which GO terms were enriched in each subontology for each cell-type, we generated enrichment plots in a way to easily compare GO terms across cell-types (**Figure 2.3, Supplemental Figure 2.4**). Some terms have identical enrichment across cell-types and thus show a tight cluster of 3 overlapping data points, such as “*vacuolar membrane*” (Down-regulated Genes) in **Figure 2.3B**. Other terms like “*endosome*” (Down-regulated Genes) show a clear difference between epidermal, cortex, and endodermal enriched DEG data points.

We also filtered clearly redundant terms. For example, enrichment statistics for CC terms “*vacuolar membrane*” and “*vacuolar part*” were identical across all cell-types, so “*vacuolar part*” was eliminated in **Figure 2.3B**. We retained the redundant terms in the complete enrichment analysis found in **Supplemental Figure 2.4**.

Epidermal cells appear to emphasize protein production over protein degradation compared to cortex and endodermal cells. Up-regulated DEGs annotated with “*ribosome*” related MF and CC terms were most enriched in epidermal cells while “*proteasome*”

A**B**

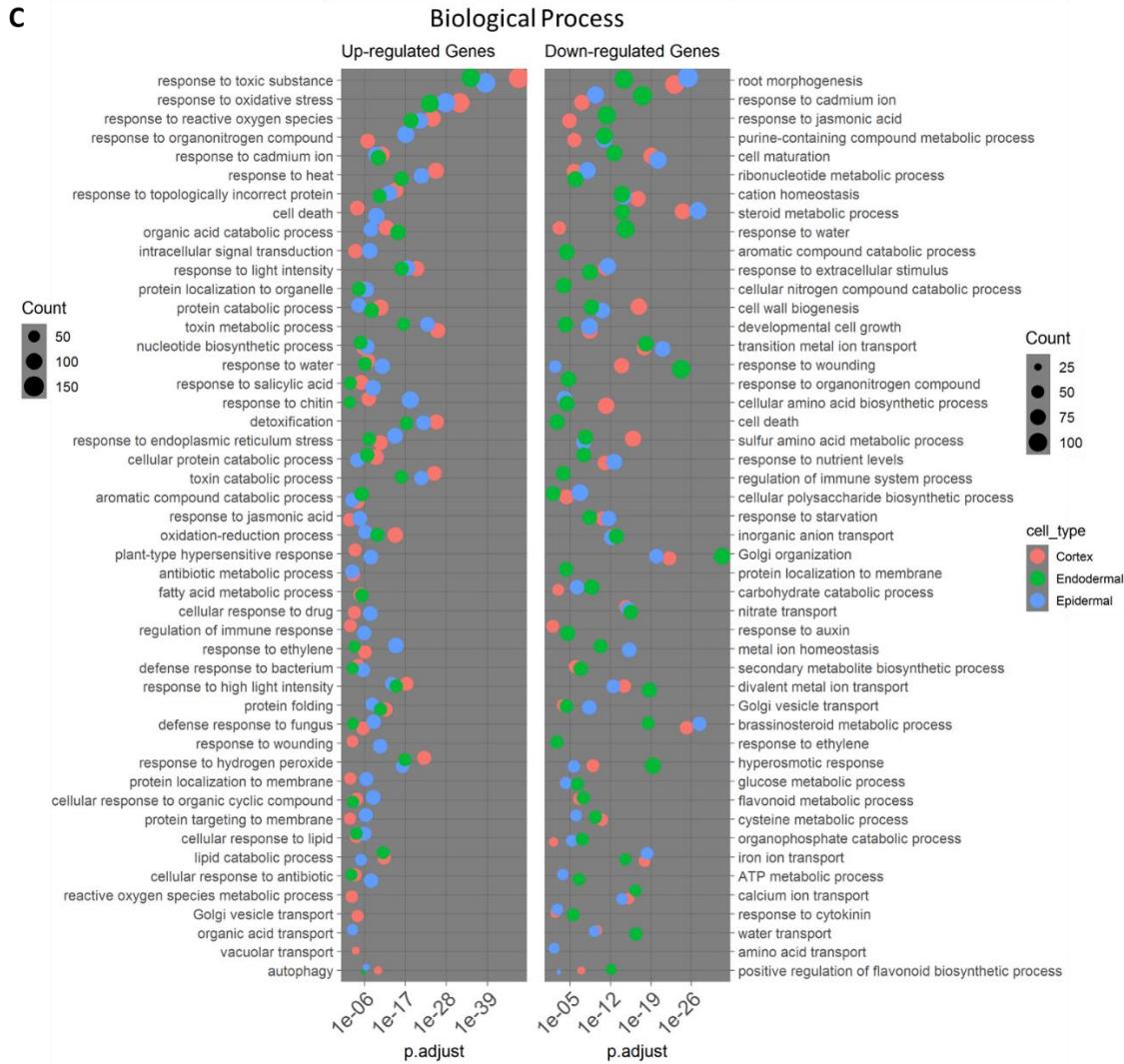


Figure 2.3: GO-Term Enrichment

Molecular Function (A), Cellular Component (B), and Biological Process (C) subontology terms enriched in response to As^{III}. Terms for each subontology arranged by adjusted p-values for up and down differentially expressed genes. Size corresponds to the number of genes (count) enriched for that term for each cell-type. Differentially expressed genes were filtered to include only $\text{abs}(\text{logFC}) > 1.3$, and RPKM > 8 in As^{III} condition for up-regulated genes or RPKM > 8 in control condition for down-regulated genes. Cell-types separated by color.

related MF and CC terms were least enriched in epidermal cells (**Figure 2.3A-B**, **Supplemental Figure 2.4A-B**).

The epidermis and cortex (to a lesser extent) strongly suppress vesicle trafficking. The CC terms related to “*vesicle*”, “*Golgi*”, and “*endosome*” are among the most significantly enriched terms for down-regulated genes in epidermal and cortex cell-types, with p.adjust values ranging from $1*10^{-20}$ to $1*10^{-30}$ (**Figure 2.3B**, **Supplemental Figure 2.4B**). These terms are enriched in down-regulated endodermal genes but are each near the p.adjust value cutoff for significance of 0.01. Instead, endodermal cells down-regulate genes with CC terms related to mitochondria.

Only the endodermal cells show enrichment of down-regulated genes involving “*mitochondria*” and “*respiratory chain*” related CC terms (**Figure 2.3B**, **Supplemental Figure 2.4B**).

Upregulated BP terms most enriched in cortex cells DEGs after As^{III} exposure include “*response to toxic substance*”, “*response to oxidative stress*”, “*response to reactive oxygen species*”, “*toxin metabolic process*”, “*detoxification*”, “*toxin catabolic process*”, “*response to hydrogen peroxide*”, and “*autophagy*” (**Figure 2.3C**, **Supplemental Figure 2.4C**). Down-regulated BP terms most significantly enriched in cortex cells include “*sulfur*” and “*cysteine*” metabolic process related terms (**Figure 2.3C**, **Supplemental Figure 2.4C**).

Cortex cells also down-regulate the most genes with “*water transmembrane transporter activity*” and “*water channel activity*” MF terms (**Figure 2.5A**, **Supplemental Figure 2.4A**).

Expression Profiles of Arsenic Detoxification Genes & Candidates

Six of nine NIP subfamily members in *A. thaliana* have been shown to be capable of bi-directional As^{III} transport; *NIP1;1*, *NIP1;2*, *NIP3;1*, *NIP5;1*, *NIP6;1*, and *NIP7;1* (Bienert et al., 2008; Isayenkov and Maathuis, 2008; Kamiya et al., 2009; Xu et al., 2015). In every sample *NIP6;1* and *NIP7;1* are not expressed (RPKM ≤ 2) while *NIP1;2*, *NIP3;1*, and *NIP5;1* have very low expression (RPKM ≤ 10) (**Figure 2.4A**). Under control conditions *NIP1;1* is principally expressed in the epidermis with RPKM of 2201, 10-fold higher than the cortex or endodermis. After As^{III} exposure, *NIP1;1* is down-regulated by log2FC -3.3/-2.7/-1.6 in epidermal/cortex/endodermal cells.

NIPs belong to a larger Major Intrinsic Proteins (MIP) superfamily of aquaporins (Kapilan et al., 2018; Bezerra-Neto et al., 2019) that also include Plasma membrane Intrinsic Proteins (PIPs), Tonoplast Intrinsic Proteins (TIPs), Small basic Intrinsic Proteins (SIPs) and X Intrinsic Proteins (XIPs). *PIP2B* shows extreme downregulation in response to As^{III} in all cell-types, with logFCs of -5.6/-4.7/-5.5 in epidermal/cortex/endodermal cells, respectively (**Figure 2.4A**). The RPKM and logFC of *PIP2B* is very similar to *NIP1;1* except the cell-type specific expression moves the opposite way, with *NIP1;1* being most highly expressed in the epidermis, then the cortex, and lowest in the endodermis. *PIP3* is not differentially expressed by the (abs)logFC > 1 threshold, but its RPKM value increases by hundreds in each cell-type after As^{III} exposure, from RPKMs of 834/1813/3150 to 1256/3617/4042 in epidermal/cortex/endodermal cells.

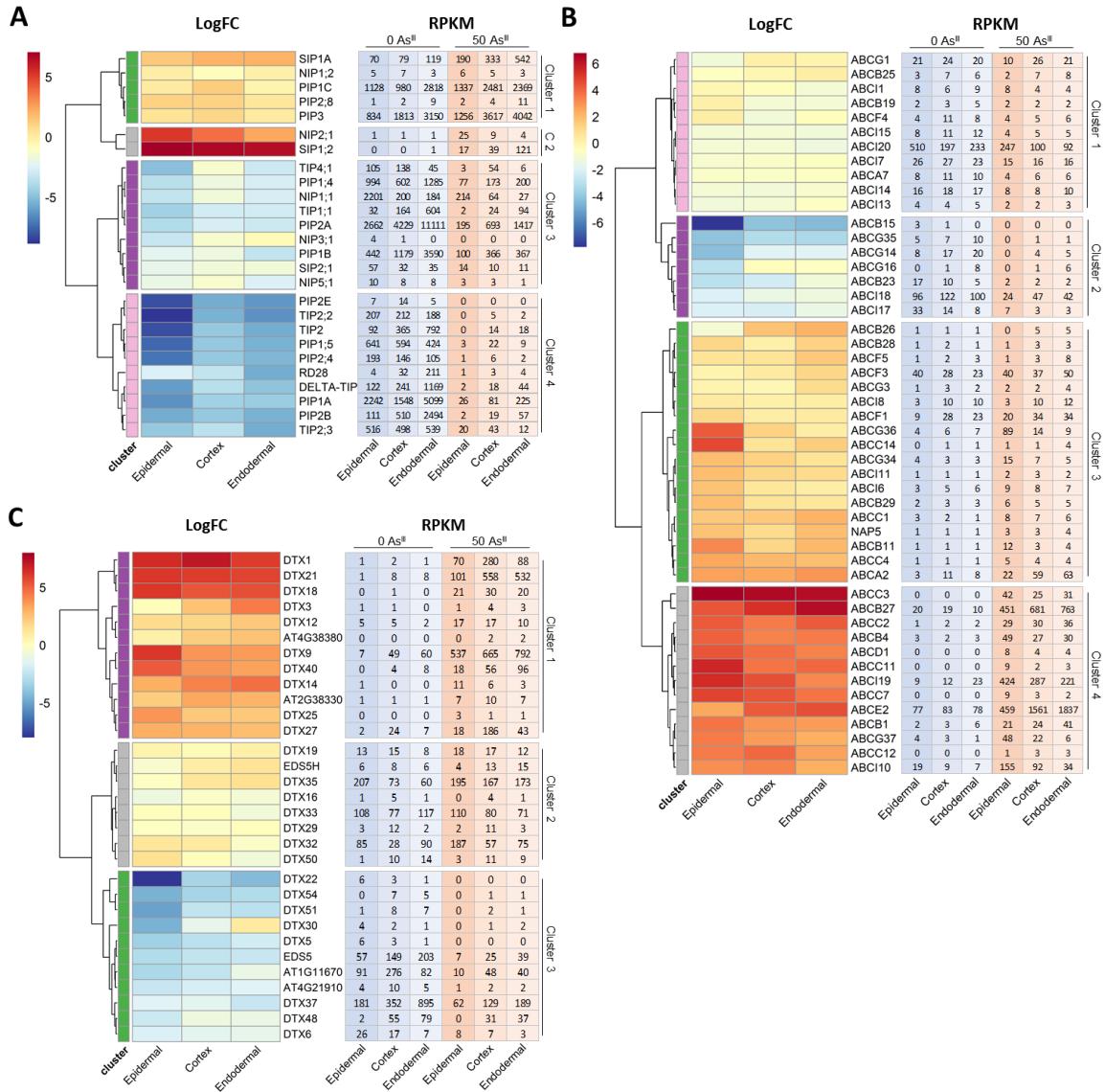


Figure 2.4: As^{III} Induced Expression by Cell-Type

Expression heatmaps showing LogFC by cell-type in response to As^{III} on left and the corresponding RPKM units arranged by cell-type and condition for each gene on right. Genes from the Major Intrinsic Proteins (MIP) superfamily of aquaporins (A), ATP-Binding Cassette (ABC) family (B), and Multidrug and Toxic Compound Extrusion (MATE) family (C). Genes that lacked a RPKM value of > 2 in at least one sample were filtered out. Heatmaps clustered by logFC.

ATP-binding cassette (ABC) proteins are known to help detoxify a wide range of toxic compounds in both prokaryotes and eukaryotes, including As^{III}. Clusters 3 and 4 of **Figure 2.4B** contain ABC genes induced by As^{III} in one or more cell-types. *ABCC1* is only slightly up-regulated under As^{III} stress in each cell-type, with logFCs of 1.6/1.6/2.1 in epidermal/cortex/endodermal cells, respectively. Under our As^{III} conditions, *ABCC1* has RPKMs of only 8/7/6 for epidermal/cortex/endodermal cells. *ABCC2*, however, is up-regulated under As^{III} stress in each cell-type, with a logFC of 4.3/3.6/4.3 and RPKMs of 29/30/36 in epidermal/cortex/endodermal cells. *ABCC3* showed the largest up-regulation of all the ABCCs, with logFCs of 6.8/6.4/6.3 and RPKMs of 42/25/31 in epidermal/cortex/endodermal cells.

No members of the ABCAs, ABCBs, ABCDs, ABCEs, ABCGs, or ABCIs subfamilies have yet been implicated in As^{III} detoxification, but some of these genes show strong induction in response to As^{III} in a cell-type specific pattern in this study (**Figure 2.4B**). *ABCB27* is strongly induced in all cell-types, but least in epidermal cells and most in endodermal cells, increasing RPKMs from 10-20 to 451/681/763 in epidermal/cortex/endodermal cells (**Figure 2.4B**). *ABCE2* is induced in all cell-types, but dramatically in cortex and endodermal cells, increasing RPKMs from 77-83 to 459/1561/1837 in epidermal/cortex/endodermal cells (**Figure 2.4B**). *ABCG36* (*At1g59870*) is induced by As^{III} specifically in epidermal cells, increasing RPKMs from 4-7 to 89/14/9 in epidermal/cortex/endodermal cells (**Figure 2.4B**). *ABCG37* shows a similar induction pattern to *ABCG36*, but less drastic, increasing RPKMs from 1-4 to 48/22/6 in epidermal/cortex/endodermal cells (**Figure 2.4B**). *ABCI19* is strongly induced in all cell-types, but particularly in epidermal cells, increasing RPKMs from 9-23 to 424/287/221 in

epidermal/cortex/endodermal cells (**Figure 2.4B**). *ABCI10* shows a similar induction pattern to *ABCI19*, but less drastic, increasing RPKMs from 7-19 to 155/92/34 in epidermal/cortex/endodermal cells (**Figure 2.4B**).

Multiple members of the Multidrug and Toxic Compound Extrusion (MATE) protein family, found in clusters 1 & 2 of **Figure 2.4C**, respond strongly to As^{III} and represent potential candidates for As^{III} transport. *DTX9* and *DTX21* are highly induced by As^{III} in each cell-type, but particularly in cortex and endodermal cells. *DTX9* RPKMs increase from 7/49/60 to 537/665/792 in epidermal/cortex/endodermal cells, respectively (**Figure 2.4C**). *DTX21* RPKMs increase from 1/8/8 to 101/558/532 in epidermal/cortex/endodermal cells (**Figure 2.4C**). *DTX1* and *DTX27* are also induced by As^{III} in each cell-type, but mostly in cortex cells. *DTX1* RPKMs increase from 1-2 to 70/280/88 in epidermal/cortex/endodermal cells (**Figure 2.4C**). *DTX27* RPKMs increase from 2-24 to 18/186/43 in epidermal/cortex/endodermal cells (**Figure 2.4C**). *DTX18*, *DTX32*, *DTX35*, and *DTX40* show different As^{III} induced, cell-type specific expression patterns, but not as dramatically as *DTX9*, *DTX21*, *DTX1*, or *DTX27*.

WRKY6 and WRKY45 transcription factors have been shown to regulate the expression of *Pht1;1* to limit As^V uptake in Arabidopsis under As^V stress. WRKY6 represses *Pht1;1* under As^V stress (Castrillo et al., 2013), while WRKY45 activates *Pht1;1* expression during Pi starvation (Wang et al., 2014). Under the As^{III} conditions in this study, WRKY6 is induced in each cell-type, but most strongly in cortex cells, with RPKMs increasing from 19/75/54 to 45/162/110 in epidermal/cortex/endodermal cells (**Supplemental Figure 2.6A**). WRKY45 expression, in contrast, remains low (< 11 RPKM) in every sample (**Supplemental Figure 2.6A**). Like As^V conditions, As^{III} also completely

reduces expression of *Pht1;1* and *Pht1;4*, another As^V/Pi transporter gene (Castrillo et al., 2013). *Pht1;1* RPKMs decrease from 23/24/14 to 1 in epidermal/cortex/endodermal cells (**Supplemental Figure 2.6C**). *Pht1;4* RPKMs decrease from 15/8/6 to 0 in epidermal/cortex/endodermal cells (**Supplemental Figure 2.6C**).

A. thaliana contains two phytochelatin synthase genes, *PCS1* (*At5g44070*) and *PCS2* (*At1g03980*). PCS enzymes synthesize PC peptides in the form of (γ -GluCys)_n-Gly ($n = 2-11$) by polymerizing glutathione (GSH, γ -GluCysGly) (Zenk, 1996; Cobbett and Goldsbrough, 2002). The cysteine thiol groups found within GSH and PCs have a high affinity for metal binding, especially for cadmium (Cd²⁺) and As^{III} (Schmoger et al., 2000). *PCS2* nearly unexpressed (RPKM ≤ 1) in every cell-type and condition. *PCS1*, however, is highly expressed in each cell type under both control and As^{III} conditions (**Supplemental Figure 2.6C**). *PCS1* shows constitutive expression in cortex/endodermal cells with RPKMs of 128/97 in control conditions and 144/94 in As^{III} conditions. In epidermal cells, however, *PCS1* is up-regulated during As^{III} stress, with a logFC of 1.2 and RPKMs increasing from 107 in control conditions to 243 in As^{III} conditions.

HAC1 (*At2g21045*), which codes for an enzyme that reduces As^V to As^{III}, is slightly down-regulated in only two cell-types, from RPKM of 1055 to 784 in epidermal cells and RPKM of 99 to 56 in endodermal cells (**Supplemental Figure 2.6C**). In cortex cells however, *HAC1* expression nearly doubles, from RPKM of 125 to 248.

The inositol transporter genes *AtINT2* and *AtINT4* which regulate loading As^{III} into phloem cells are unexpressed in these three cell-types, but other sugar transport genes like *STP7* (Sugar Transport Protein 7, *At4g02050*), *STP13* (*At5g26340*), and *NAT6* (*At5g62890*) are induced in each cell-type and show cell-type specific expression patterns

in response to As^{III}. *STP7* As^{III} induced expression pattern follows epidermal >> cortex > endodermal with RPKMs increasing from 64/17/10 to 434/144/86 in epidermal/cortex/endodermal cells (**Supplemental Figure 2.6C**). *STP13* As^{III} induced expression pattern follows endodermal >> cortex >> epidermal with RPKMs increasing from 5/60/39 to 34/192/376 (**Supplemental Figure 2.6C**). *NAT6* As^{III} induced expression pattern follows endodermal = cortex >> epidermal with RPKMs increasing from 7/24/31 to 67/215/221 (**Supplemental Figure 2.6C**).

Discussion

Arsenic found in soils used to grow crops poses a threat to human health when the toxic molecule accumulates in the developing grains of plants that humans consume in high volumes, such as rice. For arsenic to enter the above-ground tissues of these crops it must first bypass plant detoxification mechanisms in the roots which have evolved to keep arsenic below ground. Although many studies on arsenic in plants have been conducted and some important genes involved in arsenic transport have been revealed, many genes responsible for known detoxification processes, like non-aquaporin mediated As^{III} efflux out of the root cells, for example, have still not been identified. Our field has also yet to identify transcription factor or post-translational regulatory genes that regulate the cell's response to As^{III} stress specifically. Additionally, the transcriptome data we have from these studies has mostly been generated from tissue level genomic assays.

Overall, our RNA-seq analysis identified 7543 significant differentially expressed genes with $\text{abs}(\log FC) > 1$ and $\text{RPKM} > 6$ in at least one cell-type of any condition. Using

these thresholds, our Venn diagram showed 58% of up-regulated DEGs were up-regulated in every cell-type and 50% of down-regulated genes were down-regulated in every cell-type (**Figure 2.2B**). Using Ward's minimum variance method (ward.d2) we were able to show many of these DEGs that are up or down-regulated in every cell-type can still be considered specific to one or two cell-types (**Figure 2.2A**). For example, clusters 1 and 5 of **Figure 2.2A** contain genes with $\text{abs}(\log FC) > 1$ in cortex and endodermal cells, yet most of these DEGs are clearly most responsive to As^{III} stress in epidermal cells. It appears epidermal cells are most responsive to As^{III} stress, in both the number of up and down DEGs. Cluster 6 of **Figure 2.2A** also shows a large group of down-regulated DEGs specific to endodermal cells, yet we find only a small subset of up-regulated DEGs in cluster 4 is specific to endodermal cells. The small number of down-regulated DEGs specific to cortex cells is striking considering clusters 5 and 6 contain many genes specific to epidermal and endodermal cells, respectively. We also found the distribution of DEG RPKM values very similar across cell-types and conditions (**Figure 2.2C**). In all three cell-types, only As^{III} conditions show a binomial, right-skewed distribution of RPKM values for up-regulated DEGs with peak densities near 10 and 40 RPKM. We found it interesting that while most DEGs are found in clusters specific to one or two cell-types (clusters 1, 3, 4, 5, 6, 7), the overall transcriptional response of every cell-type (genome-wide changes to RPKMs) appears highly similar.

Many RNA-Seq analyses have included GO term enrichment, but often the results are limited by poor genome annotations. **Supplemental Figure 2.7** shows violin plots highlighting some limitations of GO enrichment analysis in *A. thaliana*. Terms with low gene counts (< 10) like “*eRF1 methyltransferase complex*” are too specific and terms like

“*molecular function*” with extremely high gene counts are too vague. Overly specific or vague terms create noise which effect enrichment statistics. Many terms such as “*response to cadmium ion*” are over-represented (Gene Count = 313) while many terms like “*mitotic cell cycle*” are under-represented (Gene Count = 55). The *A. thaliana* genome contains approximately 27,463 genes (UniProt) and it is reasonable to assume 313 genes are not dedicated to responding to cadmium stress while only 55 are dedicated to mitosis. There are also many redundant terms such as “*peroxidase activity*” and “*oxidoreductase activity, acting on peroxide as acceptor*” (**Supplemental Figure 2.4A**) which should be consolidated to one term. As illustrated in **Supplemental Figure 2.7**, many GO terms are either over-represented or under-represented in the genome annotation. Many terms are also either too broad or too specific in their description of a biological process, cellular component, or molecular function. Improved gene annotation is needed for this type of data to produce objective, impartial conclusions. Ranking enriched GO terms will have a larger role in future genomic studies, but only as gene annotation improves. Contrasting enriched GO terms between samples/conditions provides useful information, even with flawed annotation. To further identify cell-type specific expression patterns, we performed GO analysis contrasting terms enriched in each of our three root cell-types, after separating up and down-regulated DEGs, in response to As^{III} (**Figure 2.3, Supplemental Figure 2.4**).

GO term enrichment suggests cortex cells are the most stressed of the three cell-types. We had previously performed ICP-MS to measure the arsenic content of each cell-type under our experimental conditions and found cortex cells retained much lower concentrations of arsenic than epidermal or endodermal cells (**Supplemental Figure 2.5**). The ICP-MS results initially seemed to conflict with the cortex upregulated BP enrichment,

especially when considering the three cell-types have similar expression profiles for *ABCC1/ABCC2* and *PCS1* is more highly expressed in cortex than endodermal cells (**Figure 2.4B, Supplemental Figure 2.6C**). Downregulated BP terms enriched most significantly in cortex cells, such as “*sulfur amino acid metabolic process*” and “*cysteine metabolic process*”, however, indicate there might be lower PCS1 substrate available in cortex cells, leading to lower rates of As^{III} sequestration into the vacuole and higher cytosolic As^{III}. Cortex cells, therefore, might be less equipped to store As^{III} in the vacuole than other cell-types and might depend more on effluxing cytosolic As^{III} to neighboring cells via plasma membrane transporters. Although cortex cells down-regulate the most genes with “*water transmembrane transporter activity*” and “*water channel activity*” MF terms, cortex cells most often have the lowest abs(logFC) values for aquaporin genes downregulated under As^{III} stress (**Figure 2.4A**). After As^{III} exposure, *NIP1;1* is reduced by log2FC -3.32 in the epidermis, -2.67 in the endodermis, but only -1.59 in the cortex, which may contribute to higher As^{III} influx into cortex cells relative to other cell-types. An insufficient reduction of NIP1;1 and lower available phytochelatin could explain why cortex cells experience the most stress from As^{III}.

It should also be noted that among NIPs and PIPs with relevant expression levels, that in these three cell-types, most NIPs are generally expressed highest in the epidermis and lowest in the endodermis, while most PIPs are generally expressed highest in the endodermis and lowest in the epidermis. Our expression data suggests As^{III} transport by NIP family aquaporins in these root cells might depend solely on NIP1;1 (**Figure 2.4A**), although its role is different for each cell-type. It is possible other MIP family members are involved in As^{III} detoxification in *A. thaliana*, but to our knowledge no other *A. thaliana*

MIPs have been shown experimentally to transport As^{III}. However, heterologous expression of three rice PIP aquaporins (*OsPIP2;7*, *OsPIP2;4*, *OsPIP2;6*) in *X. laevis* oocytes resulted in 2- to 3-fold increase of influxed As^{III} compared to controls (Mosa et al., 2012), with *OsPIP2;7* showing the largest increase of As^{III}. Overexpression of these rice PIPs in *A. thaliana* increased the plants ability to transport As^{III} into and out of roots over short time intervals of 1-4 hours (Mosa et al., 2012). Over a longer four-day exposure to As^{III} these transgenic *A. thaliana* plants had increased root length and shoot biomass, which suggests the PIP aquaporins increase As^{III} tolerance and efflux in the long run. A BLAST search of these PIP protein sequences in *A. thaliana* shows *OsPIP2;7* is closest to *AthPIP3* (*At4g35100*), *OsPIP2;4* closest to *AtPIP2B* (*At2g37170*), and *OsPIP2;6* closest to *AthPIP2;8* (*At2g16850*). In this study we can see these three *Arabidopsis* PIPs share a similar expression profile, in that they all are most highly expressed in endodermal cells, followed by cortex cells, and least expressed in epidermal cells. They differ greatly, however, in their RPKM levels and response to As^{III} (**Figure 2.4A**).

PIP2;8 has low, yet stable expression, with RPKM ≤ 11 in all cell-types before and after As^{III} exposure and therefore likely has a minimal impact on As^{III} transport at best in these root cells. *PIP3* might be the best candidate for mimicking the function of the three OsPIPs above. The large increase of *PIP3* RPKM with a cell-type specific pattern suggests it may play a unique role in As^{III}, like the OsPIPs (**Figure 2.4A**). To our knowledge there is no cell-type specific expression data for the OsPIPs in rice roots under arsenic conditions.

Plants contain many ABC transporters, with around 130 ABC genes present in the *Arabidopsis* and rice genomes (Hwang et al., 2016; Lefevre and Boutry, 2018). The ABCE, ABCF, and ABCI subfamily proteins lack transmembrane domains and are all

thought to be cytosolic and not membrane bound transport proteins (Hwang et al., 2016). The ABCC subfamily contains tonoplast localized proteins critical for As^{III} sequestration into the vacuole (Shoji, 2014). Compared to expressed of other transporter families like the MATEs or MIPs, which in this study can have RPKMs in the hundreds to thousands, ABCC gene expression is more modest, never surpassing RPKM of 50 in this study.

AtABCC1 and AtABCC2 in *A. thaliana* and OsABCC1 in rice have been previously characterized to transport phytochelatin bound As^{III} (PC-As^{III}) into root vacuoles, a critical step in preventing arsenic from reaching the shoots (Song et al., 2010; Song et al., 2014). The prior studies analyzed expression of the ABCC genes from whole roots and determined *AtABCC1* and *AtABCC2* were not induced by 24 hour exposure to 110 µM As^V but were constitutively expressed (Song et al., 2010). *OsABCC1* expression, however, was shown to be induced by 24 hour exposure to 5 µM As^{III} (Song et al., 2014). This study shows that *AtABCC1* and *AtABCC2* expression, like *OsABCC1*, is regulated by arsenic and that ABCC2 is likely responsible for most As^{III} sequestration in these three cell-types (**Figure 2.4B**). Under our As^{III} conditions, *ABCC1* expression increases from RPKMs of 3/2/1 to only 8/7/6 in epidermal, cortex, and endodermal cells, which is surprisingly low considering the gene's importance. ABCC1 therefore might be more critical for As^{III} sequestration in vascular or shoot cell-types. *ABCC2* expression, in contrast, increases from RPKMs of 1/2/2 to 29/30/36 in epidermal, cortex, and endodermal cells. When considering AtABCC2 consistently transported PC-As^{III} at higher rates than AtABCC1 in multiple transgenic yeast assays (Song et al., 2010), along with *ABCC2* being expressed 3-6 times higher than *ABCC1* under As^{III} stress in this study, we find it more

likely these three root cell-types depend on ABCC2 more than ABCC1 for PC-As^{III} sequestration.

ABCC3 has been shown to transport PC-Cd into root vacuoles (Brunetti et al., 2015), but not PC-As, as far as we know. The PC-Cd transport ability, together with a highest ABCC expression in response to As^{III} in this study (**Figure 2.4B**) makes ABCC3 a great candidate for As^{III} transport and should be revisited in future assays.

This study also shows how genes regulated by As^V are expressed under As^{III} stress (**Supplemental Figure 2.6C**). *PHT1;1* and *PHT1;4* expression responds as expected, being completely down-regulated in all cell-types. Other As^V regulated genes show a more interesting expression profile. *HAC1*, which codes for an arsenate reductase enzyme that reduces As^V to As^{III}, is a critical step in As^V detoxification. Using whole *A. thaliana* roots, (Chao et al., 2014) showed *HAC1* expression is highly up-regulated under 3d of 50-300 μM As^V stress and down-regulated under 3d of 10-100 μM As^{III} stress. In this study we find *HAC1* expression is slightly down-regulated in epidermal and endodermal cells, but nearly doubles in cortex cells under our As^{III} conditions, although the overall expression pattern still follows epidermal >> cortex >> endodermal (**Supplemental Figure 2.6C**). Oxidation of As^{III} to As^V has not been shown in plants and cortex cells being the only cell-type exposed to As^V under our experimental conditions seems unlikely. Why cortex cells would induce *HAC1* when surrounding cells down-regulate the gene is puzzling.

As mentioned above, *PCS1* induction is specific to epidermal cells. Expression of *PCS1* in cortex and endodermal cells is nearly identical in both experimental conditions (**Supplemental Figure 2.6C**). This likely contributes to more As^{III} being sequestered in

epidermal than cortex and endodermal cells, consistent with our ICP-MS data in **Supplemental Figure 2.5**).

Inositol transporters AtINT2 and AtINT4, which assist in loading As^{III} into phloem cells (Duan et al., 2016), were not expressed in our three root cell-types under either conditions. *STP7*, *STP13*, and *NAT6* are sugar transporters that were highly induced by As^{III} and could be candidates for novel roles in As^{III} transport in future studies (**Supplemental Figure 2.6C**). Tiwari et al., 2014 has implicated NRAMPs as candidates for As^{III} transport, expressing *OsNRAMP1* in *A. thaliana* under control of CaMV35S promoter. These transgenic lines exhibited increased root length and seedling fresh weight compared to WT when grown 10-20 d on 5 µM and 10 µM As^{III}. No NRAMPs were differentially expressed in our study. A role for NRAMPs in arsenic tolerance in these three cell-types is therefore unlikely.

MATE family proteins have long been considered candidates for As^{III} transport. AtDTX1 was shown to increase cadmium resistance in *E. coli* (Li et al., 2002), and rice MATE proteins have some ability to modulate arsenic accumulation (Tiwari et al., 2014; Das et al., 2018), but no data exists showing any plant MATE protein directly capable of arsenic transport in yeast, oocytes, or in-vivo plant assays. Expression of *DTX1*, *DTX9*, *DTX21*, and *DTX27* are all highly up-regulated by As^{III} and are the best candidates for arsenic transport of the family (**Figure 2.4C**). *DTX18*, *DTX32*, and *DTX40* are also induced to a lesser degree and should be considered in future arsenic studies.

Our results reveal that As^{III} detoxification in plant roots is regulated at a cell-type specific level. Epidermal, cortex, and endodermal root cells in *A. thaliana* respond uniquely to promote detoxification pathways for the toxin. Previous work has identified

many genes important for As^{III} tolerance in root tissue and this study helps clarify how those genes contribute to specific detoxification mechanisms within three different root cell-types. This work also identifies candidate genes for future assays which could contribute to As^{III} tolerance specifically in epidermal, cortex, or endodermal root cells.

We need to develop healthier rice and other crop varieties that contain lower arsenic levels. We also need to develop plants that better absorb and tolerate arsenic for phytoremediation of contaminated soils. By enhancing As^{III} regulated gene expression data, this study will contribute to designing genetically engineered plants/crops which can modulate processes like arsenic uptake, efflux, or sequestration without sacrificing overall plant fitness.

Materials and Methods

Plant materials and growth conditions

Seeds for WT and GFP-reporter lines were obtained with help from the Salt Lab at The University of Aberdeen (Aberdeen, UK). The seeds were surface sterilized with a bleach and ethanol solution and allowed to dry for 3-4 hours. 2 rows of ~200 sterile seeds per row were sown onto 8 x 8 cm sterile 80-micron nylon mesh (32% open, Component Supply, /U-CMN-80B-C) which had been placed on top of solidified standard medium composed of half-strength Murashige-Skoog (MS) containing 1% sucrose and 1% agar (Type M, Sigma-Aldrich) and adjusted to pH 5.7 with KOH. Plated seeds were then stratified at 4°C in the dark for 2 days before being placed vertically in a light chamber for 7 days under standard conditions (16h of light at 22°C, 8h of dark at 18°C, 37 $\mu\text{mol m}^{-2} \text{s}^{-1}$

photon flux density. After 1 week, the mesh holding the plants was transferred onto either new standard media for control samples or standard media supplemented with 50 μ M sodium (meta)arsenite (As^{III}) (Sigma-Aldrich), then returned to the light chamber for 24 hours.

Fluorescently Activated Cell Sorting

Root tissue was harvested from 8-day old plants that had been exposed to control or As^{III} conditions for 24 hours and immediately placed in a 15 mL conical tube containing 12 mL of protoplasting solution (prepared according to (Bargmann and Birnbaum, 2010)) and rotated at room temperature for one hour. The root protoplasts were then quickly filtered through a 40 μm nylon cell strainer (Corning) using ~20 mL additional protoplast solution lacking enzymes (Cellulase and Macerozyme, Yakult) into a 50 mL conical tube and centrifuged for 10 minutes at 500 G. The supernatant was mostly removed, leaving ~ 1 mL to gently resuspend the pellet with a 7.5 mL plastic transfer pipette.

Resuspended protoplasts were immediately brought to a FACSaria (BD) cell sorter, with specifications and gating parameters similar to (Bargmann and Birnbaum, 2010). Up to 500 μL of GFP-positive cells gated into P5 (**Sup Figure 1**) were sorted directly into 2.5 mL Eppendorf tubes containing 1.5 mL cold TRIzol LS Reagent and then kept on ice for 1-2 hours before RNA extraction.

After protoplasting, WER::GFP roots produced 15-20% GFP positive epidermal cells (~80-85% non GFP cells, **Supplemental Figure 2.3B-C**), SCR::GFP roots produced

8-10% GFP positive endodermal cells (**Supplemental Figure 2.3F-G**), and Cortex::GFP roots produced only 2-4% GFP positive cortex cells (**Supplemental Figure 2.3D-E**).

RNA Extraction and Sequencing

RNA was isolated from sorted cells using the TRIzol LS Reagent User Guide protocol (Thermo Fisher). Total RNA from WT Co-0 un-sorted roots was extracted using the E.Z.N.A. Plant RNA Kit (OMEGA Bio-tek, Norcross, GA). DNase treatment was performed with the On-Column DNase I Digestion Set (Sigma). Quality and concentration of all RNA samples was analyzed using the Agilent Eukaryote Total RNA Nano Series II kit and 2100 Bioanalyzer System (**Sup Figure 2**). Samples scoring RINs > 7.5 were sent to Bob Schmitz at the University of Georgia, Department of Genetics, for sequencing. RNA sequencing was performed using the Illumina Nextseq500 system with 75bp single-read libraries. The read depth of our sequenced libraries ranged from 15 to 42 million reads per sample. TopHat was used to map read counts onto the TAIR11 genome while Cufflinks generated RPKM expression units. We used the DEseq2 R package to normalize read counts and perform differential expression analysis.

GFP-reporter lines used for cell sorting

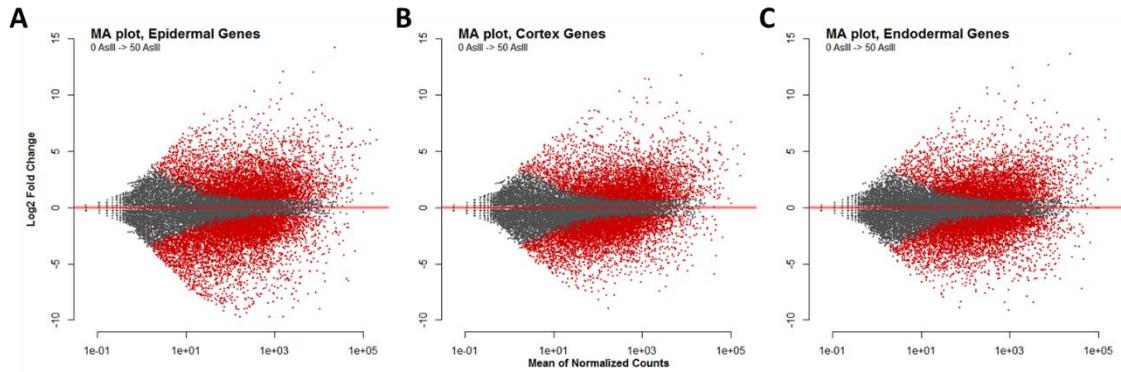
RNA profiling of the root epidermis, cortex, and endodermis was performed by using lines containing promoter-GFP constructs from the genes At5g14750 (WER::GFP),

At1g09750 (Cortex::GFP), and At3g54220 (SCR::GFP), respectively (Sena et al., 2004; Lee et al., 2006; Levesque et al., 2006; Dinneny et al., 2008).

Gene Ontology Enrichment Analysis

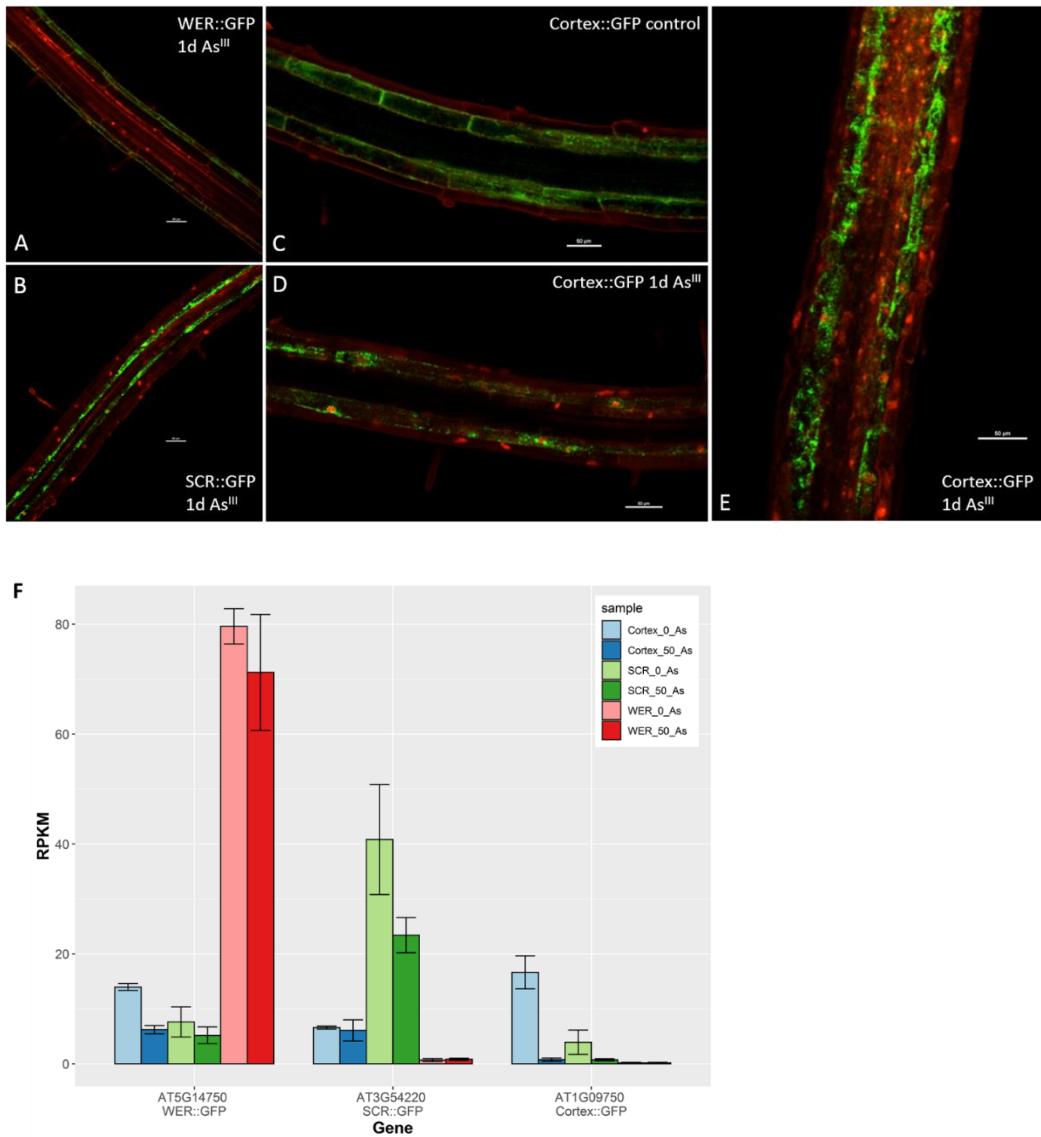
Enrichment of gene ontology categories biological process (BP), cellular component (CC), and molecular function (MF) in our data sets was determined using the “org.At.tair.db”, “DOSE”, “clusterProfiler”, and “ggplot2” R packages. The “org.At.tair.db” genome wide annotation for Arabidopsis contained a bug that required using another database (“org.Hs.eg.db”) to switch between subontology categories “MB”, “MF”, and “CC”. Prior to subontology analysis, differentially expressed gene sets were separated by up or down regulation and cell-type, then filtered. To capture physiologically relevant processes, up regulated gene sets were filtered to only include genes with a logFC of > 1.3 and a RPKM value in the $50 \mu\text{M As}^{\text{III}}$ condition > 8 . Down regulated gene sets were filtered to only include genes with a logFC of < -1.3 and a RPKM value in the $0 \mu\text{M As}^{\text{III}}$ control condition > 8 . A significance threshold for CC and MF term enrichment results was set at p-values and q-values $< 1\text{e-}02$. The significance threshold for BP term enrichment results was set at p-values and q-values $< 1\text{e-}08$.

Supplemental Data



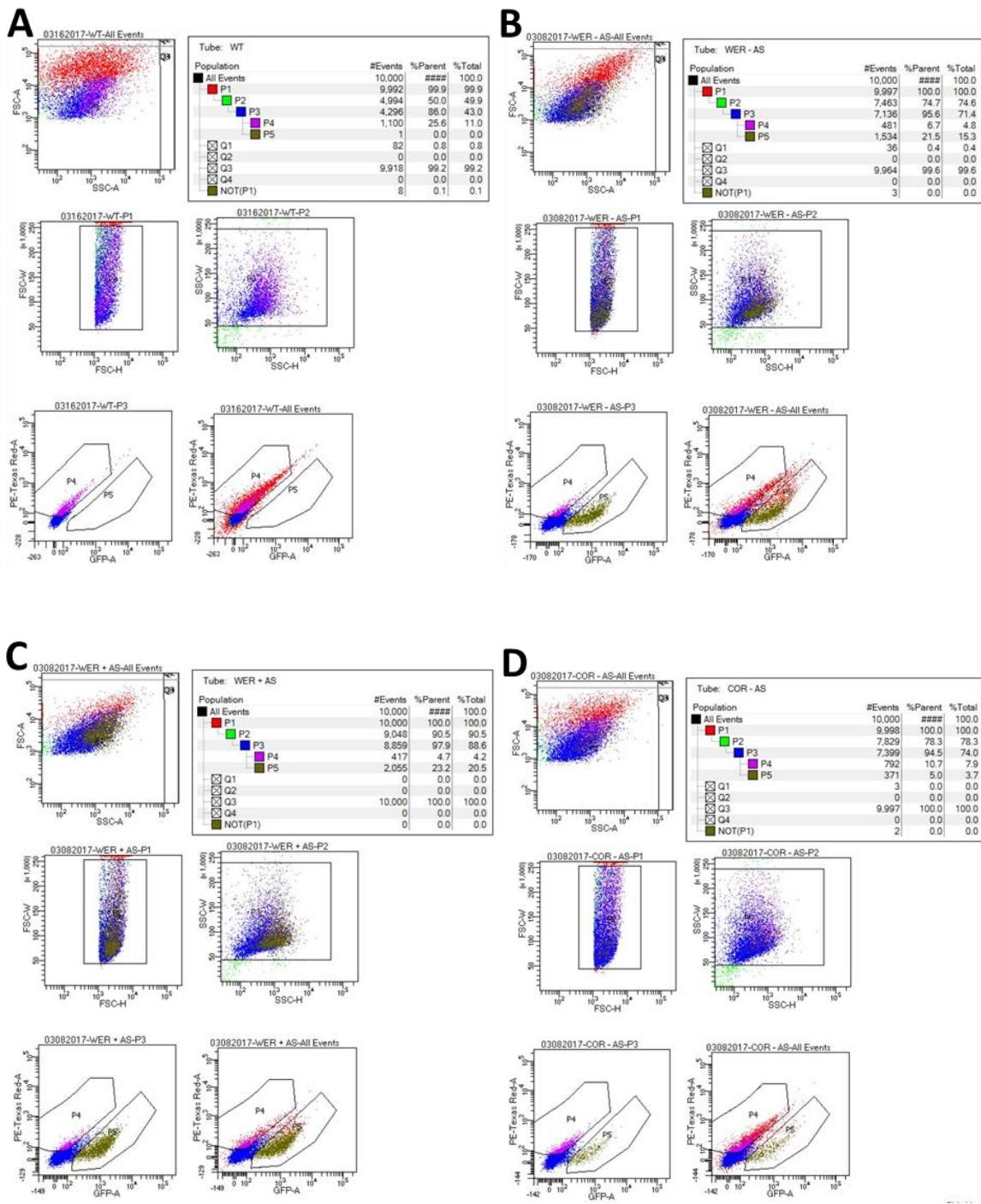
Supplemental Figure 2.1: MA Plots for Differentially Expressed Genes

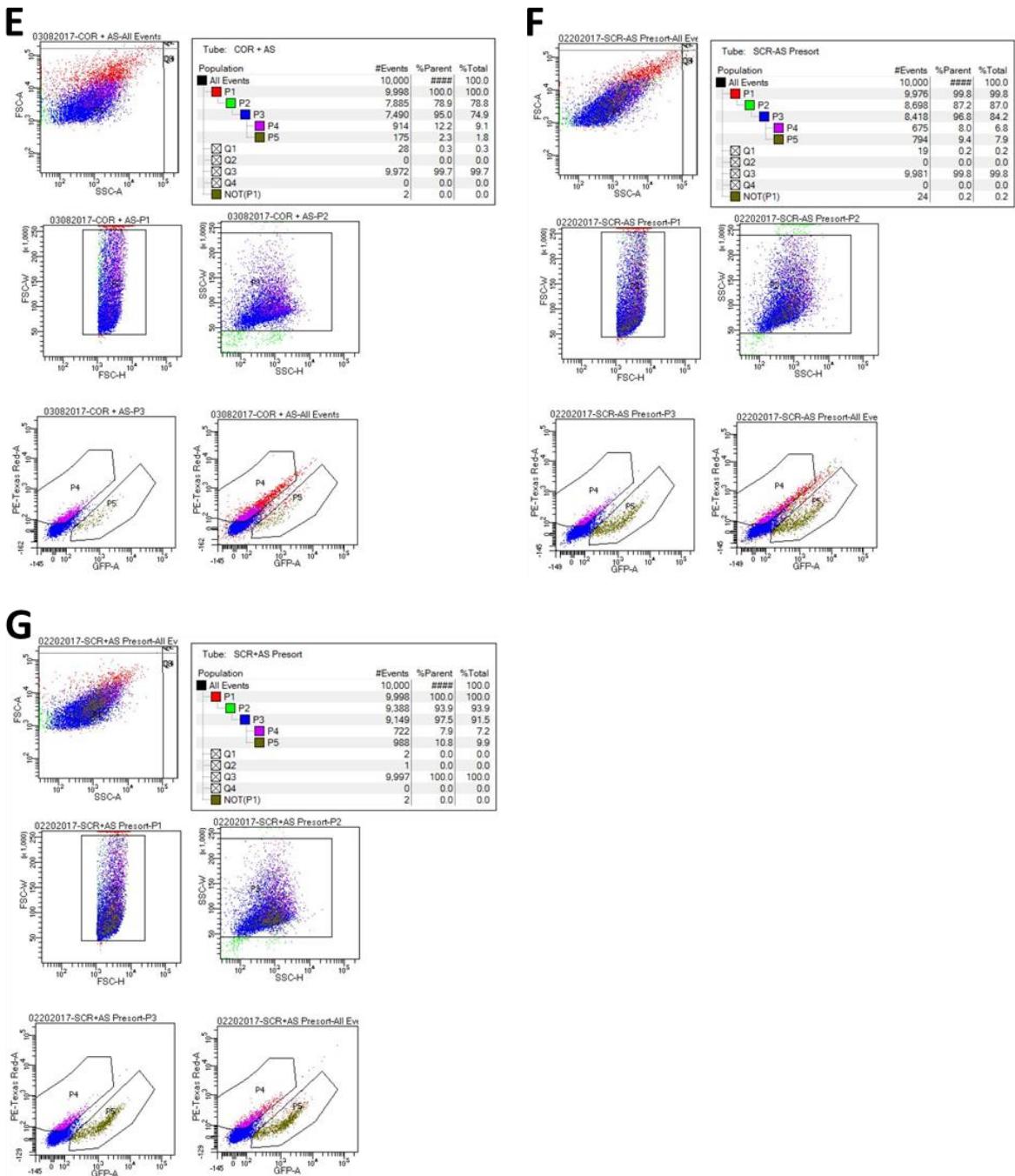
MA plots (log fold-change versus mean of normalized read counts) showing statistically significant, differentially expressed genes (red dots) in epidermal (A), cortex (B), and endodermal (C) cell-types.



Supplemental Figure 2.2: GFP-Reporter Lines

7d old GFP reporter lines after 24 hours exposure to 50 μM As^{III} show GFP fluorescence exclusive to epidermal cells (A) in WER::GFP, to endodermal cells (B) in SCR::GFP, and to cortex cells (D-E) in Cortex::GFP. GFP fluorescence of Cortex::GFP line in control conditions (C). Scale in each image is 50 μm . (F) Bar plot of RPKM units for the three GFP-reporter line genes in both experimental conditions. Promoters of *AT5G14750*, *AT3G54220*, *AT1G09750* were cloned for WER::GFP, SCR::GFP, and Cortex::GFP, respectively. Error bars represent standard error from n=3 biological replicates.

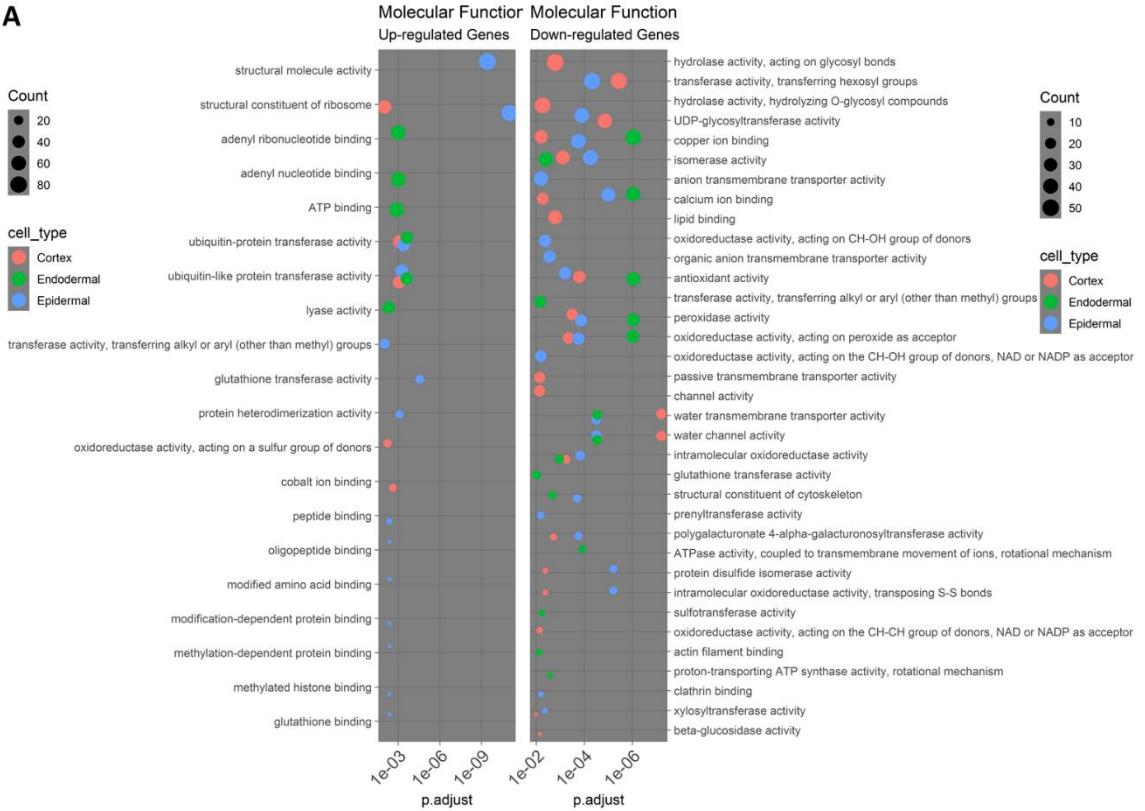
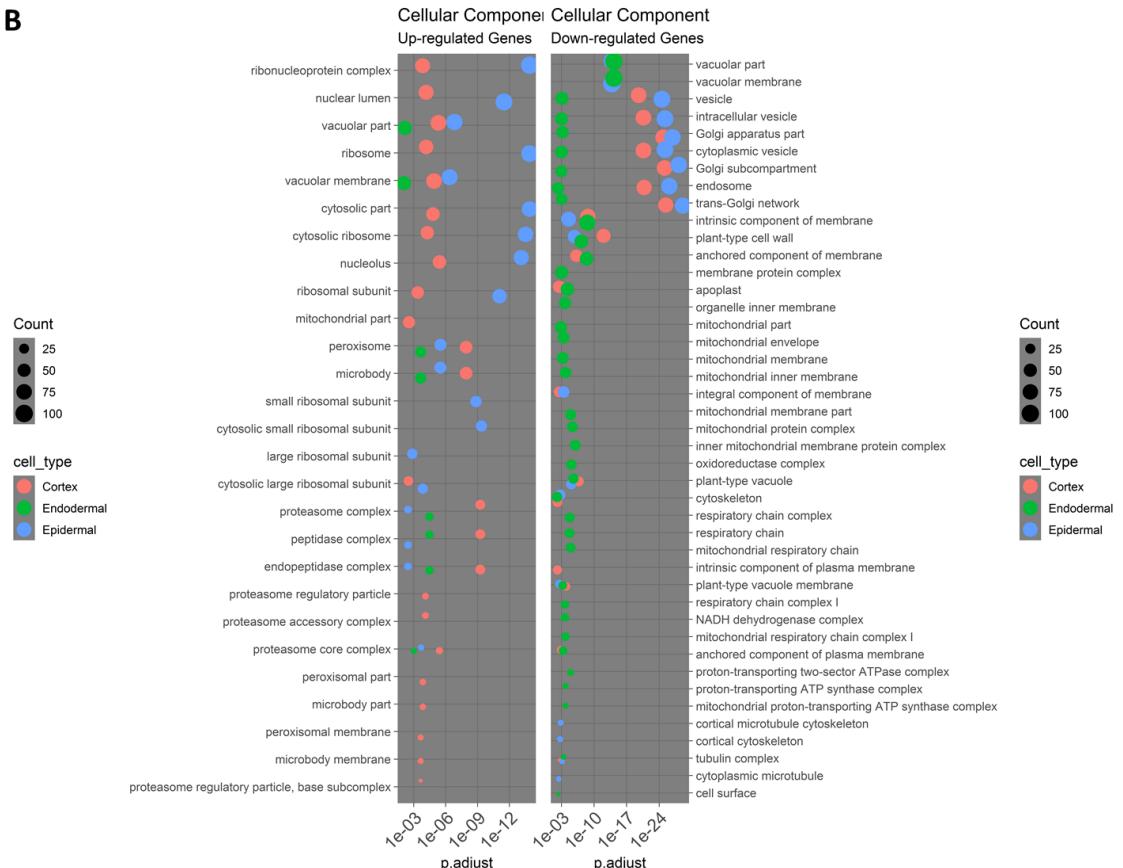




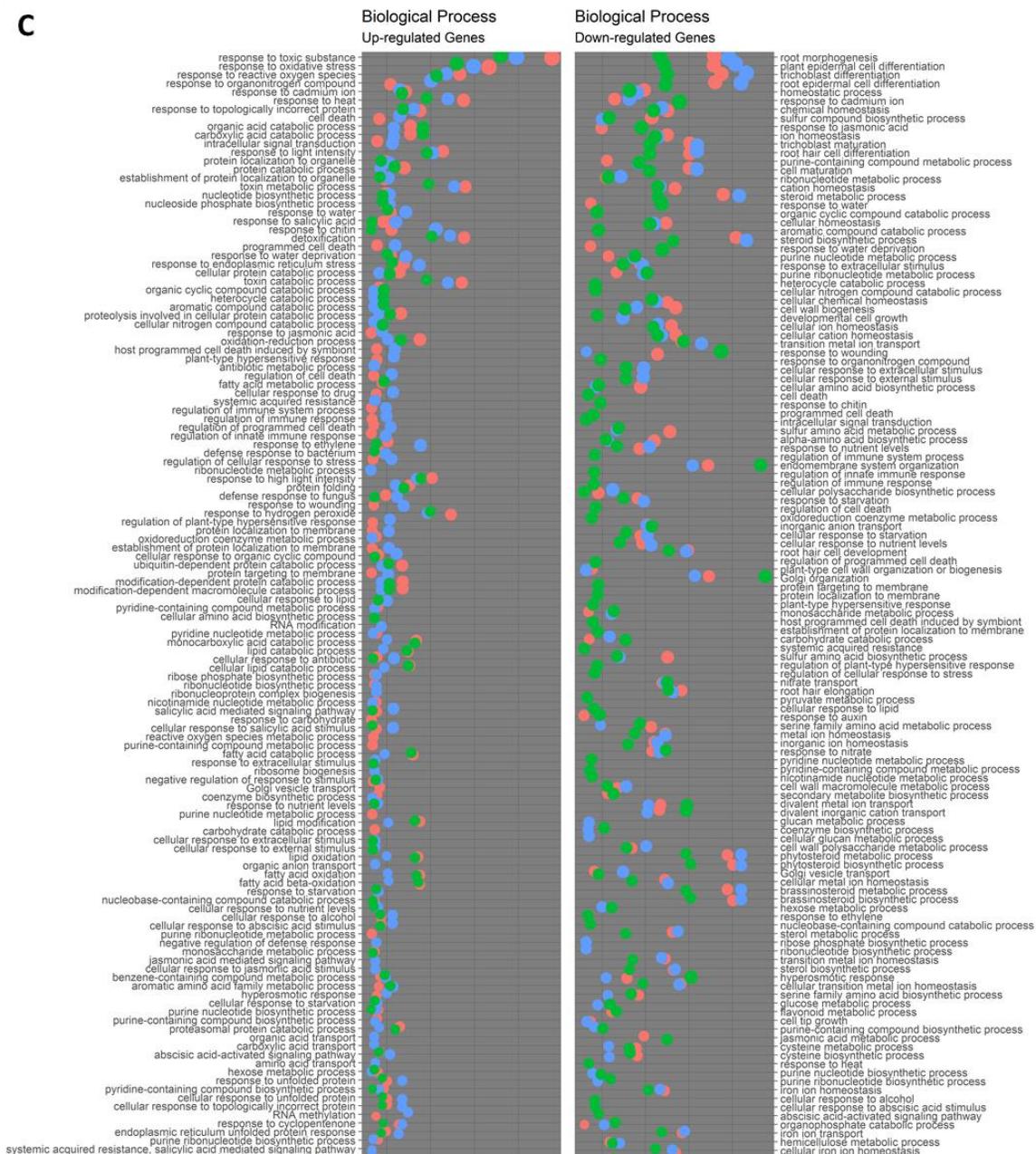
Supplemental Figure 2.3: GFP-marker lines used for FACS

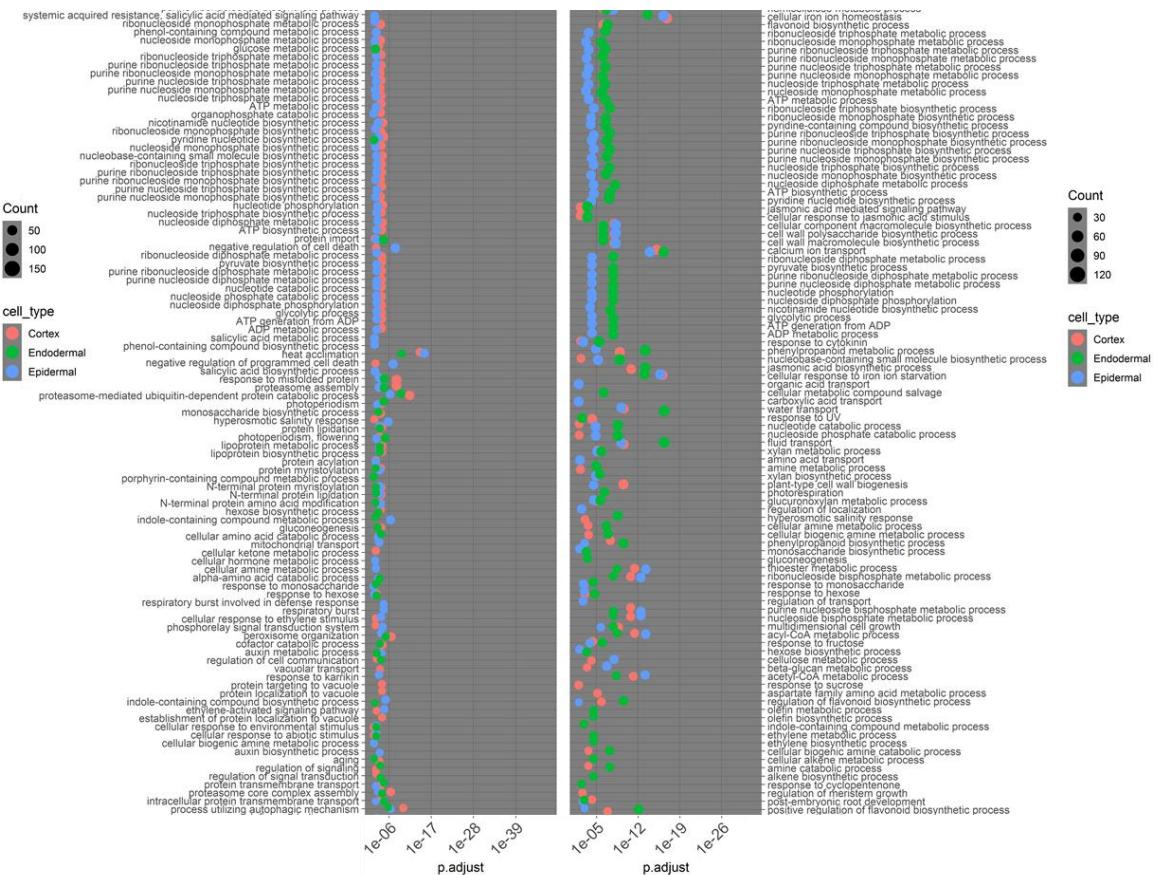
Examples of complete gating parameters and sorting statistics used in FACS for 7d old WT and GFP-reporter lines after 24 hours exposure to control or 50 μ M As^{III} conditions then root protoplasting. Protoplasts filtered through gates P1-3 for size and shape that fell within

gate P5 were collected as GFP positive events. WT control protoplasts (A), WER::GFP control protoplasts (B), WER::GFP As^{III} protoplasts (C), Cortex::GFP control protoplasts (D), Cortex::GFP As^{III} protoplasts (E), SCR::GFP control protoplasts (F), SCR::GFP As^{III} protoplasts (G).

A**B**

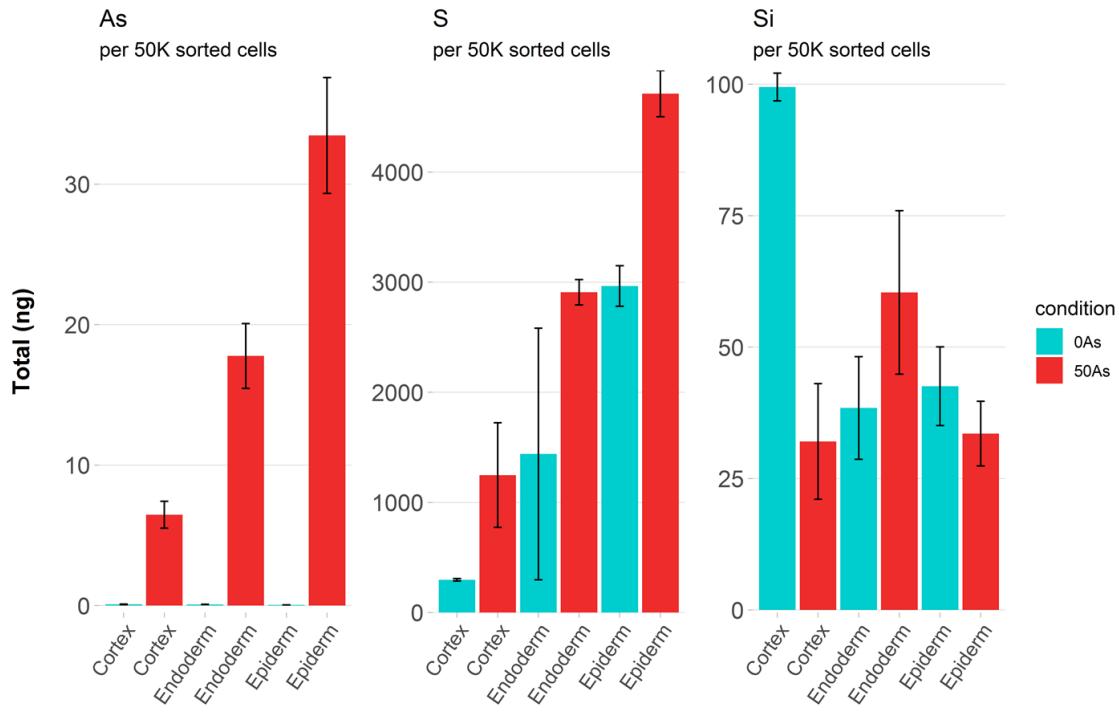
C





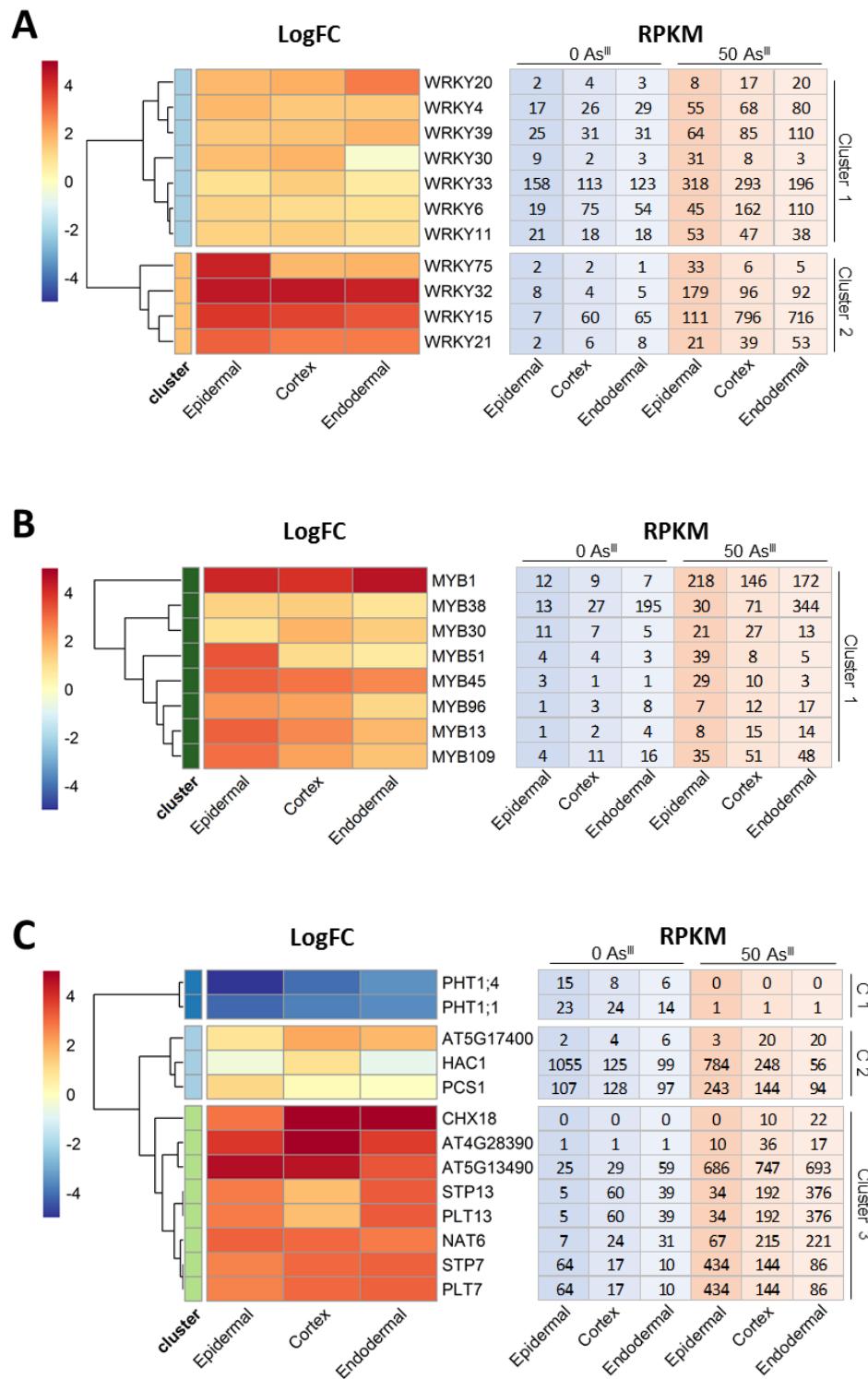
Supplemental Figure 2.4: Complete GO-Term Enrichment (with redundant terms)

Molecular Function (A), Cellular Component (B), and Biological Process (C) subontology terms enriched in response to As^{III}. Terms for each subontology arranged by adjusted p-values for up and down differentially expressed genes. Size corresponds to the number of genes (count) enriched for that term for each cell-type. Differentially expressed genes were filtered to include only $\text{abs}(\log FC) > 1.3$, and RPKM > 8 in As^{III} condition for up-regulated genes or RPKM > 8 in control condition for down-regulated genes. Cell-types separated by color.



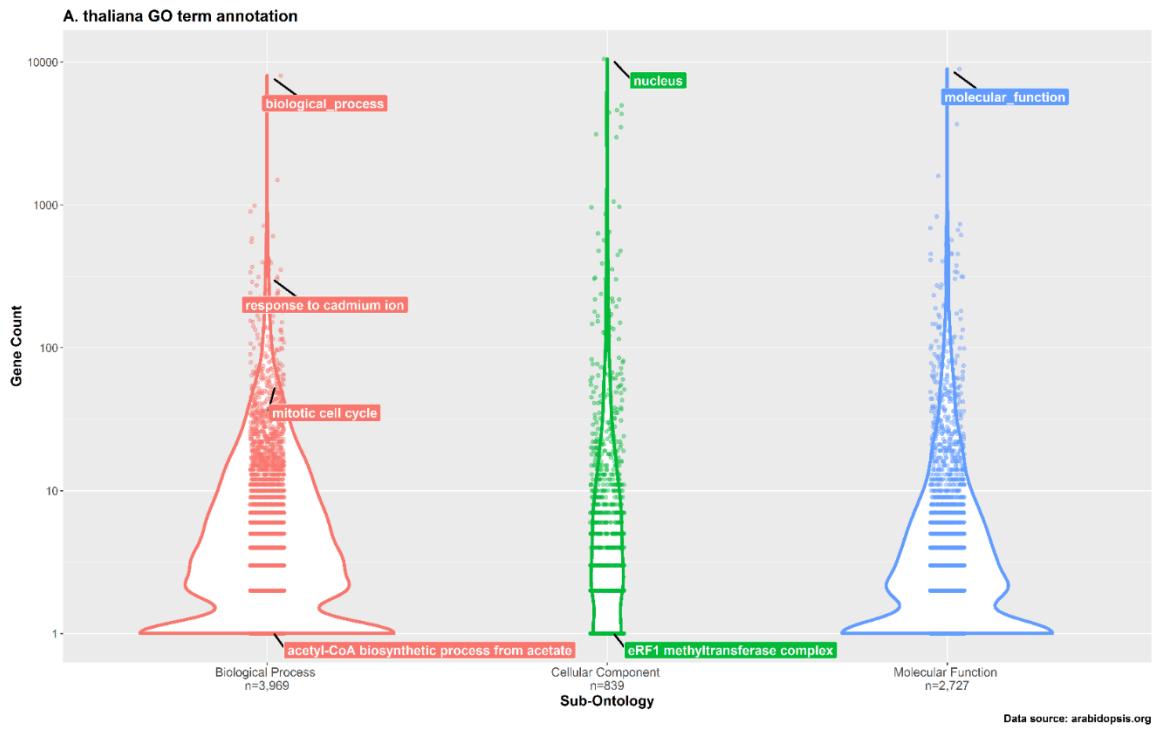
Supplemental Figure 2.5: ICP-MS by Cell-Type

Bar plots of total ng arsenic (As), sulfur (S), or silicon (Si) per 50,000 sorted cortex, endodermal, or epidermal cells in both experimental conditions.



Supplemental Figure 2.6: As^{III} Induced Expression by Cell-Type

Expression heatmaps showing LogFC by cell-type in response to As^{III} on left and the corresponding RPKM units arranged by cell-type and condition for each gene on right. Genes from the WRKY family transcription factors (A), MYB family transcription factors (B), and other genes of interest (C). Heatmaps include genes with logFC > 1 and a RPKM value of > 15 in at least one sample. Heatmaps clustered by logFC.



Supplemental Figure 2.7: *A. thaliana* GO Term Limitations

Violin plots of Biological Process (red), Cellular Component (green), and Molecular Function (blue) GO terms. Data points are unique GO terms with current (May 2020) counts of *A. thaliana* genes annotated with that term. Plots highlight some limitations to typical GO enrichment analysis. Terms with < 10 gene counts are too specific (*eRF1 methyltransferase complex*) and terms with ~ 10,000 gene counts are too vague (*molecular_function*). Many terms such as *response to cadmium ion* are over-represented (Gene Count = 313) while many terms like *mitotic cell cycle* are under-represented (Gene Count = 55).

NO	Sample	Read length nt	Raw read	Adapter trimming		Removing rRNA		Data mapping (Araport11)			
				Remained	%	Remained	%	Mapped	%	Multiple alignment	%
1	Guerinot_lab_Col_minus_AS_2	75	25,673,456	25,415,188	99.0%	25,409,134	99.98%	21,199,499	83.4%	278,023	1.3%
2	Guerinot_lab_Col_minus_AS_3	75	28,058,057	27,755,227	98.9%	27,745,114	99.96%	22,622,232	81.5%	307,404	1.4%
3	Guerinot_lab_Col_plus_AS_1	75	36,227,594	35,869,724	99.0%	35,856,446	99.96%	29,550,150	82.4%	359,265	1.2%
4	Guerinot_lab_Col_plus_AS_2	75	42,151,716	41,725,703	99.0%	41,713,751	99.97%	33,844,075	81.1%	382,656	1.1%
5	Guerinot_lab_Col_plus_AS_3	75	34,222,772	33,877,002	99.0%	33,863,223	99.96%	28,513,325	84.2%	325,763	1.1%
6	Guerinot_lab_Wer_minus_AS_3	75	18,354,728	18,179,032	99.0%	18,145,673	99.82%	15,257,357	84.1%	214,713	1.4%
7	Guerinot_lab_Wer_minus_AS_4	75	23,824,473	23,604,580	99.1%	23,574,991	99.87%	19,630,616	83.3%	286,974	1.5%
8	Guerinot_lab_Wer_minus_AS_5	75	31,876,073	31,579,485	99.1%	31,557,560	99.93%	26,939,648	85.4%	355,630	1.3%
9	Guerinot_lab_Wer_plus_AS_1	75	30,638,240	30,323,721	99.0%	30,300,102	99.92%	25,344,800	83.6%	402,553	1.6%
10	Guerinot_lab_Wer_plus_AS_2	75	22,348,791	22,116,137	99.0%	22,104,663	99.95%	18,672,588	84.5%	269,018	1.4%
11	Guerinot_lab_Wer_plus_AS_3	75	27,395,790	27,141,363	99.1%	27,129,537	99.96%	22,478,272	82.9%	367,957	1.6%
12	Guerinot_lab_Col_minus_AS	75	16,543,707	16,329,379	98.7%	16,234,861	99.42%	12,308,674	75.8%	378,084	3.1%
13	Guerinot_lab_Cortex_minus_AS_1	75	18,749,714	18,522,419	98.8%	17,851,569	98.38%	15,592,648	87.3%	175,928	1.1%
14	Guerinot_lab_Cortex_minus_AS_2	75	15,237,093	15,062,555	98.9%	14,685,489	97.50%	12,866,139	87.6%	141,356	1.1%
15	Guerinot_lab_Cortex_minus_AS_3	75	16,998,697	16,799,843	98.8%	16,497,105	98.20%	14,455,757	87.6%	163,098	1.1%
16	Guerinot_lab_Cortex_plus_AS_1	75	18,651,787	18,407,223	98.7%	18,369,287	99.79%	15,622,140	85.0%	167,128	1.1%
17	Guerinot_lab_Cortex_plus_AS_2	75	15,648,947	15,431,403	98.6%	15,234,942	98.73%	13,149,726	86.3%	140,939	1.1%
18	Guerinot_lab_Cortex_plus_AS_3	75	19,292,209	19,073,982	98.9%	18,553,785	97.27%	16,118,881	86.9%	190,074	1.2%
19	Guerinot_lab_SCR_minus_AS_1	75	14,202,031	14,031,606	98.8%	13,490,355	94.99%	11,185,504	82.9%	172,956	1.5%
20	Guerinot_lab_SCR_minus_AS_2	75	15,947,482	15,740,553	98.7%	15,211,468	96.64%	11,931,062	78.4%	268,890	2.3%
21	Guerinot_lab_SCR_minus_AS_3	75	15,187,859	14,983,546	98.7%	14,735,046	98.34%	12,328,264	83.7%	148,673	1.2%
22	Guerinot_lab_SCR_plus_AS_1	75	17,673,255	17,478,596	98.9%	17,004,379	97.29%	14,709,396	86.5%	218,173	1.5%
23	Guerinot_lab_SCR_plus_AS_2	75	17,641,086	17,419,851	98.8%	17,198,731	98.73%	15,174,807	88.2%	175,621	1.2%
24	Guerinot_lab_SCR_plus_AS_3	75	17,643,197	17,429,797	98.8%	16,967,617	97.35%	14,617,714	86.2%	153,974	1.1%

Supplemental Table 2.1: RNA Library Counts & Alignment

24 RNA samples used in this study were sequenced, trimmed, filtered, and aligned to the *A. thaliana* reference genome (Araport11). Sequencing was performed using the Illumina Nextseq500 system with 75bp fragment single-read libraries. Mapped read depth of samples ranged from ~ 11,000,000 – 33,000,000 per sample.

RPKM Range	Control			As ^{III}		
	Epidermal	Cortex	Endodermal	Epidermal	Cortex	Endodermal
0-6	1363 42.3%	1128 36.2%	1129 41.0%	0 0.0%	0 0.0%	3 0.1%
6-10	387 12.0%	410 13.2%	376 13.7%	395 12.3%	339 10.9%	355 12.9%
10-25	642 19.9%	687 22.0%	589 21.4%	696 21.6%	656 21.1%	611 22.2%
25-50	307 9.5%	408 13.1%	299 10.9%	562 17.5%	542 17.4%	520 18.9%
50-100	235 7.3%	225 7.2%	212 7.7%	473 14.7%	521 16.7%	445 16.2%
100-250	166 5.2%	170 5.5%	93 3.4%	469 14.6%	501 16.1%	417 15.2%
250-700	78 2.4%	76 2.4%	46 1.7%	369 11.5%	348 11.2%	258 9.4%
700-2000	35 1.1%	12 0.4%	8 0.3%	157 4.9%	150 4.8%	104 3.8%
2000-5000	6 0.2%	0 0.0%	0 0.0%	69 2.1%	45 1.4%	29 1.1%
5000+	0 0.0%	0 0.0%	0 0.0%	29 0.9%	14 0.4%	10 0.4%
Total	3219 100%	3116 100%	2752 100%	3219 100%	3116 100%	2752 100%

Supplemental Table 2.2: Gene Expression (RPKM) Counts from Up-Regulated DEGs by Cell-Type and Condition

Numbers in the table represent the number of up-regulated DEGs within the RPKM Range bin and the corresponding percentage in that cell-type and condition.

RPKM Range	Control			As ^{III}		
	Epidermal	Cortex	Endodermal	Epidermal	Cortex	Endodermal
0-6	1 0.0%	1 0.0%	1 0.0%	1900 57.1%	1665 56.8%	1617 48.5%
6-10	671 20.2%	667 22.7%	573 17.2%	422 12.7%	429 14.6%	474 14.2%
10-25	1079 32.4%	1034 35.3%	1035 31.1%	539 16.2%	509 17.4%	629 18.9%
25-50	660 19.8%	560 19.1%	640 19.2%	244 7.3%	190 6.5%	297 8.9%
50-100	414 12.4%	345 11.8%	451 13.5%	122 3.7%	75 2.6%	160 4.8%
100-250	288 8.7%	215 7.3%	356 10.7%	75 2.3%	48 1.6%	107 3.2%
250-700	163 4.9%	85 2.9%	163 4.9%	19 0.6%	14 0.5%	36 1.1%
700-2000	36 1.1%	20 0.7%	84 2.5%	4 0.1%	3 0.1%	13 0.4%
2000-5000	11 0.3%	6 0.2%	28 0.8%	1 0.0%	0 0.0%	0 0.0%
5000+	3 0.1%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%
Total	3326 100%	2933 100%	3333 100%	3326 100%	2933 100%	3333 100%

Supplemental Table 2.3: Gene Expression (RPKM) Counts from Down-Regulated DEGs by Cell-Type and Condition

Numbers in the table represent the number of down-regulated DEGs within the RPKM Range bin and the corresponding percentage in that cell-type and condition.

Cell-Type	Up-Reg	Down-Reg
Epidermal	2564	2465
Cortex	2459	2042
Endodermal	2145	2454

Supplemental Table 2.4: DEG Counts Per Cell-Type for GO Enrichment Analysis

Numbers in the table represent filtered DEGs using $\text{abs}(\log\text{FC}) > 1.3$ and either a RPKM > 8 in the As^{III} condition for up-regulated genes or RPKM > 8 in control condition for down-regulated genes.

**Chapter 3: Natural variation in *NIP1;1* expression determines
As^{III} tolerance in *A. thaliana* accessions**

Abstract

Arsenic is highly toxic for both humans and plants. Determining the mechanisms by which arsenic is accumulated, transported, and tolerated by plants is necessary in order to develop crops with reduced arsenic. To uncover new genetic factors controlling arsenite (As^{III}) tolerance in plants, we utilized a population of 527 *Arabidopsis thaliana* Multiparent Advanced Generation Inter-Cross (MAGIC) recombinant inbred lines (RILs) and identified two quantitative trait loci (QTL) in the population, a major QTL on chromosome 4 and a minor QTL on chromosome 3. Our analysis suggests natural variation in the expression of the nodulin-26-like protein, *NIP1;1*, among the MAGIC population is responsible for the major QTL on chromosome 4. It appears that down-regulating the expression of the plasma-membrane localized *NIP1;1* is the first line of defense against As^{III} toxicity in *Arabidopsis* and is a primary factor in determining As^{III} tolerance. Our work to identify the causative gene at the minor QTL on chromosome 3 is ongoing.

Introduction

Arsenic is a toxic metalloid ubiquitously present in soils that threatens plant and human health (Abdul et al., 2015). Pentavalent arsenate (As^{V}) is the principal arsenic redox state in aerobic soils but can be reduced to trivalent arsenite (As^{III}) under anoxic conditions, especially in the presence of anaerobic bacteria (Punshon et al., 2017; Majumder and Banik, 2019). Plants can accumulate arsenic at low levels, but rice (*Oryza sativa*) grown in flooded paddies accumulates higher levels of arsenic, mostly as As^{III} (Bakhat et al., 2017). A portion of that As^{III} is transported to and deposited in developing seeds (Carey et al., 2012), which humans consume as a staple food.

In aerobic soils, As^{V} enters root cells via the phosphate transport pathway (Sun et al., 2012; Wang et al., 2016) before being reduced to As^{III} by arsenate reductase enzymes (Chao et al., 2014). Cytosolic As^{III} is either complexed to glutathione/phytochelatin molecules and sequestered inside the vacuole via ABCC transporters (Song et al., 2010; Song et al., 2014), moved from cell to cell towards the vasculature (Tang et al., 2019), or effluxed back into the soil (Zhao et al., 2010; Chao et al., 2014). In anaerobic soils containing As^{III} , nodulin 26-like intrinsic proteins (NIPs) can act as passive channels to move As^{III} in and out of root cells.

Six of the nine NIPs in *Arabidopsis* (AtNIP1;1, AtNIP1;2, AtNIP3;1, AtNIP5;1, AtNIP6;1, and AtNIP7;1) are capable of bidirectional As^{III} transport (Bienert et al., 2008; Kamiya et al., 2009; Xu et al., 2015; Lindsay and Maathuis, 2016). CPK31, which interacts with NIP1;1 at the plasma membrane, regulates NIP1;1 As^{III} transport post-translationally (Ji et al., 2017). In the *A. thaliana* Col-0 ecotype, *NIP1;1* is highly expressed in 3-week

old plant roots, but after exposure to As^{III} *NIP1;1* and *CPK31* are down-regulated to reduce As^{III} entry into root cells. Null mutations in *AtNIP1;1* and rice *OsLsi2*, an anion transporter gene unrelated to NIP aquaporins, have been shown to increase tolerance to As^{III} by reducing the amount of As^{III} that enters the roots, resulting in less As^{III} in the roots, shoots, and seeds (Ma et al., 2008). However, knocking out these transporters is undesirable because As^{III} permeable NIPs also transport beneficial nutrients like silicon (Si), which is necessary for plant structure and disease resistance, so the overall fitness of the knockout lines becomes mitigated (Ma, 2004; Ma et al., 2006; Ma et al., 2007).

OsLsi1 is an important Si transporter that, which along with *OsLsi2*, facilitates Si transport towards the root vasculature in rice (Ma et al., 2007; Ma et al., 2008; Ma, 2010). Relative to the steel/vasculature, *OsLsi1* protein is enriched on the distal plasma membrane (PM) of exodermal and endodermal cells, whereas *OsLsi2* protein is enriched on the proximal PM of the same cells. Together these proteins coordinate the flux of Si and unfortunately As^{III}, towards the xylem and ultimately to the shoots. Approximately 60% to 80% of inorganic arsenic that enters rice roots is effluxed back to the external environment as As^{III} by the NIP aquaporin *OsLsi1* and other currently unidentified transporters (Xu et al., 2007; Zhao et al., 2010). *OsLsi1* accounts for 10-15% of total As^{III} efflux, suggesting different transporter proteins contribute to this process as well (Zhao et al., 2010).

Recently, a genome-wide association study (GWAS) based on the root growth of 133 *A. thaliana* accessions was performed to identify a locus controlling H₂O₂ tolerance (Sadhukhan et al., 2017). They found the expression level of *NIP1;1* was higher in accessions sensitive to H₂O₂ and overexpressing *NIP1;1* in Col-0, a H₂O₂ tolerant

accession, increased the sensitivity of the transgenic plants to H₂O₂ (Sadhukhan et al., 2017).

In this study, we investigated the genetic basis of natural variation in arsenic tolerance using a population of *A. thaliana* Multiparent Advanced Generation Inter-Cross (MAGIC) lines. The development of MAGIC populations allows the complementary use of both linkage and association mapping without the limitations caused by population structure (Cavanagh et al., 2008). The Arabidopsis MAGIC population encompasses the genetic variation within 19 founder accessions and consists of >500 recombinant inbred lines (RILs). Both the founder accessions and the RILs have been sequenced providing comprehensive marker coverage consisting of ~3 million individual sequence variants. This population has recently been used to uncover genetic variation and candidate genes for a number of traits including temperature dependent hypocotyl elongation growth (Box et al., 2015), cadmium accumulation within leaf tissue (Chao et al., 2012), seed storage oil composition (Menard et al., 2018), seed size and seed number (Gnan et al., 2014), and root exudate composition (Monchgesang et al., 2016).

We used inhibition of root length when plants are grown in the presence of As^{III} to carry out both QTL analysis and GWAS using software tools developed for the MAGIC population. We identified QTLs for arsenic tolerance on chromosomes 3 and 4. GWAS allowed us to identify individual SNPs associated with arsenic tolerance. Our results show that the *NIP1;1* gene locus is responsible for the major QTL on chromosome 4. More specifically, the plant's *NIP1;1* expression profile when exposed to As^{III} is a critical factor that determines whether *A. thaliana* accessions tolerate As^{III}.

Results

*Natural Variation in Arsenic Tolerance Found in *A. thaliana* MAGIC Population*

To determine whether *A. thaliana* displays variation in As^{III} tolerance, we utilized the population of *A. thaliana* Multiparent Advanced Generation Inter-Cross (MAGIC) lines developed by Korver et al. (2009) (Kover et al., 2009). We initially screened the 19 MAGIC line founder accessions for differences in tolerance to As^{III}. These MAGIC founders displayed variation in germination, shoot mass, shoot color, and root morphology when grown on As^{III} medium compared to control medium (**Supplemental Figure 3.1**), but we ultimately used the change in root length to score the MAGIC population, as it was the most consistent and easily measured metric. Root length has also previously been used to quantify As^{III} tolerance (Kamiya et al., 2009; Clemens and Ma, 2016). To collect the greatest range of root score measurements, MAGIC lines were grown on control medium for 11 days and As^{III} medium for 15 days, when roots from the fastest growing plants reached the bottom of the medium plate under each condition (**Figure 3.1A**). For every MAGIC line, we generated a root score by subtracting the average length of a plant's main root grown on As^{III} medium from the average root length when grown on control medium. For example, on control medium after 11 days, MAGIC lines #193 and #402 have similar root lengths; 3.3 and 3.0, respectively (**Figure 3.1A**). However, after 15 days on As^{III} medium, #193 and #402 have drastically different root lengths; 2.3 and 4.0, respectively (**Figure 3.1A**). Root scores for MAGIC lines #193 and #402 were therefore, 1 (3.3 - 2.3) and -1 (3.0 – 4.0), respectively.

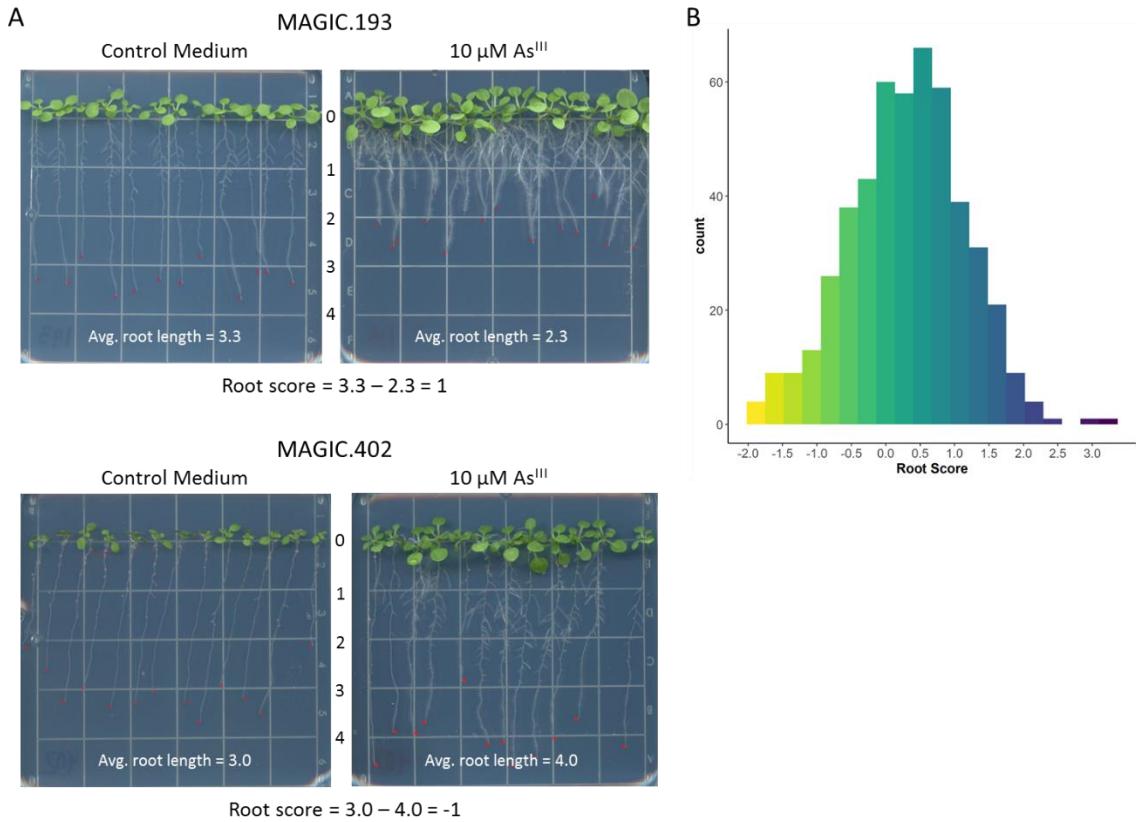


Figure 3.1: MAGIC Line Root Scoring for Arsenic Tolerance

(A) Roots for MAGIC lines #193 and #402 scored for arsenic tolerance. (B) Distribution of 492 MAGIC line root scores used for QTL mapping. MAGIC lines which tolerate As^{III} produced negative root scores; MAGIC lines sensitive to As^{III} produce positive root scores. Min = -1.94, 1st Qu. = -0.32, Median = 0.33, Mean = 0.30, 3rd Qu. = 0.85, Max = 3.18. Red dots in (A) indicate where the main root tip of each plant ended and thus where each measurement occurred.

We generated root scores for 492 of the 527 MAGIC lines (**Figure 3.1B**) that ranged from -1.9 to 3.2, representing the most tolerant and sensitive MAGIC lines, respectively. The relatively normal distribution suggests multiple loci or multiple variants at a locus are influencing As^{III} tolerance (Yang et al., 2018).

GWAS for Arsenic Tolerance

Over 3.3 million sequence variants segregate in the 19 MAGIC founders, consisting of SNPs, indels, and complex substitutions (Gan et al., 2011). These MAGIC line genomes are mosaics of those 19 MAGIC founder genomes, composed of DNA from up to 16 different founders per MAGIC line (Kover et al., 2009). 392 of the 527 MAGIC lines have been sequenced with enough coverage to be used with the 3.3 million marker library. Utilizing this robust source of natural variation, we performed genome association mapping based on the root scores of those 392 MAGIC lines to determine loci controlling As^{III} tolerance. Our analysis revealed one major QTL on chromosome 4 and one minor QTL on chromosome 3 (**Figure 3.2, Supplemental Figure 3.2**). 284 variants on chromosome 4 cross our most conservative QTL threshold of $-\log_{10}(0.05/3.3M)$. This locus spans 2,748,277 base pairs and encompasses approximately 792 genes. While it is possible multiple genes are influencing this locus association, three of the top four variants are in the promoter or just upstream of the *NIP1;1* (AT4G19030) promoter (**Figure 3.2B,C**) and *NIP1;1* is known to be involved in As^{III} transport (Kamiya et al., 2009; Ji et al., 2017). rs2412612, rs2412782, and rs2412888 are located 441 bp, 4287 bp, and 6811 bp upstream of the *NIP1;1* 5'-UTR, respectively (**Figure 3.2B,C**). *NIP1;1*, therefore, is

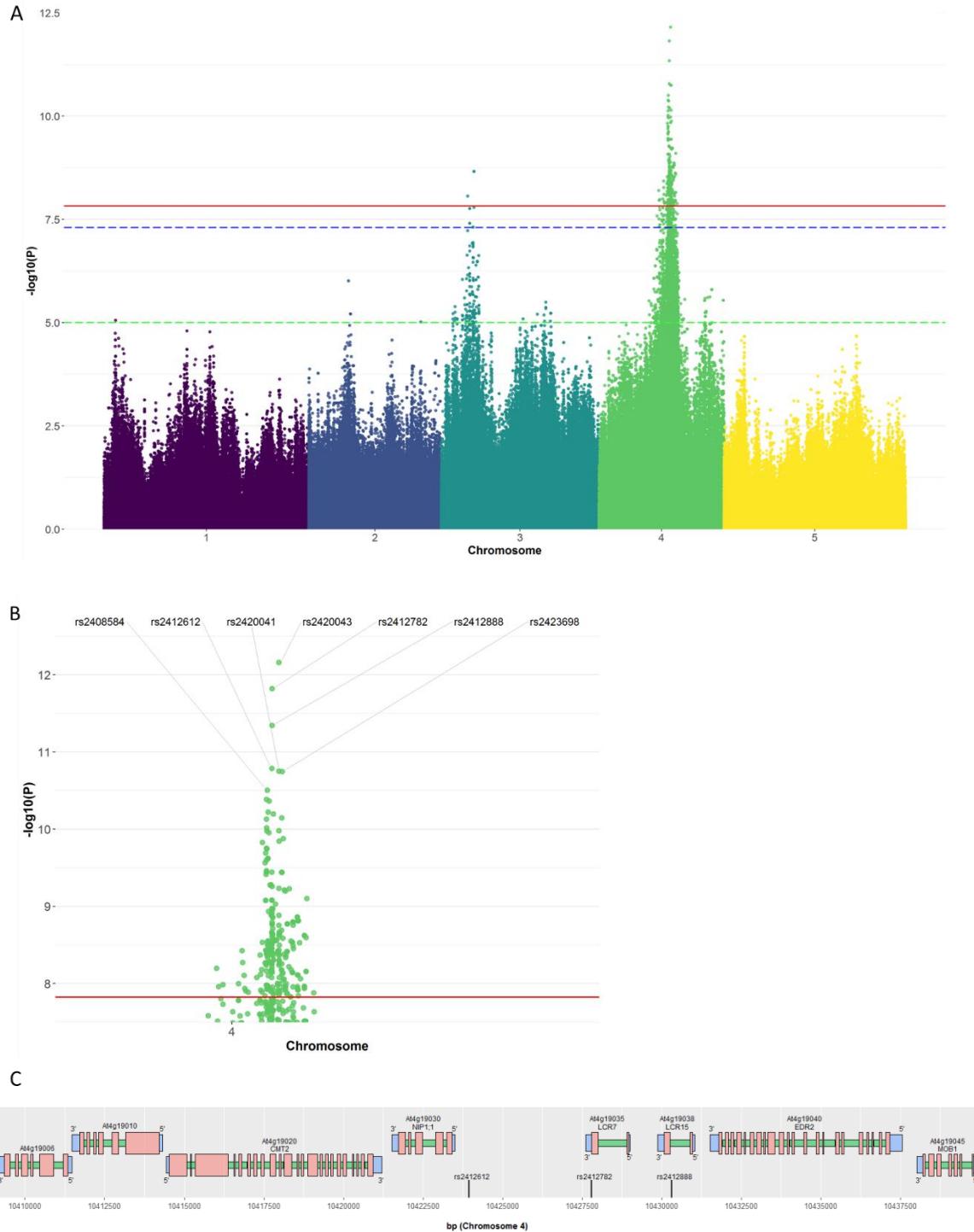


Figure 3.2: GWAS for Arsenic Tolerance Found in MAGIC Population

(A) Manhattan plot showing association of ~3.3 million sequence variants with 392 root scores for arsenic tolerance within the MAGIC population, mapped onto the five A.

thaliana chromosomes. Horizontal lines represent conventional statistical significance thresholds of $-\log_{10}(1e-5)$ (green dash) and $-\log_{10}(5*10^{-8})$ (blue dash). Red line represents the $p < 0.05$ threshold after Bonferroni correction, $-\log_{10}(0.05/3.3M)$. (B) Magnified view of major QTL on chromosome four with top seven variants associated with As^{III} tolerance labelled. (C) Location of 2nd, 3rd, and 4th most significant variants (rs2412782, and rs2412888, rs2412612) compared to *NIP1;1* locus and nearby genes. Segments: blue = UTR, pink = exon, green = intron.

our most likely candidate gene to associate with As^{III} tolerance. Six of the 9 *A. thaliana* NIP subfamily members have previously been shown to be capable of As^{III} transport (NIP1;1, NIP1;2, NIP3;1, NIP5;1, NIP6;1, and NIP7;1 (Bienert et al., 2008; Kamiya et al., 2009; Xu et al., 2015), but only *NIP1;1* and *NIP1;2* (AT4G18910), are located near a QTL in our study. The only variant within or near *NIP1;2* is rs2410604, located in the *NIP1;2* 2nd intron and has a -log10(P) value of 8.3 in our GWAS. rs2412612 is the closest variant to *NIP1;1*, located 441 bp upstream of the 5' UTR and has a -log10(P) value of 10.8 in our GWAS. *NIP1;2* expression was also nearly unexpressed in all MAGIC founder lines, regardless of condition (data not shown). The lack of *NIP1;2* expression together with the lower significance of rs2410604 compared to rs2412612 make it unlikely *NIP1;2* influences the major QTL on chromosome 4.

We also performed GWAS using a marker library of 1260 SNPs found in all 527 MAGIC lines. While the marker coverage is limited, it did allow us to perform association mapping with all 492 MAGIC scores we obtained (**Supplemental Figure 3.3A**). This analysis showed the same minor QTL on chromosome 3 and major QTL on chromosome 4, but also additional minor QTLs on chromosomes 3 and 4. There could potentially be a 2nd QTL on chromosome 4 not identified with the 3.3 million marker analysis. This analysis also shows the most significant SNP, MN4_10482087, is the closest SNP to *NIP1;1* in the library (**Supplemental Figure 3.3B**).

Each MAGIC founder was ranked by As^{III} sensitivity, 1-19 (**Supplemental Figure 1**). Can-0, being the most sensitive to As^{III}, is ranked 1st. Edi-0, the most tolerant founder line, was ranked 19th. We were able to assign MAGIC founder lineage to nearly every segment of DNA for 392 of the MAGIC lines. 384 of those MAGIC lines had a founder

haplotype assigned to the *NIP1;1* locus. To determine whether the founder haplotypes at the *NIP1;1* locus are responsible for the major QTL on chromosome 4, we plotted founder haplotype against root score for each MAGIC line at the *NIP1;1* locus and ordered haplotypes by the founder's As^{III} sensitivity phenotype rank (**Figure 3.3A**). We see strong correlation between root scores and ranked haplotype (**Figure 3.3B**), with an R² value of 0.44 and p-value of 0.0019. We found no such correlation between root scores and ranked haplotypes at the minor QTL on chromosome 3, which produces an R² of only 0.1 and p-value of 0.18. (**Figure 3.3C&D**).

Expression of NIP1;1 in MAGIC Founders

The variants in the promoter of *NIP1;1* suggest that differences in *NIP1;1* expression among the MAGIC founders might be the most important element at the major QTL on chromosome 4. Expression of *NIP1;1* in the Col-0 ecotype decreases in root tissue after exposure to As^{III} (Ji et al., 2017), but no *NIP1;1* expression data exists for other ecotypes. To determine whether natural variation in *NIP1;1* expression exists and explains As^{III} tolerance across *A. thaliana* ecotypes, we performed real-time qPCR measuring *NIP1;1* expression from the 19 MAGIC founder lines (**Figure 3.4A**). Root tissue cDNA was generated from plants grown on control medium for 7 days and plants transferred after 7 days to As^{III} medium for 1-5 days. Before As^{III} exposure (0dAs) most MAGIC founders had similar levels of *NIP1;1* expression (**Figure 3.4C**). Only Tsu-0 and Wu-0 having statistically higher expression than Col-0 and only Edi-0 having statistically lower expression than Col-0 (**Figure 3.4C**). After As^{III} exposure, as expected, we observed

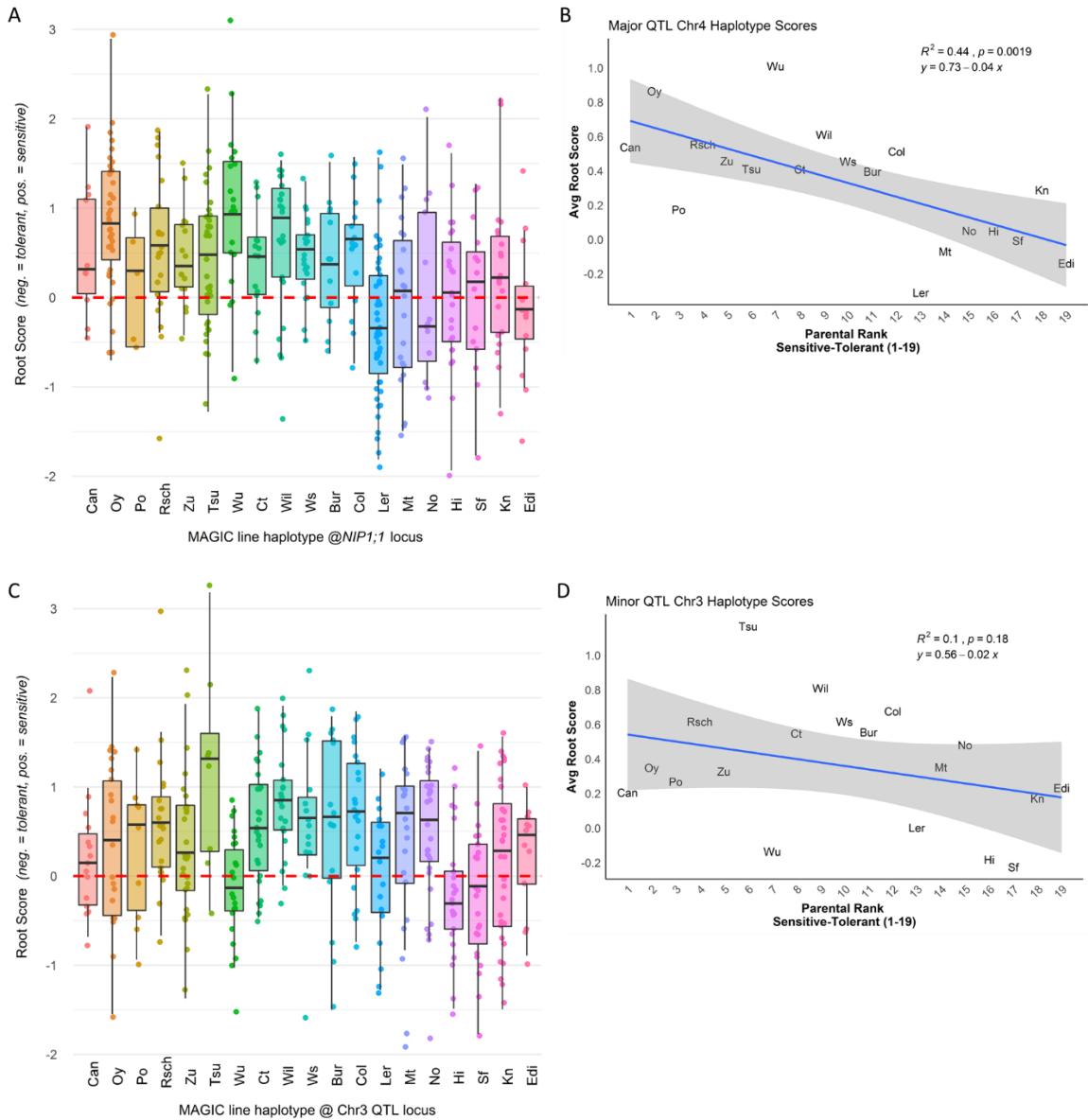


Figure 3.3: MAGIC Line Haplotypes at Major and Minor QTL

Boxplots of MAGIC line haplotypes at the *NIP1;1* (A) and Chr3 QTL (C) loci showing relationship between arsenic sensitivity phenotype rank (Parental Rank), arranged from most sensitive (Can) to most tolerant (Edi), and arsenic root scores. Each box represents the 25%, 50%, and 75% quantiles. Red dotted line represents root score of 0. Correlation of the average root scores for each MAGIC line haplotype found at the *NIP1;1* (B) and

Chr3 QTL (D) loci and phenotype rank of the 19 MAGIC line founders. *NIP1;1* locus correlates with Rank ($R^2 = 0.44$). Chr3 QTL locus does not correlate with Rank ($R^2 = 0.1$).

changes to *NIP1;1* expression across MAGIC founders (**Figure 3.4A**). We were surprised, however, to see multiple patterns of response As^{III}, over the course of 5 days As^{III} exposure (**Figure 3.4A**).

Can-0, Oy-0, and Wil-2 steadily increased *NIP1;1* expression. Po-0, Rsch-4, Ler-0, No-0, Sf-2, and Edi-0 *NIP1;1* expression remained relatively stable. Zu-0 slowly, but steadily, decreased *NIP1;1* expression. Tsu-0, Wu-0, Col-0, and to lesser degrees, Ct-1 and Kn-0, initially decreased expressed at 1dAs, but then increased *NIP1;1* back to original levels before As^{III} exposure. Cumulative *NIP1;1* expression, over the entire time course is the best indicator for overall tolerance to As^{III}. We measured a strong correlation between cumulative *NIP1;1* expression and founder arsenic sensitivity phenotype rank (Parental Rank), with an R² of 0.65 and p-value of 2.9e⁻⁰⁵ (**Figure 3.4B**). The basal *NIP1;1* expression level at 0dAs also strongly correlates with Parental Rank, with an R² of 0.61 and p-value of 7.6e⁻⁰⁵ (**Figure 3.4D**).

Finally, we expressed *NIP1;1* with its native promoter from the Can-0 ecotype in the *nip1;1* mutant background (Col-0 ecotype). Three transgenic lines (C1, C2, and C3) expressing Can-0 *NIP1;1* showed different levels of complementation when grown on 15 μM As^{III} (**Figure 3.5A**). *NIP1;1* expression in each transgenic line, however, matched the plants level of As^{III} sensitivity (**Figure 3.5B**), giving further evidence that *NIP1;1* expression is responsible for the major QTL and is a critical genetic factor in determining As^{III} tolerance for *A. thaliana*.

Discussion

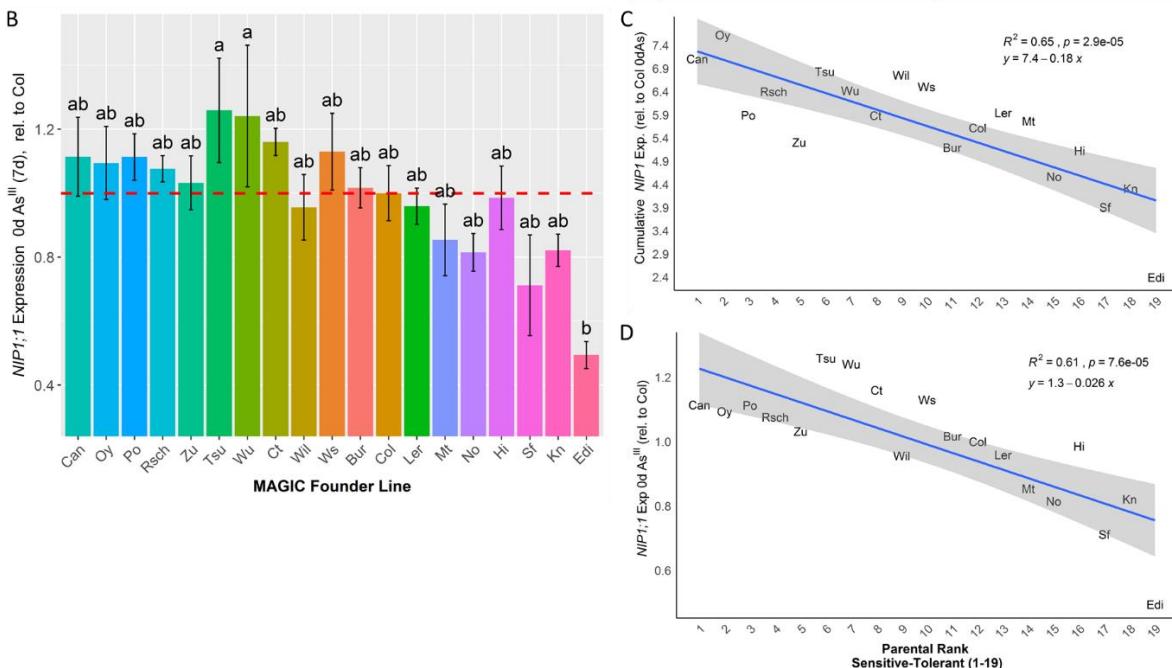
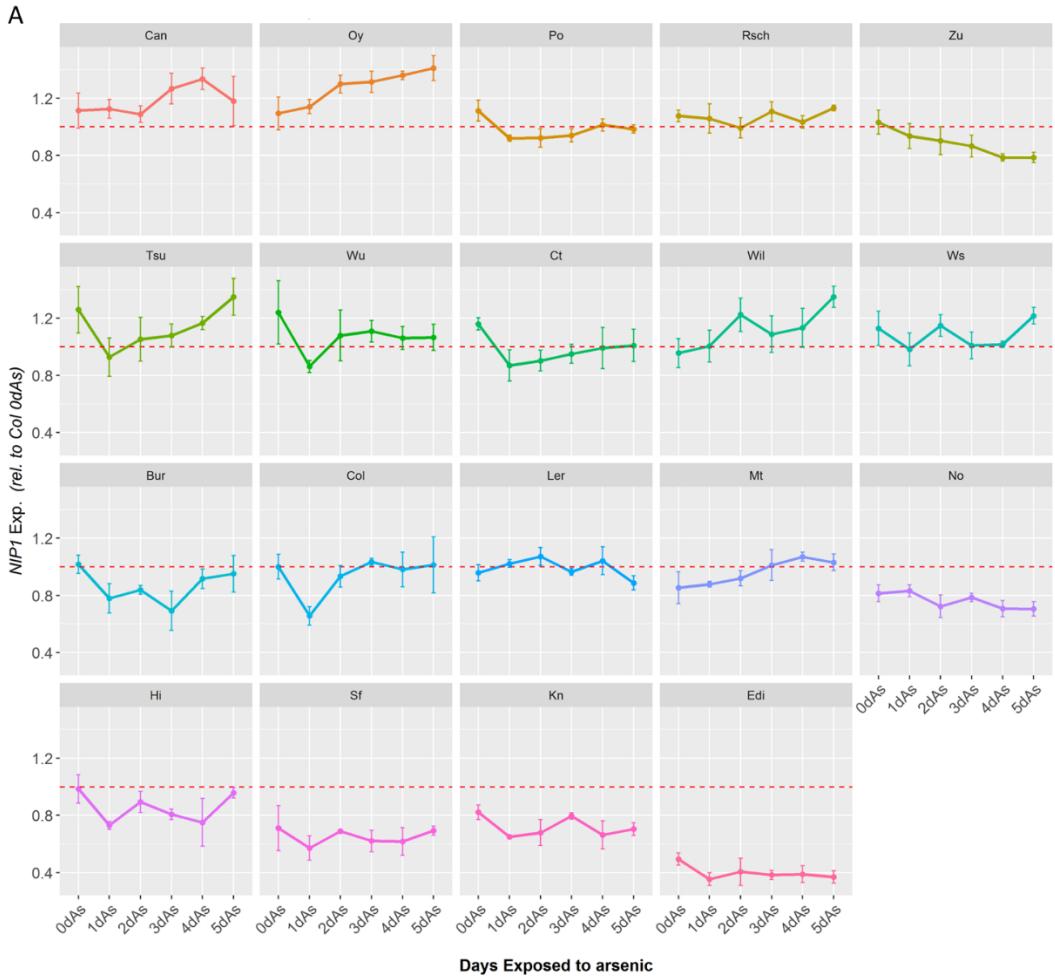
Arsenic is a naturally occurring toxic metalloid found in soils throughout the world (Wedepohl, 1995). Arsenic in the soil enters our food chain through plant roots. Most arsenic absorbed into root cells remains in the roots, but a portion of arsenic is transported to other tissues often consumed by humans like seeds (grain). Using the Arabidopsis MAGIC lines we performed a GWAS that identified two QTL that determine As^{III} tolerance within this plant population. We used two different GWAS approaches, each with their own advantages, that both identified a major QTL on chromosome 4 and a less significant minor QTL on chromosome 3.

The first approach utilized low coverage genomic maps for the MAGIC lines. Only 1260 SNP markers are found in these condensed genome mosaics. We attained 527 individual MAGIC lines that had been sequenced at the low coverage depth and were able to obtain root scores for 492 lines, which we used to measure arsenic tolerance in the population. The main advantage of the low coverage MAGIC population is having more individual genomes in the population. The high coverage MAGIC population contains fewer sequenced genomes but higher sequence depth, with a library of 3.3 million SNP, indel, and complex substitution markers, but for only 392 MAGIC lines. 384 of those high coverage MAGIC line genomes had a founder haplotype assigned to the *NIP1;1* locus. The low coverage GWAS produced three (potentially four) minor QTL; two on chromosome 3 and one (potentially two) on chromosome 4 (**Supplemental Figure 3.2**). It also produced one major QTL on chromosome 4 (**Supplemental Figure 3.2**).

The major QTL on chromosome 4 was the most significant QTL in both high and low coverage GWAS approaches. The most significant SNP in the low coverage GWAS and the 2nd, 3rd, and 4th most significant markers in the high coverage GWAS were near the *NIP1;1* locus. Finding a correlation between MAGIC line root scores and founder haplotypes (Parental Rank) at the *NIP1;1* locus further encouraged us to pursue *NIP1;1* at the major QTL (**Figure 3.3A-B**). In Korver *et al*, MAGIC line pedigrees were funnel mated so that four generations of crossing would produce MAGIC lines from a maximum of 16 distinct founder genomes; see Figure 1 of (Kover et al., 2009). However, the genome mosaics produced in this study as well as the publicly available mosaics online (see Materials and Methods) contained many examples of 17-19 founder genomes assigned to individual MAGIC lines (**Supplemental Figure 3.5**), which should not be possible. Knowing some haplotypes are incorrect should only make the correlations we found between the haplotypes and root scores more convincing. Re-sequencing more MAGIC lines with greater sequence depth and/or making improvements to the ‘reconstruction’ algorithm used to classify haplotypes to generate the genome mosaics would hopefully mitigate these errors and enhance the efficacy of these association studies. Incorporating the MAGIC line pedigree into the algorithm so that loci resulting from multiple cross over events in a relatively small window can only be assigned to haplotypes found in that MAGIC lines pedigree would correct many of these errors.

Because three of the top four most significant markers in the high coverage GWAS were in the *NIP1;1* promoter region or a few thousand bp upstream of *NIP1;1* we suspected variation in *NIP1;1* expression, specifically, could be responsible for the major QTL on chromosome 4 (**Figure 3.2C**). Measuring initial *NIP1;1* expression prior to As^{III} exposure

and expression after prolonged exposure (1-5 days) confirmed overall As^{III} tolerance is strongly correlated with *NIP1;1* expression (**Figure 3.4C-D**). No definitive expression pattern explains MAGIC founder As^{III} tolerance completely, but generally those MAGIC founders with lower *NIP1;1* prior to As^{III} exposure and/or founders that maintained low *NIP1;1* after As^{III} exposure were most tolerant. MAGIC founders with higher *NIP1;1* prior to As^{III} exposure and/or failed to down-regulate *NIP1;1* sufficiently after As^{III} exposure were most sensitive. In looking for one variable that can best predict As^{III} tolerance in this plant population, we find cumulative expression over multiple days is probably the best predictor, as it correlated the strongest with our parental ranking.



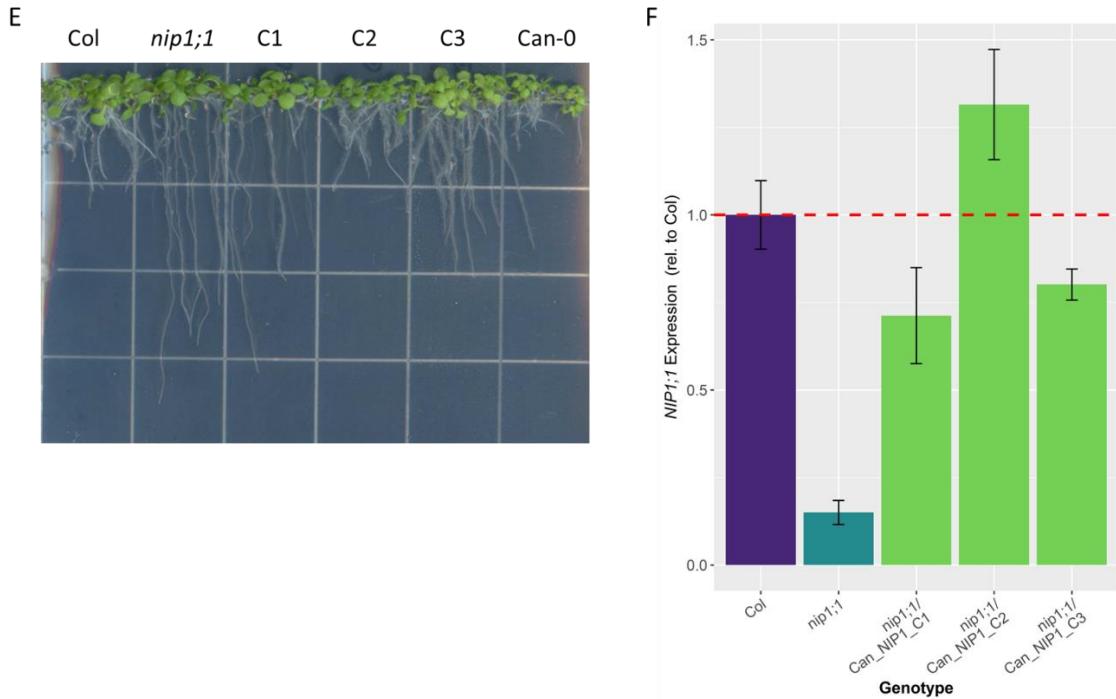


Figure 3.4: MAGIC founder *NIP1;1* expression

(A) Founder *NIP1;1* expression (qPCR), relative to Col-0 at 0dAs. Samples 0dAs – 5dAs represent plant roots harvested either after 7d on control medium (0dAs) or after 1-5 days transferred to 10 µM As^{III} media. Founders are arranged by As^{III} tolerance phenotype rank, most sensitive (Can) to most tolerant (Edi) to show correlation of phenotype with *NIP1;1* expression. (B) Founder *NIP1;1* expression (qPCR) at only 0dAs (7d old plants). Statistical significance was calculated using one-way ANOVA (*P < 0.05) with Tukey's Test. (C) Strong correlation between cumulative *NIP1;1* expression (0dAs – 5dAs) and phenotype rank of the 19 MAGIC line founders. (D) Strong correlation between *NIP1;1* expression (0dAs only) and phenotype rank. (E) Phenotype of transgenic lines (C1, C2, C3) expressing *NIP1;1* from Can-0 with native promoter in the *nip1;1* background. (F) *NIP1;1* expression, relative to Col-0, from root tissue of 7d plants on control medium.

By expressing *NIP1;1* under its native promoter, from Can-0, our most As^{III} sensitive founder, into the mutant *nip1;1* background, we were able to show how strongly *NIP1;1* expression influences As^{III} tolerance. Compared to Col-0, transgenic lines with higher *NIP1;1* levels were clearly more sensitive to As^{III} and lines with lower *NIP1;1* levels were clearly less sensitive to As^{III} (**Figure 3.4E-F**). This also supports the variation at the major QTL4 locus was not driven by differences in *NIP1;1* protein function. Variation in *NIP1;1* expression among the MAGIC founders alone sufficiently explains As^{III} tolerance. Because there are many different expression patterns for *NIP1;1* in response to As^{III} stress (**Figure 3.4A**), it is unlikely any one SNP or indel within the *NIP1;1* promoter is responsible for expression variance.

We do not expect to identify the genes responsible for most minor QTLs without new MAGIC genomes added to the population or deeper re-sequencing of the entire population. By re-sequencing the remainder of the 527 MAGIC lines that were left out in the 3.3 million marker GWAS and/or including new MAGIC lines into the gene pool for GWAS would add at least 100 new MAGIC genome mosaics to this study and vastly improve the performance of future association mapping. This would likely help us identify causative genes for the minor QTLs. We would also gain the ability to determine which haplotype combinations found at the major and minor QTLs generate the most As^{III} tolerant plants. Not all founder haplotypes are well represented at our QTLs and many haplotype combinations at our major and minor QTLs are not present (**Supplemental Figure 3.4**). Being able to determine the best combination of alleles from each QTL to produce the most As^{III} tolerant plants would incredibly valuable. Currently, our ability to assess the impact

of *NIP1;1* from certain founders is lacking. The Po-0 haplotype, for example, is present in only 5 of the 392 MAGIC lines used in our 3.3 million marker GWAS.

In conclusion, we determined variation in the expression of *NIP1;1*, a NIP aquaporin known to transport As^{III}, was the most important factor determining As^{III} tolerance in a population of *A. thaliana*. As^{III} tolerant plants have low basal levels of *NIP1;1* expression and/or maintain low *NIP1;1* expression during arsenic exposure, mitigating the entry of excess arsenic into the roots. Our study also demonstrates the capacity and some limitations of MAGIC populations for identifying genetic factors impacting traits through GWAS. As this MAGIC population expands and the sequencing depth increases, performing the same GWAS with our root score data could perhaps narrow our minor QTL loci sufficiently to uncover additional novel genetic factors contributing to As^{III} tolerance.

Materials and methods

Plant Material

Seeds for 527 MAGIC lines and 19 MAGIC founder accessions were obtained with help from the Salt Lab at The University of Nottingham and The Nottingham Arabidopsis Stock Center (Nottingham, UK). The seeds were surface sterilized with a bleach and ethanol and allowed to dry for 3-4 hours, then suspended in sterile 0.1% agarose and left in the dark at 4°C for 2-4 days. Seeds were then sown onto half-strength Murashige-Skoog (MS) medium containing 1% sucrose and 1% agar (Type M, Sigma-Aldrich), adjusted to pH 5.7 with KOH (control medium). Arsenic medium was the same as control medium

with the addition of 10 μM sodium (meta)arsenite (As^{III}) (Sigma-Aldrich), added after autoclaving.

Growth Conditions

Plates with seeds were placed vertically in a light chamber with standard conditions (16h of light at 22°C, 8h of dark at 18°C, 37 $\mu\text{mol m}^{-2} \text{s}^{-1}$ photon flux density). Plants for GWAS were grown on control medium (1/2 MS) for 11 days or arsenic medium (1/2 MS + 10 μM As^{III}) for 15 days, then imaged.

For gene expression analysis (qRT-PCR) of *NIP1;1*, seeds of the 19 MAGIC founders were sown onto sterile 80-micron nylon mesh (32% opening Component Supply, /U-CMN-80B-C) placed on plates made with control medium and grown under standard conditions for 7d. Root tissue for 0d As^{III} samples were harvested at 5 pm (ET) after the initial 7d. The other plants were transferred after 7d from control medium to 10 μM As^{III} medium by lifting the mesh with sterile tweezers and placing it onto the As^{III} medium. Those plants returned to the light chamber for 1-5d and roots were harvested at 5 pm each day. 3 replicates of root tissue, for 19 MAGIC founder lines, under each condition (0d-5d exposure to 10 μM As^{III}) were collected from 342 (3 x 19 x 6) individual plates.

Root Length Measurements

The 527 MAGIC lines were imaged and root scores for As^{III} tolerance were generated by measuring the length of the longest main root (some MAGIC lines have

multiple main roots when grown on As^{III}), from bottom of hypocotyl to root tip (vertical distance from top line to wherever the tip ended; diagonal growth did not impact length), of each individual plant. Root from 8-12 plants per plate, in triplicate reps, under both conditions (control/ As^{III}) for each MAGIC line, were measured. For each MAGIC line, the avg. root length under control conditions minus the avg. root length under arsenic conditions = final root score. 35 of the 527 MAGIC line seeds were either missing or failed to germinate. 492 total MAGIC line scores were generated. Negative scores indicate tolerance to As^{III}, positive scores indicate sensitivity to As^{III}.

Genome Association

Initial genome association was performed as described by Korver *et al.* (2009), using the HAPPY R package to construct a mosaic of MAGIC line genomes from only 1,260 single nucleotide polymorphisms (SNPs) among the 19 MAGIC founder genotypes, available with instructions at <http://mtweb.cs.ucl.ac.uk/mus/www/19genomes/MAGICsnps.htm>.

With these condensed MAGIC genotypes, QTL mapping in R which performed a genome scan with 1000 permutations to determine genome-wide thresholds for statistical significance (**Sup Figure 3**).

Improved QTL mapping was achieved by reconstructing MAGIC line mosaics from over 3 million segregating sequence variations found in the 19 MAGIC founders. Stepwise multi-locus regression with 10,000 permutation tests was then performed to evaluate genome wide significance, scanning all 3.3 million sites on the 492 MAGIC line

root scores for arsenic tolerance. Software, chromosome data, and instructions for creating the 3.3 million variant MAGIC line genome mosaics (reconstruction) and performing association mapping (genome_scan) are available at <http://mtweb.cs.ucl.ac.uk/mus/www/19genomes/MAGICseq.htm>.

Analysis of Association Markers

Genome mosaic data and association statistics generated from the reconstruction and genome_scan, respectively, were further analyzed in R. Script available at <https://github.com/TWarczak/Guerinot-Lab>.

Gene Expression Analysis

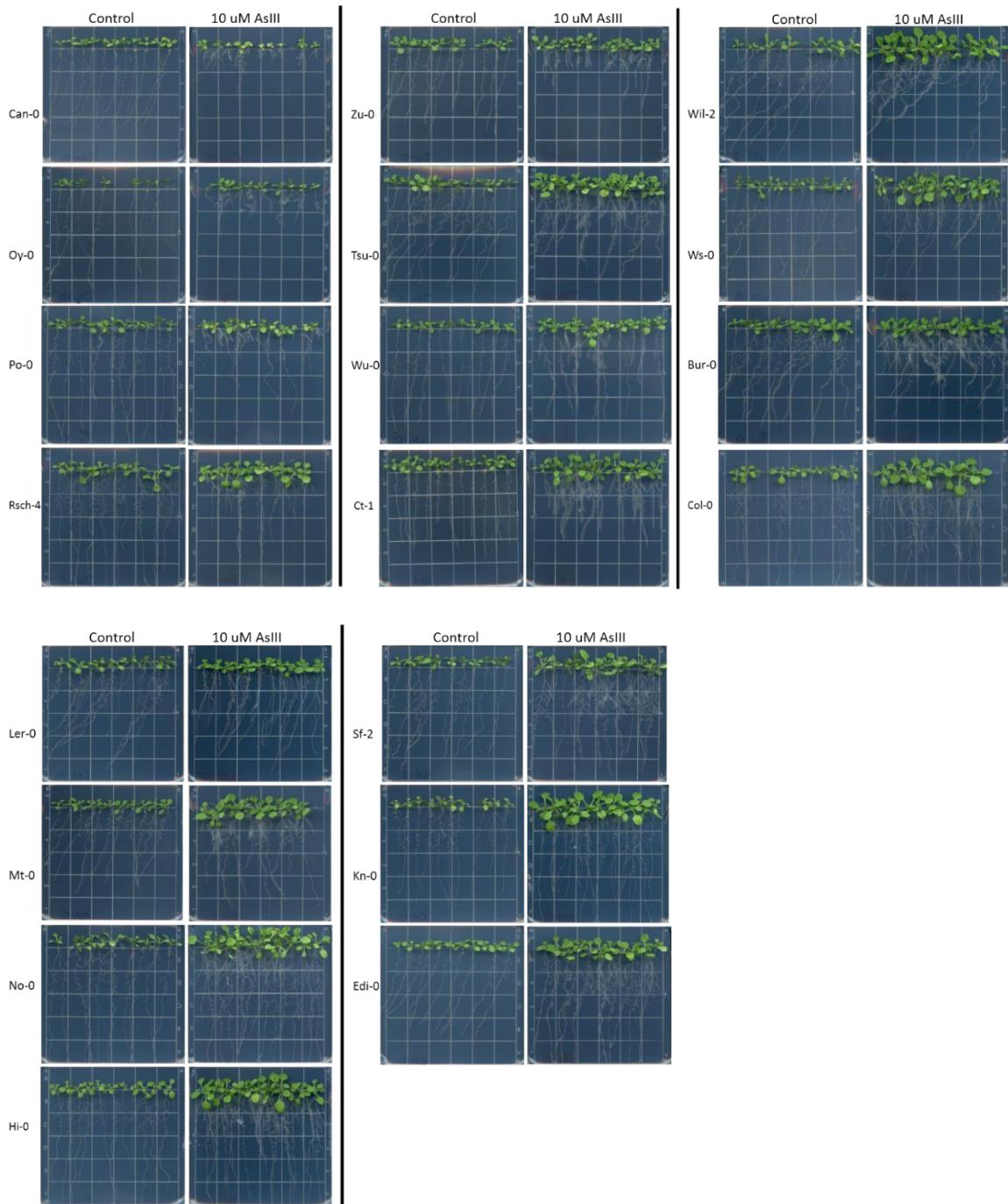
Total RNA was extracted from root tissue with the E.Z.N.A. Plant RNA Kit (OMEGA Bio-tek, Norcross, GA). DNase treatment was performed with the On-Column DNase I Digestion Set (Sigma). cDNA was generated from RNA with the iScript cDNA Synthesis Kit (Bio-Rad). Real-time PCR with iTaq Universal SYBR Green Supermix (Bio-Rad) and diluted cDNA was performed on the C1000 Touch Thermal Cycler, CFX384 Real-Time System (Bio-Rad) with Bio-Rad CFX Manager 3.0 software. Primer pairs to analyze expression for NIP1;1 were TW_106F3 (5'-GCGAGAAATCACTAAAAGTGGTTC-3') plus TW_106R2 (5'-GTAAGGTGTGGTGAAAGTTGTTAC-3') and UBQ1 (housekeeping) were TW_202F (5'-TCGTAAGTACAATCAGGATAAGATG-3') plus TW_202R (5'-

CACTGAAACAAGAAAAACAAACCC-3'). qRT-PCR data was analyzed in R, with script available at <https://github.com/TWarczak/Guerinot-Lab>.

Accession Numbers

Sequence data can be found in GenBank/EMBL/TAIR data libraries using Arabidopsis gene identifiers for At4g19030 (NIP1;1) and At3g52590 (UBQ1).

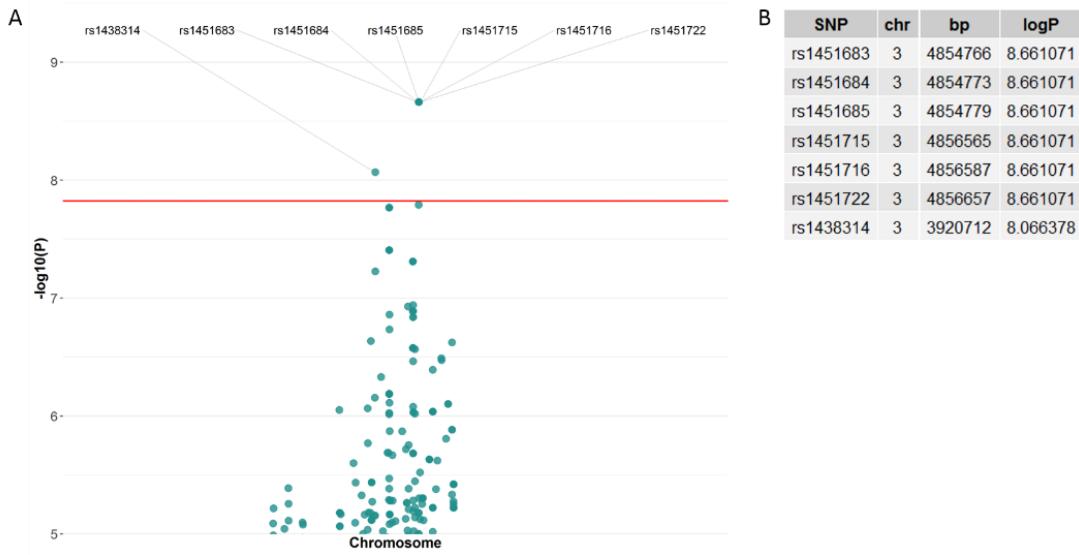
Supplemental Figures



Supplemental Figure 3.1: MAGIC Founder Lines; ordered by As^{III} sensitivity rank

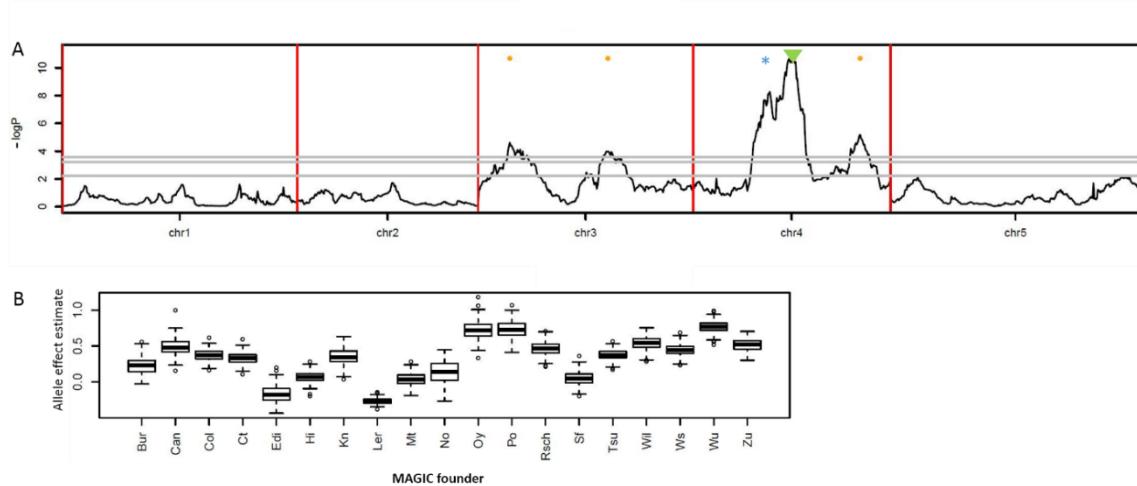
MAGIC founder lines grown on control medium (half-strength MS) for 11 days, and 10 μM As^{III} medium for 15 days. Founders are arranged from most As^{III} sensitive (Can-0) to

tolerant (Edi-0). Order used as As^{III} sensitivity rank (1-19): Can-0, Oy-0, Po-0, Rsch-4, Zu-0, Tsu-0, Wu-0, Ct-1, Wil-2, Ws-0, Bur-0, Col-0, Ler-0, Mt-0, No-0, Hi-0, Sf-2, Kn-0, Edi-0.



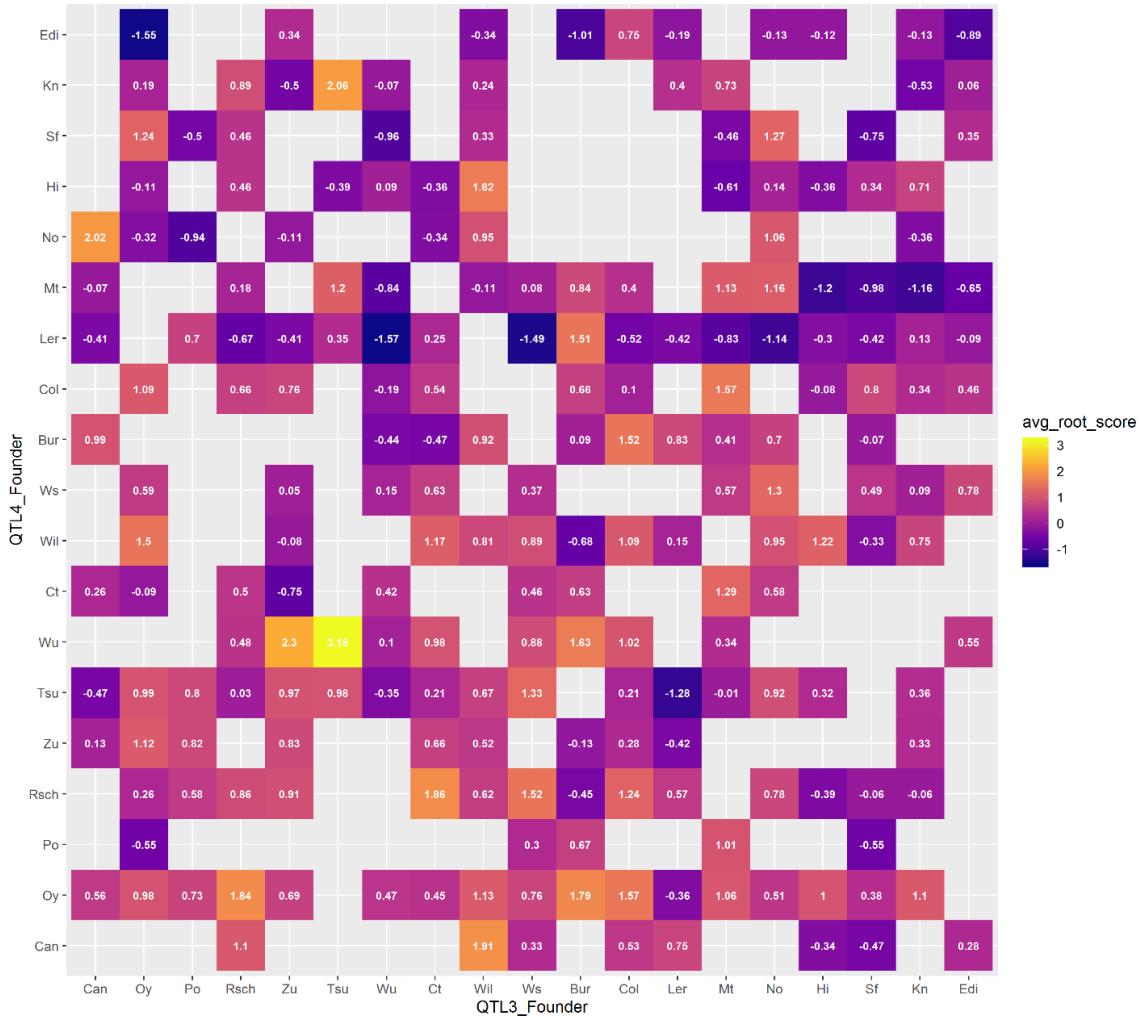
Supplemental Figure 3.2: Variants of Interest

(A) Top 7 variants at minor QTL on chromosome 3. (B) Table summary of variants and position on chromosome 3.



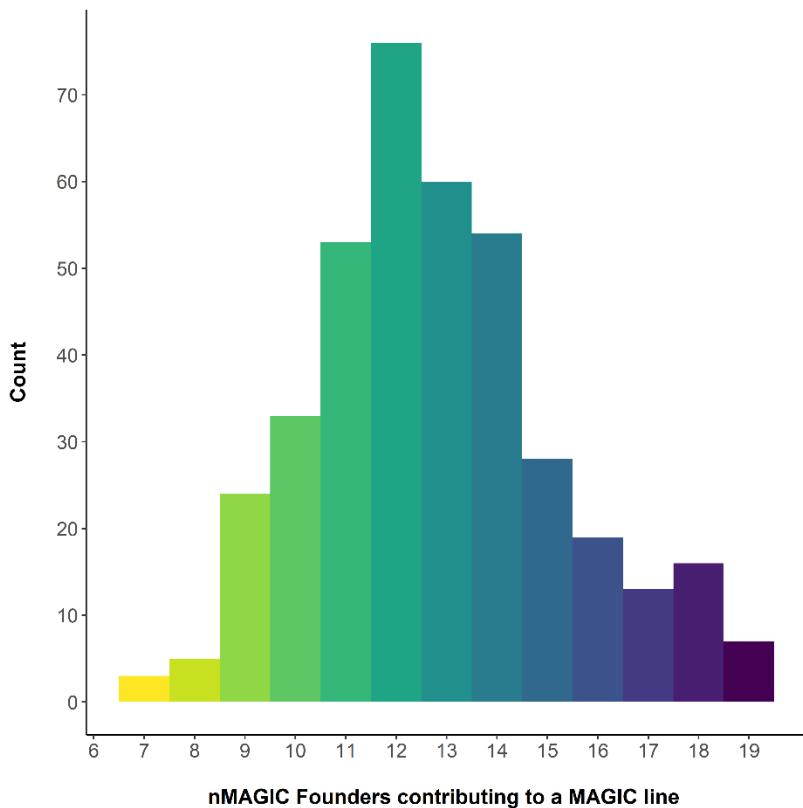
Supplemental Figure 3.3: GWAS with Condensed SNP Library Showing Association with As^{III} Tolerance

(A) Genome scan performed with the 492 root scores showing association with the 1260 SNP markers used to generate condensed genome mosaics of MAGIC lines. Two minor QTLs found on chromosome 3 and one on chromosome 4. One major QTL found on chromosome 4. A 2nd major QTL might be present upstream of the main peak. Orange dots signify peak at minor QTL, green triangle signifies major QTL, blue asterisk signifies potential QTL. -logP of 3.51 corresponds to a genome-wide p-value of <0.05. (B) Allele effect estimates for the 19 MAGIC founders at the peak SNP (MN4_10482087) of the major QTL on chromosome 4. MN4_10482087 is located 59569 bp upstream of *NIP1;1* but is the closest SNP to *NIP1;1*. The next closest SNPs to *NIP1;1* are MASC02548 and MASC01526, located 120628 bp downstream and 190423 bp upstream of the *NIP1;1* 5'UTR, respectively. Higher effect estimates associate with higher root scores and As^{III} sensitivity, lower effect estimates associate with lower root scores and As^{III} tolerance.



Supplemental Figure 3.4: Root Scores of Minor & Major QTL Haplotype Combinations

Heatmap detailing average root score of every MAGIC line founder haplotype combination at minor (QTL3) and major (QTL4) QTLs. Combinations show an uneven distribution of the two QTL loci. Many haplotype combinations are not present in the MAGIC population. For example, Ler-0 haplotype at QTL4 is well represented and found in combination with 17 of 19 other haplotypes at QTL3. The Po-0 haplotype at QTL4, in contrast, is underrepresented and only found in combination with 5 of 19 other haplotypes at QTL3, making it difficult to assess the impact of Po-0 *NIP1;1* on root scores.



Supplemental Figure 3.5: Counts of MAGIC line Founders Per Genome

Histogram counts the number of MAGIC founder haplotypes contributing to each MAGIC line of the 391 MAGIC line genomes used in Figure 2 GWAS. Compare to Figure 1 in Korver *et al.* (2009).

Chapter 4: Conclusions

At the molecular level, arsenic can influence many physiological processes in plants, including, but not limited to, proper tissue development, photosynthetic efficiency, biomass accumulation, grain filling, seed germination, root elongation and morphology, and essential nutrient acquisition.

The oxidative stress induced by arsenic, especially As^{III}, necessitates that the plant respond to rising ROS levels that can consequently damage lipids, proteins, and nucleic acids in addition to arsenic specific detoxification mechanisms for tolerating the toxic molecule itself. Determining which plant genes and proteins respond to the arsenic molecules specifically and which respond to the down-stream effects induced by arsenic stress is incredibly difficult.

Most dietary arsenic humans are exposed to comes through rice, which acquired the toxic molecules from the soil as As^{III}. To date, most transcriptomic assays for arsenic stress in plants have been conducted using As^V, which is less physiologically relevant than As^{III} for growing rice in flooded paddy fields, which is how most rice is grown around the world. Compared to these inorganic arsenic studies, even less work has been conducted using MMA, DMA, or other organic arsenic species. While As^{III} is the primary arsenical of concern, more work needs to be done to differentiate between As^{III}, As^V, MMA, or DMA specific plant responses. We must untangle this additional layer of complexity as each arsenic species is present in agricultural soils, in various concentrations.

Although many studies on arsenic in plants have been conducted and some important genes involved in arsenic transport have been revealed, many genes responsible for known processes, like non-aquaporin mediated As^{III} efflux out of the root cells, arsenic entering/exiting phloem sap, transport of arsenic into the developing grain, etc. have still

not been identified. Our field has also yet to identify any transcription factor or post-translational regulatory genes that regulate the cell's response specific to As^{III} stress and only a few genes for As^V stress. Additionally, the transcriptome data we have from these studies has mostly been generated from tissue level genomic assays.

Further studies are needed to develop crops that contain less arsenic in their grain and edible tissues, or plants capable of accumulating high levels of arsenic from arsenic contaminated soils. This thesis shows many genes within current arsenic detoxification models have cell-type specific roles in the roots. This work also elevates the importance of NIP aquaporins in determining As^{III} tolerance.

Lastly, this thesis provides opportunities for further discovery of novel genes involved in arsenic tolerance, either through the cell-type specific expression presented in Chapter 2 or by identifying the genes responsible for the minor QTLs from our GWAS in Chapter 3. We now offer our concluding thoughts from this thesis work.

Cell-Type Specific Expression in Response to As^{III}

Arsenic stress affects the expression of many genes in plants. Historically, the transcriptional response of plants to arsenic has been studied using whole tissues, such as whole roots or leaves, and most often using As^V. To begin understanding how different cell-types coordinate the tissue level transcriptional responses in previous reports, we identified 7543 significant differentially expressed genes in root epidermal, cortex, and/or endodermal cells after exposure to As^{III}, the most physiologically relevant arsenical (**Figure 2.2**).

Before narrowing the focus to individual genes, it is important to see our genome-wide expression data from multiple perspectives. **Figure 2.2** and **Supplemental Tables 2.2 & 2.3** offer a few different angles to view the 7500+ differentially expressed genes (DEGs) in this expression dataset.

In **Figure 2.2B** we set strict parameters for logFC and RPKM values to classify DEGs by individual or multiple cell-types using a Venn diagram. Our Venn diagram showed 58% of up-regulated DEGs were up-regulated in every cell-type and 50% of down-regulated genes were down-regulated in every cell-type. However, many of these genes that are “up-regulated” or “down-regulated” in all three cell-types can respond very differently in each cell. For example, one DEG with RPKM of 10 in every cell-type under control conditions might be upregulated to RPKMs of 20, 75, and 200 in epidermal, cortex, and endodermal cells, respectively. To classify DEGs using logFC and RPKM thresholds, but also by RPKM distribution across cell-types we generated **Figure 2.2A**. This allows us to identify genes that are differentially expressed in multiple cell-types, but with clear cell-type specificity. Perspective on the RPKM values of DEGs is also needed before arriving at conclusions. Through **Figure 2.2C** and **Supplemental Tables 2.2 & 2.3**, we can understand how RPKM values are distributed across cell-types under the conditions in this study. If we understand, for example, that 36-42% of up-regulated DEGs start at 0-6 RPKM in our control conditions, or that only 5-8% of up-regulated DEGs exceed 700 RPKM in our As^{III} conditions, we can begin differentiating between more important DEGs and noise. We also acknowledge RPKM value is not a true measure of biological importance. DEGs with moderate logFC and RPKM values can obviously influence overall As^{III} tolerance more than DEGs with extreme logFC or RPKM values. We hope

these multiple perspectives encourage others to further utilize this high-resolution expression data in similar or creative ways not discussed here.

Gene Ontology Analysis

As illustrated in **Supplemental Figure 2.7**, many GO terms are either over-represented or under-represented in the genome annotation. Many terms are also either too broad or too specific in their description of a biological process, cellular component, or molecular function. We also demonstrate the need for improved annotation by consolidating the many redundant terms that exist in *Arabidopsis* genome databases (**Figure 2.3** with **Supplemental Figure 2.4A**). Ranking enriched GO terms will have a larger role in future genomic studies, but only as gene annotation improves. We therefore performed GO analysis by contrasting enrichment between epidermal, cortex, and endodermal cell-types, after separating up and down-regulated DEGs, in response to As^{III} (**Figure 2.3**, **Supplemental Figure 2.4**).

Our GO term enrichment suggests cortex cells are more stressed than epidermal or endodermal cells. Cortex cells down-regulate more genes annotated by sulfur and cysteine metabolism than epidermal and endodermal cells, which might indicate less substrate is available in cortex cells for PCS1 to produce phytochelatin for As^{III} sequestration. Our ICP-MS data measuring arsenic (As), sulfur (S) in **Supplemental Figure 2.5** gives support for cortex cells sequestering less As^{III}. Cortex cells retained far less As and S (per 50,000 sorted cells) compared to endodermal or epidermal cells.

Arsenic Responsive Genes of Interest

We found that among *NIPs* and *PIPs* with relevant expression levels, that in these three cell-types, most *NIPs* are generally expressed highest in the epidermis and lowest in the endodermis, while most *PIPs* are generally expressed highest in the endodermis and lowest in the epidermis. . Our data suggests NIP1;1 is likely responsible for most As^{III} transport by NIP family aquaporins in these root cells (**Figure 2.4A**), although its role in epidermal cells is larger than in cortex and endodermal cells. Multiple *PIPs*, in contrast, might be responsible for As^{III} transport between cell-types, particularly in endodermal cells (**Figure 2.4A**). PIP1C, PIP3, PIP1;4, PIP2A, PIP1B, PIP1A, PIP2B might collectively be responsible for transporting the same substrates as NIP1;1 (including As^{III}), but in the order of endodermal >> cortex > epidermal cells.

Prior studies analyzed expression of the ABCC genes from whole roots and determined that *AtABCC1* and *ABCC2* were not induced by exposure to As^V but were constitutively expressed (Song et al., 2010). This study shows that *ABCC1* and *ABCC2* expression is regulated by As^{III} and that *ABCC2* is likely responsible for most As^{III} sequestration in these three cell-types (**Figure 2.4B**). ABCC3 has been shown to transport PC-Cd into root vacuoles (Brunetti et al., 2015), but not PC-As, as far as we know. Here ABCC3 has the highest expression of all the ABCCs in response to As^{III} (**Figure 2.4B**). The PC-Cd transport ability, along with the high expression profile, makes ABCC3 a great candidate for As^{III} transport and should be revisited in future assays.

MATE family proteins have long been considered candidates for As^{III} transport although no data exists showing any plant MATE protein performing this function.

Expression profiles of *DTX1*, *DTX9*, *DTX21*, and *DTX27* are all highly up-regulated by As^{III} and represent the best candidates in the MATE family (**Figure 2.4C**). *DTX18*, *DTX32*, and *DTX40* are also induced to a lesser degree and should be considered in future arsenic studies.

Our results reveal that As^{III} detoxification in plant roots is regulated at a cell-type specific level. Epidermal, cortex, and endodermal root cells in *A. thaliana* respond uniquely to promote detoxification pathways for the toxin. Previous work has identified many genes important for As^{III} tolerance in root tissue and this study helps clarify how those genes contribute to specific detoxification mechanisms within three different root cell-types. This work also identifies candidate genes for future assays which could contribute to As^{III} tolerance specifically in epidermal, cortex, or endodermal root cells.

MAGIC Population Identifies Multiple QTL for As^{III} Tolerance

Using the *Arabidopsis* MAGIC lines and two GWAS approaches with either low or high marker coverage, we identified two QTL (and other potential QTL) that determine As^{III} tolerance within this plant population (**Figure 3.2**, **Supplemental Figure 3.2**). The most significant (major) QTL is found on chromosome 4 and the less significant (minor) QTL is found on chromosome 3. The major QTL on chromosome 4 was the most significant QTL in both high and low coverage GWAS approaches. The most significant SNP in the low coverage GWAS and the 2nd, 3rd, and 4th most significant markers in the high coverage GWAS were near the *NIP1;1* locus (**Figure 3.2B-C**). Three of the top four

most significant markers in the high coverage GWAS were in the *NIP1;1* promoter region or a few thousand bp upstream of the *NIP1;1* promoter.

We do not expect to identify the genes responsible for most minor QTLs without new MAGIC genomes added to the population or deeper re-sequencing of the entire population. Fewer MAGIC lines were sequenced for the high coverage GWAS than the low coverage GWAS. Ultimately 100 fewer root scores were used in the higher coverage GWAS (392) than the low coverage GWAS (492), which identified additional minor QTLs on chromosome 3 and 4. We encourage researchers to re-sequence the remainder of the 527 MAGIC lines that were left out in the 3.3 million marker GWAS. Doing so would add over 100 new MAGIC genome mosaics to this study and our high resolution GWAS could then add 100 additional root scores, vastly improving the performance of our association mapping. This would likely help identify the causative genes for the minor QTLs.

***NIP1;1* Expression Responsible for Major QTL**

We found MAGIC line root scores correlate with founder haplotypes (Parental Rank) at the *NIP1;1* locus (**Figure 3.3A-B**), while no correlation existed between the minor QTL locus on chromosome 3 and Parental Rank (**Figure 3.3C-D**). To determine if variation in expression of *NIP1;1* within the MAGIC population was responsible the major QTL we measured expression of the 19 MAGIC founders prior to As^{III} exposure and after prolonged As^{III} exposure (1-5 days). Both initial *NIP1;1* expression (0dAs) and cumulative *NIP1;1* expression (0-5dAs) strongly correlated with Parental Rank (**Figure 3.4A-D**). No definitive expression pattern explains MAGIC founder As^{III} tolerance completely, but

generally those MAGIC founders with low *NIP1;1* expression prior to As^{III} exposure and/or founders that maintained low *NIP1;1* after As^{III} exposure were most tolerant. MAGIC founders with higher *NIP1;1* prior to As^{III} exposure and/or failed to down-regulate *NIP1;1* sufficiently after As^{III} exposure were most sensitive.

To confirm *NIP1;1* expression was sufficient to explain the major QTL, we cloned *NIP1;1* from Can-0, the most As^{III} sensitive founder, along with its native promoter into the As^{III} tolerant *nip1;1* mutant background (Col-0). Transgenic lines showed a variety of *NIP1;1* expression compared to WT. Transgenic lines with higher *NIP1;1* expression are more sensitive to As^{III} while lines with lower *NIP1;1* are less sensitive to As^{III} (**Figure 3.4E-F**). This confirms *NIP1;1* expression in the MAGIC founders alone can sufficiently explain As^{III} tolerance and is the most important factor determining As^{III} tolerance in the *A. thaliana* MAGIC population. It also confirmed the variation at the major QTL4 locus was not driven by differences in NIP1;1 protein function. The many different expression patterns for *NIP1;1* in response to As^{III} stress (**Figure 3.4A**) makes it unlikely any one SNP or indel within the *NIP1;1* promoter is responsible for expression variance.

Appendix I: MATE family of proteins in plants

Multidrug and Toxic Compound Extrusion (MATE) family proteins have long been considered candidates for As^{III} transport. Members of the distant clade on the bottom left of **Figure A1.1** have been implicated in aluminum (Al) and iron (Fe) tolerance via their ability to efflux citrate and malate across the plasma membrane into the apoplast (Zhou et al., 2014). Only a few members in the other clades been characterized. MATE family protein AtDTX1 was shown to increase cadmium resistance in *E. coli* (Li et al., 2002), and rice MATE proteins have some ability to modulate arsenic accumulation (Tiwari et al., 2014; Das et al., 2018), but no data exists showing any plant MATE protein directly capable of arsenic transport in yeast, oocytes, or in-vivo plant assays. Our root cell-type RNA-seq data shows that multiple MATE proteins are highly induced when plants are exposed to As^{III} (**Figure A1.2**). To determine if MATE proteins could transport As^{III} across membranes we cloned five Arabidopsis MATE genes (*DTX1*, *DTX9*, *DTX19*, *DTX21*, *DTX40*) induced by arsenic into budding yeast (**Figure A1.3**). Budding yeast strain BY4743 (Brachmann et al., 1998) and a mutant strain *acr3* (Ghosh et al., 1999) lacking the As^{III} effluxer ACR3 were controls. SM4 is an arsenic sensitive budding yeast mutant strain lacking *YCF1* and the three other vacuolar ABCC-type ABC transporters, *YHL035c*, *YLL015w*, and *YLL048c* (Ghosh et al., 1999; Leslie et al., 2004; Song et al., 2010). SM7 has the mutant background of SM4, but also expresses the wheat PCS *TaPCS1*, which can regain arsenic tolerance when complimented with an PC-As^{III} transport protein, such as AtABCC1 or AtABCC2 (Clemens et al., 1999; Song et al., 2010). ScACR3 increased tolerance to As^{III} in *acr3*, SM4, and SM7 lines, but no Arabidopsis MATE increased tolerance to any budding yeast strain (**Figure A1.3**). Plant MATEs are thought to localize to the plasma membrane or vacuole. It is possible these plant MATE proteins do not

localize to the yeast plasma membrane or vacuolar membrane, which would prevent them from transporting their normal substrates.

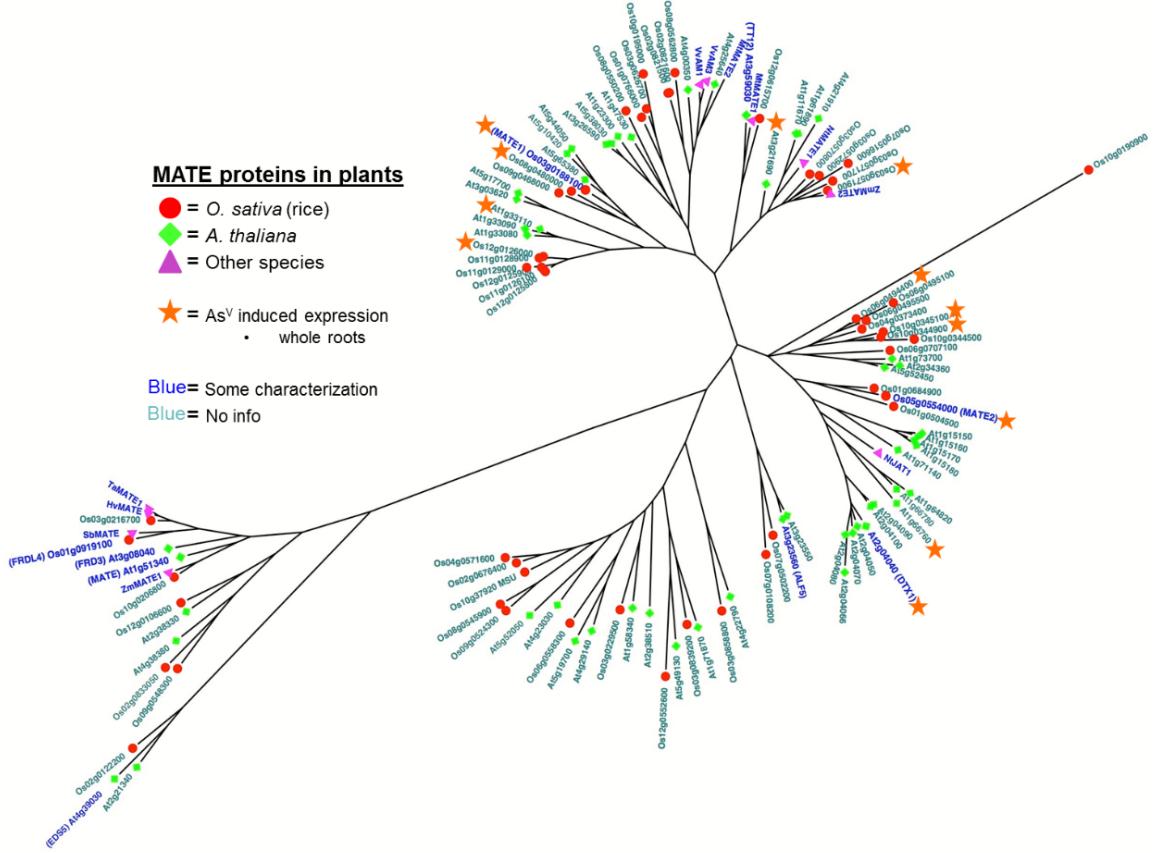


Figure A1.1: MATE protein phylogeny in selected plants

Multidrug and toxic compound extrusion (MATE) family proteins in *O. sativa* (rice), *A. thaliana*, and other selected plants. Members of the distant clade on the bottom left have been implicated in aluminum (Al) and iron (Fe) tolerance via citrate and malate efflux across the plasm membrane into the apoplast. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (-82305.76) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with

branch lengths measured in the number of substitutions per site. This analysis involved 118 amino acid sequences. There were a total of 1539 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

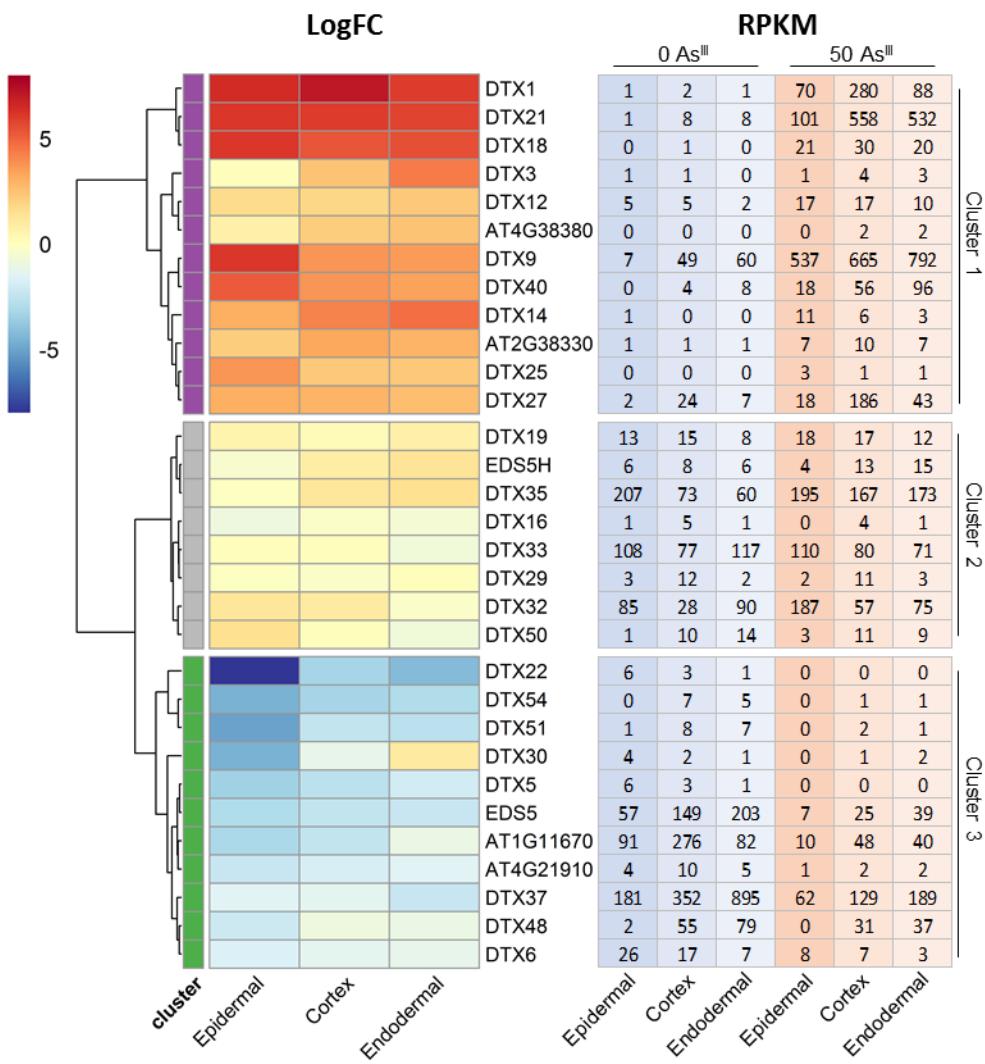
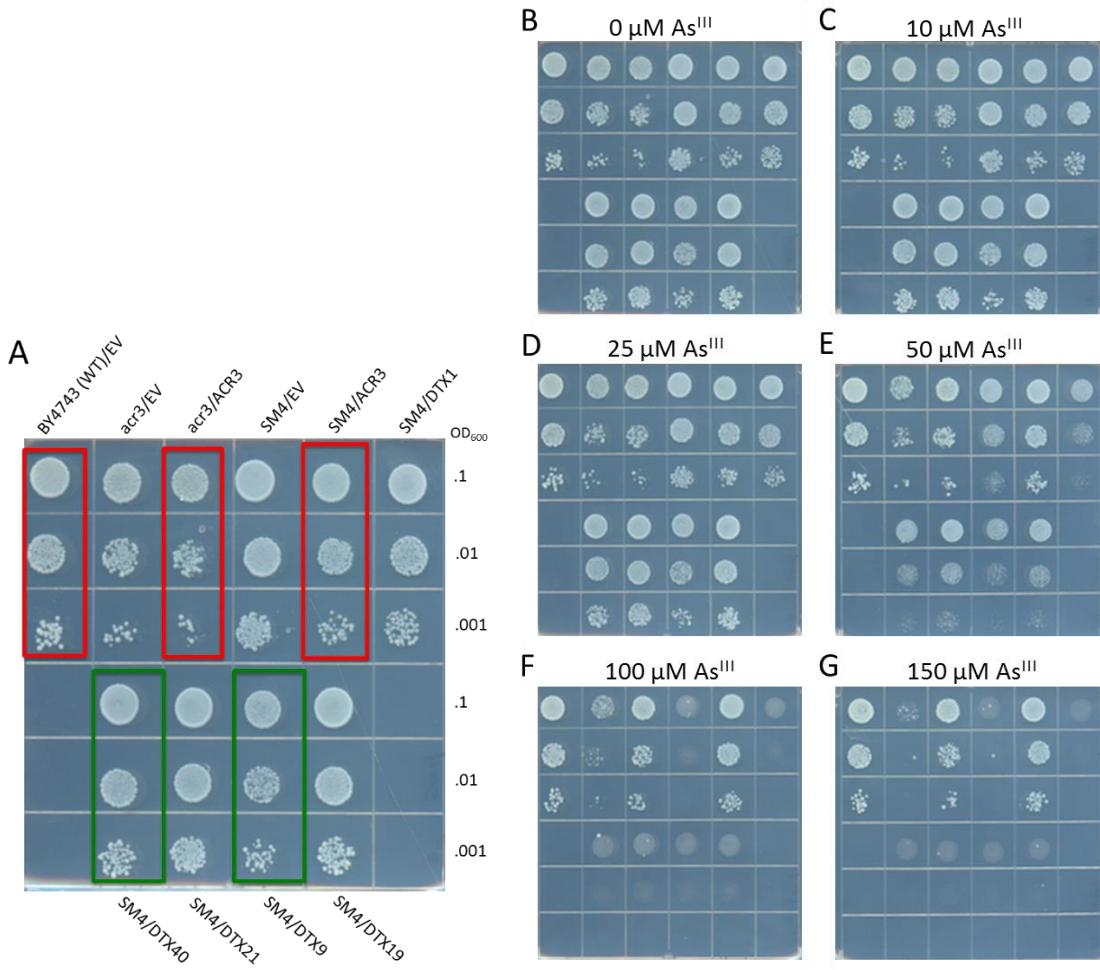


Figure A1.2: As^{III} Induced Expression by Cell-Type (MATEs)

Expression heatmaps showing differentially expressed MATE family genes. LogFC by cell-type in response to As^{III} on left and the corresponding RPKM units arranged by cell-type and condition for each gene on right. Heatmaps include genes with logFC > 1 and a RPKM value of > 2 in at least one sample. Heatmaps clustered by logFC.



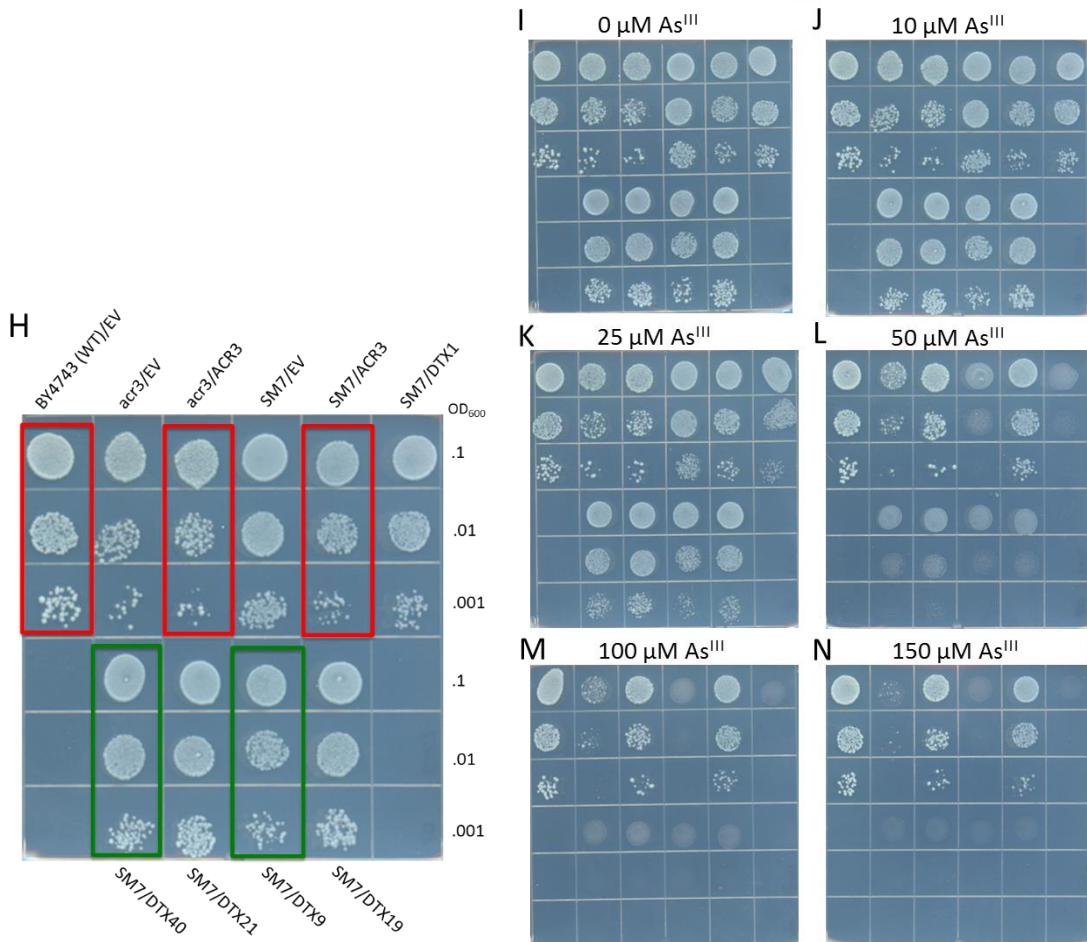


Figure A1.3: MATE yeast spot assay on As^{III}

MATE proteins DTX1, DTX40, DTX21, DTX9, and DTX19 all fail to enhance arsenic tolerance or accumulation in SM4 (A-G) budding yeast or SM7 (H-N) budding yeast expressing PCS (phytochelatin synthase) when grown on increasing concentrations of As^{III}. EV, empty vector control. ACR3, positive control for *acr3*, SM4, and SM7 yeast. Cells were grown in half-strength SD-Ura (synthetic dextrose lacking uracil) medium supplemented with 0-150 µM As^{III} for 60 hr @ 30°C. 0.1, 0.01, and 0.001 dilutions of O.D.600 (optical density at 600 nm of the yeast suspension) were initially spotted on the plates. Other evidence suggests MATE proteins fail to localize to the plasma membrane in budding yeast (data not shown). Sequences were cloned into pAG426GPD.

pAG426GPD-ccdB was a gift from Susan Lindquist (Addgene plasmid # 14156 ;
<http://n2t.net/addgene:14156> ; RRID:Addgene_14156).

Appendix II: *A. thaliana* mutants screened for As^{III} sensitivity

Sarah Jennewein contributed to genotyping and screening of the mutant lines listed in **Table A2.1** for As^{III} tolerance. Karina Lopes contributed to genotyping and screening of mutant *abc* and *pip* lines found in **Figure A2.1** and **Figure A2.2**.

Our root cell-type RNA-seq data from Chapter 2 produced many candidate genes. We were most interested in identifying transcription factors and transport proteins that showed up-regulation in the presence of As^{III}, especially in the epidermal cells as evidence suggests plant roots efflux As^{III} back into the external medium independent of the obvious NIP aquaporins. Approximately 60% to 80% of inorganic arsenic that enters rice roots is effluxed back to the external environment as As^{III} by the NIP aquaporin OsLsi1 (**Figure 1.1**) (Xu et al., 2007; Zhao et al., 2010), but OsLsi1 accounts for only 10-15% of total As^{III} efflux. Other currently unidentified transporters, therefore, must also contribute to As^{III} efflux, most likely in the outermost epidermal and exodermal root cell-types. Very few transcription factors have been implicated in responding to As^V stress and we have found no publicly available data implicating any transcription factor specifically in As^{III} detoxification mechanisms. WRKY6 and WRKY45 transcription factors regulate the expression of *Pht1;1* by limiting As^V uptake in Arabidopsis (Castrillo et al., 2013; Wang et al., 2014). In response to As^V, WRKY6 represses *Pht1;1* expression, which is accompanied by endocytosis of the PHT1;1 protein from the plasma membrane (Castrillo et al., 2013). WRKY45, in contrast, activates *Pht1;1* expression (Wang et al., 2014).

To identify novel transcription factors and transporters that contribute to As^{III} detoxification in plants, we screened more than 130 *A. thaliana* lines with single mutations in approximately 80 different genes that were up-regulated by As^{III}, most of which were preferentially up-regulated in epidermal cells (**Table A2.1**). All mutant lines were grown on 10-15 µM As^{III} along with WT (Col-0) and imaged to determine arsenic tolerant or sensitive phenotypes. No single mutant line showed an altered arsenic phenotype as clear as the *nip1;1* mutant line, which has been well documented as tolerant to arsenic. We were

able to see slight tolerance in *abcb27* and *abci19* single mutant lines (**Figure A2.1**). ABCB27 (At5g39040) is a vacuole localized half-size ABC transporter shown to be involved in aluminum resistance (Huang et al., 2012; Lei et al., 2017). ABCI19 is a soluble cytosolic protein recently implicated in modulating cytokinin responses during early seedling development (Kim et al., 2020).

Plasma membrane Intrinsic Proteins (PIPs) might also be involved with arsenic transport. Overexpression of PIP aquaporins *OsPIP2;4*, *OsPIP2;6*, or *OsPIP2;7* enhanced rice tolerance to As^{III} while increasing plant biomass (Mosa et al., 2012). We identified several *PIPs* from **Figure 2.4A** that were down-regulated by As^{III}, most in an endodermal > cortex > epidermal expression pattern. *NIP1;1* expression was also down-regulated at a similar scale, but in an epidermal > cortex > endodermal expression pattern. It is possible multiple PIP proteins collectively perform a similar As^{III} transport function as NIP1;1, but in different cell-types. *NIP1;1* expression suggests it might be more responsible for As^{III} transport in epidermal cells than other cell-types, whereas *PIP* expression suggests PIPs might be most important in endodermal cells. We obtained several double, triple, and one quadruple *pip* mutant lines to test for altered arsenic sensitivity/tolerance. We found the quadruple *pip* mutant line (*pip1A*, *pip1B*, *pip2A*, *pip2B*) to be more tolerant to As^{III} than WT (Col-0) as well as the double or triple *pip* mutant lines (**Figure A2.2**). The quadruple *pip* mutant line showed increased shoot biomass and increased root length and density compared to WT, double *pip* mutants, or triple *pip* mutants. More work must be undertaken, however, to confirm if any of these PIP proteins are capable of transporting arsenic.

Gene_ID	Gene_Name	rpkms epi_0As	rpkms epi_50As	rpkms cor_0As	rpkms cor_50As	rpkms end_0As	rpkms end_50As	logFC_epi	logFC_cor	logFC_end	conf	As mutant	pheno
AT5G08790	ANAC081, ATAF2	1471	8941	1029	3876	2652	3616	2.7	2	0.5	n	NA	
AT4G33940	RING/U-box.2	88	2570	30	440	51	374	4.9	3.8	3	y	n	
AT3G54620	BZIP25	334	1959	146	1259	208	914	2.5	3	2.2	y	n	
AT1G72175	DUF1232.2	172	1008	41	360	32	168	2.6	2.9	2.5	n	NA	
AT5G66070	RING/U-box.5	189	948	87	384	146	175	2.3	2.2	0.3	y	n	
AT5G64750	ABR1	12	825	71	471	78	437	6.2	2.9	2.6	y	n	
AT5G48655	RING/U-box.4	18	806	20	469	28	366	5.7	4.7	3.8	n	NA	
AT2G47520	ERF71	7	500	3	277	2	374	6.1	6.7	7.6	y	n	
AT3G02790	MBS1	65	469	25	73	45	74	2.9	1.7	0.9	y	n	
AT3G28120	SAP12	66	426	68	243	94	163	2.8	2	0.9	y	n	
AT3G46080	C2H2.2	156	354	21	28	19	11	1.1	0.5	-0.6	y	n	
AT2G36930	C2H2.1	12	330	7	127	9	142	4.8	4.2	4.1	y	n	
AT5G23090	NF-YB13	79	280	17	60	22	49	2	1.8	1.3	y	n	
AT5G65050	AGL31	45	272	16	117	19	87	2.7	2.9	2.4	y	n	
AT3G11200	AL2	29	270	26	181	23	163	3.3	2.9	2.9	y	n	
AT3G52800	A20/AN1-like.1	82	268	75	111	101	77	1.8	0.7	-0.3	y	n	
AT3G19860	BHLH121, URI	75	266	22	81	27	51	1.9	1.8	1.1	n	NA	
AT4G02640	BZIP10	78	241	27	150	27	75	1.7	2.4	1.5	n	NA	
AT3G09230	MYB1	14	239	7	121	5	141	4.2	4.2	4.9	y	n	
AT4G38140	RING/U-box.3	98	237	14	28	5	11	1.4	1	1.2	n	NA	
AT5G49450	BZIP1	22	224	18	83	21	54	3.2	2.3	1.4	y	n	
AT4G17500	ERF-1	44	220	43	105	78	88	2.4	1.4	0.2	y	n	
AT5G48560	BHLH78	19	207	6	83	8	73	3.3	3.9	3.3	n	NA	
AT4G30935	WRKY32	8	195	3	79	4	75	4.7	4.7	4.3	y	n	
AT5G47230	ERF5	18	191	21	107	23	72	3.5	2.4	1.6	y	n	
AT3G12480	NF-YC11	51	162	26	59	26	36	1.7	1.2	0.6	y	n	
AT2G6150	HSFA2	3	155	3	79	3	51	6	5	4	y	n	
AT2G13960	Homeodomain-like	1	143	1	29	1	30	6.7	4.9	4.3	y	n	
AT3G24050	GATA1	55	140	19	53	24	44	1.4	1.5	0.9	y	n	
AT4G18880	HSF-A4A	5	139	3	30	5	18	5	3.5	1.9	y	n	
AT5G28770	BZIP63	33	128	11	47	15	40	2.1	2.1	1.6	y	n	
AT2G24500	FZF	20	121	24	73	21	72	2.6	1.7	1.9	y	n	
AT3G16870	GATA17	5	118	4	36	5	39	4.7	3.1	3.2	y	n	
AT4G18890	BEH3	8	116	4	36	4	24	3.9	3.2	2.8	y	n	
AT5G16470	MBS2	13	114	8	35	13	36	3.2	2.2	1.6	n	NA	
AT1G22510	DUF1232.1	51	107	18	20	26	23	1.2	0.2	0	n	NA	
AT2G38880	HAP3	11	100	6	16	9	15	3.3	1.5	1.1	y	n	
AT1G14510	ALT	9	77	9	44	13	30	3.1	2.4	1.3	y	n	
AT1G50440	RING/FYVE/PHD	15	52	13	13	12	8	1.9	0.2	-0.4	y	n	
AT1G07980	NF-YC10	37	137	19	65	13	43	1.9	1.8	1.8	y	n	
AT2G40340	DREBC	7	395	6	277	5	223	6	5.7	5.7	y	n	
AT4G36930	BHLH24/SPT	13	236	4	118	4	105	4.2	4.9	5	y	n	
AT4G36780	BEH2	23	145	58	349	60	409	2.8	2.7	2.9	y	n	
AT2G23320	WRKY15	8	121	44	659	33	585	4.1	4.1	4.3	y	n	
AT2G41835	C2H2-type	7	116	3	78	2	92	3.7	4.8	5.4	y	n	
AT1G62300	WRKY6	20	49	55	134	33	90	1.3	1.3	1.6	y	n	
AT2G38470	WRKY33	174	347	84	242	95	160	1.1	1.6	0.9	n	NA	
AT3G19910	RING/U-box.6	37	100	30	225	29	167	1.5	3	2.6	y	n	
AT3G17100	BHLH147	51	508	48	840	52	739	3.4	4.2	3.9	n	NA	
AT1G34370	STOP1	38	377	37	626	28	714	3.4	4.2	4.8	y	n	
AT1G06850	BZIP52	50	353	28	515	23	362	3	4.4	4.2	n	NA	
AT5G20910	AIP2	70	299	25	424	37	431	2.2	4.2	3.7	y	n	
AT3G24120	PHL2	43	257	26	245	24	118	2.7	3.3	2.4	y	n	
AT5G04410	C2	13	234	20	293	22	331	4.3	4	4	y	n	
AT4G17490	ERF6	22	214	32	307	27	264	3.3	3.1	3.2	y	n	
AT2G24260	BHLH66	18	139	16	224	36	371	3	3.9	3.4	n	NA	
AT5G67580	TBP3, TRB2	6	112	8	91	9	52	4.3	3.5	2.7	y	n	
AT5G54680	BHLH105/ILR3	15	97	21	142	22	141	2.8	2.8	2.7	n	NA	
AT3G46600	GRAS TF	14	89	26	135	33	167	2.6	2.5	2.5	n	NA	
AT4G36920	AP2	29	74	17	126	21	87	1.4	3	2.3	y	n	
AT3G04670	WRKY39	27	70	23	71	23	90	1.5	1.8	2.1	n	NA	
AT3G10490	SGS1	10	70	6	94	7	152	2.9	4	4.6	y	n	
AT1G53320	TLPL	21	66	43	144	33	110	1.8	1.9	1.9	y	n	
AT1G13960	WRKY4	18	60	19	56	21	66	1.8	1.7	1.8	y	n	
AT1G14685	BPC TF	11	56	12	60	12	73	2.3	2.3	2.6	y	n	
AT1G04990	ZINC FINGER	27	28	33	149	24	106	0.2	2.2	2.3	y	n	
AT4G29100	BHLH68	3	18	16	120	21	240	2.7	2.9	3.6	y	n	
AT4G00150	HAM3	4	5	30	258	18	421	0.5	3.1	4.6	n	NA	
AT1G78700	BES1/BZR1	59	97	58	172	27	126	0.7	1.6	2.3	n	NA	
AT4G36990	HSF4, HSFB1, TBF1	14	85	75	206	212	102	2.7	1.6	-0.9	y	n	
AT5G08130	BHLH46/BIM1	25	16	36	95	22	144	-0.6	1.5	2.7	n	NA	
AT2G04040	DTX1	1	76	1	232	1	72	6.8	7.8	6.3	y	n	
AT1G33110	DTX21	1	110	6	462	5	435	6.6	6.5	6.7	y	n	
AT1G66760	DTX9	8	586	37	551	42	648	6.2	4.2	4.1	y	n	
AT1G59870	ABCG36, PDR8, PEN3	4	97	4	12	7	7	4.3	1.7	0.3	y	n	
AT3G53480	ABCG37, PDR9, PIS1	5	52	2	18	1	5	3.5	3.1	2.3	y	n	
AT4G19030	NIP1;1	2425	234	148	53	185	22	-3.3	-1.4	-3	y	y	
AT2G34390	NIP2;1	1	28	1	7	1	3	5.7	4.2	2.5	n	NA	
AT1G01620	PIP1;3	1242	1459	727	2053	2453	1936	0.2	1.4	-0.3	y	n	
AT3G04090	SIP1;1	77	208	59	276	76	443	1.5	2.2	2.6	y	n	
AT5G39040	ABC-B2, ALS1	23	491	14	564	7	623	4.5	5.5	6.7	n	NA	

Table A2.1: Single mutant As^{III} phenotyping

Table of genes selected based on cell-type expression profile. Columns 3-8 are RPKM values for epidermal, cortex, and endodermal cell-types after 24 hours exposure to 0 µM As^{III} or 50 µM As^{III}, as described in Chapter 2 of this thesis. Columns 9-11 are logFoldChange values for epidermal, cortex, and endodermal cell-types after As^{III} exposure. Column 12 answers whether a homozygous mutant in the coding region of gene was attained and confirmed by PCR. Column 13 answers whether confirmed single homozygous mutants displays an arsenic sensitive or tolerant phenotype when grown for 14 days on 0.5x MS medium supplemented with 10 µM As^{III}. No arsenic sensitive or tolerant single mutant phenotypes were found besides the *nip1;1* mutant.

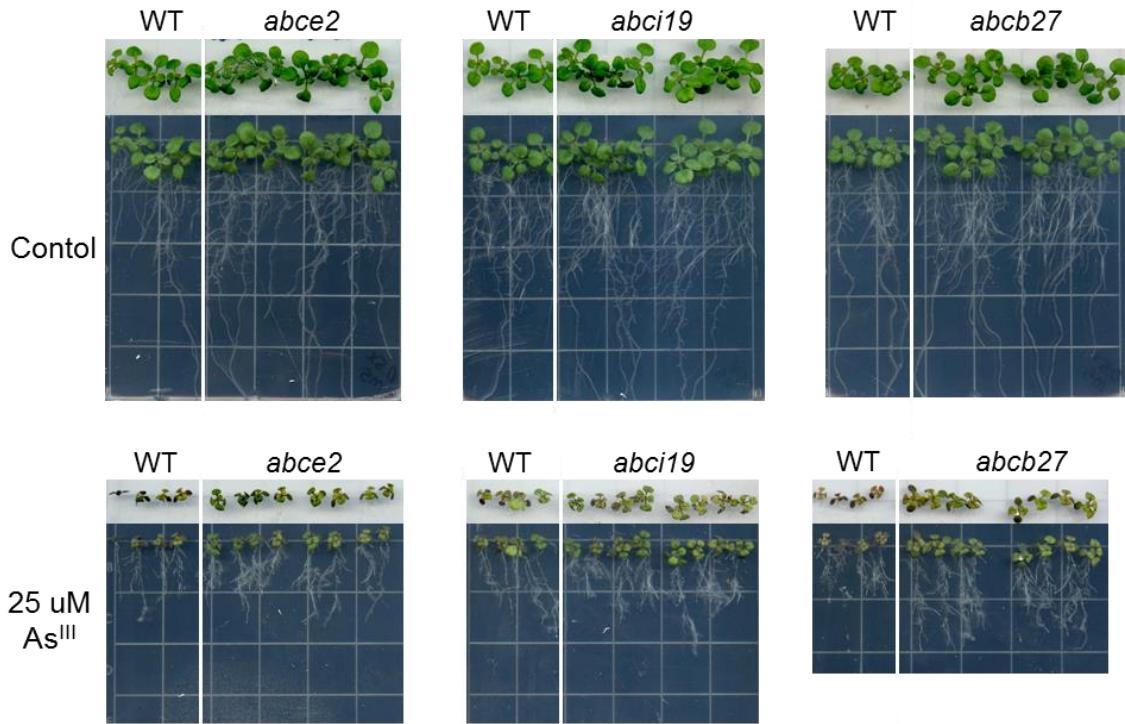


Figure A2.1: Single *abc* mutant lines show elevated As^{III} tolerance

WT (Col-0) and *abc* mutant plants were grown for 14 days on 0.5x MS medium (Control) or medium supplemented with 25 μ M As^{III}.

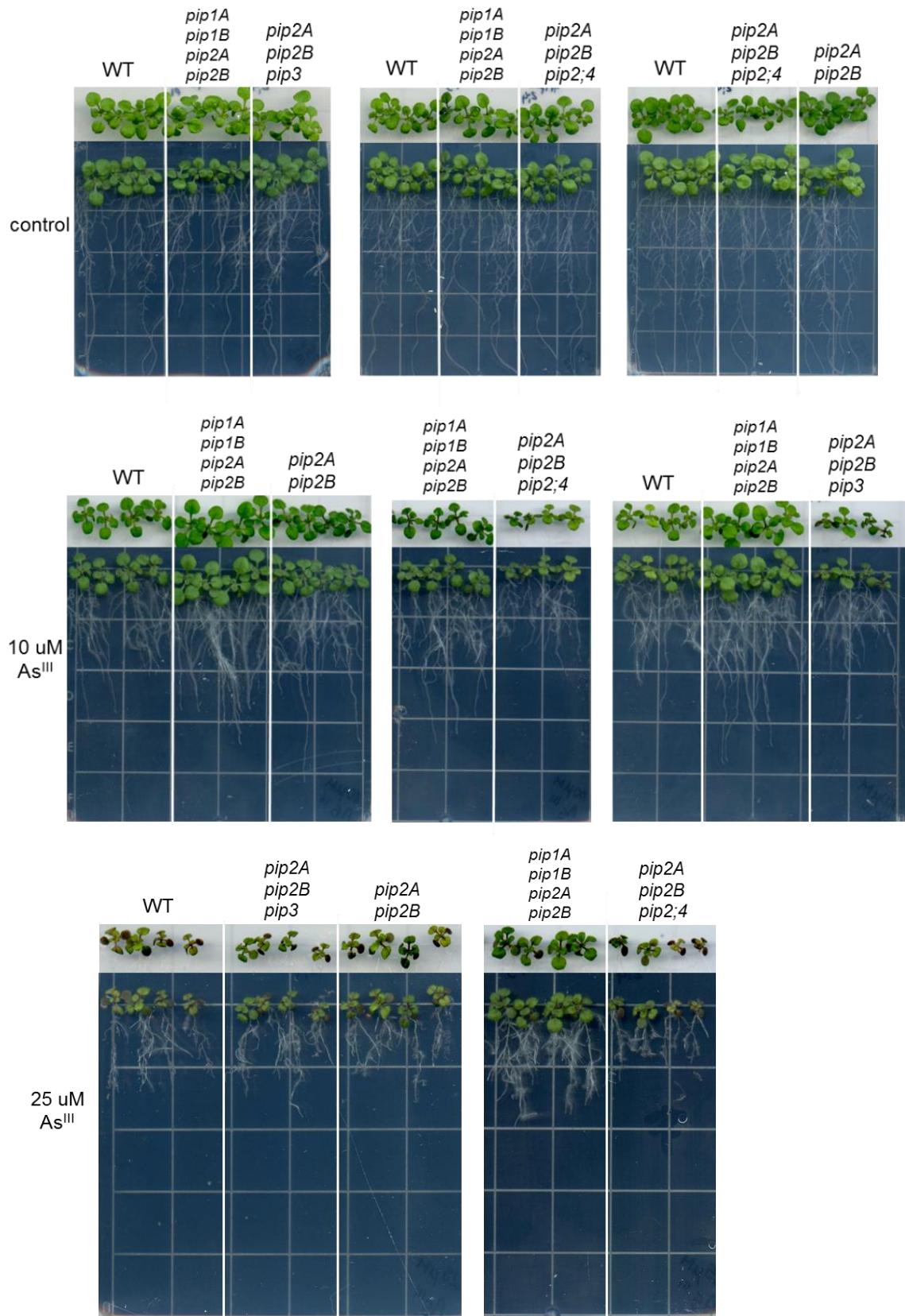


Figure A2.2: Quadruple *pip* mutant line shows elevated As^{III} tolerance

WT (Col-0) and *pip* mutant plants were grown for 14 days on 0.5x MS medium (Control) or medium supplemented with 10 or 25 μ M As^{III}.

Appendix III: Understanding arsenic dynamics in agronomic systems to predict and prevent uptake by crop plants

All authors contributed to the writing, with Tracy Punshon serving as corresponding author.



Understanding arsenic dynamics in agronomic systems to predict and prevent uptake by crop plants



Tracy Punshon ^{a,*}, Brian P. Jackson ^b, Andrew A. Meharg ^c, Todd Warczack ^a, Kirk Scheckel ^d, Mary Lou Guerinot ^a

^a Dartmouth College, Department of Biology, 78 College Street, Hanover, NH 03755, USA

^b Dartmouth College, Department of Earth Sciences, Hanover, NH 03755, USA

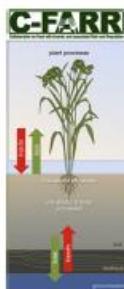
^c Institute for Global Food Security, Queen's University Belfast, Belfast BT9 5HN, United Kingdom

^d USEPA Office of Research and Development, National Risk Management Laboratory, 26 West Martin Luther King Drive, Cincinnati, OH 45224, USA

HIGHLIGHTS

- Consumption of staple foods such as rice, apple juice and vegetables grown in contaminated soil is now recognized as a tangible route of human exposure to arsenic
- Arsenic occurs in food because it is present in the soil and water and is taken up by crop plants.
- Understanding the sources of arsenic to crop plants and influence the dynamics of the agronomic arsenic cycle are key to reducing crop uptake of arsenic now, and preventing exposure in future.
- This review considers natural and anthropogenic sources of arsenic to the soil, biogeochemical cycling, rhizosphere processes, plant processes, and mitigation strategies
- This review recommends: mobilizing existing soil data so that it is readily accessible to commercial and private growers; expanding detailed soil monitoring; reconsideration, unification and enforcement of action levels for agricultural soil arsenic based on updated science, community outreach and education about the potential for arsenic in the soil, as necessary steps to protect valuable soil resources.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 19 August 2016

Received in revised form 16 December 2016

ABSTRACT

This review is on arsenic in agronomic systems, and covers processes that influence the entry of arsenic into the human food supply. The scope is from sources of arsenic (natural and anthropogenic) in soils, biogeochemical

* Corresponding author.

E-mail addresses: tracy.punshon@dartmouth.edu (T. Punshon), bjp@dartmouth.edu (B.P. Jackson), aa.meharg@qub.ac.uk (A.A. Meharg), Todd.M.Warczak.GR@dartmouth.edu (T. Warczack), scheckel.kirk@epa.gov (K Scheckel), mary.lou.guerinot@dartmouth.edu (M.L. Guerinot).

Accepted 16 December 2016
Available online 30 December 2016

Keywords:
Arsenic
Sources
Soil
Agriculture
Plants
Mitigation

and rhizosphere processes that control arsenic speciation and availability, through to mechanisms of uptake by crop plants and potential mitigation strategies. This review makes a case for taking steps to prevent or limit crop uptake of arsenic, wherever possible, and to work toward a long-term solution to the presence of arsenic in agronomic systems. The past two decades have seen important advances in our understanding of how biogeochemical and physiological processes influence human exposure to soil arsenic, and this must now prompt an informed reconsideration and unification of regulations to protect the quality of agricultural and residential soils.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Consumption of staple foods such as rice, beverages such as apple juice, or vegetables grown in historically arsenic-contaminated soils are now recognized as tangible routes of arsenic exposure. The presence of elevated concentrations of arsenic in the soil is not a pre-requisite for dietary arsenic exposure; seen in the accumulation of arsenic by rice grown in uncontaminated soils (Norton et al., 2012). When drinking-water arsenic concentrations are low, dietary arsenic can be a significant exposure (Carlin et al., 2015). Understanding the sources of arsenic to crop plants and the factors that influence them is key to reducing human exposure now and preventing exposure in future. In addition to the abundant natural sources of arsenic, there are a large number of industrial and agricultural sources of arsenic to the soil; from mining wastes, coal fly ash, glass manufacturing, pesticide application, wastewater sludge, pharmaceutical waste, livestock dips, smelting activities to phosphate fertilizers. Plant uptake of arsenic was previously assumed to be too low to merit setting limits for arsenic in food crops, but given that measurable biological effects occur at arsenic levels below the current maximum contaminant level (MCL) for drinking water (Bodwell et al., 2004), these low levels can still translate into significant exposures, particularly in children (Davis et al., 2012) and presumably in adults who consume a lot of rice. In response, the World Health Organization (WHO) set an advisory MCL for inorganic arsenic in white (polished) rice of 0.2 mg/kg (WHO, 2016) along with the limit of 10 µg/L in water, and the European Union set similar standards that included a lower MCL (0.1 mg/kg) for rice-containing baby foods (European Union, 2015). Currently, dietary arsenic exposure is suspected to play a role in cardiovascular disease in adults (Moon et al., 2012), and to disrupt the glucocorticoid system (involved in learning and memory) to those exposed in utero (Caldwell et al., 2014). An in-depth review of the current findings on the relationship between dietary arsenic exposure and human health is provided by Davis et al. (this issue).

In the United States, regulations on arsenic are distributed to several agencies. The Environmental Protection Agency (EPA) developed the MCL for arsenic in drinking water (10 µg/L) in 2006; a level supported by the World Health Organization, Canada and the European Union. In the state of New Jersey (USA) the limit is 5 µg/L, and in Australia, 7 µg/L. Many other nations still adopt a level of 50 µg/L (Bahrain, Bangladesh, Bolivia, China, Egypt, India, Indonesia, Oman, Philippines, Saudi Arabia, Sri Lanka, Vietnam, Zimbabwe) (Yamamura et al., 2001), with the exception of Mexico (35 µg/L). In the USA, the Food and Drug Administration (FDA) is responsible for setting action levels for arsenic in food, which includes apple and pear juice at 10 µg/L, in line with EPA's drinking water MCL. In Canada, the Canadian Food Inspection Agency issued alerts on excessive arsenic in rice and pear products in 2014. Consistent with the European Commission's limit for arsenic in rice used in food production for infants and young children, the FDA is proposing an action level of 0.1 mg/kg for inorganic arsenic in infant rice cereal (FDA, U., 2016). Foods in Australia and New Zealand may not contain >1 mg/kg dry mass of arsenic, and salt for food use must not contain >0.5 mg/kg. Japan has a limit of 15 mg/kg of arsenic in paddy soils (Japan, 2016). Likewise, Thailand has an agricultural arsenic soil quality standard of 3.9 mg/kg. Within the USA, states differ widely in their action levels for arsenic in soil, for instance New Jersey has a cleanup

criterion of 20 mg/kg and Florida has a cleanup target level of 2.1 mg/kg and 12 mg/kg for industrial sites (Henke, 2009).

Arsenic occurs in food because it is present in soil and water and is taken up by plants. This review article brings together the latest scientific information on arsenic in agronomic systems, describing its sources in soils and the processes that influence the uptake of arsenic by crop plants. The intention of this review is to prompt a reconsideration and unification of government regulations on action levels for arsenic in agricultural soil; raise awareness of how both former and ongoing inputs of arsenic to soil can result in food contamination and impacts to human health and finally, to indicate the way forward for mitigation strategies that safeguard valuable soil resources.

2. Natural sources of arsenic in soil

Below toxic concentrations, the higher the total soil arsenic concentration (the sum of all arsenic species, regardless of bioavailability) the higher the crop uptake of arsenic. This is true of anaerobic cultivation systems such as rice (Adomako et al., 2009; Lu et al., 2009; Williams et al., 2007), aerobic horticultural systems (Norton et al., 2013) as well as conventional (aerobic) agriculture (Williams et al., 2007). The global average total soil arsenic concentration is 5 mg/kg, (equivalent to parts per million), but there is large variation between and within geographical regions (Koljonen et al., 1989). Where soils have formed on arsenic-rich bedrock, or downstream of these bedrocks, very high concentrations of natural arsenic can result. Concentrations of up to 4000 mg/kg arsenic have been measured in soils from the arsenopyrite belt (iron arsenic sulfide, FeAsS) in Styria, Austria (Geislinger et al., 2002), for instance. There are approximately 568 known minerals that contain arsenic as a critical component (IMA, 2014). Arsenic is present in many rock-forming minerals because it can chemically substitute for phosphorus (V), silicate (IV), aluminum (III), iron (III) and titanium (IV) in mineral structures. Global mapping data of total arsenic concentrations in topsoil is not available, although large-scale regional maps are available for soil arsenic concentrations in Europe (Lado et al., 2008) and the USA (Shacklette and Boerngen, 1984). European data predicts that most soils range <7.5–20 mg/kg arsenic, with a median of 6 mg/kg (Lado et al., 2008). This prediction comes from block regression-kriging, a spatial prediction technique based on regressing soil arsenic concentrations against auxiliary variables, and is useful because it uses a particularly high resolution (block size of 5 km²). On a continental scale, large zones of soils with approximately 30 mg/kg arsenic have been found in southern France, the north-eastern Iberian Peninsula and south-west England, with the two latter being zones of extensive natural mineralization associated with base and precious metal mining activities. The United States Geological Survey (USGS) soil sampling of the contiguous USA reports a mean soil arsenic concentration of approximately 5 mg/kg with 5 and 95 percentile values of approximately 1.3 and 13 mg/kg respectively (Smith et al., 2014). Large regional patterns are apparent in the data, for example the soils of New Hampshire have soil arsenic concentrations of approximately 10 mg/kg arsenic, and Florida, 3.5 mg/kg. The sampling density goal for the USA surface soils and stream sediments database is 1 per 289 km² (USGS, 2016), but is currently at only 1 sample per 1600 km². This contrasts with smaller regional surveys such as the recently published Tellus database for

Northern Ireland that has a sampling density of 2 km² (Young and Donald, 2013) (median total soil arsenic concentration 8.7 mg/kg). At this sampling density, fine-scale data for factors shown to affect soil arsenic, such as bedrock type, altitude and organic matter for instance, can be observed, providing the opportunity to make predictions about arsenic bioavailability and mobility.

Soil or sediment arsenic concentrations are the result of the complex and dynamic interplay between inputs and outputs (Smedley and Kinniburgh, 2002). Natural sources of arsenic to agronomic catchments are dominated by bedrock weathering (mechanical, chemical and biological) and depositional inputs, with the ultimate sinks at the base of catchments often being a significant distance from sources (Saunders et al., 2005). Outputs include leaching into water bodies (vertically and horizontally), soil erosion (Smedley and Kinniburgh, 2002) and bioturbation (Mestrot et al., 2011). In arid regions surface evaporation of water can lead to arsenic enrichment from the draw up of subsurface water (Smedley and Kinniburgh, 2002) and from waters used in crop irrigation (Lawgali and Meharg, 2011). Mass-balances (accounting for all inputs and outputs for a particular ecosystem) are rarely conducted for arsenic fluxes within catchment areas, but a good example is from a mining-impacted catchment area (Melegy et al., 2011), where chemical weathering followed by mechanical weathering dominated arsenic inputs, which were primarily from arsenopyrite. Similarly, in a gold-mining region, weathering contributed an estimated 95% of the arsenic (Drahota et al., 2006). In a forested catchment area, where atmospheric arsenic inputs were the dominant source to highly organic soil (soils with >10% organic matter), inputs of arsenic via precipitation were ~6 g/ha/y (Huang and Matzner, 2007), and organic soils were a net source of arsenic, while mineral soils (<10% organic matter) were a sink. This agrees with depositional inputs of arsenic measured in the UK, which ranged from ~1 to ~10 g arsenic/ha/y (CEH, 2008). UK regional scale maps show that arsenic deposition is highest at altitude and in the west of the country; the least polluted regions with air masses originating in the Atlantic. This suggests a marine source of arsenic. Depositional maps relate well to soil arsenic maps such as in maps of Northern Ireland and England (UKSO, 2016) that show highest arsenic concentrations in peat soils at higher altitude, along with bedrock geological anomalies. Peat soils at higher altitude are sinks for arsenic, and become sources if the peat is mineralized or eroded. The topic of upland organic soils acting as sinks and sources of arsenic is receiving more research attention (Mikutta and Rothwell, 2016), and could be important on a regional scale as a source of arsenic to downstream sediments.

In large catchment areas of continental importance, such as the deltas that form to the south and east of the Himalayas, plate tectonic-derived mechanical weathering is thought to be the most important source of arsenic. One theory is that the mechanical weathering caused by Pleistocene tectonic uplift in the Himalayas is the key to understanding why arsenic is so elevated in Holocene aquifers, such as those of SE Asia, and in the glacial tills of Europe and North America (Saunders et al., 2005). Mechanical weathering of bedrock exposes previously inaccessible mineral surfaces, and the finer grinding leads to enhanced surface areas for chemical and microbial weathering to take place, causing greater solubilisation of arsenic (Smedley and Kinniburgh, 2002; Saunders et al., 2005; Mailloux et al., 2009). Chemical and microbial weathering can take place at or near the source, or in sediment sinks. For instance, bacteria isolated from Bay of Bengal aquifers can mobilize arsenic from apatite (Mailloux et al., 2009) (see Section 5). Invariably, the arsenic loadings into soil will be dependent on arsenic in the bedrock, and the extent of the weathering of that bedrock-derived material along the route from source to sink. Soils with basalt bedrock had the lowest median arsenic content, while those with psammite, semipelite, and lithic arsenite bedrocks had the highest. Interpretation of such fine-scale mapping can ultimately lead to predictions of soil arsenic concentrations where detailed maps are not available. Combined with an understanding of soil chemistry, this will enhance the ability to predict elevated concentrations of arsenic in crops (Williams et al., 2011).

3. Anthropogenic sources of arsenic to soil

Many anthropogenic activities have increased soil arsenic concentrations above the natural, background levels mentioned in Section 2 above, and they have the potential to increase the arsenic concentration in food. This is especially the case in the USA where the widespread use of arsenic-based herbicides, pesticides and livestock antibiotics throughout the 20th century has ultimately increased the arsenic concentrations of current productive USA agricultural soils (Murphy and Autcott, 1998; Lasky et al., 2004; Nachman et al., 2013).

3.1. Base and precious metal mining

The dominant mineral source of arsenic is thought to be pyrite (iron sulfide, FeS₂) (Fendorf and Kocar, 2009), an economically important ore deposit. High arsenic concentrations are found in many oxide minerals and hydrous metal oxides, either part of their structure or as sorbed and occluded species (Bowell et al., 2014a). Iron oxides accumulate arsenic up to concentrations of several weight percent (1 wt% being equivalent to 10,000 mg/kg), and arsenic tends to bind to iron (III) hydroxides whenever they are present. Arsenic is found predominantly as arsenopyrite but also can occur as orpiment (arsenic trisulfide As₂S₃), realgar (α -As₄S₄) and other arsenic sulfide minerals (Craw and Bowell, 2014; Bowell et al., 2014b). Arsenic is a byproduct of most mining operations and is present at high concentrations in the mine waste, and, because arsenic sulfides are particularly prone to oxidation in surface environments, in mining wastewaters (Craw and Bowell, 2014; Bowell et al., 2014c). Arsenic can constitute 1% or more of the ore and solid waste, and wastewaters and impacted streams often contain dissolved arsenic concentrations ranging from 0.01 to over 10 mg/L. Because mining and smelting operations are localized, arsenic contamination of soils exists around the mine site with the concentration decreasing with distance from the source. Windblown dispersion of fine particulate material is a particular problem, spreading contamination greater distances from the mine site. This fine material – which is not completely removed by washing (Norton et al., 2013) – can directly contaminate plant material; especially leafy material with high surface area. This presents a tangible risk to residents and home gardeners in the vicinity of areas with significant surface soil arsenic contamination. A comparison of arsenic concentrations in vegetables grown in SW England (the site of historic mining activities) with those from a pristine site in NW Scotland found a generally good correlation between total plant arsenic and soil arsenic concentrations. Increased arsenic concentrations were measured in produce from SW England where soil arsenic concentrations ranged from 120 to 1130 mg/kg. Arsenic concentrations were high in leafy greens (kale, spinach, lettuce) and some unpeeled vegetables (potatoes, swedes, carrots) were higher than when peeled, which, in both cases, points to contamination from windblown soil particles and soil adhesion to below ground biomass, rather than from root uptake. In this particular study, the majority of arsenic was present as the inorganic form (Norton et al., 2013). Similar results were obtained from home gardens near the Iron King Mine Superfund Site in Arizona, USA (Ramirez-Andreotta et al., 2013a; Ramirez-Andreotta et al., 2013b). Here the tailings had arsenic concentrations of 3710 mg/kg and residential soil sampled adjacent to the site ranged from 120 to 633 mg/kg. Edible plant tissue concentrations ranged from <0.01 to 1.96 mg/kg (plant concentrations are expressed as dry weight throughout), and were generally positively correlated with soil arsenic concentrations. Leafy and high surface area vegetables such as lettuce, kale, broccoli and cabbage accumulated higher arsenic concentrations than beans, tomatoes, cucumbers and peppers. Arsenic in mine-affected vineyard soils in Italy ranged from 4 to 283 mg/kg and positive correlations were observed between soil concentrations and arsenic levels in vine leaves and grapes, however, levels in wine were low (<1.62 µg/L) (Bertoldi et al., 2013). In the Hunan province, China, the high levels of inorganic arsenic in rice have been traced

to mining activities in the area (Ma et al., 2016; Williams et al., 2009; Zhu et al., 2008).

3.2. Coal combustion for energy

The concentration of arsenic in USA coal ranges from 1 to 71 mg/kg with an average concentration of 24 mg/kg (Kolker et al., 2006). Fly ash, the major byproduct of the coal combustion process, consists of fine particles that are driven out with the flue gases, and is a major source of arsenic to the wider environment. Coal ash is one of the most abundant of industrial wastes; close to 130 million tons (EPA, 2016) of coal fly was generated in the USA in 2014, with 100 million tons estimated from the European Union in 2011 (Feuerborn, 2011). Arsenic concentrates in the fly ash during combustion of coal for energy; the median arsenic concentration in USA fly ash is 71 mg/kg (Yager et al., 2015). Fly ash is often sliced into settling basins, and because arsenic in fresh ash is quite soluble, wastewater arsenic concentrations can consequently be quite high. Arsenic can build up in the sediments of coal fly ash settling basins and reach concentrations of over 1000 mg/kg. Catastrophic failures of these settling basins have caused severe environmental problems and contaminated surface waters with arsenic (Schmidt, 2010). There is a well-founded concern that arsenic from coal combustion wastes can contaminate soil and enter the food supply. The use of coal fly ash as a soil amendment can lead to elevated arsenic concentrations in crops (as well as boron, selenium and molybdenum), although its lack of soil macronutrients and the potential for arsenic toxicity prevents the sole application of coal fly ash as a soil amendment (Schumann and Sumner, 1999; Jackson et al., 1999). Formulating ash/organic waste mixtures that conform to USEPA regulations for total arsenic application and meet soil and plant fertility requirements has been shown to safe and effective for agronomic use (Schumann and Sumner, 2004).

3.3. Pesticides

Perhaps the largest anthropogenic input of arsenic to agricultural soils in the USA is from the agricultural use of arsenic-based pesticides and herbicides for most of the 20th century. Calcium arsenate and lead arsenate were used extensively up to the 1950s, mostly on orchard soils to combat the codling moth. At peak, 132,000 metric tons of each pesticide compound was applied annually between 1930 and 1940 (Murphy and Aucott, 1998). In addition to apples, inorganic arsenic pesticides were used on a range of crops including essentially all fruit trees, vine berries, sweet potatoes, white potatoes, most vegetables and cotton (Murphy and Aucott, 1998). Both lead and arsenate have long residence times in soils and high concentrations (often > 100 mg/kg) of these two elements have been reported in old orchard soils in Washington (Yokel and Delistraty, 2003), North Carolina (Embrick et al., 2005), New Hampshire (Renshaw et al., 2006), New Jersey (Peryea and Creger, 1994) and Virginia (Robinson et al., 2007). There is some evidence of greater mobility for arsenic (than lead) (Renshaw et al., 2006; Peryea and Kammerer, 1997), and retention of both elements depends on soil type and other environmental factors but most of this legacy contaminant remains in the soil (Peryea and Creger, 1994). Use of lead arsenate decreased after 1950s and was finally banned in 1988. The organic arsenic compounds dimethylarsinic acid (DMA) and monomethylarsonic acid (MMA) were used as pesticides on cotton and herbicides for golf courses and right-of-ways until they too were withdrawn from use in 2013. High levels of MMA were reported in transient surface waters adjacent to a crop sprayer operation (Bednar et al., 2002). Legacy soil arsenic contamination resulting from organic arsenical pesticides plays a major role in straighthead disease of rice (Rahman et al., 2008) (see Section 6.1). It may be that arsenical pesticides have leached to groundwater, as has been suggested for the Texas High Plains Aquifer (Hudak, 2000), although a study of the Ogallala aquifer in the High Plains in Texas found no evidence of anthropogenic arsenic in

the groundwater (Reedy et al., 2007). Similarly, there was no relationship between groundwater arsenic and past (inorganic) arsenic pesticide usage in a comprehensive study of New Hampshire groundwater sources (Ayotte et al., 2006). About 10%, depending on soil substrate, of monosodium methyl arsenate applied to sandy soils (simulated golf course greens) leached into percolating water. Demethylation and methylation occurred because both inorganic arsenic species and DMA were also detected in the percolating water (Feng et al., 2005). As with mining-impacted soils, plants grown on soils that are high in arsenic from arsenical pesticide contamination take up higher levels of arsenic into their edible tissues, observed for example in potatoes (Codling et al., 2016), carrots (Codling et al., 2015) and leafy green vegetables (McBride et al., 2015; Lim and McBride, 2015).

Former pesticide application has been suggested be a factor in the presence of higher levels of total arsenic found in rice grown in the south-central regions of the USA (Zavala and Duxbury, 2008; Zavala et al., 2008) compared to other areas of the USA and to other countries, such as Bangladesh (Williams et al., 2005a). Evidence on varietal differences in arsenic uptake, speciation and distribution within rice grain (see also Section 6) strongly suggest that soil arsenic concentration is not the sole, nor particularly the main driver of this phenomenon. Factors likely to be influential may also include the differences in the soil microbial community composition between geographical regions that affect arsenic methylation, considering that plants themselves cannot methylate arsenic (Lomax et al., 2012).

3.4. Wood preservatives

Chromated copper arsenate (CCA) is used as a wood preservative and was extensively used on decking and other residential usages until a voluntary manufacturer withdrawal in 2003. The primary health concern is for young children in direct contact with CCA-treated wood, but localized leaching of arsenic (as well as chromium and copper) also occurs to surrounding soil. Soil arsenic concentrations of 37–250 mg/kg have been reported for soils sampled near CCA-treated utility poles ($N = 12$) (Poushat and Zagury, 2006) and mean arsenic concentrations for soils collected below decks and footbridges in Florida, USA was reported to be 28.5 mg/kg compared with a control concentration of 1.3 mg/kg ($N = 65$) (Townsend et al., 2003). Arsenic from CCA contaminated soils appears to be more bioavailable than from other anthropogenic sources to soil (Juhasz et al., 2011).

3.5. Organic manures

Land application of sewage sludge (biosolids) in the USA is regulated by Environmental Protection Agency Part 503 Biosolids rule; which set the maximum arsenic concentration of the sludge at 75 mg/kg, an annual pollutant-loading rate of 2.0 kg arsenic/ha (kg/ha) and a cumulative pollution-loading rate of 41 kg/ha over the lifetime of applications. Assuming a plow layer of 17 cm, application at the maximum annual rate implies an approximate 1.2 mg/kg maximum increase in soil arsenic, while the cumulative maximum loading rate could increase soil arsenic concentrations by approximately 24 mg/kg over the lifetime of application and assuming no loss from the soil profile. This cumulative loading rate of 24 mg/kg is significant when considered against an average soil arsenic concentration of 5 mg/kg (see Section 2), however, relative to mine-impacted or inorganic arsenic pesticide impacted soils where arsenic concentrations are frequently > 100 mg/kg, it is of lesser concern. Also, sewage sludge is often high in aluminum and iron oxide phases, used in the flocculation process, which are efficient scavengers of inorganic arsenic thus lowering the arsenic bioavailability (Carbonell-Barrachina et al., 2000).

Arsenic occurs in animal wastes primarily because of the former use of arsenic antibiotics in poultry and turkey feed; until 2015 four drugs, roxarsone, p-arsanilic acid, carbarsone and nitrosone, were regulated for use, with roxarsone being the most prevalent. As of 2016 all four of

these compounds have all been withdrawn from use (USFDA, n.d.). All four are organic arsenic compounds with an arsenate functional group attached to a benzene ring, and differ by other substituents on the ring. The compounds are not readily adsorbed or metabolized and so occur at concentrations up to 40 mg/kg in animal manures. This provides three points of entry to the human food chain; directly through arsenic in chicken and turkey meat (Lasky et al., 2004; Nachman et al., 2013; Taylor, 2004), from plant uptake after land application of manure, and runoff to surface water or groundwater. A number of studies have shown that these organic arsenic compounds can be degraded by both photolytic (Bednar et al., 2003) and microbial (Stolz et al., 2007) processes and that this degradation happens both during composting of stockpiled litter (Garbarino et al., 2003; Jackson et al., 2003) and after land application (Jackson et al., 2006; Rutherford et al., 2003). Long term application of poultry litter to Upper Coastal Plain soils increased soil arsenic concentrations from 2.7 to 8.4 mg/kg after 25 years of application (Han et al., 2004). Similar increases have been reported for other southern states of the USA (Fisher et al., 2015). There is evidence to suggest that other soluble constituents of the litter, for example phosphate and dissolved organic carbon compounds, facilitate arsenic solubility and leaching (Jackson et al., 2006; Fisher et al., 2015; Shah et al., 2009).

3.6. Seaweed fertilizers

Seaweeds can contain far higher concentrations of arsenic than crop plants; up to 100 mg/kg (Taylor et al., this issue). In most cases the arsenic is present as arsenosugars, which are of low toxicity to humans (Taylor et al., this issue). However, as in the case of poultry litter, these compounds degrade (ultimately) to inorganic arsenic after land application (Castlehouse et al., 2003). Although seaweeds are a 'niche' soil amendment, their use in agriculture is increasing and has been adopted by many organic farms as a soil fertilizer as well as a feed additive in organic dairy farming (Antaya et al., 2015).

4. Biogeochemical cycling within terrestrial agronomic ecosystems

Arsenic cycles within the soil surface and near-surface environment (Bowell, 1994), influenced by mineralogy, abiotic factors such as pH and redox potential (E_H), and biotic factors such as microbially-mediated biomethylation.

4.1. Redox regulation

The most important biogeochemical step in the exposure of humans to arsenic is its release from soils and sediments into pore water; the water contained within soil pores and/or rock (Fendorf and Kocar, 2009). With the exception of extreme pH conditions (<4 or >9), or high concentrations of competing ions (e.g. phosphate, silicic acid or silicate (Luxton et al., 2006)) the release of arsenic from its strong bonds with soil particles depends upon redox potential (E_H); the extent of aeration of the soil (Fendorf and Kocar, 2009). As E_H falls, electron acceptors are depleted and anoxic conditions develop, causing iron oxides and oxyhydroxides to be reduced and dissolve, releasing sorbed arsenic into the soil solution (Meharg and Zhao, 2012) where it can be taken up by plant roots, or leached into groundwater.

Agronomic cropping systems can be divided with respect to arsenic mobilization on the basis of their redox status. Dominant biogeochemical processes influencing aerobic systems, specifically cereals, upland rice, fruit tree orchards, and community gardens, differ from those that dominate in anaerobic systems, predominantly in flooded rice paddies. In aerobic soils, arsenic speciation is predominantly arsenate (arsenic (V)), and is tightly bound to soil particles. Under anaerobic or flooded conditions, arsenic is reduced, and arsenite (arsenic (III)) is the dominant species (Fendorf and Kocar, 2009). Arsenite is less stably bound to aluminum hydroxides and aluminosilicate clay minerals in

the soil than arsenate, for which they exhibit a much stronger binding preference (Fendorf and Kocar, 2009). With few exceptions (such as under conditions of sulfur release), transition of arsenic speciation from arsenate to arsenite is the most influential factor to arsenic bioavailability; and it is under anaerobic conditions where arsenic becomes an imminent human health concern. Influential biogeochemical processes in aerobic systems are ageing and accumulation of arsenic in soil, and in anaerobic systems reductive dissolution of iron-bearing minerals is the dominant process.

4.2. Biotransformation to methylated and volatile species

Volatile arsenicals are arsenic species with a boiling point below 150 °C; the most volatile of which is arsine gas (AsH_3), followed by monomethylarsine ($MeAsH_2$), dimethylarsine (Me_2AsH_2) and finally completely methylated trimethylarsine (TMA).

Volatile arsenic species can be formed either biotically – by fungi, bacteria and algae (Mestrot et al., 2011; Wang et al., 2014; Turpeinen et al., 2002) or abiotically (Wang et al., 2014). In natural systems arsines readily react with oxygen to form non-volatile oxidation products, with AsH_3 most rapidly oxidized and challenging to detect in environmental samples. Oxidation of the arsine gases to inorganic arsenic species completes the arsenic cycle, with arsenic returned to the soil by rain or dry deposition (Pongratz, 1998).

Arsenic methylation in soils increases with decreasing redox potential (Frohne et al., 2011), and addition of organic matter. Increased arsenic volatilization was measured in soil after the addition of rice straw (Jia et al., 2013a), and animal waste products (Mestrot et al., 2013). Inoculation of fungi (*Penicillium* and *Ulocladium* spp.) increased arsenic volatilization up to 8 fold in heavily contaminated and spiked soils (Edvontoro et al., 2004). Microbially mediated arsenic volatilization remains very inefficient, which hinders attempts to use it in soil remediation. Gaseous arsines are volatilized from arsenic contaminated soils into the atmosphere at very low rates; a microcosm study found 0.5–70 µg of arsenic kg⁻¹ soil year⁻¹ was volatilized from a range of soils and a range of arsenic levels (Mestrot et al., 2011), and field measurements of arsenic volatilization are 1–2 orders of magnitude lower than those made in laboratory mesocosms (Meharg and Zhao, 2012). Genetic transformation of bacteria, using genes encoding for the protein product arsenite S-adenosyl methyltransferase (*arsM*) is an attempt to enhance arsenic methylation and volatilization. The *arsM* from *Rhodopseudomonas palustris* was expressed in *Sphingomonas desiccabilis* and *Bacillus idriensis* grown in an aqueous system, resulting in a 10-fold increase in arsenic volatilization compared to the wild type strains. In a soil-based system, 2.2–4.5% of arsenic was removed via microbially-mediated volatilization over an incubation period of 30 days (Liu et al., 2011) (see also Section 5.1).

4.3. Changes in soil arsenic bioavailability due to ageing

Although arsenic in aerobic soils has a lower bioavailability and presents less of an immediate concern for crop uptake, aerobic soil can accumulate arsenic from human inputs, retain them for long periods of time, and release them when redox conditions change (see Section 3.3). Human inputs of arsenic, as discussed in Section 3, are diverse; biosolids, sewage sludge, coal fly ash, poultry litter, industrial waste, arsenical pesticides and from irrigation with naturally arsenic-enriched groundwater. For aerobic soils, ageing – where binding stability of arsenic to soil particles increases over time, is a particularly important part of arsenic cycling. Factors controlling ageing of arsenic include soil type, organic matter content and arsenic species. Both inorganic and organic arsenic species are subject to ageing, with studies indicating a slow oxidation process from arsenite to arsenate over time (Yang et al., 2005).

5. Rhizosphere processes

Processes occurring in the rhizosphere (the boundary layer of soil under the influence of plant roots) dramatically influence arsenic concentrations and bioavailability because they involve local alterations in redox potential, pH and organic matter content. Rhizosphere acidification occurs during iron uptake by all plant species during cation uptake and charge balance, when protons are released into the rhizosphere. Plants release anywhere from 10 to 250 mg of carbon per gram of root tissue into the rhizosphere; about 10–40% of their total photosynthetically fixed carbon (Newman, 1985), making the rhizosphere particularly rich in organic carbon compared to bulk soil, which in turn exerts an influence on arsenic solubility by stimulating microbially-mediated reductive dissolution of soil minerals. Large differences have been found in the arsenic concentration of rhizosphere soils compared with bulk soils in highly arsenic-contaminated areas, with higher concentrations of arsenic in rhizosphere soils compared to bulk soils (Acosta et al., 2015).

In anaerobic soils, the iron plaque that develops on the submerged stem and roots of rice plants dominates rhizosphere dynamics of arsenic. In flooded environments such as paddy fields, plants oxygenate the rhizosphere through specialized tissues called aerenchyma, which are found in many aquatic plants and emergent macrophytes such as rice. This radial oxygen loss creates an oxidized layer around plant tissue that stimulates aerobic microbial activity and the oxidation of iron, which precipitates and forms a visible iron plaque on the root surface (Wu et al., 2016; Seyfferth, 2015; Yamaguchi et al., 2014; Lee et al., 2013; Hossain et al., 2009). Formation of an iron (oxyhydr)oxide plaque on root surfaces can alter the uptake of arsenic by rice, acting as a sorbent for excess nutrients such as ferrous iron (reduced iron) as well as arsenic and aluminum (Wu et al., 2012). Rates of oxygen loss influence iron plaque formation (Wu et al., 2012), and vary between rice cultivars (Li et al., 2015; Mei et al., 2012). Studies conducted over the last forty years are inconsistent on whether iron plaque prevents or enhances arsenic uptake by plants (Seyfferth, 2015), and the hypothesis that arsenic influences the quality and amount of the iron plaque (Lee et al., 2013). Profound differences in mineral composition and quantity of laboratory-created iron plaques has been demonstrated experimentally (Seyfferth, 2015), which may have contributed to these inconsistencies.

5.1. Microbial activity

Microbes directly and indirectly influence arsenic speciation in rhizosphere soil, and are widely considered to play a key role in arsenic biogeochemistry (Gadd, 2010). Under certain nutrient-limited conditions, microbes actively weather minerals to access nutrients for cellular growth, which releases arsenic (Mailloux et al., 2009), as well as creating abiotic conditions that induce changes in arsenic speciation via production of organic acids, polysaccharides and ligands. Soil microorganisms can strongly affect soil redox, regulating arsenic release into pore water (Wang et al., 2016a). A number of strains of bacteria have also been shown to contribute to the formation of arsenic minerals by using arsenic as a terminal electron acceptor, such as *Desulfovibrio auripigmentum* (Newman et al., 1997), *Desulfovibrio* strain Ben-RB (Macy et al., 2000), *Shewanella oneidensis* (Wang et al., 2016b) and *S. putrefaciens* CN32 (Smeaton et al., 2012). These microorganisms also differ in their capabilities for liberating arsenic from specific arsenate-bearing minerals (Wang et al., 2016a).

Microbial transformation can mobilize arsenic by converting inorganic to organic forms, including MMA and DMA (Jia et al., 2013b; Xu et al., 2016). Plants translocate organic arsenicals from roots to the (frequently edible) above-ground parts more efficiently than inorganic arsenic (Carey et al., 2011a; Carey et al., 2011b; Carey et al., 2010) (see Section 6), therefore microbial transformation to organic arsenicals can increase human dietary exposure.

Plants, green algae and microbes can all enzymatically transform arsenic species (Jia et al., 2013b; Baker et al., 1983), but methylated forms of arsenic detected in plants are a product of rhizosphere bacteria; plants cannot methylate arsenic (Lomax et al., 2012; Jia et al., 2013b; Hansen et al., 2011; Arao et al., 2011). The genomes of >85 arsenic-metabolizing archaea and bacteria have been sequenced for genes involved in arsenic metabolism (Andres and Bertin, 2016). In bacteria, archaea and fungi, arsenic methylation is catalyzed by homologs of *arsM*, (see Section 4) (Jia et al., 2013b). Resistance to arsenite and arsenate exists in nearly all microbes, which also confers the ability to transform arsenate into volatile arsenic gases (Zhang et al., 2011), a particularly effective way of removing arsenic.

Profiling the transcriptome, proteome and metabolome of arsenic contaminated soils offers way of understanding microbially-mediated rhizosphere arsenic processes (Andres and Bertin, 2016). This approach measures the presence and expression of specific genes, rather than attempting to isolate and study the microbes that carry them, 98% of which – it is estimated – do not grow in culture (Stewart, 2012). Microbially mediated arsenic metabolic processes that play a major role in arsenic cycling in agronomic systems include arsenite oxidation (via the *aio* genes), arsenate respiration (via the *arr* genes), arsenate reduction (via the *ars* genes) and arsenite methylation (via the *arsM* genes) (Xiao et al., 2016). Interested readers are referred to the recent excellent work of Andres and Bertin (Andres and Bertin, 2016) for a comprehensive review of this subject. Microbially mediated redox processes strongly influence arsenic uptake in rice, involving *aioA*, *arsC* and *arrA* (Jia et al., 2013b), with pH emerging as an important factor in the distribution of microbes in paddy soils. Testing a variety of soils has shown that bacteria possessing the *arsM* gene for methylating arsenic are widespread and phylogenetically diverse, and even in paddy soils with low concentrations of arsenic, genes for arsenic metabolism are abundant (Jia et al., 2013b).

6. Arsenic and crop plants

Much of our understanding about the physiological mechanisms of arsenic uptake in plants comes from the study of a limited number of species. Called *model plant* species, they are extensively studied, well described, easy to grow, and the results can be compared between studies. The understanding is that the information gained from studying model plants is applicable to other plant species. From a genetic perspective, orthologous genes exist in different plant species that have evolved from a common ancestral gene, and they usually retain the same function. Characterization of arsenic-related genes in a model plant strongly suggests that they exist and perform similar functions in other species. Caveats to this are their levels of expression, which makes some plants more adept at accumulating arsenic than others. In this section, much of the knowledge gained on arsenic uptake and metabolism of plants comes from the study of mouse-eared pennycress (also called thale cress or rockcress) (*Arabidopsis thaliana* Heynh.) and rice (*Oryza sativa* L.); model plants with fully sequenced genomes. These species represent dicotyledonous (e.g. flowers, vegetables, deciduous trees) and monocotyledonous plant species (e.g. grasses, palm trees) respectively, thereby representing much of the edible crop species. An exception to this is the study of the arsenic hyperaccumulating fern (Chinese Brake fern, *Pteris vittata*), a seedless plant that is able to accumulate up to 22,630 mg/kg (dry weight) arsenic in its fronds (Ma et al., 2001).

6.1. Phytotoxicity of arsenicals

Arsenic is toxic to plants (Sharma, 2012). Despite lower acute human toxicity of the organic arsenicals (median lethal dose is 700–1600 mg/kg and 700–2600 mg/kg for MMA and DMA respectively compared to 10–20 mg/kg for inorganic forms) (Le et al., 2000) no one form of arsenic is consistently more toxic to plants (Finnegan and Chen,

2012a). Soybean yields are affected when tissue arsenic levels exceed 1 mg/kg, and 4 mg/kg limits cotton yields (Deuel and Swoboda, 1972), whereas in barley tissue concentrations of 20 mg/kg inhibited growth (Davis et al., 1978). Higher yield-limiting arsenic levels have been recorded in rice: 20–100 mg/kg in above ground biomass, and 1000 mg/kg in root tissue (Adriano, 2001). By contrast, potatoes (*Solanum tuberosum* L.) suffered no growth inhibition in soils containing 290 mg/kg arsenic (Codling et al., 2016). In some plants species, organic forms are more toxic than inorganic, for example in rice (order of toxicity: MMA > arsenite > arsenate = DMA) (Marin et al., 1992a), and in smooth cordgrass (*Spartina alterniflora* Loisel) (DMA = MMA > arsenite > arsenate) (Carbonell-Barachina et al., 1998).

Plants vary in their tolerance to arsenic, and the stress response differs for each arsenic species (Choudhury et al., 2011; Shri et al., 2009; Chakrabarty et al., 2009). The chemical similarities between arsenate and phosphate means that arsenic can replace phosphate in biomolecules like ATP (adenosine triphosphate, a molecule used for intercellular energy transfer), with negative impacts on growth and metabolism (Ullrich-Eberius et al., 1989). In rice in particular, DMA and MMA induce straighthead disease (arsenic-associated straighthead disease), significantly lowering yield of certain rice varieties (Rahman et al., 2008). Straighthead is a physiological disorder of rice characterized by sterile florets, which remain upright at maturity instead of bending over under the weight of the filled grain. The exact cause of straighthead is unknown, but consistent flooding, low soil pH, high iron availability and high organic matter content have all been implicated in naturally-occurring straighthead disease (Rahman et al., 2008). Arsenic's suspected role in straighthead comes from observations of more frequent outbreaks in rice grown in soil where arsenical herbicides such as monosodium methanearsonate (MSMA) – used in cotton production in the USA – have been historically applied.

6.2. Arsenic uptake mechanisms

In magnitude, plants take up arsenicals from the soil in the order arsenite > arsenate > DMA > MMA (Raab et al., 2007; Finnegan and Chen, 2012b), with the various arsenic species entering via different root membrane transport proteins in the root plasma membrane that allow ions and molecules to cross with varying levels of selectivity, or target specificity. Similarities in chemical structure between arsenate and phosphate, and between arsenite and silicic acid, govern their entry into root cells. Arsenate enters root cells through phosphate transporters (the Phosphate Transporter 1 family of proteins; PHT1) in both the model plant *Arabidopsis thaliana* (Shin et al., 2004; Remy et al., 2012) and in rice (Jia et al., 2011; Ye et al., 2015; Sun et al., 2012) (Fig. 1). In rice, Low Silicon 1 (OsLsi1) and OsLsi2 are silicic acid transporters and arsenite, MMA^V, and DMA^V are among their unintended targets (Li et al., 2009a; Ma et al., 2008). These Nodulin 26-like Intrinsic Proteins (NIPs) (Abedin et al., 2002), which are members of the aquaporin water channel superfamily of proteins (Pommerenig et al., 2015) embedded in the exodermal cell membranes of rice roots, move arsenic from the soil into the vascular system for distribution to the stem and leaves. OsLsi2 works in tandem with OsLsi1 to transport arsenite inward toward the xylem (Mitani et al., 2009; Ma et al., 2007a) (vascular tissue that conducts water and dissolved nutrients up from the roots). The arsenic uptake specificity of OsLsi1 is arsenite ≫ MMA > DMA (Abedin et al., 2002). These bidirectional NIP transport proteins also efflux arsenite back into the soil, but since OsLsi1 effluxes only 15–20% of the arsenite in roots cells (Zhao et al., 2010), there may be other unidentified arsenite efflux transporters contributing to this process.

6.3. Arsenic transport and metabolism in plants

Transport of arsenite into the xylem for delivery to the shoot is less well characterized than its uptake from the soil. Arsenic is transported to the grain mainly via the phloem (Carey et al., 2011a; Carey et al.,

2011b; Carey et al., 2010) (vascular tissue that conducts sugars and metabolic products from the leaves), by transporters in the nodes (Song et al., 2014a), but their characterization is still in the early stages. Transporters for myo-inositol (Inositol Transporter 2 and 4); an important sugar for developing rice grains, also transport arsenite into the phloem companion cells (Schneider et al., 2006; Schneider et al., 2007). In *Arabidopsis*, INT2 or AtINT4 load about 45–64% arsenite into the grain (Duan et al., 2015). The identity of transporters that move arsenite out of the phloem and into the grain are also unknown, but manipulating the target specificity of the INT genes might show promise in molecular genetic or plant breeding mitigation efforts as a way to prevent arsenite from reaching the grain.

Despite having a lower affinity for transporters into the plant than the inorganic forms, organic arsenic species are more efficiently transported toward the shoot than inorganic forms (Raab et al., 2007; Finnegan and Chen, 2012b) because they are not complexed by phytochelatins (PCs); sulphhydryl-rich glutathione (GSH) polymers (Zenk, 1996; Schmoger et al., 2000). Likewise, in broad beans (*Vicia faba* L.) grown in a soil containing 90% inorganic arsenic, DMA and MMA were the dominant arsenic forms in the bean (68%) (Sadee et al., 2016). In root vegetables, carrot (*Daucus carota* L.) and beet (*Beta vulgaris* L.) grown on arsenic-contaminated soils, arsenic forms were predominantly inorganic, but for beets in particular were not readily identified using the typical standards (arsenate, arsenite, MMA and DMA) (Pizarro et al., 2016).

The arsenic species composition of rice grain is influenced by the arsenic transport rate of the particular cultivar (Seyfferth et al., 2011; Williams et al., 2005b). Rice cultivars currently grown in the USA have an arsenic speciation split approximately equally between inorganic arsenic and DMA, while cultivars grown in Bangladesh contain mostly inorganic arsenic (Williams et al., 2005b). While lower inorganic arsenic in rice grain seems favorable for avoiding human health effects, the assumed safety of DMA is contentious (Hughes, 2002), being based on acute toxicity data, and not on genotoxicity or carcinogenicity, which are equally relevant in long term safety considerations.

Arsenic detoxification inside cells uses a multi-step process beginning with reduction of arsenate to arsenite using an arsenate reductase enzyme (Sanchez-Bermudo et al., 2014; Chao et al., 2014). In *Arabidopsis*, the protein High Arsenic Content1 (HAC1; also called Arsenate Reductase QTL1; ARQ1) reduces arsenate (Chao et al., 2014). Even though arsenite is more toxic than arsenate (Abedin et al., 2002; Styblo et al., 2000; Marin et al., 1992b), it is hypothesized that ancestral organisms to plants were exposed almost exclusively to arsenite before atmospheric oxygen enabled arsenate formation (Rosen, 1999), and this mechanism persisted through natural selection. Arsenite is then complexed by PCs, and transported into the vacuole (Zenk, 1996) via ATP Binding Cassette (ABC-type) transporters (Song et al., 2014b; Song et al., 2010a). This process depletes glutathione availability, rendering the plant more susceptible to other oxidative stresses, which inhibits photosynthesis, pigment production, and the integrity of cell membranes (Hartley-Whitaker et al., 2001; Lee et al., 2012; Mascher et al., 2002; Stoeva et al., 2003).

7. Limiting arsenic uptake by crops

7.1. Water management

Although the traditional method for cultivating rice involves flooding leveled, tilled fields before or shortly after planting germinated seedlings, flooded soil is not a biological requirement of rice plants. Flooding is used for weed and vermin control, for mobilization of key nutrients such as iron, phosphate and zinc, and importantly, flooding discourages the buildup of root nematodes over multiple years of rice growth. As mentioned earlier, flooded conditions mobilize soil-bound arsenic through reductive dissolution of Fe (oxyhydr)oxides, and the reduction of arsenate to the more mobile arsenite (Dixit and Hering,

2003). Water management strategies that involve periods of oxic soil conditions can decrease arsenic uptake in rice by limiting dissolution of arsenic. Rice grown in non-flooded or aerobic conditions has a lower yield than intermittently or constantly flooded rice (Li et al., 2009b; Grassi et al., 2009; Arao et al., 2009). Intermittent flooding (flooding maintained until full tillering, followed by intermittent irrigation) is a promising management technique to reduce arsenic levels, and can potentially produce higher grain yields than either non-flooded or constantly flooded conditions (Hu et al., 2013a). However, oxic conditions increase cadmium concentrations in the grain when grown in acidic soils (Hu et al., 2013a; Hu et al., 2013b; Honma et al., 2016), and cadmium is also a highly toxic metal. The observed increases in cadmium were also a cultivar-specific trait, but the increase in cadmium uptake between rice grown under aerobic conditions were approximately an order of magnitude greater than their flooded counterparts. Pot experiments suggest that water management strategies implemented during the heading period of rice growth (when the rice panicle has emerged from the stem and is fully visible, just before flowering) can

regulate both arsenic and cadmium concentration in the grain (Arao et al., 2009; Hu et al., 2013b).

7.2. Amendment and fertilization practices

Soil amendment involves incorporating substances into the plow layer that either add missing nutrients, reduce the bioavailability of existing potentially toxic substances (to prevent crop uptake), or both. Soil amendments that have shown potential in reducing arsenic uptake by plants include iron-, and silica-based additives. The use of iron-based amendments increases in the concentration of free iron oxide in the soil, retarding the release of arsenite from the solid phase into soil solution, (mentioned in Section 4.1 and discussed in Section 5), whereas silica fertilization inhibits arsenic uptake by competitive inhibition at the plant root surface while adding an essential nutrient.

Zero valent iron powder (90% iron) and iron oxide (56% iron) incorporation prevented uptake of arsenic into the grain of rice grown on soil containing 39.5 mg/kg total arsenic by approximately 45%, and

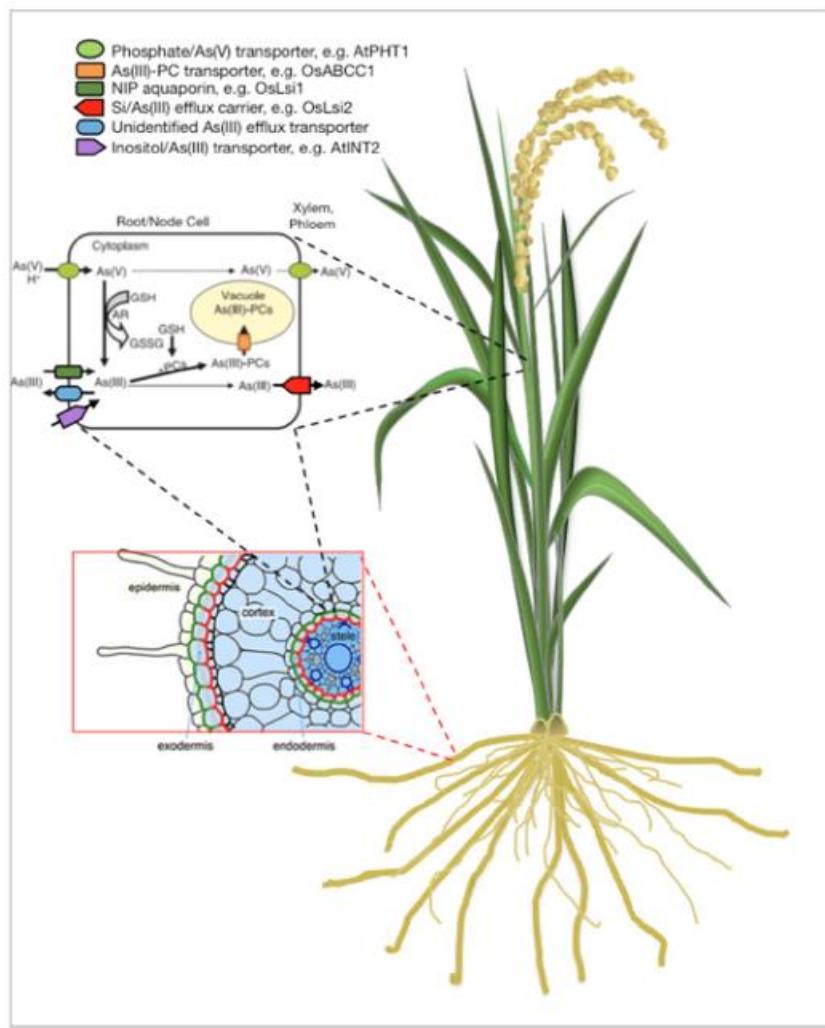


Fig. 1. Generalized diagram of arsenic uptake, transport and metabolism in plants. GSH, glutathione; AR, arsenate reductase; GSSG, oxidized glutathione; PC, phytochelatin. Modified from Zhao et al. (2009) and Ma et al. (2007b).

corresponded with a reduction in bioavailable arsenic in the soil (Matsumoto et al., 2016). Amendment with iron oxides (at a rate of 2%) was also more effective at reducing grain arsenic than phosphate amendment (Farrow et al., 2015). Amendments have also been used in combination with water management strategies to try and reduce both arsenic and cadmium concentrations in rice simultaneously (Honma et al., 2016), without success. Reduction of arsenic in the grain was achieved with iron oxide addition and constant flooding, whereas cadmium reduction was achieved with converter furnace slag addition and rain water management (no irrigation after midseason drainage until harvest).

A combination of ethylenediaminetetraacetic acid ferric sodium salt (iron EDTA) and calcium peroxide was effective for reducing arsenic uptake by vegetable crops (lettuce, Chinese cabbage and radish) from soils containing 14 mg/kg total arsenic (Chou et al., 2016), again by increasing amorphous aluminum and iron oxides. It is likely that this this level of arsenic contamination would be deemed too high for commercial vegetable production, so these amendments may only be feasible for use in private vegetable gardens. Questions remain about whether iron oxide amendment application only temporarily reduces arsenic bioavailability (Tiberg et al., 2016). In addition, the suitability for arsenic immobilization is highest at the lower soil pH range, and is strongly affected by soil phosphorus concentration, which strongly competes with arsenic.

Rice plants take up high concentrations of silica, constituting up to 10% of dry matter in the straw and husk of the plant (Penido et al., 2016). As mentioned earlier (Section 6) the silicon membrane transporter (Lsi1) is the main route of arsenite entry in to rice root cells, and provision of silicon causes competitive inhibition of arsenite uptake. Increasing silicon availability in the soil also reduces the expression of the Lsi1 transporter in the plant, which further decreases the potential for arsenic uptake. Fertilization of rice paddy soils with silicon is a potential mitigation strategy for preventing or reducing arsenic uptake by rice through competitive inhibition of arsenite uptake (Meharg and Meharg, 2015). The use of synthetic silicon fertilizers, such as calcium silicate or silica gel is prohibitively expensive for smallholder farmers in developing countries, however reusing the silicon-rich parts of the rice plant that remain after harvesting and grain processing may provide a sustainable solution that also addresses the ongoing issue of silicon depletion of the soil (Penido et al., 2016). Soil incorporation of fresh rice husks, or the ash that remains after burning the husk and straw for energy (which is a common practice for smallholder farmers), can provide silicon without increasing methane production and decreases either total or inorganic arsenic in rice grain (Seyfferth et al., 2016).

Despite the potential of soil amendment with iron oxides or silica to reduce arsenic bioavailability or prevent plant uptake of arsenic, the high cost of these amendments inevitably prevents their use, especially by small holder farmers. Large rice producers in the US or Europe have not so far adopted widespread use of these soil amendments to reduce rice grain arsenic concentrations. It is also reasonable to assume that use of expensive soil amendments would drive up the cost of rice. Their lack of use may also be attributable to the fact that iron amendments are essentially untested in a diverse range of large-scale agricultural settings and their performance will vary between soil types. In non-rice agricultural systems arsenic is tightly bound to the solid phase; significant crop uptake from oxidized soil is likely to be a result of extreme contamination, in which case effective mitigation is restricted to redirecting land use away from edible crops. In systems subject to periodic flooding, improving drainage remains the best mitigation strategy.

7.3. Mitigation using plant breeding approaches

The development of crops that accumulate high levels of arsenic and yet remain healthy, while preventing arsenic from reaching the edible grain is thought to hold great potential as a strategy for reducing human exposure to dietary arsenic. The use of molecular genetics

techniques such as alterations in gene expression characteristics, gene editing to alter target specificity, or alternately, using traditional plant breeding techniques are both tangible approaches. Both use knowledge of the arsenic uptake and tolerance characteristics of plants to develop varieties with desired characteristics. These characteristics include lower arsenic uptake (Norton et al., 2009), higher arsenite efflux (Duan et al., 2012) and increased vacuolar arsenic sequestration (Song et al., 2010b). For instance, many rice cultivars have now been screened to identify those that accumulate lower levels of arsenic in their grain and efforts are underway to identify the genes underlying this trait (Norton et al., 2009; Norton et al., 2014). Overexpressing Arabidopsis ABC-type transporters that sequester arsenite-PC complexes in the cell vacuole results in plants able to grow in otherwise toxic concentrations of arsenic (Song et al., 2010a). Conversely, knocking out the function of the related rice ABC transporter OsABCC1 results in higher levels of grain arsenic. The OsABCC1 transporter limits arsenic transport to grains by sequestering arsenic in the vacuoles of the phloem companion cells directly connected to the grain. By combining what we have learned from the overexpression studies in Arabidopsis and the loss-of-function study in rice, overexpression of OsABCC1 can be used as a strategy to breed arsenic tolerance and low-arsenic accumulating rice cultivars. Another promising strategy is based on expressing the arsenate efflux transporter from yeast (*Saccharomyces cerevisiae*) in rice, which can reduce arsenic accumulation in brown rice by 20%. A less successful idea to methylate sodium arsenite to DMA by expressing an algal arsM gene in Arabidopsis resulted in lethal phytotoxicity (Tang et al., 2016), suggesting that arsenic methylation in plants can only be an effective detoxification strategy if volatile arsines are the end point of the methylation.

8. Conclusions

The discovery of arsenic in staple foods, beverages and other products has increased awareness and stimulated research on the sources and the processes involved. The information brought together here illustrates the numerous geochemical and biological processes that influence the movement of arsenic into the food supply.

It is clear there must be strategies for preventing arsenic exposure, that operate in both the short term – to protect consumers from existing contamination – and in the long term, to prevent further contamination. This requires government regulation on the permissible levels of arsenic food, with lower levels for infant foods (see Nachman et al., this issue), which must work in tandem with long term goals to address arsenic in agricultural soils, actively prevent further inputs and identify contaminated areas for mitigation. Our recommendations are that the information in this review is used to inform a reconsideration and a unification of regulations on the action levels of agricultural soil arsenic, which in the USA for example, exist only at the state level, vary widely from state to state, and have no formal channels of enforcement. We recommend that educating the community and garnering their support and involvement for lowering exposure to arsenic through food is an approach already shown to hold enormous potential. Direct involvement of the commercial rice growing community in research and development of arsenic mitigation strategies and amendments is needed. Much effort has been given to short-term, greenhouse-scale testing of amendment formulations that will ultimately be too expensive, impractical, or ineffective in the long term. Community-based participatory research should extend to the agricultural community, leading to partnerships that will make longer-term field-scale testing of mitigation strategies accessible. Feasibility should be a primary consideration in arsenic mitigation research. Community outreach efforts targeted to commercial growers or the home gardener specifically must raise awareness of the significance and potential impacts of former land uses, encouraging testing for the presence of arsenic in the soil and educating growers on crops shown to accumulate arsenic in their edible parts. Information gathering on former arsenic input into the soil from pesticides and from proximity to various waste sites is of paramount

importance, and will allow monitoring and mitigation to be targeted to where it is needed most. Currently there is no readily available source of soil arsenic concentration information at a sufficient resolution to inform commercial producers or homeowners; this information needs to be accessible to everyone, everywhere. Going forward, management and remediation of arsenic contaminated soils is essential both for human health and food security, and innovative technologies are urgently needed that will expedite this process. Innovative solutions such as the use of rice husks to add silicon to the soil to offset arsenic uptake, and the use of cultivars with low-arsenic accumulating characteristics point the way forward for sustainable solutions.

Acknowledgements

This paper, a product of the Collaborative on Food with Arsenic and associated Risk and Regulation (C-FARR), is supported by the Dartmouth College Toxic Metals Superfund Research Program through funds from the National Institute of Environmental Health Sciences of the National Institutes of Health under Award Number 1R13ES026493-01 to C. Chen and Award Number P42ES007373 to B. Stanton, and the Children's Environmental Health and Disease Prevention Research Center at Dartmouth through funds from the National Institute of Environmental Health Sciences of the National Institutes of Health under Award Number P01ES022832 to M. Karagas. Although EPA contributed to this article, the research presented was not performed by or funded by EPA and was not subject to EPA's quality system requirements. Consequently, the views, interpretations, and conclusions expressed in this article are solely those of the authors and do not necessarily reflect or represent EPA's views or policies. TP was supported by funds from the National Institute of General Medical Sciences Center for Biomedical Research Excellence (P20GM104416). The authors disclose that there are no actual or potential conflicts of interest.

References

- Abedin, M.J., Feldmann, J., Meharg, A.A., 2002. Uptake kinetics of arsenic species in rice plants. *Plant Physiol.* 128 (3), 1120–1128.
- Acosta, J.A., Arcosena, J.M., Faz, A., 2015. Speciation of arsenic in bulk and rhizosphere soils from artisanal cooperative mines in Bolivia. *Chemosphere* 138, 1014–1020.
- Adomako, E.E., et al., 2009. Enhanced transfer of arsenic to grain for Bangladesh grown rice compared to US and EU. *Environ. Int.* 35 (3), 476–479.
- Adriano, D.C., 2001. Trace Elements in Terrestrial Environments, second ed. Springer Verlag, New York, p. 867.
- Andres, J., Bertin, P.N., 2016. The microbial genomics of arsenic. *FEMS Microbiol. Rev.* 40 (2), 299–322.
- Antaya, N.T., et al., 2015. Incremental amounts of *Aspergillus nodosum* meal do not improve animal performance but do increase milk iodine output in early lactation dairy cows fed high-forage diets. *J. Dairy Sci.* 98 (3), 1991–2004.
- Arao, T., et al., 2009. Effects of water management on cadmium and arsenic accumulation and dimethylarsinic acid concentrations in Japanese rice. *Environ. Sci. Technol.* 43 (24), 9361–9367.
- Arao, T., et al., 2011. Effects of arsenic compound amendment on arsenic speciation in rice grain. *Environ. Sci. Technol.* 45 (4), 1291–1297.
- Ayotte, J.D., et al., 2006. Modeling the probability of arsenic in groundwater in New England as a tool for exposure assessment. *Environ. Sci. Technol.* 40 (11), 3578–3585.
- Baker, M.D., et al., 1983. Methylation of arsenic by fresh-water green-algae. *Can. J. Fish. Aquat. Sci.* 40 (8), 1254–1257.
- Bednar, A.J., et al., 2002. Presence of organoarsenicals used in cotton production in agricultural water and soil of the southern United States. *J. Agric. Food Chem.* 50 (25), 7340–7344.
- Bednar, A.J., et al., 2003. Photodegradation of roxarsone in poultry litter leachates. *Sci. Total Environ.* 302 (1–3), 237–245.
- Bertoldi, D., et al., 2013. Arsenic present in the soil-vine-wine chain in vineyards situated in an old mining area in Trentino, Italy. *Environ. Toxicol. Chem.* 32 (4), 773–779.
- Bodwell, J.E., Kingsley, L.A., Hamilton, J.W., 2004. Arsenic at very low concentrations alters glucocorticoid receptor (GR)-mediated gene activation but not GR mediated gene repression; complex dose response effects are closely correlated with levels of activated GR and require a functional GR DNA binding domain. *Chem. Res. Toxicol.* 17, 1064–1076.
- Bowell, R.J., 1994. Sorption of arsenic by iron-oxides and oxyhydroxides in soils. *Appl. Geochem.* 9 (3), 279–286.
- Bowell, R.J., et al., 2014a. The environmental geochemistry of arsenic - an overview. *Arsenic: Environmental Geochemistry, Mineralogy, and Microbiology.* 79, pp. 1–16.
- Bowell, R.J., et al., 2014b. *Rev. Mineral. Geochem.* 79, 635.
- Bowell, R.J., et al., 2014c. The environmental geochemistry of arsenic - an overview. In: Bowell, R.J., et al. (Eds.), *Arsenic: Environmental Geochemistry, Mineralogy, and Microbiology*, pp. 1–16.
- Caldwell, K.E., et al., 2014. Prenatal arsenic exposure alters the programming of the glucocorticoid signaling system during embryonic development. *Neurotoxicol. Teratol.* 47C, 66–79.
- Carbone-Barrachina, A.A., et al., 1998. The influence of arsenic chemical form and concentration on *Spartina patens* and *Spartina alterniflora* growth and tissue arsenic concentration. *Plant Soil* 198, 33–43.
- Carbone-Barrachina, A.A., et al., 2000. Arsenic chemistry in municipal sewage sludge as affected by redox potential and pH. *Water Res.* 34 (1), 216–224.
- Carey, A.M., et al., 2010. Grain unloading of arsenic species in rice. *Plant Physiol.* 152 (1), 309–319.
- Carey, A.M., et al., 2011a. Phloem transport of arsenic species from flag leaf to grain during grain filling. *New Phytol.* 192 (1), 87–98.
- Carey, A.M., et al., 2011b. A review of recent developments in the speciation and location of arsenic and selenium in rice grain. *Anal. Bioanal. Chem.* 402, 3275–3286.
- Carlin, D.J., et al., 2015. Arsenic and environmental health: state of the science and future research opportunities. *Environ. Health Perspect.*
- Castlehouse, H., et al., 2003. Biotransformation and accumulation of arsenic in soil amended with seaweed. *Environ. Sci. Technol.* 37 (5), 951–957.
- CEH, 2008. United Kingdom Pollutant Deposition Maps. 2016 Database. (Available from: <http://www.pollutantdeposition.ceh.ac.uk/content/heavy-metals>).
- Chakrabarty, D., et al., 2009. Comparative transcriptome analysis of arsenate and arsenite stresses in rice seedlings. *Chemosphere* 74 (5), 688–702.
- Chao, D.Y., et al., 2014. Genome-wide association mapping identifies a new arsenate reductase enzyme critical for limiting arsenic accumulation in plants. *PLoS Biol.* 12 (12), e1002009.
- Chou, M.L., et al., 2016. Inhibition of ethylenediaminetetraacetic acid ferric sodium salt (EDTA-Fe) and calcium peroxide (CaO₂) on arsenic uptake by vegetables in arsenic-rich agricultural soil. *J. Geochem. Explor.* 163, 19–27.
- Choudhury, B., Chowdhury, S., Biswas, A.K., 2011. Regulation of growth and metabolism in rice (*Oryza sativa* L.) by arsenic and its possible reversal by phosphate. *J. Plant Interact.* 6 (1), 15–24.
- Colding, E.E., Chaney, R.L., Green, C.E., 2015. Accumulation of lead and arsenic by carrots grown on lead-arsenate contaminated orchard soils. *J. Plant Nutr.* 38 (4), 509–525.
- Colding, E.E., Chaney, R.L., Green, C.E., 2016. Accumulation of lead and arsenic by potato grown on lead-arsenate-contaminated orchard soils. *Commun. Soil Sci. Plant Anal.* 47 (6), 799–807.
- Craw, D., Bowell, R.J., 2014. The characterization of arsenic in mine waste. In: Bowell, R.J., et al. (Eds.), *Arsenic: Environmental Geochemistry, Mineralogy, and Microbiology*, pp. 473–505.
- Davis, R.D., Beckett, P.H.T., Wollan, E., 1978. Critical levels of twenty potentially toxic elements in young spring barley. *Plant Soil* 49, 395–408.
- Davis, M.A., et al., 2012. Rice consumption and urinary arsenic concentrations in US children. *Environ. Health Perspect.* 120, 1418–1424.
- Deuel, L.E., Swoboda, A.R., 1972. Arsenic toxicity to cotton and soybeans. *J. Environ. Qual.* 1 (3), 317–320.
- Dixit, S., Hering, J.G., 2003. Comparison of arsenic(V) and arsenic(III) sorption onto iron oxide minerals: implications for arsenic mobility. *Environ. Sci. Technol.* 37 (18), 4182–4189.
- Drahota, P., et al., 2006. Weathering and erosion fluxes of arsenic in watershed mass budgets. *Sci. Total Environ.* 372 (1), 306–316.
- Duan, G., et al., 2012. Expressing SeACR3 in rice enhances arsenite efflux and reduced arsenic accumulation in rice grains. *Plant Cell Physiol.* 53 (1), 154–163.
- Duan, G.-L., et al., 2015. Inositol transporters AtNT2 and AtNT4 regulate arsenic accumulation in *Arabidopsis* seeds. *Nat. Plants* 2, 15202.
- Edvinton, B.B., et al., 2004. Microbial formation of volatile arsenic in cattle dip site soils contaminated with arsenic and DDT. *Appl. Soil Ecol.* 25 (3), 207–217.
- Embrick, L.L., et al., 2005. Characterization of lead and arsenic contamination at Barber Orchard, Haywood County, NC. *Microchem. J.* 81 (1), 117–121.
- EPA, 2016. Coal Ash Basics [Web Page] EPA June 7, 2016. (Available from: <https://www.epa.gov/coalash/coal-ash-basics>).
- European Union, 2015. Commission Regulation (EU) 2015/1006 of 25 June 2015 amending Regulation No. 1881/2006 as regards maximum levels of inorganic arsenic in foodstuffs. Free access to EU law (treaties, directives, regulations, consolidated legislation), preparatory acts, case-law, international agreements, EFTA documents, summaries of EU legislation and other documents cited 2016 November 22, 2016. (Available from: <http://eur-lex.europa.eu/eli/reg/2015/1006/oj>).
- Farrow, E.M., et al., 2015. Reducing arsenic accumulation in rice grain through iron oxide amendment. *Ecotoxicol. Environ. Saf.* 118, 55–61.
- U. FDA, 2016. In: Sucher, L. (Ed.), *FDA Proposes Limit for Inorganic Arsenic in Infant Rice Cereal*. US Food and Drug Administration.
- Fendorf, S., Kocar, B.D., 2009. Chapter 3 biogeochemical processes controlling the fate and transport of arsenic: implications for south and Southeast Asia. *Advances in Agronomy*. Academic Press, pp. 137–164.
- Feng, M., et al., 2005. Arsenic transport and transformation associated with MSMA application on a golf course green. *J. Agric. Food Chem.* 53 (9), 3556–3562.
- Feuerborn, H.-J., 2011. Coal combustion products in Europe - an update on production and utilisation standardisation and regulation. World of Coal Ash (WOCA) Conference (Denver, CO, USA).
- Finnegan, P.M., Chen, W., 2012a. Arsenic toxicity: the effects on plant metabolism. *Front. Physiol.* 3, 182.
- Finnegan, P.M., Chen, W.H., 2012b. Arsenic toxicity: the effects on plant metabolism. *Front. Physiol.* 3.
- Fisher, D.J., Yonkos, L.T., Staver, K.W., 2015. Environmental concerns of roxarsone in broiler poultry feed and litter in Maryland, USA. *Environ. Sci. Technol.* 49 (4), 1999–2012.

- Frohne, T., et al., 2011. Controlled variation of redox conditions in a floodplain soil: impact on metal mobilization and biomethylation of arsenic and antimony. *Geoderma* 160 (3–4), 414–424.
- Gadd, G.M., 2010. Metals, minerals and microbes: geomicrobiology and bioremediation. *Microbiology-Sgm* 156, 609–643.
- Garbarino, J.R., et al., 2003. Environmental fate of roxarsone in poultry litter. I. Degradation of roxarsone during composting. *Environ. Sci. Technol.* 37 (8), 1509–1514.
- Geiszinger, A., Goessler, W., Kosmus, W., 2002. Organotarsinic compounds in plants and soil on top of an ore vein. *Appl. Organomet. Chem.* 16 (5), 245–249.
- Grassi, C., et al., 2009. Aerobic rice crop performance and water use efficiency. *J. Agric. Environ. Int. Dev.* 103 (4), 259–270.
- Han, F.X., et al., 2004. Arsenic solubility and distribution in poultry waste and long-term amended soil. *Sci. Total Environ.* 320 (1), 51–61.
- Hansen, H.R., et al., 2011. Identification of tetramethylarsonium in rice grains with elevated arsenic content. *J. Environ. Monit.* 13 (1), 32–34.
- Hartley-Whitaker, J., Ainsworth, G., Meharg, A.A., 2001. Copper- and arsenate-induced oxidative stress in *Alocasia lanatus* L. clones with differential sensitivity. *Plant Cell Environ.* 24 (7), 713–722.
- Henke, K., 2009. Appendix E: regulation of arsenic: a brief survey and bibliography. In: Henke, K. (Ed.), *Arsenic*. John Wiley & Sons, Ltd, pp. 545–557.
- Honna, T., et al., 2016. Effects of soil amendments on arsenic and cadmium uptake by rice plants (*Oryza sativa* L. cv. Koshihikari) under different water management practices. *Soil Sci. Plant Nutr.* 62 (4), 349–356.
- Hossain, M.B., et al., 2009. The effects of iron plaque and phosphorus on yield and arsenic accumulation in rice. *Plant Soil* 317 (1–2), 167–176.
- Hu, P.J., et al., 2013a. Water management affects arsenic and cadmium accumulation in different rice cultivars. *Environ. Geochem. Health* 35 (6), 767–778.
- Hu, P.J., et al., 2013b. Effect of water management on cadmium and arsenic accumulation by rice (*Oryza sativa* L.) with different metal accumulation capacities. *J. Soils Sediments* 13 (5), 916–924.
- Huang, J.H., Matzner, E., 2007. Biogeochemistry of organic and inorganic arsenic species in a forested catchment in Germany. *Environ. Sci. Technol.* 41 (5), 1564–1569.
- Hudak, P.F., 2000. Distribution and sources of arsenic in the southern High Plains Aquifer, Texas, USA. *J. Environ. Sci. Health, Part A: Tox. Hazard. Subst. Environ. Eng.* 35 (6), 899–913.
- Hughes, M.F., 2002. Arsenic toxicity and potential mechanisms of action. *Toxicol. Lett.* 133 (1), 1–16.
- IMA, 2014. The new International Mineralogical Association list of minerals. A work in progress. International Mineralogical Society. <http://www.ima-mineralogy.org/Minlist.htm>.
- Jackson, B.P., Seaman, J.C., Bertsch, P.M., 2006. Fate of arsenic compounds in poultry litter upon land application. *Chemosphere* 65 (11), 2028–2034.
- Jackson, B.P., et al., 1999. Trace element solubility from land application of fly ash/organic waste mixtures. *J. Environ. Qual.* 28 (2), 639–647.
- Jackson, B.P., et al., 2003. Trace element speciation in poultry litter. *J. Environ. Qual.* 32 (2), 535–540.
- Japan, M.o.E.G.o., 2016. Environmental quality standards for soil pollution. Water/Soil/Ground Environment (cited 2016 12 December).
- Jia, H., et al., 2011. The phosphate transporter gene OsPht1;8 is involved in phosphate homeostasis in rice. *Plant Physiol.* 156 (3), 1164–1175.
- Jia, Y., et al., 2013a. Microbial arsenic methylation in soil and rice rhizosphere. *Environ. Sci. Technol.* 47 (7), 3141–3148.
- Jia, Y., et al., 2013b. Microbial arsenic methylation in soil and rice rhizosphere. *Environ. Sci. Technol.* 47 (7), 3141–3148.
- Juhász, A.I., Weber, J., Smith, E., 2011. Predicting arsenic relative bioavailability in contaminated soils using meta analysis and relative bioavailability-bioaccessibility regression models. *Environ. Sci. Technol.* 45 (24), 10676–10683.
- Koljonen, T., et al., 1988. 12th international geochemical exploration symposium and the 4th symposium on methods of geochemical prospecting geochemical atlas of Finland: preliminary aspects. *J. Geochim. Explor.* 32 (1), 231–242.
- Kolker, A., et al., 2004. Arsenic in coal Fly Ash Sheet, p. 4.
- Lado, L.R., Hengl, T., Reuter, H.I., 2008. Heavy metals in European soils: a geostatistical analysis of the FOREGS Geochemical database. *Geoderma* 148 (2), 189–199.
- Lasky, T., et al., 2004. Mean total arsenic concentrations in chicken 1989–2000 and estimated exposures for consumers of chicken. *Environ. Health Perspect.* 112 (1), 18–21.
- Lawgall, V.F., Meharg, A.A., 2011. Levels of arsenic and other trace elements in southern Libyan agricultural irrigated soil and non-irrigated soil projects. *Water Qual. Expo. Health* 3 (2), 79–90.
- Le, X.C., et al., 2000. Determination of monomethylarsonous acid, a key arsenic methylation intermediate, in human urine. *Environ. Health Perspect.* 108 (11), 1015–1018.
- Lee, K.W., et al., 2012. Overexpression of alfalfa mitochondrial HSP23 in prokaryotic and eukaryotic model systems confers enhanced tolerance to salinity and arsenic stress. *Biootechnol. Lett.* 34 (1), 167–174.
- Lee, C.-H., et al., 2013. Iron plaque formation and its effect on arsenic uptake by different genotypes of paddy rice. *Plant Soil* 363 (1–2), 231–241.
- Li, R.Y., et al., 2009a. The rice aquaporin Lsi1 mediates uptake of methylated arsenic species. *Plant Physiol.* 150 (4), 2071–2080.
- Li, R.Y., et al., 2009b. Mitigation of arsenic accumulation in rice with water management and silicon fertilization. *Environ. Sci. Technol.* 43 (10), 3778–3783.
- Li, R.Y., et al., 2015. Uptake and accumulation characteristics of arsenic and iron plaque in rice at different growth stages. *Commun. Soil Sci. Plant Anal.* 46 (19), 2509–2522.
- Lim, M.P., McBride, M.B., 2015. Arsenic and lead uptake by brassicas grown on an old orchard site. *J. Hazard. Mater.* 299, 656–663.
- Liu, S., et al., 2011. Arsenic removal from contaminated soil via biovolatilization by genetically engineered bacteria under laboratory conditions. *J. Environ. Sci.* 23 (9), 1544–1550.
- Lomax, C., et al., 2012. Methylated arsenic species in plants originate from soil microorganisms. *New Phytol.* 193 (3), 665–672.
- Lu, Y., et al., 2009. Baseline soil variation is a major factor in arsenic accumulation in Bengal delta paddy rice. *Environ. Sci. Technol.* 43 (6), 1724–1729.
- Luxton, T.P., Tadanier, C.J., Eick, M.J., 2006. Mobilization of arsenite by competitive interaction with silicic acid. *Soil Sci. Soc. Am. J.* 70 (1), 204–214.
- Ma, L.Q., et al., 2001. A fern that hyperaccumulates arsenic (vol 409, pg 579, 2001). *Nature* 411 (6836), 438–438.
- Ma, J.F., et al., 2007a. An efflux transporter of silicon in rice. *Nature* 448 (7150), 209–212.
- Ma, J.F., et al., 2007b. An efflux transporter of silicon in rice. *Nature* 448, 209–212.
- Ma, J.F., et al., 2008. Transporters of arsenite in rice and their role in arsenic accumulation in rice grain. *Proc. Natl. Acad. Sci. U. S. A.* 105 (29), 9931–9935.
- Ma, L., et al., 2016. Arsenic speciation in locally grown rice grains from Hunan Province, China: spatial distribution and potential health risk. *Sci. Total Environ.* 557–558, 438–444.
- Macy, J.M., et al., 2000. Two new arsenate/sulfate-reducing bacteria: mechanisms of arsenite reduction. *Arch. Microbiol.* 173 (1), 49–57.
- Mailoux, B.J., et al., 2009. Microbial mineral weathering for nutrient acquisition releases arsenic. *Appl. Environ. Microbiol.* 75 (8), 2558–2565.
- Marin, A.R., Masschelein, P.H., Patrick, W.H., 1992a. The influence of chemical form and concentration of arsenic on rice growth and tissue arsenic concentration. *Plant Soil* 139, 175–183.
- Marin, A.R., Masschelein, P.H., Patrick, W.H., 1992b. The influence of chemical form and concentration of arsenic on rice growth and tissue arsenic concentration. *Plant Soil* 139 (2), 175–183.
- Mascher, R., et al., 2002. Arsenate toxicity: effects on oxidative stress response molecules and enzymes in red clover plants. *Plant Sci.* 163 (5), 961–969.
- Matsumoto, S., et al., 2016. Evaluation of the effects of application of iron materials on the accumulation and speciation of arsenic in rice grain grown on uncontaminated soil with relatively high levels of arsenic. *Environ. Exp. Bot.* 125, 42–51.
- McBride, M.B., et al., 2015. Arsenic and lead uptake by vegetable crops grown on an old orchard site amended with compost. *Water Air Soil Pollut.* 226 (8).
- Meharg, C., Meharg, A.A., 2015. Silicon, the silver bullet for mitigating biotic and abiotic stress, and improving grain quality, in rice? *Environ. Exp. Bot.* 120, 8–17.
- Meharg, A.A., Zhao, F.J., 2012. *Arsenic & Rice*. Springer Verlag, Dordrecht (xi, 171 p).
- Mei, X.Q., et al., 2012. The effects of radial oxygen loss on arsenic tolerance and uptake in rice and on its rhizosphere. *Environ. Pollut.* 165, 109–117.
- Melegy, A., et al., 2011. Weathering fluxes of arsenic from a small catchment in Slovakia Republic. *Environ. Earth Sci.* 64 (2), 549–555.
- Mestrot, A., et al., 2011. Field fluxes and speciation of arsines emanating from soils. *Environ. Sci. Technol.* 45 (5), 1798–1804.
- Mestrot, A., et al., 2013. Arsenic volatilization in model anaerobic biogas digesters. *Appl. Geochem.* 33, 294–297.
- Mikutta, C., Rothwell, J.J., 2016. Peat bogs as hotspots for organoarsenical formation and persistence. *Environ. Sci. Technol.* 50 (8), 4314–4323.
- Mitani, N., et al., 2009. Identification and characterization of maize and barley Lsi1-like silicon efflux transporters reveals a distinct silicon uptake system from that in rice. *Plant Cell* 21 (7), 2133–2142.
- Moon, K., Guallar, E., Navas-Acien, A., 2012. Arsenic exposure and cardiovascular disease: an updated systematic review. *Curr. Atheroscler. Rep.* 14 (6), 542–555.
- Murphy, E.A., Autcott, M., 1998. An assessment of the amounts of arsenical pesticides used historically in a geographical area. *Sci. Total Environ.* 218 (2–3), 89–101.
- Nachman, K.E., et al., 2013. Roxarsone, inorganic arsenic, and other arsenic species in chicken: a US-based market basket sample. *Environ. Health Perspect.* 121 (7), 818–824.
- Newman, E.L., 1985. The rhizosphere: carbon sources and microbial populations. In: Fitter, A.H., Atkinson, D., Read, D.J. (Eds.), *Ecological Interactions in Soil: Plants, Microbes and Animals*. Blackwell Scientific, Oxford, UK, pp. 107–121.
- Newman, D.K., et al., 1997. Dissimilatory arsenate and sulfate reduction in *Desulfovomaculum auripigmentum* sp. nov. *Arch. Microbiol.* 168 (5), 380–388.
- Norton, G.J., et al., 2009. Identification of low inorganic and total grain arsenic rice cultivars from bangladesh. *Environ. Sci. Technol.* 43 (15), 6070–6075.
- Norton, G.J., et al., 2012. Variation in grain arsenic assessed in a diverse panel of rice (*Oryza sativa*) grown in multiple sites. *New Phytol.* 193 (3), 650–664.
- Norton, G., et al., 2013. Arsenic speciation and localization in horticultural produce grown in a Historically impacted mining region. *Environ. Sci. Technol.* 47 (12), 6164–6172.
- Norton, G.J., et al., 2014. Genome wide association mapping of grain arsenic, copper, molybdenum and zinc in rice (*Oryza sativa* L.) grown at four international field sites. *PLoS One* 9 (2), e89685.
- Penedo, E.S., et al., 2016. Biogeochemical impacts of silicon-rich rice residue incorporation into flooded soils: implications for rice nutrition and cycling of arsenic. *Plant Soil* 399 (1–2), 75–87.
- Perryea, F.J., Creger, T.J., 1994. Vertical-distribution of lead and arsenic in soils Contaminated with lead arsenate pesticide-residues. *Water Air Soil Pollut.* 78 (3–4), 297–306.
- Perryea, F.J., Kammerer, R., 1997. Phosphate-enhanced movement of arsenic out of lead arsenate-contaminated topsoil and through uncontaminated subsoil. *Water Air Soil Pollut.* 93 (1–4), 243–254.
- Pizarro, I., et al., 2016. Bioavailability and arsenic speciation in carrots, beets and quinoa from a contaminated area of Chile. *Sci. Total Environ.* 565, 557–563.
- Pommerenig, B., Diehn, T.A., Bienert, G.P., 2015. Metalloido-porins: essentiality of Nodulin 26-like intrinsic proteins in metalloid transport. *Plant Sci.* 238, 212–227.
- Pongratz, R., 1998. Arsenic speciation in environmental samples of contaminated soil. *Sci. Total Environ.* 224 (1–3), 133–141.
- Poushat, P., Zagury, G.J., 2006. In vitro gastrointestinal bioavailability of arsenic in soils collected near CCA-treated utility poles. *Environ. Sci. Technol.* 40 (13), 4317–4323.

- Raab, A., et al., 2007. Uptake and translocation of inorganic and methylated arsenic species by plants. *Environ. Chem.* 4 (3), 197–203.
- Rahman, M.A., et al., 2008. Straighthead disease of rice (*Oryza sativa L.*) induced by arsenic toxicity. *Environ. Exp. Bot.* 62 (1), 54–59.
- Ramirez-Andreotta, M.D., et al., 2013a. Home gardening near a mining site in an arsenic-endemic region of Arizona: assessing arsenic exposure dose and risk via ingestion of home garden vegetables, soils, and water. *Sci. Total Environ.* 454, 373–382.
- Ramirez-Andreotta, M.D., et al., 2013b. A green house and field-based study to determine the accumulation of arsenic in common homegrown vegetables grown in mining-affected soils. *Sci. Total Environ.* 443, 299–306.
- Reedy, R.C., et al., 2007. Unsaturated zone arsenic distribution and implications for groundwater contamination. *Environ. Sci. Technol.* 41 (20), 6914–6919.
- Remy, E., et al., 2012. The Pht1;9 and Pht1;8 transporters mediate inorganic phosphate acquisition by the *Arabidopsis thaliana* root during phosphorus starvation. *New Phytol.* 195 (2), 356–371.
- Renshaw, C.E., et al., 2006. Impact of land disturbance on the fate of arsenical pesticides. *J. Environ. Qual.* 35 (1), 61–67.
- Robinson, G.R., et al., 2007. Assessment of contamination from arsenical pesticide use on orchards in the Great Valley region, Virginia and West Virginia, USA. *J. Environ. Qual.* 36 (3), 654–663.
- Rosen, B.P., 1999. Families of arsenic transporters. *Trends Microbiol.* 7 (5), 207–212.
- Rutherford, D.W., et al., 2003. Environmental fate of roxarsone in poultry litter, part II. Mobility of arsenic in soils amended with poultry litter. *Environ. Sci. Technol.* 37 (8), 1515–1520.
- Sadee, B.A., Foulkes, M.E., Hill, S.J., 2016. A study of arsenic speciation in soil, irrigation water and plant tissue: a case study of the broad bean plant, *Vicia faba*. *Food Chem.* 210, 362–370.
- Sanchez-Bermudo, E., et al., 2014. Natural variation in arsenate tolerance identifies an arsenate reductase in *Arabidopsis thaliana*. *Nat. Commun.* 5, 4617.
- Saunders, J.A., et al., 2005. Natural arsenic contamination of Holocene alluvial aquifers by linked tectonic, weathering, and microbial processes. *Geochim. Geophys. Geosyst.* 6.
- Schmidt, C., 2010. Coal ash spill still a problem. *Chem. Eng. News Arch.* 88 (49), 13.
- Schmoger, M.E., Owen, M., Grill, E., 2000. Detoxification of arsenic by phytochelatins in plants. *Plant Physiol.* 122 (3), 793–801.
- Schneider, S., et al., 2006. *Arabidopsis INOSTOL TRANSPORTER4* mediates high-affinity H(+) symport of myoinositol across the plasma membrane. *Plant Physiol.* 141 (2), 565–577.
- Schneider, S., et al., 2007. *Arabidopsis INOSTOL TRANSPORTER2* mediates H(+)-symport of different inositol epimers and derivatives across the plasma membrane. *Plant Physiol.* 145 (4), 1395–1407.
- Schumann, A.W., Summer, M.E., 1999. Plant nutrient availability from mixtures of fly ashes and biosolids. *J. Environ. Qual.* 28 (5), 1651–1657.
- Schumann, A.W., Summer, M.E., 2004. Formulation of environmentally soundwaste mixtures for land application. *Water Air Soil Pollut.* 152 (1–4), 195–217.
- Seyffarth, A.L., 2015. Abiotic effects of dissolved oxyanions on iron plaque quantity and mineral composition in a simulated rhizosphere. *Plant Soil* 397 (1–2), 43–61.
- Seyffarth, A., et al., 2011. Defining the distribution of arsenic species and plant nutrients in rice (*Oryza sativa L.*) from the root to the grain. *Geochim. Cosmochim. Acta* 75, 6655–6671.
- Seyffarth, A.L., et al., 2016. Soil incorporation of silica-rich rice husk decreases inorganic arsenic in rice grain. *J. Agric. Food Chem.* 64 (19), 3760–3766.
- Shacklette, H.T., Boerner, J.G., 1984. Element Concentrations in soils and other surficial materials of the Contiguous United States. 105. US Geological Survey. US Government Printing Office, Washington, DC.
- Shah, S.B., et al., 2009. Leaching of nutrients and trace elements from stockpiled turkey litter into soil. *J. Environ. Qual.* 38 (3), 1063–1065.
- Sharma, I., 2012. Arsenic induced oxidative stress in plants. *Biology* 67 (3), 447–453.
- Shin, H., et al., 2004. Phosphate transport in *Arabidopsis*: Pht1;1 and Pht1;4 play a major role in phosphate acquisition from both low- and high-phosphate environments. *Plant J.* 39 (4), 629–642.
- Shri, M., et al., 2009. Effect of arsenic on growth, oxidative stress, and antioxidant system in rice seedlings. *Ecotoxicol. Environ. Saf.* 72 (4), 1102–1110.
- Smeaton, C.M., et al., 2012. Simultaneous Release of Fe and As during the Reductive Dissolution of Pb-As Jarosite by *Shewanella putrefaciens* CN32. *Environ. Sci. Technol.* 46 (23), 12823–12831.
- Smedley, P.L., Kinniburgh, D.G., 2002. A review of the source, behaviour and distribution of arsenic in natural waters. *Appl. Geochem.* 17 (5), 517–568.
- Smith, D.B., et al., 2014. In: U.G. Survey (Ed.), Geochemical and Mineralogical Maps for Soils of the Contiguous United States: U.S., p. 386.
- Song, W.Y., et al., 2010a. Arsenic tolerance in *Arabidopsis* is mediated by two ABC-type phytochelatin transporters. *Proc. Natl. Acad. Sci. U.S.A.* 107 (49), 21187–21192.
- Song, W.Y., et al., 2010b. Arsenic tolerance in *Arabidopsis* is mediated by two ABC-type phytochelatin transporters. *Proc. Natl. Acad. Sci. U.S.A.* 107 (49), 21187–21192.
- Song, W.Y., et al., 2014a. A rice ABC transporter, OsABC1, reduces arsenic accumulation in the grain. *Proc. Natl. Acad. Sci. U.S.A.* 111 (44), 15699–15704.
- Song, W.Y., et al., 2014b. A rice ABC transporter, OsABC1, reduces arsenic accumulation in the grain. *Proc. Natl. Sci. U.S.A.* 111 (44), 15699–15704.
- Stewart, E.J., 2012. Growing Unculturable bacteria. *J. Bacteriol.* 194 (16), 4151–4160.
- Stoeva, N., Berova, M., Zlatev, Z., 2003. Physiological response of maize to arsenic contamination. *Biol. Plant.* 47 (3), 449–452.
- Stolz, J.F., et al., 2007. Biotransformation of 3-nitro-4-hydroxybenzene aronic acid (roxarsone) and release of inorganic arsenic by *Clostridium* species. *Environ. Sci. Technol.* 41 (3), 818–823.
- Styblo, M., et al., 2000. Comparative toxicity of trivalent and pentavalent inorganic and methylated arsenicals in rat and human cells. *Arch. Toxicol.* 74 (6), 289–299.
- Sun, S., et al., 2012. A constitutive expressed phosphate transporter, OsPht1;1, modulates phosphate uptake and translocation in phosphate-replete rice. *Plant Physiol.* 159 (4), 1571–1581.
- Tang, Z., et al., 2016. Arsenic methylation in *Arabidopsis thaliana* expressing an algal arsenite methyltransferase gene increases arsenic phytotoxicity. *J. Agric. Food Chem.* 64 (13), 2674–2681.
- Taylor, D.A., 2004. Funky chicken - consumers exposed to arsenic in poultry. *Environ. Health Perspect.* 112 (1), A50.
- Tiberg, C., et al., 2016. Immobilization of Cu and As in two contaminated soils with zero-valent iron - long-term performance and mechanisms. *Appl. Geochem.* 67, 144–152.
- Townsend, T.G., et al., 2003. Chromium, copper and arsenic concentrations in soil underneath CCA-treated wood structures. *Soil Sediment Contam.* 12, 779–798.
- Turpeinen, R., Päntsä-Kallio, M., Kairesalo, T., 2002. Role of microbes in controlling the speciation of arsenic and production of arsines in contaminated soils. *Sci. Total Environ.* 285 (1–3), 133–145.
- UKSO, 2016. NSI topsoil Arsenic map. Topsoil arsenic map of England and Wales cited 2016 August 11. [Available from: <http://www.ukso.org/nsi/Arsenic.html>].
- Ulrich-Ebner, C.I., Sanz, A., Novacky, A.J., 1989. Evaluation of arsenate-associated and vanadate-associated changes of electrical membrane potential and phosphate transport in Lemma gibba. *G1.J. Exp. Bot.* 40 (210), 119–128.
- USFDA, d. Arsenic-based Animal Drugs and Poultry. [Available from: <http://www.fda.gov/AnimalVeterinary/SafetyHealth/ProductSafetyInformation/umc257540.htm>].
- USGS, 2016. Mineral Resources On-Line Spatial Data: National Geochemical Survey database. September 3 2014 [cited 2016 July 18]; National scale geochemical analysis of stream sediments and soils in the US, from existing data, reanalysis of existing samples and new sampling [Available from: <http://mrdata.usgs.gov/geochemistry/ngs.html>] [Website].
- Wang, P., et al., 2014. A review on completing arsenic biogeochemical cycle: microbial volatilization of arsines in environment. *J. Environ. Sci.* 26 (2), 371–381.
- Wang, Y., et al., 2016a. Release and transformation of arsenic from As-bearing iron minerals by Fe-reducing bacteria. *Chem. Eng. J.* 295, 29–38.
- Wang, J., et al., 2016b. Biotransformation and biomethylation of arsenic by *Shewanella oneidensis* MR-1. *Chemosphere* 145, 329–335.
- WHO, 2016. Arsenic. International Programme on Chemical Safety. Health Impacts of Chemicals cited 2016 November 22, 2016. [Available from: http://www.who.int/ipcs/assessment/public_health/arsenic/en].
- Williams, P.N., et al., 2005a. Variation in arsenic speciation and concentration in paddy rice related to dietary exposure. *Environ. Sci. Technol.* 39 (15), 5531–5540.
- Williams, P.N., et al., 2005b. Variation in arsenic speciation and concentration in paddy rice related to dietary exposure. *Environ. Sci. Technol.* 39 (15), 5531–5540.
- Williams, P.N., et al., 2007. Greatly enhanced arsenic shoot assimilation in rice leads to elevated grain levels compared to wheat and barley. *Environ. Sci. Technol.* 41 (19), 6854–6859.
- Williams, P.N., et al., 2009. Occurrence and partitioning of Cadmium, Arsenic and Lead in mine impacted paddy rice. Hunan, China. *Environ. Sci. Technol.* 43 (3), 637–642.
- Williams, P.N., et al., 2011. Organic matter-solid phase interactions are critical for predicting arsenic release and plant uptake in Bangladesh paddy soils. *Environ. Sci. Technol.* 45 (14), 6080–6087.
- Wu, C., et al., 2012. Do radial oxygen loss and external aeration affect iron plaque formation and arsenic accumulation and speciation in rice? *J. Exp. Bot.* 63 (8), 2961–2970.
- Wu, C., et al., 2016. The effect of silicon on iron plaque formation and arsenic accumulation in rice genotypes with different radial oxygen loss (ROL). *Environ. Pollut.* 212, 27–33.
- Xiao, K.-Q., et al., 2016. Metagenomic analysis revealed highly diverse microbial arsenic metabolism genes in paddy soils with low-arsenic contents. *Environ. Pollut.* 211, 1–8.
- Xu, L., et al., 2016. Speciation change and redistribution of arsenic in soil under anaerobic microbial activities. *J. Hazard. Mater.* 301, 538–546.
- Yager, J.W., Greene, T., Schoof, R.A., 2015. Arsenic relative bioavailability from diet and airborne exposures: implications for risk assessment. *Sci. Total Environ.* 536, 368–381.
- Yamaguchi, N., et al., 2014. Arsenic distribution and speciation near rice roots influenced by iron plaques and redox conditions of the soil matrix. *Environ. Sci. Technol.* 48 (3), 1549–1556.
- Yamamura, S., et al., 2001. Drinking water guidelines and standards. United Nations Synthesis Report on Arsenic in Drinking water. United Nations, United Nation, Geneva, p. 18.
- Yang, J.K., et al., 2005. Adsorption, oxidation, and bioaccessibility of As(III) in soils. *Environ. Sci. Technol.* 39 (18), 7102–7110.
- Ye, Y., et al., 2015. The phosphate transporter Gene OsPht1;4 is involved in phosphate homeostasis in rice. *PLoS One* 10 (5), e0126186.
- Yokel, J., Delistraty, D.A., 2003. Arsenic, lead, and other trace elements in soils contaminated with pesticide residues at the Hanford site (USA). *Environ. Toxicol.* 18 (2), 104–114.
- Young, M., Donald, A., 2013. A Guide to the Tellus Data. Geological Survey of Northern Ireland, Belfast, UK, p. 233.
- Zavalav, Y.J., Duxbury, J.M., 2008. Arsenic in rice: I. Estimating normal levels of total arsenic in rice grain. *Environ. Sci. Technol.* 42 (10), 3856–3860.
- Zavalav, Y.J., et al., 2008. Arsenic in rice: II. Arsenic speciation in USA grain and implications for human health. *Environ. Sci. Technol.* 42 (10), 3861–3866.
- Zenk, M.H., 1996. Heavy metal detoxification in higher plants—a review. *Gene* 179 (1), 21–30.
- Zhang, S., et al., 2011. Cometabolic biotransformation of fenpropatrin by *Clostridium* species strain ZP3. *Biodegradation* 22 (5), 869–875.
- Zhao, F.J., et al., 2009. Arsenic uptake and metabolism in plants. *New Phytol.* 181, 777–794.
- Zhao, F.J., et al., 2010. The role of the rice aquaporin Lsi1 in arsenite efflux from roots. *New Phytol.* 186 (2), 392–399.
- Zhu, Y.G., et al., 2008. High percentage inorganic arsenic content of mining impacted and nonimpacted Chinese rice. *Environ. Sci. Technol.* 42 (13), 5008–5013.

Appendix IV: Opportunities and Challenges for Dietary Arsenic Intervention

All authors contributed to the writing, with Keeve Nachman serving as corresponding author.

Brief Communication

A Section 508-conformant HTML version of this article
is available at <https://doi.org/10.1289/EHP3997>.

Opportunities and Challenges for Dietary Arsenic Intervention

Keeve E. Nachman,^{1,2,3,4} Tracy Punshon,^{5,6,7} Laurie Rardin,⁶ Antonio J. Signes-Pastor,^{6,7,9} Carolyn J. Murray,⁷ Brian P. Jackson,^{6,8} Mary Lou Guerinot,⁵ Thomas A. Burke,^{1,3} Celia Y. Chen,^{5,6} Habibul Ahsan,¹⁰ Maria Argos,¹¹ Kathryn L. Cottingham,^{5,7} Francesco Cubadda,¹² Gary L. Ginsberg,¹³ Britton C. Goodale,^{6,14} Margaret Kurzus-Spencer,^{15,16} Andrew A. Meharg,¹⁷ Mark D. Miller,¹⁸ Anne E. Nigra,¹⁹ Claire B. Pendergrast,²⁰ Andrea Raab,²¹ Ken Reimer,²² Kirk G. Scheckel,²³ Tanja Schwerdtle,²⁴ Vivien F. Taylor,^{6,8} Erik J. Tokar,²⁵ Todd M. Warczak,⁵ and Margaret R. Karagas^{6,7,9}

¹Risk Sciences and Public Policy Institute, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA

²Department of Environmental Health and Engineering, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA

³Department of Health Policy and Management, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA

⁴Johns Hopkins Center for a Livable Future, Johns Hopkins University, Baltimore, Maryland, USA

⁵Department of Biological Sciences, Dartmouth College, Hanover, New Hampshire, USA

⁶Dartmouth Superfund Research Program, Hanover, New Hampshire, USA

⁷Dartmouth Children's Environmental Health and Disease Prevention Research Center, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire, USA

⁸Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire, USA

⁹Department of Epidemiology, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA

¹⁰Department of Public Health Sciences, University of Chicago, Chicago, Illinois, USA

¹¹Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, Illinois, USA

¹²Department of Food Safety, Nutrition and Veterinary Public Health, Istituto Superiore di Sanità – Italian National Institute of Health, Rome, Italy

¹³Yale School of Public Health, 60 College St, New Haven, Connecticut, USA

¹⁴Department of Microbiology and Immunology, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA

¹⁵Department of Pediatrics, College of Medicine, University of Arizona, Tucson, Arizona, USA

¹⁶Department of Community, Environment and Policy, Mel & Enid College of Public Health, University of Arizona, Tucson, Arizona, USA

¹⁷Institute for Global Food Security, Queen's University Belfast, David Keir Building, Malone Road, Belfast, BT9 5BN, Northern Ireland, UK

¹⁸Western States Pediatric Environmental Health Specialty Unit, University of California, San Francisco, San Francisco, California, USA

¹⁹Department of Environmental Health Sciences, Columbia University Mailman School of Public Health, New York, New York, USA

²⁰University of Washington School of Public Health, Seattle, Washington, USA

²¹Department of Chemistry, University of Aberdeen, Aberdeen, UK

²²Royal Military College, Kingston, Ontario, Canada

²³Land and Materials Management Division, National Risk Management Research Laboratory, United States Environmental Protection Agency, Cincinnati, Ohio, USA

²⁴Institute of Nutritional Sciences, University of Potsdam, Germany

²⁵National Toxicology Program Laboratory, National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

SUMMARY: The diet is emerging as the dominant source of arsenic exposure for most of the U.S. population. Despite this, limited regulatory efforts have been aimed at mitigating exposure, and the role of diet in arsenic exposure and disease processes remains understudied. In this brief, we discuss the evidence linking dietary arsenic intake to human disease and discuss challenges associated with exposure characterization and efforts to quantify risks. In light of these challenges, and in recognition of the potential longer-term process of establishing regulation, we introduce a framework for shorter-term interventions that employs a field-to-plate food supply chain model to identify monitoring, intervention, and communication opportunities as part of a multisector, multiagency, science-informed, public health systems approach to mitigation of dietary arsenic exposure. Such an approach is dependent on coordination across commodity producers, the food industry, nongovernmental organizations, health professionals, researchers, and the regulatory community. <https://doi.org/10.1289/EHP3997>

Background

Emerging data suggest that arsenic, even at relatively low levels of exposure, may affect human health, particularly during early life (NRC 2014). In consideration of this issue, regulatory efforts have been mobilized to address exposures, primarily from drinking water (Nachman et al. 2017). As successes are increasingly documented in these exposure-reduction efforts (Nigra et al. 2017; Welch et al. 2018), we view the diet as a driving source of exposure in populations with low drinking water arsenic. Despite

its contributions to aggregate arsenic exposures, efforts to tackle arsenic in food have been relatively sparse. The Collaborative on Food with Arsenic and associated Risk and Regulation (C-FARR) brought arsenic and food scientists together with policy stakeholders for a workshop focusing on knowledge gaps and policy questions in recognition of this lagging focus. The resulting five papers address this issue from soil to plate to policy (Cubadda et al. 2017; Davis et al. 2017; Nachman et al. 2017; Punshon et al. 2017; Taylor et al. 2017). Moving beyond the C-FARR workshop, and considering scientific and policy hurdles, we discuss here an array of immediate-term opportunities that exist among multiple stakeholder groups to intervene on dietary arsenic exposures.

Arsenic exposure through drinking water has established health impacts (NRC 2014), and although existing science supports the assumption that effects will be similar from food, the role of diet in population arsenic exposures has been less studied. Most epidemiologic studies describing arsenic's effects have relied on drinking water concentrations or urine measurements as exposure measures, which are not designed to disentangle the role of dietary arsenic in the occurrence of disease (Nachman et al. 2017). Further, in comparison with drinking water, where arsenic occurs solely in inorganic forms, the species profile for

Address correspondence to K.E. Nachman, 615 N. Wolfe St., W7010-E, Baltimore, MD 21205, USA. Telephone: (410) 223-1811; Fax: (410) 502-7579. Email: k.nachman@jhu.edu

The authors declare they have no actual or potential competing financial interests.

Received 3 June 2018; Revised 16 July 2018; Accepted 20 July 2018;
Published 31 August 2018.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

food is complex, comprising inorganic arsenic (arsenite and arsenate) and more than 100 organic arsenic species that vary within and across foods (Cubadda et al. 2017). Although epidemiologic studies have provided clear evidence for the carcinogenicity of inorganic arsenic (iAs) and link it to myriad noncancer outcomes, our understanding of the chronic toxicity of organic species is emerging, further complicating characterization of dietary arsenic exposures (Taylor et al. 2017).

Although drinking water may be an easier medium for epidemiologists to use to examine arsenic exposure, it is not the dominant pathway for most of the U.S. population. An earlier assessment of public water in 25 states estimated that <4% of people were served by systems with arsenic concentrations >1 µg/L, and just over 1% were above the 10 µg/L U.S. EPA Maximum Contaminant Level (MCL) (Mushak et al. 2000). Since then, public water exposures have declined, following the implementation of the MCL (Nigra et al. 2017). Evaluations of modeled exposure across dietary and drinking water pathways suggest that in those with drinking water arsenic concentrations below the MCL, diet is responsible for the majority of both total and iAs intake (Kurzius-Spencer et al. 2014; Xue et al. 2010).

Despite the importance of the dietary pathway, sparse data exist about the role of dietary arsenic exposure in disease processes. Among these investigations, most have focused on rice, and very few have jointly considered drinking water exposures along with other potential dietary sources. Additional challenges in characterizing the role of diet include difficulty in separating out the effects of arsenic in the water used to cook rice in high water arsenic regions, estimating arsenic levels in rice and rice products and adjusting for the effects of dietary and other confounders (e.g., carbohydrate content of rice). To date, few studies raise the possibility of adverse health impacts of rice consumption among those with relatively low drinking water arsenic concentrations. These studies found a trend in cardiovascular disease risk with white (but not brown or overall) rice consumption in the prospective Nurses' Health Study among participants living in low arsenic regions (<3 µg/L), but not among those living in higher arsenic regions (Muraki et al. 2015). In Bangladesh, Melkonian et al. (2013) found a relationship between steamed rice consumption and premalignant and malignant skin lesions in both cross-sectional and prospective analyses of the Health Effects of Arsenic Longitudinal Study only among those with drinking water arsenic concentrations <100 µg/L. In the United States, an association between rice consumption and squamous cell carcinoma of the skin was observed in a population-based study, largely among those with tap water concentrations of arsenic <1 µg/L (Gossai et al. 2017). Using data from the Nurses' Health Study and the Health Professionals Follow-up Study, Zhang et al. (2016) found no significant increases in cancer risk from rice consumption but identified a borderline significant increased bladder cancer risk when comparing the highest rice-consumption group to the lowest; the study did not involve consideration of drinking water arsenic. Thus, although limited, epidemiologic evidence currently suggests that the effects of rice consumption may mirror those of arsenic in drinking water.

Dietary exposure characterization poses unique methodological challenges, as few studies describe distributions of arsenic across foods, and far fewer involve speciation (Cubadda et al. 2017). These data gaps complicate efforts to quantify dietary exposures and limit the ability to target foods for intervention based on their relative contributions to aggregate intake.

Though risks from food cannot be precisely quantified, existing evidence supports efforts to limit exposure wherever possible, especially for vulnerable populations (Naujokas et al. 2013). Despite the dominance of diet in population arsenic

exposures, arsenic in food remains unregulated in the United States (Nachman et al. 2017). Under normal circumstances, regulatory agencies would be expected to play a key role in drafting enforceable interventions aimed at minimizing population exposures. In the case of arsenic, no enforceable regulatory standards exist for any foods, though the U.S. Food and Drug Administration (U.S. FDA) has proposed draft action levels and accompanying industry guidance for apple juice (U.S. FDA 2013) and infant rice cereal (U.S. FDA 2016). Legislative attempts to force stringent regulatory action (U.S. Congress 2012, 2015, 2017) have thus far been unsuccessful, though in its 2018 evaluation of U.S. FDA and U.S. Department of Agriculture (USDA) action on arsenic in rice, the U.S. Government Accountability Office (U.S. GAO) recommended FDA finalize its guidance on arsenic in infant rice cereal and called for improvements in interagency coordination on analytical methods and risk-assessment approaches for arsenic and other food contaminants (U.S. GAO 2018).

We previously described an approach that could be used to reduce dietary iAs exposure in the U.S. population. It employs existing monitoring data (and proposes newer data streams) to facilitate an iterative reevaluation of population dietary iAs exposures, and subsequently prioritizes specific foods for intervention, both in the form of measures to reduce arsenic concentrations in key foods (via action levels and regulatory standards to compel producer efforts to reduce arsenic content) and dietary advice when those reductions are more difficult (Nachman et al. 2017). If adequately protective standards are implemented and properly enforced (Signes-Pastor et al. 2017b), this type of holistic approach, which recognizes the pervasiveness of iAs in the food supply (Schoof et al. 1999), may hold promise for reducing population exposures. To complement this approach, especially in circumstances where regulatory action may occur more slowly (Krisberg 2017; Samet et al. 2017), other opportunities exist that can be implemented in the short term and may yield public health benefits stemming from arsenic exposure reductions.

Discussion

In the absence of regulations specific to arsenic in food, we argue that there are many opportunities for nonregulatory stakeholders to mitigate dietary exposures based on our current understanding of the food supply chain. Furthermore, we believe it is possible to reduce exposure to arsenic through control of source releases, and demonstration of ways to reduce the arsenic content of retail foods may pave the way for future regulatory changes at the federal level.

Figure 1 depicts intervention and monitoring points along the food supply that may be meaningfully employed by various stakeholders to reduce population dietary exposures. Given that rice has been the subject of considerable intervention research, we use it in demonstrating opportunities. We recognize, however, as we have previously specified (Nachman et al. 2017), that a coordinated, iterative monitoring intervention strategy that considers the contribution of all foods and beverages would begin to address the current needs of reducing exposure.

Rice cultivar selection and agronomic practices dominate as driving factors of grain arsenic concentrations (Norton et al. 2012; Yang et al. 2017). The influence of genetic variation on grain arsenic indicates that breeding low-arsenic rice cultivars is a viable approach and is currently underway (Norton et al. 2009). Manipulating flooding cycles in rice production, tied to sustainable agricultural practices for reducing water use, can also reduce grain arsenic (Li et al. 2009; Yang et al. 2017). Fertilization strategies, particularly using silicon, have potential as an effective strategy to reduce arsenic, and prevent certain plant diseases (e.g., straighthead disease in rice) (Limmer et al. 2018). Equally

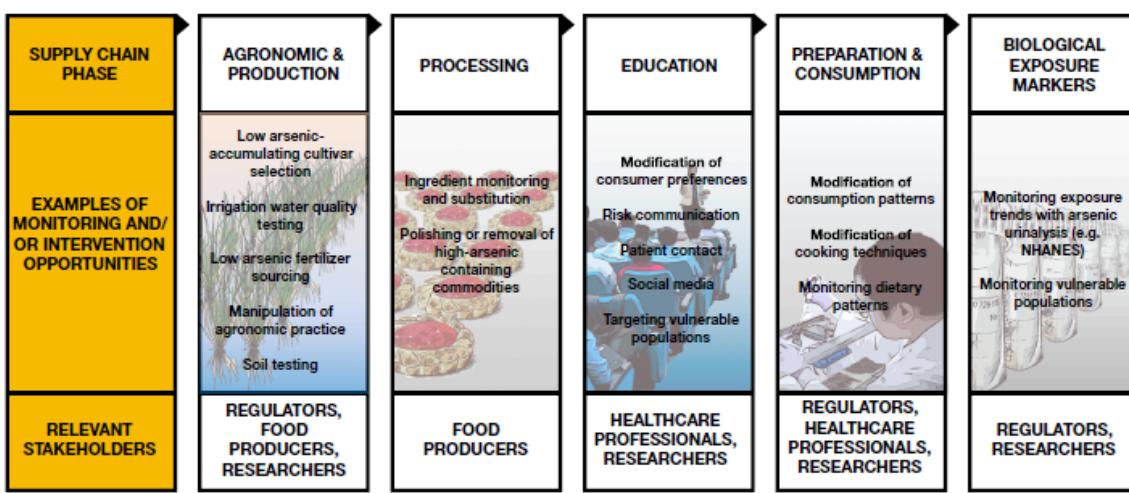


Figure 1. The food supply chain and opportunities for dietary arsenic exposure monitoring and intervention. A field-to-plate food supply chain model can aid in identifying monitoring, intervention and communication opportunities as part of a multisector, multiagency, science-informed, public health systems approach to mitigation of dietary arsenic exposure.

important to demonstrating efficacy of these approaches, is making them accessible and attractive to small scale rural farming operations; for example, re-using rice husk as a silicon fertilizer for rice paddy soils [shown to lower total and iAs (Penido et al. 2016; Seyfferth et al. 2016; Teasley et al. 2017)] has great potential because silicon fertilizers are too expensive to be implemented at many smallholding farms, and, if used, would potentially drive up the price of rice.

Recent Codex Alimentarius international food standards codes of practice (Codex 2017) lay out logical steps for the prevention and reduction of arsenic contamination in rice, grouped into: source-directed measures (determining that the arsenic concentrations of soils and irrigation water are not elevated before use; cognizance of proximity to point sources of arsenic contamination; avoiding use of arsenical pesticides and veterinary drugs, feed, soil amendments, and fertilizer), agricultural measures (use of alternative intermittent flooding strategies and low-arsenic cultivars), and monitoring and risk communication programs. Adherence to this code has the potential to dramatically reduce arsenic concentrations, particularly if food-product labeling were used to raise consumer awareness.

Recognizing that the arsenic content of many finished foods is dependent on the raw commodities that serve as ingredients, opportunities exist regarding both the selection of ingredients and quality assurance measures related to the arsenic concentrations of those constituents. For example, the replacement of high-fructose corn syrup with organic brown-rice syrup in products like toddler formula and energy bars (Jackson et al. 2012) and the use of the algae-derived gelatin substitute carrageenan in a variety of foods (Díaz et al. 2012; U.S. FDA 2017b) may contribute to dietary arsenic intake, although the health impacts of constituents require further testing. Currently, no programs exist to monitor the arsenic content of additives or ingredients, and recommended international standards for some of these inputs may not be based upon the most recent evidence (Food and Agriculture Organization 2001). Processor-level preventive actions, such as ingredient monitoring or ingredient substitution with others less likely to accumulate arsenic, may be beneficial.

Additional reductions may be achievable during food and ingredient processing by taking advantage of our understanding of

the nature of iAs accumulation within commodities. For example, rice-processing procedures designed to remove the husk (which has been shown to accumulate iAs at levels as high as 1 ppm) and further polish the grain can achieve substantial iAs reductions (Carey et al. 2015; Meharg et al. 2008; Signes-Pastor et al. 2017a; Signes et al. 2008). Further, lower iAs content in rice-based products can be achieved by dilution of rice with other gluten-free grain (Carey et al. 2018).

In some instances when adequate exposure reduction cannot be accomplished by lowering the arsenic content of foods, alternative approaches to reduce dietary exposures may be warranted, such as modifying consumer preferences to influence purchasing behaviors (which would exert pressure on producers to address the arsenic content of their foods) (Nachman et al. 2017).

Interventions to communicate information on arsenic in food are somewhat limited and often rice-centric, though some helpful sources exist, including a comprehensive website focused on arsenic in rice and rice products by Consumer Reports (2015). The U.S. FDA (2017a) also maintains a communications and data-sharing website for arsenic in food. Research-based consumer education resources highlighting arsenic health risks and dietary (and other) exposure-reduction options are available online in a comprehensive website (Dartmouth College Superfund Research Program 2017) and an interactive graphic (Children's Environmental Health and Disease Prevention Research Center at Dartmouth 2016). Beyond these vetted resources, social media sites and blogs from gluten-free dietary information resources and Celiac disease organizations (e.g., Celiac Disease Foundation 2017) also provide exposure-reduction resources. Research is needed to determine the effectiveness of these tools on individual behavior change and consumer decision-making. Additional communication efforts need to be tailored and targeted to populations at higher risk of adverse health effects from arsenic exposure, such as infants, children, and pregnant women, as well as to populations more highly exposed due to cultural preference and dietary restrictions (Lai et al. 2015).

Communication efforts successful in changing consumer behaviors may have collateral benefits along the food supply chain. Research published by Jackson et al. (2012) identifying elevated iAs in organic brown-rice syrup used in toddler formula resulted in a media cycle and public outcry that prompted some

food manufacturers to take action to address arsenic in their products (Lundberg Family Farms 2016; Nature's One 2018), illustrating how media attention and consumer opinion can drive industry action to reduce arsenic content in food products even without regulation.

Recent research has highlighted post-retail opportunities for consumers to reduce the arsenic content of finished foods. Several studies have shown that some of the iAs content is readily leached from rice by washing and soaking, and even more can be removed by increasing the volume of cooking water that rice encounters, followed by draining off the excess iAs-containing water generated during cooking (Raab et al. 2009; Sengupta et al. 2006; Signes et al. 2008). To this end, a new rice cooking approach has been proposed based on a continual stream of percolating near-boiling water; when low-arsenic water has been used, this method has been demonstrated to remove ~57% of iAs from cooked rice at water-to-rice cooking ratio of 12:1, and a removal of ~70% of iAs with a ratio of 150:1 or higher in cooked rice bran (Signes-Pastor et al. 2017a). Proponents of this approach have suggested its feasibility both in industrial and domestic settings, and possibly even when low arsenic cooking water is not available by percolating freshly recycled distilled water (Carey et al. 2015). Such an approach may provide a near-term, suitable solution to mitigate high iAs levels in rice and rice bran, though further research is needed to assess potential loss of key vitamins and other water-soluble compounds that could be refortified if necessary.

Existing population-level health surveys, especially those that describe dietary patterns and collect biomarker data, can help target dietary interventions, and to a lesser extent, assist in evaluating existing exposure-reduction efforts. The What We Eat in America food survey, conducted as part of the National Health and Nutrition Examination Survey (USDA 2018), enables the characterization of dietary patterns across key population subgroups, including both vulnerable life stages regarding arsenic exposure and population subgroups with higher consumption of arsenic-accumulating foods. Urinary arsenic concentrations have been consistently related to rice consumption among populations throughout the world, and in both interventional and observational studies (Davis et al. 2017). These studies encompass populations of various ages, ranging from infancy and early childhood to pregnancy and later adulthood. Arsenic is known to pass through the placenta (e.g., Hall et al. 2007); in a prospective cohort study, intake of rice during pregnancy was related to arsenic concentration in infant toenails collected within a few months of birth (Davis et al. 2014), and a longitudinal study found an increase of infants' urinary arsenic during transition to solid food, including rice cereals, fruits, and vegetables (Signes-Pastor et al. 2018). Evaluations of arsenic biomarker trends among these subpopulations may allow for identification of interventions aimed at specific subpopulations or foods. Further, recent efforts to disentangle dietary from drinking water exposures within biomarker measurements may enhance sensitivity of evaluation of dietary interventions (Jones et al. 2016).

Conclusions

Dietary arsenic exposures are an increasingly recognized public health concern, and this concern calls for interventions to reduce exposure, especially to vulnerable populations, such as pregnant women, infants, and children. Regulatory advances will be essential for effective intervention to lower dietary arsenic exposure, but opportunities also exist for various stakeholders, including commodity producers, the food industry, physicians, and the public, to take steps along the supply chain. For instance, we have proposed a proactive scheme designed to compel the production

of low-arsenic foods (Nachman et al. 2017). Such approaches require better monitoring efforts, such as those implemented in the Italian Total Diet Study (Cubadda et al. 2016), that employ the latest measurement techniques and push the field of practice forward. Moving further upstream, public health benefits may be derived from policies aimed at creating producer recommendations or setting production standards for soils and irrigation water.

Building upon GAO's (2018) recommendations of interagency coordination on arsenic in food, future regulatory policy interventions must consider the relative contributions of different dietary and other pathways to aggregate arsenic exposures. Further, enhanced efforts to monetize the avoidance of non-cancer health outcomes (e.g., cardiovascular, developmental, and many others) associated with arsenic exposure would be of great value. The effect of reducing the occurrence of these outcomes was not quantitatively considered in the U.S. EPA's economic justification of the arsenic in drinking water rule (U.S. EPA 2000). An economic understanding of the public health benefits of preventing these outcomes would likely justify more stringent measures to reduce exposure.

An urgent component of moving regulatory policy forward is updating the iAs toxicological assessment. The EPA IRIS's endeavor to update its iAs risk assessment is currently underway but has yet to be finalized. GAO called for U.S. FDA to update its rice risk assessment (U.S. GAO 2018), however, such efforts may be redundant with, and less comprehensive than that of the EPA IRIS program. In its 2018 review of updates to the IRIS program, the National Research Council (NRC), expressed satisfaction with the program's adoption of systematic review methods in pursuit of hazard characterization and dose-response assessment (NRC 2018). With much effort already invested in the development of the iAs assessment, and ongoing coordination with the NRC, it is anticipated that the toxicity values produced by this effort will be key in informing future FDA rice (and other food) risk assessment endeavors. With these toxicity values, interventions will be able to set sights on risk-based exposure limits, rather than limiting efforts to general exposure reduction in the absence of any target based on actual dose-response relationships.

Certain gaps in understanding of the science exist that, if filled, would bolster policy efforts aimed at exposure reduction. Specifically, primary research, as well as systematic reviews and meta-analyses of "other priority outcomes" (e.g., diabetes, neurodevelopmental toxicity and immune effects) and "other endpoints to consider" (e.g., renal disease, liver and pancreatic cancer, and hypertension), as specified by the NRC (2014), are needed. In addition, research to better characterize life stage-specific windows of vulnerability to arsenic exposure, as well as their interactions with genetic factors, epigenetic alterations, and other contaminant exposures (Cardenas et al. 2015), as well as the timing of gene-expression changes (Wright and Christiani 2010) is warranted. Further, broadening consideration of arsenic to include oral exposure to organic arsenic species commonly found in foods (including methylated forms, arenosugars, and arsenolipids) would be of value, especially for understanding dietary exposures. Focusing interventions narrowly on inorganic species may not account for the true health burdens resulting from exposure. Although uncertainty remains over population-wide health effects of dietary arsenic, focus on susceptible populations is critical. The mitigating role of nutrients involved in arsenic metabolism, such as folate and other B vitamins (Kurzus-Spencer et al. 2017; Spratlen et al. 2017) strongly suggests that susceptible populations will also include those with gastrointestinal disorders that inhibit the intestinal absorption of these vitamins, such as individuals with Celiac

disease, a population with increased dietary arsenic exposure from the prevalence of a rice-based, gluten-free diet (Bulka et al. 2017; Punshon and Jackson 2018).

Dietary arsenic exposure is an important public health challenge that contributes to population risk of a broad number of adverse health impacts. Current regulatory approaches for arsenic are limited, and the statutory basis for potential controls is fragmented across multiple agencies. As a result, the patchwork of state and federal efforts has had limited success in reducing dietary exposures and often results in mixed messages to consumers about risks. It is time to rethink our current approaches and develop a systems approach to dietary arsenic exposures — from source to ingestion. A renewed, inclusive approach to problem formulation, including the social and behavioral sciences (The National Academies 2018), can provide a renewed perspective on both efforts to characterize the scope of the problem and the most productive opportunities to intervene. Our greatest success in reducing dietary exposures can be achieved only by a multisector, multiagency, science-informed, public health systems approach coordinated across regulators and the industry (Burke et al. 2017). When those are not enough to ensure minimal arsenic contributions through diet, a well-informed public empowered to make the right food choices is essential.

Acknowledgments

This paper, a product of the Collaborative on Food with Arsenic and associated Risk and Regulation (C-FARR), is supported by the Dartmouth College Toxic Metals Superfund Research Program through funds from the National Institute of Environmental Health Sciences (NIEHS) of the National Institutes of Health (NIH) under Award Number 1R13ES026493-01 to C. C. and Award Number P42ES007373 to Bruce Stanton, and the Children's Environmental Health and Disease Prevention Research Center at Dartmouth through funds from the NIEHS of the NIH under Award Number P01ES022832 and from the U.S. Environmental Protection Agency RD-83544201 to M. K. The views expressed in this paper are the those of the authors and do not necessarily reflect the official views of any agency of the United States or other government. We are grateful to A. Seyfert for helpful discussions regarding silica and rice production.

References

- Bulka CM, Davis MA, Karagas MR, Ahsan H, Argos M. 2017. The unintended consequences of a gluten-free diet. *Epidemiology* 28(3):e24–e25, PMID: 28166100, <https://doi.org/10.1097/EDE.0000000000000640>.
- Burke TA, Cascio WE, Costa DL, Deener K, Fontaine TD, Fulk FA, et al. 2017. Rethinking environmental protection: meeting the challenges of a changing world. *Environ Health Perspect* 125(3):A43–A49, PMID: 28248180, <https://doi.org/10.1289/EHP1465>.
- Carden AS, Koestler DC, Houseman EA, Jackson BP, Kile ML, Karagas MR, et al. 2015. Differential DNA methylation in umbilical cord blood of infants exposed to mercury and arsenic in utero. *Epigenetics* 10(6):508–515, PMID: 25923418, <https://doi.org/10.1080/15592294.2015.1046026>.
- Carey M, Jiujin X, Farias JG, Meharg AA. 2015. Rethinking rice preparation for highly efficient removal of inorganic arsenic using percolating cooking water. *PloS One* 10(7):e0131608, PMID: 26200355, <https://doi.org/10.1371/journal.pone.0131608>.
- Carey M, Donaldson E, Signes-Pastor AJ, Meharg AA. 2018. Dilution of rice with other gluten free grains to lower inorganic arsenic in foods for young children in response to European Union regulations provides impetus to setting stricter standards. *PloS One* 13(3):e0194700, PMID: 29547635, <https://doi.org/10.1371/journal.pone.0194700>.
- Celiac Disease Foundation. 2017. "Arsenic in the Gluten-Free Diet: Are You at Risk?" <https://celiac.org/blog/2017/05/arsenic-gluten-free-diet-risk/> [accessed 8 April 2018].
- Children's Environmental Health and Disease Prevention Research Center at Dartmouth. 2016. "Arsenic." <http://www.dartmouth.edu/~childrenshealth/arsenic/> [accessed 23 March 2018].
- Codex Alimentarius. 2017. "Code of Practice for the Prevention and Reduction of Arsenic Contamination in Rice." http://www.fao.org/fao-who-codexalimentarius/sh-proxy/en/?lnk=1&url=https%253A%252F%252Fworksfa.ce.fao.org%252Fsites%252Fcodex%252Fstandards%252FCAC%2BRCP%2B77-2017%252FCXC_077e.pdf [accessed 18 April 2018].
- Consumer Reports. 2015. "How Much Arsenic Is in Your Rice?" <https://www.consumerelectronics.org/cro/magazine/2015/01/how-much-arsenic-is-in-your-rice/index.htm> [accessed 23 March 2018].
- Cubadda F, D'Amato M, Aureli F, Raggi A, Mantovani A. 2016. Dietary exposure of the Italian population to inorganic arsenic: The 2012–2014 Total Diet Study. *Food Chem Toxicol* 98(Pt B):148–158, PMID: 27756704, <https://doi.org/10.1016/j.fct.2016.10.015>.
- Cubadda F, Jackson BP, Cottingham KL, Van Horne YO, Kurzius-Spencer M. 2017. Human exposure to dietary inorganic arsenic and other arsenic species: state of knowledge, gaps and uncertainties. *Sci Total Environ* 579:1228–1239, PMID: 27914647, <https://doi.org/10.1016/j.scitotenv.2016.11.108>.
- Dartmouth College Superfund Research Program. 2017. "Arsenic and You." <http://www.arsenicandyou.org> [accessed 23 March 2018].
- Davis MA, Li Z, Gilbert-Diamond D, Mackenzie TA, Cottingham KL, Jackson BP, et al. 2014. Infant toenails as a biomarker of in utero arsenic exposure. *J Expo Sci Environ Epidemiol* 24(5):467, PMID: 24896769, <https://doi.org/10.1039/jes.2014.38>.
- Davis MA, Signes-Pastor AJ, Argos M, Slaughter F, Pendergrast C, Punshon T, et al. 2017. Assessment of human dietary exposure to arsenic through rice. *Sci Total Environ* 586:1237–1244, PMID: 28233618, <https://doi.org/10.1016/j.scitotenv.2017.02.119>.
- Díaz O, Tapia Y, Muñoz O, Montoro R, Velez D, Almela C. 2012. Total and inorganic arsenic concentrations in different species of economically important algae harvested from coastal zones of Chile. *Food Chem Toxicol* 50(3–4):744–749, PMID: 2238359, <https://doi.org/10.1016/j.jct.2011.11.024>.
- Food and Agriculture Organization. 2001. "Carrageenan." <http://www.fao.org/ag/agn-jecfa-additives/specs/Monograph/Additive-117.pdf> [accessed 23 March 2018].
- Gossai A, Zens MS, Punshon T, Jackson BP, Perry AE, Karagas MR. 2017. Rice consumption and squamous cell carcinoma of the skin in a United States population. *Environ Health Perspect* 125(9):097005, PMID: 28934722, <https://doi.org/10.1289/EHP1065>.
- Hall M, Gamble M, Slavkovich V, Liu X, Levy D, Cheng Z, et al. 2007. Determinants of arsenic metabolism: blood arsenic metabolites, plasma folate, cobalamin, and homocysteine concentrations in maternal-newborn pairs. *Environ Health Perspect* 115(10):1503–1509, PMID: 17938743, <https://doi.org/10.1289/ehp.9906>.
- Jackson BP, Taylor VF, Karagas MR, Punshon T, Cottingham KL. 2012. Arsenic, organic foods, and brown rice syrup. *Environ Health Perspect* 120(5):623–626, PMID: 2236149, <https://doi.org/10.1289/ehp.1104619>.
- Jones MR, Tellez-Plaza M, Vaidya D, Grau M, Francesconi KA, Goessler W, et al. 2016. Estimation of inorganic arsenic exposure in populations with frequent seafood intake: evidence from MESA and NHANES. *Am J Epidemiol* 184(8):590–602, PMID: 2702745, <https://doi.org/10.1093/aje/kwv097>.
- Krisberg K. 2017. "Regulatory changes ignore science, threaten public health: new administration eroding progress." <http://thenationsthalth.apublications.org/content/47/8/1.3> [accessed 16 March 2018].
- Kurzius-Spencer M, Burgess JL, Harris RB, Hartz V, Robarge J, Huang S, et al. 2014. Contribution of diet to aggregate arsenic exposures—an analysis across populations. *J Expo Sci Environ Epidemiol* 24(2):156–162, PMID: 23860400, <https://doi.org/10.1038/jes.2013.37>.
- Kurzius-Spencer M, da Silva V, Thomson CA, Hartz V, Hsu C-H, Burgess JL, et al. 2017. Nutrients in one-carbon metabolism and urinary arsenic methylation in the National Health and Nutrition Examination Survey (NHANES) 2003–2004. *Sci Total Environ* 607–608:381–390, PMID: 28697391, <https://doi.org/10.1016/j.scitotenv.2017.07.019>.
- Lai PY, Cottingham KL, Steinmaus C, Karagas MR, Miller MD. 2015. Arsenic and rice: translating research to address health care providers' needs. *J Pediatr* 167(4):797–803, PMID: 26253210, <https://doi.org/10.1016/j.jpeds.2015.07.003>.
- Li RY, Stroud JL, Ma JF, McGrath SF, Zhao FJ. 2009. Mitigation of arsenic accumulation in rice with water management and silicon fertilization. *Environ Sci Technol* 43(10):3778–3783, PMID: 19544887.
- Limmer MA, Wise P, Dykes GE, Seyfert AL. 2018. Silicon decreases dimethylarsinic acid concentration in rice grain and mitigates straighthead disorder. *Environ Sci Technol* 52(8):4809–4816, PMID: 29608840, <https://doi.org/10.1021/acs.est.8b00300>.
- Lundberg Family Farms. 2016. "A Letter from the CEO." <http://www.lundberg.com/info/arsenic-in-food/> [accessed 23 March 2018].
- Meharg AA, Lombi E, Williams PN, Scheckel KG, Feldmann J, Raab A, et al. 2008. Speciation and localization of arsenic in white and brown rice grains. *Environ Sci Technol* 42(4):1051–1057, PMID: 18351071.
- Melkonian S, Argos M, Hall MN, Chen Y, Parvez F, Pierce B, et al. 2013. Urinary and dietary analysis of 18,470 Bangladeshis reveal a correlation of rice consumption with arsenic exposure and toxicity. *PLoS One* 8(11):e80691, PMID: 24260455, <https://doi.org/10.1371/journal.pone.0080691>.

- Muraki I, Wu H, Imamura F, Laden F, Rimm EB, Hu FB, et al. 2015. Rice consumption and risk of cardiovascular disease: results from a pooled analysis of 3 U.S. cohorts. *Am J Clin Nutr* 101(1):164–172, PMID: 25527760, <https://doi.org/10.3945/ajcn.114.087551>.
- Mushak P, McKinzie M, Olson E, Metrick A. 2000. Arsenic and old laws: a scientific and public health analysis of arsenic occurrence in drinking water, its health effects, and EPA's outdated arsenic tap water standard. In: *Arsenic and Old Laws: A Scientific and Public Health*. New York, NY:NRDC Publications Department.
- Nachman KE, Ginsberg GL, Miller MD, Murray CJ, Nigra AE, Pendergrast CB. 2017. Mitigating dietary arsenic exposure: current status in the United States and recommendations for an improved path forward. *Sci Total Environ* 581:582:221–236, PMID: 28065543, <https://doi.org/10.1016/j.scitotenv.2016.12.112>.
- Nature's One. 2018. "Nature Knows What's Best for Your Baby." <https://www.natureseone.com/our-story> [accessed 23 March 2018].
- Naujokas MF, Anderson B, Ahsan H, Aposhian HV, Graziano JH, Thompson C, et al. 2013. The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environ Health Perspect* 121(3):295–302, PMID: 23458756, <https://doi.org/10.1289/ehp.1205875>.
- Nigra AE, Sanchez TR, Nachman KE, Harvey DE, Chilrud SN, Graziano JH, et al. 2017. The effect of the Environmental Protection Agency maximum contaminant level on arsenic exposure in the USA from 2003 to 2014: an analysis of the National Health and Nutrition Examination Survey (NHANES). *Lancet Public Health* 2(1):e513–e521, PMID: 29250608, [https://doi.org/10.1016/S2468-2667\(17\)30195-0](https://doi.org/10.1016/S2468-2667(17)30195-0).
- Norton GJ, Islam MR, Deacon CM, Zhao F-J, Stroud JL, McGrath SP, et al. 2009. Identification of low inorganic and total grain arsenic rice cultivars from Bangladesh. *Environ Sci Technol* 43(15):6070–6075, PMID: 19731720.
- Norton GJ, Pinson SR, Alexander J, Mackay S, Hansen H, Duan GL, et al. 2012. Variation in grain arsenic assessed in a diverse panel of rice (*Oryza sativa*) grown in multiple sites. *New Phytol* 193(3):650–664, PMID: 2214224, <https://doi.org/10.1111/j.1469-8137.2011.03903.x>.
- NRC (National Research Council). 2014. "Critical Aspects of EPA's IRIS Assessment of Inorganic Arsenic: Interim Report." Washington, DC: National Academies Press. <http://www.nap.edu/catalog/18594/critical-aspects-of-epas-iris-assessment-of-inorganic-arsenic-interim> [accessed 3 August 2018].
- NRC. 2018. "Progress toward Transforming the Integrated Risk Information System (IRIS) Program: A 2018 Evaluation (2018)." Washington, DC: National Academies Press. <http://www.nap.edu/download/25086> [accessed 18 April 2018].
- Penido ES, Bennett AJ, Hanson TE, Seyfert AL. 2016. Biogeochemical impacts of silicon-rich rice residue incorporation into flooded soils: implications for rice nutrition and cycling of arsenic. *Plant Soil* 399(1–2):75–87, <https://doi.org/10.1007/s11104-015-2682-3>.
- Punshon T, Jackson BP, Meharg AA, Warczack T, Scheckel K, Guerinot ML. 2017. Understanding arsenic dynamics in agronomic systems to predict and prevent uptake by crop plants. *Sci Total Environ* 581–582:209–220, PMID: 28043702, <https://doi.org/10.1016/j.scitotenv.2016.12.111>.
- Punshon T, Jackson B. 2018. Essential micronutrient and toxic trace element concentrations in gluten containing and gluten-free foods. *Food Chem* 252:258–264, PMID: 29478539, <https://doi.org/10.1016/j.foodchem.2018.01.120>.
- Raab A, Baskaran C, Feldmann J, Meharg AA. 2009. Cooking rice in a high water to rice ratio reduces inorganic arsenic content. *J Environ Monit* 11(1):41–44, PMID: 19137137, <https://doi.org/10.1039/B816906C>.
- Samet JM, Burke TA, Goldstein BD. 2017. The Trump administration and the environment — heed the science. *N Engl J Med* 336(12):1182–1188, PMID: 28249122, <https://doi.org/10.1056/NEJMMS1615242>.
- Schoof R, Yost L, Eickhoff J, Creelius E, Cragni D, Meacher D, et al. 1999. A market basket survey of inorganic arsenic in food. *Food Chem Toxicol* 37(8):839–846, PMID: 10506007.
- Sengupta M, Hossain M, Mukherjee A, Ahmed S, Das B, Nayak B, et al. 2006. Arsenic burden of cooked rice: traditional and modern methods. *Food Chem Toxicol* 44(11):1823–1829, PMID: 16876928, <https://doi.org/10.1016/j.fct.2006.06.003>.
- Seyfert AL, Morris AH, Gill R, Kearns KA, Mann JN, Paukett M, et al. 2016. Soil incorporation of silica-rich rice husk decreases inorganic arsenic in rice grain. *J Agric Food Chem* 64(19):3760–3766, PMID: 27109244, <https://doi.org/10.1021/acs.jafc.6b01201>.
- Signes-Pastor AJ, Carey M, Meharg AA. 2017a. Inorganic arsenic removal in rice bran by percolating cooking water. *Food Chem* 234:76–80, PMID: 28551270, <https://doi.org/10.1016/j.foodchem.2017.04.140>.
- Signes-Pastor AJ, Woodside JV, McMullan P, Mullan K, Carey M, Karagas MR, et al. 2017b. Levels of infants' urinary arsenic metabolites related to formula feeding and weaning with rice products exceeding the EU inorganic arsenic standard. *PLoS One* 12(5):e0176923, PMID: 28472079, <https://doi.org/10.1371/journal.pone.0176923>.
- Signes-Pastor AJ, Cottingham KL, Carey M, Sayarath V, Palys T, Meharg AA, et al. 2018. Infants' dietary arsenic exposure during transition to solid food. *Sci Rep* 8(1):7114, PMID: 29739998, <https://doi.org/10.1038/s41598-018-25372-1>.
- Signes A, Mitra K, Burlió F, Carbonell-Barrachina A. 2008. Effect of two different rice dehusking procedures on total arsenic concentration in rice. *Eur Food Res Technol* 226(3):561–567, <https://doi.org/10.1007/s00217-007-0571-6>.
- Sprattes MJ, Gamble MV, Grau-Perez M, Kuo C-C, Best LG, Yracheta J, et al. 2017. Arsenic metabolism and one-carbon metabolism at low-moderate arsenic exposure: evidence from the Strong Heart Study. *Food Chem Toxicol* 105:387–397, PMID: 28479390, <https://doi.org/10.1016/j.fct.2017.06.004>.
- Taylor V, Goodale B, Raab A, Schwertle T, Reimer K, Conklin S, et al. 2017. Human exposure to organic arsenic species from seafood. *Sci Total Environ* 580:266–282, PMID: 28024743, <https://doi.org/10.1016/j.scitotenv.2016.12.113>.
- Teasley WA, Limmer MA, Seyfert AL. 2017. How rice (*Oryza sativa* L.) responds to elevated Si under different Si-rich soil amendments. *Environ Sci Technol* 51(18):10335–10343, PMID: 28795805, <https://doi.org/10.1021/acs.est.7b01740>.
- The National Academies. 2018. "The Environmental Health Matters Initiative." <http://nas-sites.org/environhealthmatters/background> [accessed 30 April 2018].
- U.S. Congress. 2012. H.R.3984 – Apple Juice Act of 2012. <https://www.congress.gov/bill/112th-congress/house-bill/3984> [accessed 16 March 2018].
- U.S. Congress. 2015. H.R.2529 – Reducing Food-Based Inorganic Compounds Exposure Act of 2015. <https://www.congress.gov/bill/114th-congress/house-bill/2529/text> [accessed 16 March 2018].
- U.S. Congress. 2017. H.R.4535 – Rice Act. <https://www.congress.gov/bill/115th-congress/house-bill/4535> [accessed 18 April 2018].
- USDA (U.S. Department of Agriculture). 2018. WWEIA/NHANES Overview. <https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/wweianhanes-overview/> [accessed 4 May 2018].
- U.S. EPA (U.S. Environmental Protection Agency). 2000. "Arsenic in Drinking Water Rule Economic Analysis (EPA 815-R-00-026)." <https://heisep.epa.gov/Exe/ZyPdf.cgi?Dockey=20001YQT> [accessed 9 July 2018].
- U.S. FDA (U.S. Food and Drug Administration). 2013. "Arsenic in Apple Juice." <https://www.fda.gov/Food/FoodborneIllnessContaminants/Metals/ucm280209.htm> [accessed 15 March 2018].
- U.S. FDA. 2016. "Arsenic in Rice and Rice Products." <https://www.fda.gov/Food/FoodborneIllnessContaminants/Metals/ucm319870.htm> [accessed 15 March 2018].
- U.S. FDA. 2017a. "Arsenic." <https://www.fda.gov/Food/FoodborneIllnessContaminants/Metals/ucm280202.htm> [accessed 23 March 2018].
- U.S. FDA. 2017b. "Combination Metals Testing." <https://www.fda.gov/Food/FoodborneIllnessContaminants/Metals/ucm521427.htm> [accessed 23 March 2018].
- U.S. GAO (Government Accountability Office). 2018. "Federal Efforts to Manage the Risk of Arsenic in Rice." <https://www.gao.gov/assets/700/690701.pdf> [accessed 18 April 2018].
- Welch B, Smit E, Cardenas A, Hystad P, Kile ML. 2018. Trends in urinary arsenic among the U.S. population by drinking water source: results from the National Health and Nutrition Examination Survey 2003–2014. *Environ Res* 162:8–17, PMID: 29272814, <https://doi.org/10.1016/j.envres.2017.12.012>.
- Wright RO, Christiani D. 2010. Gene-environment interaction and children's health and development. *Curr Opin Pediatr* 22(2):197, PMID: 20090521, <https://doi.org/10.1097/MOP.0b013e328336ebf9>.
- Xue J, Zartarian V, Wang S-W, Liu SV, Georgopoulos P. 2010. Probabilistic modeling of dietary arsenic exposure and dose and evaluation with 2003–2004 NHANES data. *Environ Health Perspect* 118(3):345, PMID: 20194069, <https://doi.org/10.1289/ehp.0901205>.
- Yang J, Zhou Q, Zhang J. 2017. Moderate wetting and drying increases rice yield and reduces water use, grain arsenic level, and methane emission. *Crop J* 5(2):151–158, <https://doi.org/10.1016/j.cj.2016.06.002>.
- Zhang R, Zhang X, Wu K, Wu H, Sun Q, Hu FB, et al. 2016. Rice consumption and cancer incidence in US men and women. *Int J Cancer* 138(3):555–564, PMID: 26219234, <https://doi.org/10.1002/ijc.29704>.

References

- Abbas MHH, Meharg AA** (2008) Arsenate, arsenite and dimethyl arsinic acid (DMA) uptake and tolerance in maize (*Zea mays L.*). In *Plant and Soil*, Vol 304, pp 277-289
- Abdul KSM, Jayasinghe SS, Chandana EPS, Jayasumana C, De Silva PMCS** (2015) Arsenic and human health effects: A review. *Environmental Toxicology and Pharmacology* **40**: 828-846
- Abedin MJ, Feldmann J, Meharg AA** (2002) Uptake kinetics of arsenic species in rice plants. *Plant Physiology* **128**: 1120-1128
- Abercrombie JM, Halfhill MD, Ranjan P, Rao MR, Saxton AM, Yuan JS, Stewart CN, Jr.** (2008) Transcriptional responses of *Arabidopsis thaliana* plants to As (V) stress. *BMC Plant Biol* **8**: 87
- Akai J, Izumi K, Fukuhara H, Masuda H, Nakano S, Yoshimura T, Ohfuji H, Md Anawar H, Akai K** (2004) Mineralogical and geomicrobiological investigations on groundwater arsenic enrichment in Bangladesh. *Applied Geochemistry* **19**: 215-230
- Alam MGM, Snow ET, Tanaka A** (2003) Arsenic and heavy metal contamination of vegetables grown in Samta village, Bangladesh. *Science of the Total Environment* **308**: 83-96
- Alava P, Tack F, Laing GD, de Wiele TV** (2012) HPLC-ICP-MS method development to monitor arsenic speciation changes by human gut microbiota. *Biomed Chromatogr* **26**: 524-533
- Ali W, Isner JC, Isayenkov SV, Liu W, Zhao FJ, Maathuis FJ** (2012) Heterologous expression of the yeast arsenite efflux system ACR3 improves *Arabidopsis thaliana* tolerance to arsenic stress. *New Phytol* **194**: 716-723

Andres J, Bertin PN (2016) The microbial genomics of arsenic. *Fems Microbiology Reviews* **40:** 299-322

Bakhat HF, Zia Z, Fahad S, Abbas S, Hammad HM, Shahzad AN, Abbas F, Alharby H, Shahid M (2017) Arsenic uptake, accumulation and toxicity in rice plants: Possible remedies for its detoxification: A review. *Environ Sci Pollut Res Int* **24:** 9142-9158

Bargmann BO, Birnbaum KD (2010) Fluorescence activated cell sorting of plant protoplasts. *J Vis Exp*

Bertoldi D, Villegas TR, Larcher R, Santato A, Nicolini G (2013) Arsenic present in the soil-vine-wine chain in vineyards situated in an old mining area in Trentino, Italy. *Environmental Toxicology and Chemistry* **32:** 773-779

Bezerra-Neto JP, de Araujo FC, Ferreira-Neto JRC, da Silva MD, Pandolfi V, Aburjaile FF, Sakamoto T, de Oliveira Silva RL, Kido EA, Barbosa Amorim LL, Ortega JM, Benko-Iseppon AM (2019) Plant Aquaporins: Diversity, Evolution and Biotechnological Applications. *Curr Protein Pept Sci* **20:** 368-395

Bienert GP, Thorsen M, Schussler MD, Nilsson HR, Wagner A, Tamas MJ, Jahn TP (2008) A subgroup of plant aquaporins facilitate the bi-directional diffusion of As(OH)₃ and Sb(OH)₃ across membranes. *BMC Biol* **6:** 26

Bowell RJ, Alpers CN, Jamieson HE, Nordstrom DK, Majzlan J (2014) The Environmental Geochemistry of Arsenic - An Overview -. Arsenic: Environmental Geochemistry, Mineralogy, and Microbiology **79:** 1-16

Box MS, Huang BE, Domijan M, Jaeger KE, Khattak AK, Yoo SJ, Sedivy EL, Jones DM, Hearn TJ, Webb AAR, Grant A, Locke JCW, Wigge PA (2015) ELF3 Controls Thermoresponsive Growth in Arabidopsis. *Current Biology* **25:** 194-199

Brachmann CB, Davies A, Cost GJ, Caputo E, Li JC, Hieter P, Boeke JD (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful

set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**: 115-132

Brautigam A, Schaumloffel D, Preud'homme H, Thondorf I, Wesenberg D (2011) Physiological characterization of cadmium-exposed Chlamydomonas reinhardtii. *Plant Cell and Environment* **34**: 2071-2082

Brulle F, Cocquerelle C, Wamalah AN, Morgan AJ, Kille P, Lepretre A, Vandenbuicke F (2008) cDNA cloning and expression analysis of Eisenia fetida (Annelida : Oligochaeta) phytochelatin synthase under cadmium exposure. *Ecotoxicology and Environmental Safety* **71**: 47-55

Brunetti P, Zanella L, De Paolis A, Di Litta D, Cecchetti V, Falasca G, Barbieri M, Altamura MM, Costantino P, Cardarelli M (2015) Cadmium-inducible expression of the ABC-type transporter AtABCC3 increases phytochelatin-mediated cadmium tolerance in Arabidopsis. *J Exp Bot* **66**: 3815-3829

Buchet JP, Lauwerys R, Roels H (1981) Comparison of the Urinary-Excretion of Arsenic Metabolites after a Single Oral Dose of Sodium Arsenite, Monomethylarsonate, or Dimethylarsinate in Man. *International Archives of Occupational and Environmental Health* **48**: 71-79

Cao Y, Sun D, Ai H, Mei HY, Liu X, Sun SB, Xu GH, Liu YG, Chen YS, Ma LNQ (2017) Knocking Out OsPT4 Gene Decreases Arsenate Uptake by Rice Plants and Inorganic Arsenic Accumulation in Rice Grains. *Environmental Science & Technology* **51**: 12131-12138

Carbonell-Barrachina AA, Aarabi MA, DeLaune RD, Gambrell RP, Patrick WH (1998) The influence of arsenic chemical form and concentration on Spartina patens and Spartina alterniflora growth and tissue arsenic concentration. *Plant and Soil* **198**: 33-43

Carey AM, Lombi E, Donner E, de Jonge MD, Punshon T, Jackson BP, Guerinot ML, Price AH, Meharg AA (2012) A review of recent developments in the

speciation and location of arsenic and selenium in rice grain. *Anal Bioanal Chem* **402:** 3275-3286

Carey AM, Norton GJ, Deacon C, Scheckel KG, Lombi E, Punshon T, Guerinot

ML, Lanzirotti A, Newville M, Choi YS, Price AH, Meharg AA (2011)

Phloem transport of arsenic species from flag leaf to grain during grain filling.

New Phytologist **192:** 87-98

Castrillo G, Sanchez-Bermejo E, de Lorenzo L, Crevillen P, Fraile-Escanciano A, Tc

M, Mouriz A, Catarecha P, Sobrino-Plata J, Olsson S, Leo Del Puerto Y,

Mateos I, Rojo E, Hernandez LE, Jarillo JA, Pineiro M, Paz-Ares J, Leyva A

(2013) WRKY6 transcription factor restricts arsenate uptake and transposon

activation in Arabidopsis. *Plant Cell* **25:** 2944-2957

Catarecha P, Segura MD, Franco-Zorrilla JM, Garcia-Ponce B, Lanza M, Solano R,

Paz-Ares J, Leyva A (2007) A mutant of the Arabidopsis phosphate transporter

PHT1;1 displays enhanced arsenic accumulation. *Plant Cell* **19:** 1123-1133

Cavanagh C, Morell M, Mackay I, Powell W (2008) From mutations to MAGIC:

resources for gene discovery, validation and delivery in crop plants. *Curr Opin*

Plant Biol **11:** 215-221

Ceglowski M, Schroeder G (2015) Removal of heavy metal ions with the use of

chelating polymers obtained by grafting pyridine-pyrazole ligands onto

polymethylhydrosiloxane. *Chemical Engineering Journal* **259:** 885-893

Chao DY, Chen Y, Chen J, Shi S, Chen Z, Wang C, Danku JM, Zhao FJ, Salt DE

(2014) Genome-wide association mapping identifies a new arsenate reductase

enzyme critical for limiting arsenic accumulation in plants. *PLoS Biol* **12:**

e1002009

Chao DY, Silva A, Baxter I, Huang YS, Nordborg M, Danku J, Lahner B,

Yakubova E, Salt DE (2012) Genome-Wide Association Studies Identify Heavy

Metal ATPase3 as the Primary Determinant of Natural Variation in Leaf Cadmium in *Arabidopsis thaliana*. Plos Genetics **8**

Chen Y, Xu W, Shen H, Yan H, Xu W, He Z, Ma M (2013) Engineering arsenic tolerance and hyperaccumulation in plants for phytoremediation by a PvACR3 transgenic approach. Environ Sci Technol **47**: 9355-9362

Chen YS, Hua CY, Chen JX, Rathinasabapathi B, Cao Y, Ma LQ (2019) Expressing Arsenite Antiporter PvACR3;1 in Rice (*Oryza sativa L.*) Decreases Inorganic Arsenic Content in Rice Grains. Environmental Science & Technology **53**: 10062-10069

Chen YS, Hua CY, Jia MR, Fu JW, Liu X, Han YH, Liu YG, Rathinasabapathi B, Cao Y, Ma LQ (2017) Heterologous Expression of *Pteris vittata* Arsenite Antiporter PvACR3;1 Reduces Arsenic Accumulation in Plant Shoots. Environmental Science & Technology **51**: 10387-10395

Clemens S, Kim EJ, Neumann D, Schroeder JI (1999) Tolerance to toxic metals by a gene family of phytochelatin synthases from plants and yeast. EMBO J **18**: 3325-3333

Clemens S, Ma JF (2016) Toxic Heavy Metal and Metalloid Accumulation in Crop Plants and Foods. Annual Review of Plant Biology, Vol 67 **67**: 489-512

Cobbett C, Goldsbrough P (2002) Phytochelatins and metallothioneins: Roles in heavy metal detoxification and homeostasis. Annual Review of Plant Biology **53**: 159-182

Craw D, Bowell RJ (2014) The Characterization of Arsenic in Mine Waste. Arsenic: Environmental Geochemistry, Mineralogy, and Microbiology **79**: 473-505

Cubadda F, Jackson BP, Cottingham KL, Van Horne YO, Kurzius-Spencer M (2017) Human exposure to dietary inorganic arsenic and other arsenic species: State of knowledge, gaps and uncertainties. Science of the Total Environment **579**: 1228-1239

Cui J, Jing C (2019) A review of arsenic interfacial geochemistry in groundwater and the role of organic matter. *Ecotoxicol Environ Saf* **183**: 109550

Das N, Bhattacharya S, Bhattacharyya S, Maiti MK (2018) Expression of rice MATE family transporter OsMATE2 modulates arsenic accumulation in tobacco and rice. *Plant Mol Biol* **98**: 101-120

De Chaudhuri S, Ghosh P, Sarma N, Majumdar P, Sau TJ, Basu S, Roychoudhury S, Ray K, Giri AK (2008) Genetic variants associated with arsenic susceptibility: Study of purine nucleoside phosphorylase, arsenic (+3) methyltransferase, and glutathione S-transferase omega genes. *Environmental Health Perspectives* **116**: 501-505

de la Fuente C, Clemente R, Alburquerque JA, Velez D, Bernal MP (2010) Implications of the Use of As-Rich Groundwater for Agricultural Purposes and the Effects of Soil Amendments on As Solubility. *Environmental Science & Technology* **44**: 9463-9469

Deng F, Yamaji N, Ma JF, Lee SK, Jeon JS, Martinoia E, Lee Y, Song WY (2018) Engineering rice with lower grain arsenic. *Plant Biotechnology Journal* **16**: 1691-1699

Denver DEH (2004) <Historic Arsenical Pesticide Research.pdf>. In DDoE Health, ed, **Dhankher OP, Li YJ, Rosen BP, Shi J, Salt D, Senecoff JF, Sashti NA, Meagher RB** (2002) Engineering tolerance and hyperaccumulation of arsenic in plants by combining arsenate reductase and gamma-glutamylcysteine synthetase expression. *Nature Biotechnology* **20**: 1140-1145

Dinneny JR, Long TA, Wang JY, Jung JW, Mace D, Pointer S, Barron C, Brady SM, Schiefelbein J, Benfey PN (2008) Cell identity mediates the response of *Arabidopsis* roots to abiotic stress. *Science* **320**: 942-945

Dittmar J, Voegelin A, Maurer F, Roberts LC, Hug SJ, Saha GC, Ali MA, Badruzzaman AB, Kretzschmar R (2010) Arsenic in soil and irrigation water

affects arsenic uptake by rice: complementary insights from field and pot studies. Environ Sci Technol **44**: 8842-8848

Drobna Z, Waters SB, Devesa V, Harmon AW, Thomas DJ, Styblo M (2005)

Metabolism and toxicity of arsenic in human urothelial cells expressing rat arsenic (+3 oxidation state)-methyltransferase. Toxicology and Applied Pharmacology **207**: 147-159

Duan G, Liu W, Chen X, Hu Y, Zhu Y (2013) Association of arsenic with nutrient elements in rice plants. Metallomics **5**: 784-792

Duan GL, Hu Y, Schneider S, McDermott J, Chen J, Sauer N, Rosen BP, Daus B, Liu Z, Zhu YG (2016) Inositol transporters AtINT2 and AtINT4 regulate arsenic accumulation in Arabidopsis seeds. Nat Plants **2**: 15202

Duan GL, Kamiya T, Ishikawa S, Arao T, Fujiwara T (2012) Expressing ScACR3 in Rice Enhanced Arsenite Efflux and Reduced Arsenic Accumulation in Rice Grains. Plant and Cell Physiology **53**: 154-163

Duan YH, Gan YQ, Wang YX, Liu CX, Yu K, Deng YM, Zhao K, Dong CJ (2017) Arsenic speciation in aquifer sediment under varying groundwater regime and redox conditions at Jianghan Plain of Central China. Science of the Total Environment **607**: 992-1000

Duncan EG, Maher WA, Foster SD, Krikowa F, O'Sullivan CA, Roper MM (2017) Dimethylarsenate (DMA) exposure influences germination rates, arsenic uptake and arsenic species formation in wheat. Chemosphere **181**: 44-54

Dunivin TK, Yeh SY, Shade A (2019) A global survey of arsenic-related genes in soil microbiomes. Bmc Biology **17**

Duxbury JM, Mayer AB, Lauren JG, Hassan N (2003) Food chain aspects of arsenic contamination in Bangladesh: effects on quality and productivity of rice. J Environ Sci Health A Tox Hazard Subst Environ Eng **38**: 61-69

Epstein E (1994) The Anomaly of Silicon in Plant Biology. Proceedings of the National Academy of Sciences of the United States of America **91:** 11-17

Fetter K, Van Wilder V, Moshelion M, Chaumont F (2004) Interactions between plasma membrane aquaporins modulate their water channel activity. Plant Cell **16:** 215-228

Finnegan PM, Chen WH (2012) Arsenic toxicity: the effects on plant metabolism. Frontiers in Physiology **3**

Foyer CH, Noctor G (2005) Redox homeostasis and antioxidant signaling: A metabolic interface between stress perception and physiological responses. Plant Cell **17:** 1866-1875

Frisbie WS (1936) Federal Control of Spray Residues on Fruits and Vegetables. Am J Public Health Nations Health **26:** 369-373

Fu SF, Chen PY, Nguyen QT, Huang LY, Zeng GR, Huang TL, Lin CY, Huang HJ (2014) Transcriptome profiling of genes and pathways associated with arsenic toxicity and tolerance in Arabidopsis. BMC Plant Biol **14:** 94

Gan XC, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, Kahles A, Bohnert R, Jean G, Derwent P, Kersey P, Belfield EJ, Harberd NP, Kemen E, Toomajian C, Kover PX, Clark RM, Ratsch G, Mott R (2011) Multiple reference genomes and transcriptomes for Arabidopsis thaliana. Nature **477:** 419-423

Garcia-Sanchez A, Alonso-Rojo P, Santos-Frances F (2010) Distribution and mobility of arsenic in soils of a mining area (Western Spain). Science of the Total Environment **408:** 4194-4201

Ghosh M, Shen J, Rosen BP (1999) Pathways of As(III) detoxification in *Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences of the United States of America **96:** 5001-5006

Gilbert-Diamond D, Cottingham KL, Gruber JF, Punshon T, Sayarath V, Gandolfi AJ, Baker ER, Jackson BP, Folt CL, Karagas MR (2011) Rice consumption contributes to arsenic exposure in US women. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 20656-20660

Gnan S, Priest A, Kover PX (2014) The Genetic Basis of Natural Variation in Seed Size and Seed Number and Their Trade-Off Using *Arabidopsis thaliana* MAGIC Lines. *Genetics* **198:** 1751-+

Gonzalez E, Solano R, Rubio V, Leyva A, Paz-Ares J (2005) PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR1 is a plant-specific SEC12-related protein that enables the endoplasmic reticulum exit of a high-affinity phosphate transporter in *Arabidopsis*. *Plant Cell* **17:** 3500-3512

Gresser MJ (1981) ADP-arsenate. Formation by submitochondrial particles under phosphorylating conditions. *J Biol Chem* **256:** 5981-5983

Guo JB, Dai XJ, Xu WZ, Ma M (2008) Overexpressing GSH1 and AsPCS1 simultaneously increases the tolerance and accumulation of cadmium and arsenic in *Arabidopsis thaliana*. *Chemosphere* **72:** 1020-1026

Ha SB, Smith AP, Howden R, Dietrich WM, Bugg S, O'Connell MJ, Goldsbrough PB, Cobbett CS (1999) Phytochelatin synthase genes from *arabidopsis* and the yeast *Schizosaccharomyces pombe*. *Plant Cell* **11:** 1153-1163

Hammami A, Gonzalez F, Ballester A, Blazquez ML, Munoz JA (2003) Simultaneous uptake of metals by activated sludge. *Minerals Engineering* **16:** 723-729

Hartley-Whitaker J, Ainsworth G, Meharg AA (2001) Copper- and arsenate-induced oxidative stress in *Holcus lanatus* L. clones with differential sensitivity. *Plant Cell and Environment* **24:** 713-722

He ZY, Yan HL, Chen YS, Shen HL, Xu WX, Zhang HY, Shi L, Zhu YG, Ma M (2016) An aquaporin PvTIP4;1 from *Pteris vittata* may mediate arsenite uptake. *New Phytologist* **209:** 746-761

Heckwolf M, Pater D, Hanson DT, Kaldenhoff R (2011) The *Arabidopsis thaliana* aquaporin AtPIP1;2 is a physiologically relevant CO₂ transport facilitator. *Plant Journal* **67:** 795-804

Heikens A, Panaullah GM, Meharg AA (2007) Arsenic behaviour from groundwater and soil to crops: impacts on agriculture and food safety. *Rev Environ Contam Toxicol* **189:** 43-87

Huang CF, Yamaji N, Chen Z, Ma JF (2012) A tonoplast-localized half-size ABC transporter is required for internal detoxification of aluminum in rice. *Plant J* **69:** 857-867

Huang LL, Pyc M, Alseekh S, McCarty DR, de Crecy-Lagard V, Gregory JF, Henry CS, Fernie AR, Mullen RT, Hanson AD (2018) A plastidial pantoate transporter with a potential role in pantothenate synthesis. *Biochemical Journal* **475:** 813-825

Huang TL, Nguyen QT, Fu SF, Lin CY, Chen YC, Huang HJ (2012) Transcriptomic changes and signalling pathways induced by arsenic stress in rice roots. *Plant Mol Biol* **80:** 587-608

Hwang JU, Song WY, Hong D, Ko D, Yamaoka Y, Jang S, Yim S, Lee E, Khare D, Kim K, Palmgren M, Yoon HS, Martinoia E, Lee Y (2016) Plant ABC Transporters Enable Many Unique Aspects of a Terrestrial Plant's Lifestyle. *Molecular Plant* **9:** 338-355

Indriolo E, Na G, Ellis D, Salt DE, Banks JA (2010) A Vacuolar Arsenite Transporter Necessary for Arsenic Tolerance in the Arsenic Hyperaccumulating Fern *Pteris vittata* Is Missing in Flowering Plants. *Plant Cell* **22:** 2045-2057

Isayenkov SV, Maathuis FJ (2008) The *Arabidopsis thaliana* aquaglyceroporin AtNIP7;1 is a pathway for arsenite uptake. *FEBS Lett* **582:** 1625-1628

Islam FS, Gault AG, Boothman C, Polya DA, Charnock JM, Chatterjee D, Lloyd JR (2004) Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature* **430:** 68-71

Ji R, Zhou L, Liu J, Wang Y, Yang L, Zheng Q, Zhang C, Zhang B, Ge H, Yang Y, Zhao F, Luan S, Lan W (2017) Calcium-dependent protein kinase CPK31 interacts with arsenic transporter AtNIP1;1 and regulates arsenite uptake in *Arabidopsis thaliana*. *PLoS One* **12**: e0173681

Jia Y, Huang H, Zhong M, Wang FH, Zhang LM, Zhu YG (2013) Microbial Arsenic Methylation in Soil and Rice Rhizosphere. *Environmental Science & Technology* **47**: 3141-3148

Jiang JS, Daniels BV, Fu D (2006) Crystal structure of AqpZ tetramer reveals two distinct Arg-189 conformations associated with water permeation through the narrowest constriction of the water-conducting channel. *Journal of Biological Chemistry* **281**: 454-460

Johanson U, Karlsson M, Johansson I, Gustavsson S, Sjovall S, Fraysse L, Weig AR, Kjellbom P (2001) The complete set of genes encoding major intrinsic proteins in *Arabidopsis* provides a framework for a new nomenclature for major intrinsic proteins in plants. *Plant Physiol* **126**: 1358-1369

Jones DT, Taylor WR, Thornton JM (1992) The Rapid Generation of Mutation Data Matrices from Protein Sequences. *Computer Applications in the Biosciences* **8**: 275-282

Kabir MS, Salam MA, Paul DN, Hossain MI, Rahman NM, Aziz A, Latif MA (2016) Spatial Variation of Arsenic in Soil, Irrigation Water, and Plant Parts: A Microllevel Study. *ScientificWorldJournal* **2016**: 2186069

Kamiya T, Fujiwara T (2009) *Arabidopsis NIP1;1 Transports Antimonite and Determines Antimonite Sensitivity*. *Plant and Cell Physiology* **50**: 1977-1981

Kamiya T, Islam MR, Duan GL, Uraguchi S, Fujiwara T (2013) Phosphate deficiency signaling pathway is a target of arsenate and phosphate transporter OsPT1 is involved in As accumulation in shoots of rice. *Soil Science and Plant Nutrition* **59**: 580-590

Kamiya T, Tanaka M, Mitani N, Ma JF, Maeshima M, Fujiwara T (2009) NIP1;1, an aquaporin homolog, determines the arsenite sensitivity of *Arabidopsis thaliana*. *J Biol Chem* **284**: 2114-2120

Kapilan R, Vaziri M, Zwiazek JJ (2018) Regulation of aquaporins in plants under stress. *Biol Res* **51**: 4

Khan KA, Stroud JL, Zhu YG, McGrath SP, Zhao FJ (2010) Arsenic bioavailability to rice is elevated in Bangladeshi paddy soils. *Environ Sci Technol* **44**: 8515-8521

Kim A, Chen J, Khare D, Jin JY, Yamaoka Y, Maeshima M, Zhao Y, Martinoia E, Hwang JU, Lee Y (2020) Non-intrinsic ATP-binding cassette proteins ABCI19, ABCI20 and ABCI21 modulate cytokinin response at the endoplasmic reticulum in *Arabidopsis thaliana*. *Plant Cell Rep* **39**: 473-487

Kinniburgh DG, Smedley PL (2001) Arsenic contamination of groundwater in Bangladesh. British Geological Survey, Keyworth

Koljonen T, Gustavsson N, Noras P, Tanskanen H (1989) Geochemical Atlas of Finland: preliminary aspects. *Journal of Geochemical Exploration* **32**: 231-242

Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R (2009) A Multiparent Advanced Generation Inter-Cross to Fine-Map Quantitative Traits in *Arabidopsis thaliana*. *Plos Genetics* **5**

Kumar B, Smita K, Flores LC (2017) Plant mediated detoxification of mercury and lead. *Arabian Journal of Chemistry* **10**: S2335-S2342

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* **35**: 1547-1549

Kumar V, Vogelsang L, Seidel T, Schmidt R, Weber M, Reichelt M, Meyer A, Clemens S, Sharma SS, Dietz KJ (2019) Interference between arsenic-induced toxicity and hypoxia. *Plant Cell Environ* **42**: 574-590

Kumari P, Rastogi A, Shukla A, Srivastava S, Yadav S (2018) Prospects of genetic engineering utilizing potential genes for regulating arsenic accumulation in plants. *Chemosphere* **211**: 397-406

Kurosawa K, Egashira K, Tani M, Jahiruddin M, Moslehuddin AZ, Rahman ZM (2008) Groundwater-soil-crop relationship with respect to arsenic contamination in farming villages of Bangladesh--a preliminary study. *Environ Pollut* **156**: 563-565

Lado LR, Hengl T, Reuter HI (2008) Heavy metals in European soils: A geostatistical analysis of the FOREGS Geochemical database. *Geoderma* **148**: 189-199

Lai VW, Sun Y, Ting E, Cullen WR, Reimer KJ (2004) Arsenic speciation in human urine: are we all the same? *Toxicol Appl Pharmacol* **198**: 297-306

LeBlanc MS, McKinney EC, Meagher RB, Smith AP (2013) Hijacking membrane transporters for arsenic phytoextraction. *Journal of Biotechnology* **163**: 1-9

Lee JS, Lee SW, Chon HT, Kim KW (2008) Evaluation of human exposure to arsenic due to rice ingestion in the vicinity of abandoned Myungbong Au-Ag mine site, Korea. *Journal of Geochemical Exploration* **96**: 231-235

Lee JY, Colinas J, Wang JY, Mace D, Ohler U, Benfey PN (2006) Transcriptional and posttranscriptional regulation of transcription factor expression in Arabidopsis roots. *Proc Natl Acad Sci U S A* **103**: 6055-6060

Lefevre F, Boutry M (2018) Towards Identification of the Substrates of ATP-Binding Cassette Transporters. *Plant Physiology* **178**: 18-39

Lei GJ, Yokosho K, Yamaji N, Fujii-Kashino M, Ma JF (2017) Functional characterization of two half-size ABC transporter genes in aluminium-accumulating buckwheat. *New Phytol* **215**: 1080-1089

Leslie EM, Haimeur A, Waalkes MP (2004) Arsenic transport by the human multidrug resistance protein 1 (MRP1/ABCC1) - Evidence that a tri-glutathione conjugate is required. *Journal of Biological Chemistry* **279:** 32700-32708

Levesque MP, Vernoux T, Busch W, Cui H, Wang JY, Blilou I, Hassan H, Nakajima K, Matsumoto N, Lohmann JU, Scheres B, Benfey PN (2006) Whole-genome analysis of the SHORT-ROOT developmental pathway in Arabidopsis. *PLoS Biol* **4:** e143

Li L, He Z, Pandey GK, Tsuchiya T, Luan S (2002) Functional cloning and characterization of a plant efflux carrier for multidrug and heavy metal detoxification. *J Biol Chem* **277:** 5360-5368

Li RY, Ago Y, Liu WJ, Mitani N, Feldmann J, McGrath SP, Ma JF, Zhao FJ (2009) The Rice Aquaporin Lsi1 Mediates Uptake of Methylated Arsenic Species. *Plant Physiology* **150:** 2071-2080

Li YJ, Dhankher OP, Carreira L, Balish RS, Meagher RB (2005) Arsenic and mercury tolerance and cadmium sensitivity in Arabidopsis plants expressing bacterial gamma-glutamylcysteine synthetase. *Environmental Toxicology and Chemistry* **24:** 1376-1386

Lindsay ER, Maathuis FJ (2016) Arabidopsis thaliana NIP7;1 is involved in tissue arsenic distribution and tolerance in response to arsenate. *FEBS Lett* **590:** 779-786

Liu WJ, Wood BA, Raab A, McGrath SP, Zhao FJ, Feldmann J (2010) Complexation of arsenite with phytochelatins reduces arsenite efflux and translocation from roots to shoots in Arabidopsis. *Plant Physiol* **152:** 2211-2221

Liu XL, Zhang SZ, Shan XQ, Zhu YG (2005) Toxicity of arsenate and arsenite on germination, seedling growth and amylolytic activity of wheat. *Chemosphere* **61:** 293-301

- Lomax C, Liu WJ, Wu LY, Xue K, Xiong JB, Zhou JZ, McGrath SP, Meharg AA, Miller AJ, Zhao FJ** (2012) Methylated arsenic species in plants originate from soil microorganisms. *New Phytologist* **193**: 665-672
- Love MI, Huber W, Anders S** (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550
- Luan M, Liu J, Liu Y, Han X, Sun G, Lan W, Luan S** (2018) Vacuolar Phosphate Transporter 1 (VPT1) Affects Arsenate Tolerance by Regulating Phosphate Homeostasis in Arabidopsis. *Plant Cell Physiol* **59**: 1345-1352
- Ma JF** (2004) Role of silicon in enhancing the resistance of plants to biotic and abiotic stresses. *Soil Science and Plant Nutrition* **50**: 11-18
- Ma JF** (2010) Silicon Transporters in Higher Plants. Mips and Their Role in the Exchange of Metalloids **679**: 99-110
- Ma JF, Tamai K, Yamaji N, Mitani N, Konishi S, Katsuhara M, Ishiguro M, Murata Y, Yano M** (2006) A silicon transporter in rice. *Nature* **440**: 688-691
- Ma JF, Yamaji N** (2006) Silicon uptake and accumulation in higher plants. *Trends in Plant Science* **11**: 392-397
- Ma JF, Yamaji N, Mitani N, Tamai K, Konishi S, Fujiwara T, Katsuhara M, Yano M** (2007) An efflux transporter of silicon in rice. *Nature* **448**: 209-212
- Ma JF, Yamaji N, Mitani N, Xu XY, Su YH, McGrath SP, Zhao FJ** (2008) Transporters of arsenite in rice and their role in arsenic accumulation in rice grain. *Proc Natl Acad Sci U S A* **105**: 9931-9935
- Ma L, Wang L, Jia YY, Yang ZG** (2016) Arsenic speciation in locally grown rice grains from Hunan Province, China: Spatial distribution and potential health risk. *Science of the Total Environment* **557**: 438-444

Ma R, Shen JL, Wu JS, Tang Z, Shen QR, Zhao FJ (2014) Impact of agronomic practices on arsenic accumulation and speciation in rice grain. Environmental Pollution **194:** 217-223

Majumder S, Banik P (2019) Geographical variation of arsenic distribution in paddy soil, rice and rice-based products: A meta-analytic approach and implications to human health. J Environ Manage **233:** 184-199

Marafante E, Vahter M, Norin H, Envall J, Sandstrom M, Christakopoulos A, Ryhage R (1987) Biotransformation of Dimethylarsinic Acid in Mouse, Hamster and Man. Journal of Applied Toxicology **7:** 111-117

Marin AR, Masscheleyn PH, Patrick WH (1992) The Influence of Chemical Form and Concentration of Arsenic on Rice Growth and Tissue Arsenic Concentration. Plant and Soil **139:** 175-183

Meharg AA, Rahman M (2003) Arsenic contamination of Bangladesh paddy field soils: Implications for rice contribution to arsenic consumption. Environmental Science & Technology **37:** 229-234

Menard GN, Bryant FM, Kelly AA, Craddock CP, Lavagi I, Hassani-Pak K, Kurup S, Eastmond PJ (2018) Natural variation in acyl editing is a determinant of seed storage oil composition. Scientific Reports **8**

Monchgesang S, Strehmel N, Schmidt S, Westphal L, Taruttis F, Muller E, Herklotz S, Neumann S, Scheel D (2016) Natural variation of root exudates in *Arabidopsis thaliana*-linking metabolomic and genomic data. Scientific Reports **6**

Moore SA, Moennich DM, Gresser MJ (1983) Synthesis and hydrolysis of ADP-arsenate by beef heart submitochondrial particles. J Biol Chem **258:** 6266-6271

Morton J, Leese E (2011) Arsenic speciation in clinical samples: urine analysis using fast micro-liquid chromatography ICP-MS. Analytical and Bioanalytical Chemistry **399:** 1781-1788

Mosa KA, Kumar K, Chhikara S, Mcdermott J, Liu ZJ, Musante C, White JC, Dhankher OP (2012) Members of rice plasma membrane intrinsic proteins subfamily are involved in arsenite permeability and tolerance in plants. *Transgenic Research* **21**: 1265-1277

Moyano A, Garcia-Sanchez A, Mayorga P, Anawar HM, Alvarez-Ayuso E (2009) Impact of irrigation with arsenic-rich groundwater on soils and crops. *J Environ Monit* **11**: 498-502

Murphy EA, Aucott M (1998) An assessment of the amounts of arsenical pesticides used historically in a geographical area. *Science of the Total Environment* **218**: 89-101

Naseri E, Nde-Tchoupe AI, Mwakabona HT, Nanseu-Njiki CP, Noubatep C, Njau KN, Wydra KD (2017) Making Fe-0-Based Filters a Universal Solution for Safe Drinking Water Provision. *Sustainability* **9**

Nelson WC, Lykins MH, Mackey J, Newill VA, Finklea JF, Hammer DI (1973) Mortality among orchard workers exposed to lead arsenate spray: A cohort study. *Journal of Chronic Diseases* **26**: 105-118

Nickson RT, McArthur JM, Ravenscroft P, Burgess WG, Ahmed KM (2000) Mechanism of arsenic release to groundwater, Bangladesh and West Bengal. *Applied Geochemistry* **15**: 403-413

Nishida S, Duan GL, Ohkama-Ohtsu N, Uraguchi S, Fujiwara T (2016) Enhanced arsenic sensitivity with excess phytochelatin accumulation in shoots of a SULTR1;2 knockout mutant of *Arabidopsis thaliana* (L.) Heynh. *Soil Science and Plant Nutrition* **62**: 367-372

Noctor G (2006) Metabolic signalling in defence and stress: the central roles of soluble redox couples. *Plant Cell and Environment* **29**: 409-425

Norra S, Berner ZA, Agarwala P, Wagner F, Chandrasekharam D, Stuben D (2005) Impact of irrigation with As rich groundwater on soil and crops: A geochemical

case study in West Bengal Delta Plain, India. Applied Geochemistry **20**: 1890-1906

Norton GJ, Lou-Hing DE, Meharg AA, Price AH (2008) Rice-arsenate interactions in hydroponics: whole genome transcriptional analysis. J Exp Bot **59**: 2267-2276

Norton GJ, Nigar M, Williams PN, Dasgupta T, Meharg AA, Price AH (2008) Rice-arsenate interactions in hydroponics: a three-gene model for tolerance. J Exp Bot **59**: 2277-2284

Perez-Palacios P, Funes-Pinter I, Agostini E, Talano MA, Ibanez SG, Humphry M, Edwards K, Rodriguez-Llorente ID, Caviedes MA, Pajuelo E (2019) Targeting Acr3 from *Ensifer medicae* to the plasma membrane or to the tonoplast of tobacco hairy roots allows arsenic extrusion or improved accumulation. Effect of acr3 expression on the root transcriptome. Metallomics **11**: 1864-1886

Podgorski J, Berg M (2020) Global threat of arsenic in groundwater. Science **368**: 845-850

Pouschat P, Zagury GJ (2006) In vitro gastrointestinal bioavailability of arsenic in soils collected near CCA-treated utility poles. Environ Sci Technol **40**: 4317-4323

Punshon T, Carey AM, Ricachenevsky FK, Meharg AA (2018) Elemental distribution in developing rice grains and the effect of flag-leaf arsenate exposure. Environmental and Experimental Botany **149**: 51-58

Punshon T, Jackson BP, Meharg AA, Warczack T, Scheckel K, Guerinot ML (2017) Understanding arsenic dynamics in agronomic systems to predict and prevent uptake by crop plants. Sci Total Environ **581-582**: 209-220

Qiao JT, Li XM, Li FB (2018) Roles of different active metal-reducing bacteria in arsenic release from arsenic-contaminated paddy soil amended with biochar. J Hazard Mater **344**: 958-967

Raab A, Williams PN, Meharg A, Feldmann J (2007) Uptake and translocation of inorganic and methylated arsenic species by plants. *Environmental Chemistry* **4:** 197-203

Rahman MA, Hasegawa H, Rahman MM, Miah MAM, Tasmin A (2008) Straighthead disease of rice (*Oryza sativa L.*) induced by arsenic toxicity. *Environmental and Experimental Botany* **62:** 54-59

Renshaw CE, Bostick BC, Feng X, Wong CK, Winston ES, Karimi R, Folt CL, Chen CY (2006) Impact of land disturbance on the fate of arsenical pesticides. *J Environ Qual* **35:** 61-67

Rhee J, Horie T, Sasano S, Nakahara Y, Katsuhara M (2017) Identification of an H₂O₂ permeable PIP aquaporin in barley and a serine residue promoting H₂O₂ transport. *Physiologia Plantarum* **159:** 120-128

Rigouin C, Nylin E, Cogswell AA, Schaumloffel D, Dobritzsch D, Williams DL (2013) Towards an understanding of the function of the phytochelatin synthase of *Schistosoma mansoni*. *PLoS Negl Trop Dis* **7:** e2037

Rosas-Castor JM, Guzman-Mar JL, Hernandez-Ramirez A, Garza-Gonzalez MT, Hinojosa-Reyes L (2014) Arsenic accumulation in maize crop (*Zea mays*): A review. *Science of the Total Environment* **488:** 176-187

Ruiz-Huerta EA, de la Garza Varela A, Gomez-Bernal JM, Castillo F, Avalos-Borja M, SenGupta B, Martinez-Villegas N (2017) Arsenic contamination in irrigation water, agricultural soil and maize crop from an abandoned smelter site in Matehuala, Mexico. *J Hazard Mater* **339:** 330-339

Sadhukhan A, Kobayashi Y, Nakano Y, Iuchi S, Kobayashi M, Sahoo L, Koyama H (2017) Genome-wide Association Study Reveals that the Aquaporin NIP1;1 Contributes to Variation in Hydrogen Peroxide Sensitivity in *Arabidopsis thaliana*. *Mol Plant* **10:** 1082-1094

Sanchez-Bermejo E, Castrillo G, del Llano B, Navarro C, Zarco-Fernandez S, Martinez-Herrera DJ, Leo-del Puerto Y, Munoz R, Camara C, Paz-Ares J, Alonso-Blanco C, Leyva A (2014) Natural variation in arsenate tolerance identifies an arsenate reductase in *Arabidopsis thaliana*. *Nat Commun* **5**: 4617

Sasaki A, Yamaji N, Ma JF (2016) Transporters involved in mineral nutrient uptake in rice. *J Exp Bot* **67**: 3645-3653

Schmoger MEV, Oven M, Grill E (2000) Detoxification of arsenic by phytochelatins in plants. *Plant Physiology* **122**: 793-801

Sena G, Jung JW, Benfey PN (2004) A broad competence to respond to SHORT ROOT revealed by tissue-specific ectopic expression. *Development* **131**: 2817-2826

Seo HM, Jung Y, Song S, Kim Y, Kwon T, Kim DH, Jeung SJ, Yi YB, Yi G, Nam MH, Nam J (2008) Increased expression of OsPT1, a high-affinity phosphate transporter, enhances phosphate acquisition in rice. *Biotechnol Lett* **30**: 1833-1838

Shi S, Wang T, Chen Z, Tang Z, Wu Z, Salt DE, Chao DY, Zhao FJ (2016) OsHAC1;1 and OsHAC1;2 Function as Arsenate Reductases and Regulate Arsenic Accumulation. *Plant Physiol* **172**: 1708-1719

Shoji T (2014) ATP-binding cassette and multidrug and toxic compound extrusion transporters in plants: a common theme among diverse detoxification mechanisms. *Int Rev Cell Mol Biol* **309**: 303-346

Shukla T, Khare R, Kumar S, Lakhwani D, Sharma D, Asif MH, Trivedi PK (2018) Differential transcriptome modulation leads to variation in arsenic stress response in *Arabidopsis thaliana* accessions. *Journal of Hazardous Materials* **351**: 1-10

Singh N, Ma LQ, Srivastava M, Rathinasabapathi B (2006) Metabolic adaptations to arsenic-induced oxidative stress in *Pteris vittata* L and *Pteris ensiformis* L. *Plant Science* **170**: 274-282

Smedley PL, Kinniburgh DG (2002) A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry* **17**: 517-568

Smith AH, Lingas EO, Rahman M (2000) Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. *Bull World Health Organ* **78**: 1093-1103

Smith DB, Cannon WF, Woodruff LG, Solano F, Ellefsen KJ (2014) Geochemical and Mineralogical Maps for Soils of the Conterminous United States. U.S. Geological Survey: 23-24

Song WY, Park J, Mendoza-Cózatl DG, Suter-Grottemeyer M, Shim D, Hörtensteiner S, Geisler M, Weder B, Rea PA, Rentsch D, Schroeder JI, Lee Y, Martinoia E (2010) Arsenic tolerance in *Arabidopsis* is mediated by two ABCC-type phytochelatin transporters. *Proc. Natl. Acad. Sci. U.S.A.* **107**: 21187-21192

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

Su T, Xu Q, Zhang FC, Chen Y, Li LQ, Wu WH, Chen YF (2015) WRKY42 Modulates Phosphate Homeostasis through Regulating Phosphate Translocation and Acquisition in *Arabidopsis*. *Plant Physiology* **167**: 1579-U1717

Su YH, McGrath SP, Zhao FJ (2010) Rice is more efficient in arsenite uptake and translocation than wheat and barley. *Plant and Soil* **328**: 27-34

Sun S, Gu M, Cao Y, Huang X, Zhang X, Ai P, Zhao J, Fan X, Xu G (2012) A constitutive expressed phosphate transporter, OsPht1;1, modulates phosphate uptake and translocation in phosphate-replete rice. *Plant Physiol* **159**: 1571-1581

Sun SK, Chen Y, Che J, Konishi N, Tang Z, Miller AJ, Ma JF, Zhao FJ (2018) Decreasing arsenic accumulation in rice by overexpressing OsNIP1;1 and

OsNIP3;3 through disrupting arsenite radial transport in roots. *New Phytologist* **219**: 641-653

Tamai K, Ma JF (2003) Characterization of silicon uptake by rice roots. *New Phytologist* **158**: 431-436

Tang Z, Chen Y, Chen F, Ji YC, Zhao FJ (2017) OsPTR7 (OsNPF8.1), a Putative Peptide Transporter in Rice, is Involved in Dimethylarsenate Accumulation in Rice Grain. *Plant and Cell Physiology* **58**: 904-913

Tang Z, Chen Y, Miller AJ, Zhao FJ (2019) The C-type ATP-Binding Cassette Transporter OsABCC7 Is Involved in the Root-to-Shoot Translocation of Arsenic in Rice. *Plant Cell Physiol* **60**: 1525-1535

Tang Z, Wang Y, Gao A, Ji Y, Yang B, Wang P, Tang Z, Zhao FJ (2020) Dimethylarsinic acid is the causal agent inducing rice straighthead disease. *J Exp Bot*

Tang Z, Zhao F-J (2020) The roles of membrane transporters in arsenic uptake, translocation and detoxification in plants. *Critical Reviews in Environmental Science and Technology*: 1-36

Tawfik DS, Viola RE (2011) Arsenate Replacing Phosphate: Alternative Life Chemistries and Ion Promiscuity. *Biochemistry* **50**: 1128-1134

Thomas DJ, Li JX, Waters SB, Xing WB, Adair BM, Drobna Z, Devesa V, Styblo M (2007) Arsenic (+3 oxidation state) methyltransferase and the methylation of arsenicals. *Experimental Biology and Medicine* **232**: 3-13

Thomas DJ, Nava GM, Cai SY, Boyer JL, Hernandez-Zavala A, Gaskins HR (2010) Arsenic (+3 Oxidation State) Methyltransferase and the Methylation of Arsenicals in the Invertebrate Chordate *Ciona intestinalis*. *Toxicological Sciences* **113**: 70-76

Tian S, Wang XB, Li P, Wang H, Ji HT, Xie JY, Qiu QL, Shen D, Dong HS (2016)

Plant Aquaporin AtPIP1;4 Links Apoplastic H₂O₂ Induction to Disease

Immunity Pathways. *Plant Physiology* **171**: 1635-1650

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 Environ* **37**:

140-152

Tiwari M, Sharma D, Singh M, Tripathi RD, Trivedi PK (2014) Expression of

OsMATE1 and OsMATE2 alters development, stress responses and pathogen

susceptibility in Arabidopsis. *Sci Rep* **4**: 3964

Townsend T, Solo-Gabriele H, Tolaymat T, Stook K, Hosein N (2003) Chromium,

copper, and arsenic concentrations in soil underneath CCA-treated wood

structures. *Soil & Sediment Contamination* **12**: 779-798

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg

SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression

analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**: 562-578

Tripathi RD, Srivastava S, Mishra S, Singh N, Tuli R, Gupta DK, Maathuis FJ

(2007) Arsenic hazards: strategies for tolerance and remediation by plants. *Trends Biotechnol* **25**: 158-165

Tsuji N, Hirayanagi N, Okada M, Miyasaka H, Hirata K, Zenk MH, Miyamoto K

(2002) Enhancement of tolerance to heavy metals and oxidative stress in

Dunaliella tertiolecta by Zn-induced phytochelatin synthesis. *Biochemical and*

Biophysical Research Communications **293**: 653-659

Tsuji N, Nishikori S, Iwabe O, Shiraki K, Miyasaki H, Takagi M, Hirata K,

Miyamoto K (2004) Characterization of phytochelatin synthase-like protein

encoded by alr0975 from a prokaryote, Nostoc sp PCC 7120. Biochemical and Biophysical Research Communications **315**: 751-755

Ullrichbeberius CI, Sanz A, Novacky AJ (1989) Evaluation of Arsenate-Associated and Vanadate-Associated Changes of Electrical Membrane-Potential and Phosphate-Transport in Lemna-Gibba-G1. Journal of Experimental Botany **40**: 119-128

Van de Wiele T, Gallawa CM, Kubachka KM, Creed JT, Basta N, Dayton EA, Whitacre S, Du Laing G, Bradham K (2010) Arsenic metabolism by human gut microbiota upon in vitro digestion of contaminated soils. Environ Health Perspect **118**: 1004-1009

Vatamaniuk OK, Bucher EA, Ward JT, Rea PA (2001) A new pathway for heavy metal detoxification in animals - Phytochelatin synthase is required for cadmium tolerance in *Caenorhabditis elegans*. Journal of Biological Chemistry **276**: 20817-20820

Verma PK, Verma S, Pande V, Mallick S, Tripathi RD, Dhankher OP, Chakrabarty D (2016) Overexpression of Rice Glutaredoxin OsGrx_C7 and OsGrx_C2.1 Reduces Intracellular Arsenic Accumulation and Increases Tolerance in *Arabidopsis thaliana* (vol 7, 740, 2016). Frontiers in Plant Science **7**

Vernoux T, Wilson RC, Seeley KA, Reichheld JP, Muroy S, Brown S, Maughan SC, Cobbett CS, Van Montagu M, Inze D, May MJ, Sung ZR (2000) The ROOT MERISTEMLESS1/CADMIUM SENSITIVE2 gene defines a glutathione-dependent pathway involved in initiation and maintenance of cell division during postembryonic root development. Plant Cell **12**: 97-109

Wang H, Xu Q, Kong YH, Chen Y, Duan JY, Wu WH, Chen YF (2014) Arabidopsis WRKY45 transcription factor activates PHOSPHATE TRANSPORTER1;1 expression in response to phosphate starvation. Plant Physiol **164**: 2020-2029

Wang P, Zhang W, Mao C, Xu G, Zhao FJ (2016) The role of OsPT8 in arsenate uptake and varietal difference in arsenate tolerance in rice. *J Exp Bot* **67**: 6051-6059

Wang Y, Li R, Li D, Jia X, Zhou D, Li J, Lyi SM, Hou S, Huang Y, Kochian LV, Liu J (2017) NIP1;2 is a plasma membrane-localized transporter mediating aluminum uptake, translocation, and tolerance in *Arabidopsis*. *Proc Natl Acad Sci U S A* **114**: 5047-5052

Wedepohl KH (1995) The Composition of the Continental-Crust. *Geochimica Et Cosmochimica Acta* **59**: 1217-1232

Weig AR, Jakob C (2000) Functional identification of the glycerol permease activity of *Arabidopsis thaliana* NLM1 and NLM2 proteins by heterologous expression in *Saccharomyces cerevisiae*. *FEBS Lett* **481**: 293-298

Williams PN, Islam S, Islam R, Jahiruddin M, Adomako E, Soliaman ARM, Rahman GKMM, Lu Y, Deacon C, Zhu YG, Meharg AA (2009) Arsenic limits Trace Mineral Nutrition (Selenium, Zinc, and Nickel) in Bangladesh Rice Grain. *Environmental Science & Technology* **43**: 8430-8436

Williams PN, Lei M, Sun GX, Huang Q, Lu Y, Deacon C, Meharg AA, Zhu YG (2009) Occurrence and Partitioning of Cadmium, Arsenic and Lead in Mine Impacted Paddy Rice: Hunan, China. *Environmental Science & Technology* **43**: 637-642

Williams PN, Villada A, Deacon C, Raab A, Figuerola J, Green AJ, Feldmann J, Meharg AA (2007) Greatly enhanced arsenic shoot assimilation in rice leads to elevated grain levels compared to wheat and barley. *Environ Sci Technol* **41**: 6854-6859

World Health Organization (2018) <https://www.who.int/en/news-room/fact-sheets/detail/arsenic>. In Fact Sheets: Arsenic, Vol 2020

Wu ZC, Ren HY, McGrath SP, Wu P, Zhao FJ (2011) Investigating the Contribution of the Phosphate Transport Pathway to Arsenic Accumulation in Rice. *Plant Physiology* **157**: 498-508

Xu JM, Shi SL, Wang L, Tang Z, Lv TT, Zhu XL, Ding XM, Wang YF, Zhao FJ, Wu ZC (2017) OsHAC4 is critical for arsenate tolerance and regulates arsenic accumulation in rice. *New Phytologist* **215**: 1090-1101

Xu LY, Wu X, Wang SF, Yuan ZD, Xiao F, Ming Y, Jia YF (2016) Speciation change and redistribution of arsenic in soil under anaerobic microbial activities. *Journal of Hazardous Materials* **301**: 538-546

Xu W, Dai W, Yan H, Li S, Shen H, Chen Y, Xu H, Sun Y, He Z, Ma M (2015) Arabidopsis NIP3;1 Plays an Important Role in Arsenic Uptake and Root-to-Shoot Translocation under Arsenite Stress Conditions. *Mol Plant* **8**: 722-733

Xu XY, McGrath SP, Meharg AA, Zhao FJ (2008) Growing rice aerobically markedly decreases arsenic accumulation. *Environmental Science & Technology* **42**: 5574-5579

Xu XY, McGrath SP, Zhao FJ (2007) Rapid reduction of arsenate in the medium mediated by plant roots. *New Phytol* **176**: 590-599

Yaneff A, Sigaut L, Marquez M, Alleva K, Pietrasanta LI, Amodeo G (2014) Heteromerization of PIP aquaporins affects their intrinsic permeability. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 231-236

Yang JL, Yeh CT, Ramamurthy RK, Qi XS, Fernando RL, Dekkers JCM, Garrick DJ, Nettleton D, Schnable PS (2018) Empirical Comparisons of Different Statistical Models To Identify and Validate Kernel Row Number-Associated Variants from Structured Multi-parent Mapping Populations of Maize. *G3-Genes Genomes Genetics* **8**: 3567-3575

- Ye Y, Li P, Xu T, Zeng L, Cheng D, Yang M, Luo J, Lian X** (2017) OsPT4
Contributes to Arsenate Uptake and Transport in Rice. *Front Plant Sci* **8**: 2197
- Yin N, Zhang Z, Cai X, Du H, Sun G, Cui Y** (2015) In Vitro Method To Assess Soil Arsenic Metabolism by Human Gut Microbiota: Arsenic Speciation and Distribution. *Environ Sci Technol* **49**: 10675-10681
- Yokohira M, Arnold LL, Pennington KL, Suzuki S, Kakiuchi-Kiyota S, Herbin-Davis K, Thomas DJ, Cohen SM** (2010) Severe systemic toxicity and urinary bladder cytotoxicity and regenerative hyperplasia induced by arsenite in arsenic (+3 oxidation state) methyltransferase knockout mice. A preliminary report. *Toxicology and Applied Pharmacology* **246**: 1-7
- Yoon Y, Lee WM, An YJ** (2015) Phytotoxicity of arsenic compounds on crop plant seedlings. *Environmental Science and Pollution Research* **22**: 11047-11056
- Yu H, Wu B, Zhang XX, Liu S, Yu J, Cheng S, Ren HQ, Ye L** (2016) Arsenic Metabolism and Toxicity Influenced by Ferric Iron in Simulated Gastrointestinal Tract and the Roles of Gut Microbiota. *Environ Sci Technol* **50**: 7189-7197
- Yu LJ, Luo YF, Liao B, Xie LJ, Chen L, Xiao S, Li JT, Hu SN, Shu WS** (2012) Comparative transcriptome analysis of transporters, phytohormone and lipid metabolism pathways in response to arsenic stress in rice (*Oryza sativa*). *New Phytol* **195**: 97-112
- Zenk MH** (1996) Heavy metal detoxification in higher plants - A review. *Gene* **179**: 21-30
- Zhang J, Martinoia E, Lee Y** (2018) Vacuolar Transporters for Cadmium and Arsenic in Plants and their Applications in Phytoremediation and Crop Development. *Plant and Cell Physiology* **59**: 1317-1325
- Zhao FJ, Ago Y, Mitani N, Li RY, Su YH, Yamaji N, McGrath SP, Ma JF** (2010) The role of the rice aquaporin Lsi1 in arsenite efflux from roots. *New Phytol* **186**: 392-399

Zhao FJ, Ma JF, Meharg AA, McGrath SP (2009) Arsenic uptake and metabolism in plants. *New Phytol* **181**: 777-794

Zhao FJ, Stroud JL, Khan MA, McGrath SP (2012) Arsenic translocation in rice investigated using radioactive As-73 tracer. *Plant and Soil* **350**: 413-420

Zhou G, Pereira JF, Delhaize E, Zhou M, Magalhaes JV, Ryan PR (2014) Enhancing the aluminium tolerance of barley by expressing the citrate transporter genes SbMATE and FRD3. *J Exp Bot* **65**: 2381-2390

Zhu YG, Sun GX, Lei M, Teng M, Liu YX, Chen NC, Wang LH, Carey AM, Deacon C, Raab A, Meharg AA, Williams PN (2008) High percentage inorganic arsenic content of mining impacted and nonimpacted Chinese rice. *Environmental Science & Technology* **42**: 5008-5013

Zhu YG, Yoshinaga M, Zhao FJ, Rosen BP (2014) Earth Abides Arsenic Biotransformations. *Annu Rev Earth Planet Sci* **42**: 443-467