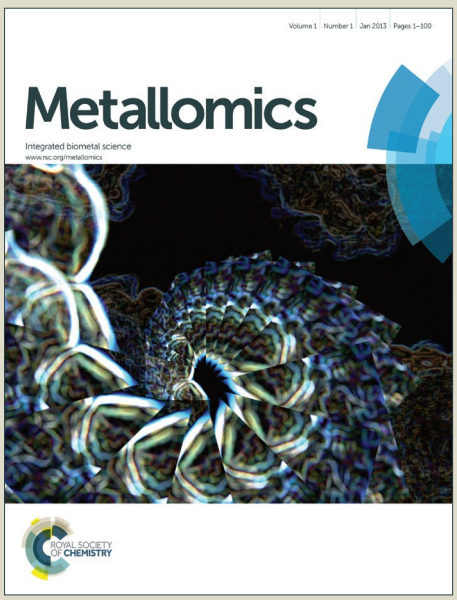
Metallomics

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/metallomics

A novel arsenic methyltransferase gene of *Westerdykella aurantiaca* isolated from arsenic contaminated soil: phylogenetic, physiological, biochemical studies and its role in arsenic bioremediation

Shikha Verma[1,3], Pankaj Kumar Verma[1,3], Alok Kumar Meher[2], Sanjay Dwivedi[4], Amit Kumar Bansiwal[2], Veena Pande[3], Pankaj Kumar Srivastava[4], Praveen Chandra Verma[1], Rudra Deo Tripathi[4], Debasis Chakrabarty[1,*]

[1]Genetics and Molecular Biology Division, CSIR-National Botanical Research Institute, India

[2]Environmental Material Division, CSIR-National Environmental Engineering Research Institute, India

[3]Department of Biotechnology, Kumaun University, India

[4]Plant Ecology and Environmental Science Division, CSIR-National Botanical Research Institute, India

[*]Corresponding author (chakrabartyd@nbri.res.in)

A table of contents entry: The study explores new insights in arsenic metabolism by *WaarsM* and provides a potential approach for bioremediation process.

Abstract

Elevated arsenic in the environment and agricultural soil is a serious concern to crop production and human health. Among the different detoxification mechanism, methylation of arsenic is a widespread phenomenon in nature. A number of microorganisms are able to methylate arsenic, but less is known about arsenic metabolism in fungi. We identified a novel arsenic methyltransferase (*WaarsM*) gene from a soil fungus, *Westerdykella aurantiaca*. *WaarsM* showed sequence homology with all known arsenic methyltransferases having three conserved SAM binding motifs. Expression of *WaarsM* enhanced arsenic resistance in *E. coli* (Δ*ars*) and *S. cerevisiae* (Δ*acr2*) strains by biomethylation and required endogenous reductants, preferably GSH, for methyltransferase activity. The purified WaarsM catalyzes the production of methylated arsenicals from both AsIII and AsV, also displays AsV reductase activity.. It displayed higher methyltransferase activity and lower K_M 0.1945±0.021 mM, K_M 0.4034±0.078 mM for AsIII and AsV, respectively. *S. cerevisiae* (Δ*acr2*) cells expressing *WaarsM* produced 2.2 ppm volatile arsenic and 0.64 ppm DMA(V) with 0.58 ppm volatile arsenicals when exposed to 20 ppm AsV and 2 ppm AsIII, respectively. Arsenic tolerance in rice after co-culture with genetically engineered yeast suggested its potential role in arsenic bioremediation. Thus, characterization of *WaarsM* provides a potential strategy to reduce arsenic concentration in soil with reduced arsenic accumulation in crops grown in arsenic contaminated areas, and thereby alleviating human health risks.

Introduction

Elevated arsenic concentration in farmland soil is the major source of arsenic accumulation in crops [1]. Staple food crops, like rice (*Oryza sativa*), grown in arsenic contaminated fields have lower yield and high arsenic accumulation in their edible parts. Higher accumulation of arsenic in edible plants cause increased human health risks, including cancer, especially in Southeast Asia [2]. Arsenic exists in two forms in the environment: the inorganic arsenate (AsV) and arsenite (AsIII), and the organic mono, di or trimethyl arsines. AsIII is more toxic than AsV whereas the organic arsenicals, trimethyl arsine [TMA(III)] is almost non-toxic and volatile[3]. Arsenic biotransformation was considered as a major pathway for arsenic detoxification, which includes the processes of oxidation, reduction and methylation[4-7]. Arsenic methyltransferase is a key enzyme that catalyzes the transfer of a methyl group to the acceptor (arsenic) in the presence of the methyl group donor like the methylcobalamin, S-adenosylmethionine [8]. Arsenic methylation is a sequential reaction, where toxic inorganic arsenic methylates into less toxic pentavalent mono, di and tri-methylated arsenicals. Trivalent methylarsonous acid [MAs(III)] and dimethylarsinous acid [DMAs(III)] are highly toxic intermediates of methylation reaction and are easily oxidized into less toxic methylarsenate [MAs(V)], dimethylarsenate [DMAs(V)] and trimethylarsine oxide [TMAs(V)O] [9].

In microbes, TMA(III), a volatile metabolite, elaborates the arsenic detoxification pathway by converting the inorganic arsenic into organic arsenic[10]. Soil microbes thus play a crucial role in environmental arsenic detoxification by biotransformation of inorganic arsenicals into innocuous organic forms[11]. In our previous study, a fungal strain *Westerdykella aurantiaca* isolated from arsenic contaminated (9.45–15.63 mg kg^{-1}) agricultural soils from West Bengal,

India, was shown to possess high tolerance against arsenic stress[12]. However, not much is known about arsenic metabolism in fungi by rapid methylation and volatilization.

Bioremediation of arsenic contaminated soils by indigenous microorganisms has limited application ,whereas genetically engineered (GE) microorganisms are a better choice for arsenic bioremediation as they are able to generate the high amount of volatile arsenic. Recently, *arsM* gene was expressed in *Sphingomonas desiccabilis*, *Bacillus idriensis* and *Pseudomonas putida* , resulting in improved arsenic volatilization from polluted soil[13-14]. In earlier studies yeast was reported to be used for plant growth promotion and as a bio-control agent[15-17]. Development of new GE yeast with higher arsenic methylation and volatalization abilities can be a new and alternate strategy for bioremediation of arsenic contaminated soil and therebyenabling low arsenic accumulation and increased yield in crops cultivated in arsenic affected areas.

In this study, a novel arsenic methyltransferase gene (*WaarsM*) that catalyzes methylation of inorganic arsenicals was identified from *W. aurantiaca*. *WaarsM* was expressed in mutants of *Escherichia coli* and *Saccharomyces cerevisiae* for functional characterization of the gene. Here, the main focus of our study was to understand the molecular and enzymatic mechanism of arsenic methylation by *WaarsM* of *W. aurantiaca* and to assess the ability of *WaarsM* in arsenic bioremediation and its ecological significance as well.

MATERIALS AND METHODS

Microorganisms and culture conditions

W. aurantiaca was isolated from arsenic contaminated agricultural soil and grown in mycological broth consisting of 1% (w/v) peptone, 4% (w/v) dextrose, pH 7 at 28°C in dark. For solid medium, mycological broth with 1.5% (w/v) agar was boiled and cooled to room

temperature and autoclaved at 121°C for 20 min at 15 psi pressure. The media was poured in petriplates and allowed to solidify before use. *W. aurantiaca* was used as a source of arsenic methyltransferase and RNA. The microorganisms used in study and their sources are listed in ESI, Table S2.

E. coli cells were grown at 37°C in LB media. Strain DH5α was used for plasmid replication. Strain AW3110 (DE3) was used for functional verification of *WaarsM* gene. Strain BL21 (DE3) was used for protein purification. *S. cerevisiae* was cultured in synthetic minimal (SD) liquid medium composed of 0.67% (w/v) yeast nitrogen base without amino acids (Sigma, USA) and 0.5% (w/v) D-glucose with an amino acid mixture and 1.5% (w/v) agar (for solid medium). *S. cerevisiae* strains, *Invsc1* and *Δacr2* were also used for the functional verification of *WaarsM* gene. *Invsc1* was also used to determine the arsenic bioremediation ability of *WaarsM*. Unless specified otherwise all chemicals were procured from Sigma, and all reagents used in this study were of analytical or better grade.

Identification of arsenic methyltransferase of *W. aurantiaca*

From the leads from our previous study on arsenic volatilization by *W. aurantiaca*, we identified an arsenic methyltransferase gene inform this soil fungus. For expression analysis, mRNA were purified from *W. aurantiaca* grown in mycological broth treated with 10 ppm AsV (Na_3AsO_4) at 28°C for 3, 6, 9 and 12 days. The first-strand cDNA was synthesized by RevertAid First Strand cDNA synthesis Kit (Thermo Scientific™, USA) and further used for qRT-PCR analysis of arsenic methyltransferase gene under AsV stress in a time-dependent manner.

Full-length amplification of *WaarsM* gene from *W. aurantiaca*

Full-length genes encoding arsenic methyltransferase were searched and identified by keyword, domain name (arsenic methyltransferase) and BLASTp searches available at NCBI. Degenerate primers were designed with the help of software Hyden[18] and partial fragment was amplified using these primers. The sequence of the partial fragment was used to design the gene-specific primers for 5' and 3' RACE using 5' and 3' RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (Invitrogen™, USA). For full-length amplification of *WaarsM* gene, cDNA from the 9th day treated sample was used as the template. The 5' and 3' RACE amplified fragments were cloned into pTZ57R/T vector and were sequenced. The Full-length sequence of *WaarsM* was established by aligning the sequences of partial fragment as well as 5' RACE and 3' RACE fragments. Finally, a fragment encompassing the complete open reading frame (ORF) of *WaarsM* was amplified by gene-specific primers (ESI, Table S1).

In silico* analysis of *WaarsM

The homologous sequences of various arsenic methyltransferase were searched using BLASTp program available at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The sequence alignment of full-length proteins from different organisms was done using MUSCLE software (v3.8)[19] and a phylogenetic tree was constructed using MEGA 6.0.1 via the Neighbor-Joining (NJ) method with the following parameters: WAG protein substitution model, gamma distribution, and bootstrap (1000).

Cloning and expression of *WaarsM* gene in *E. coli* and *S. cerevisiae*

A 0.876-kb fragment of *WaarsM* was amplified using the gene specific primers. The PCR product was digested with *Eco*RI and *Xba*I, and inserted into pYES2 shuttle vector under *GAL1* promoter to construct plasmids of pYES2-*WaarsM* and into pET28b(+) expression vector at

*Nde*I and *Bam*HI for protein expression under T7 promoter. The plasmids were subsequently transformed in strains of *S. cerevisiae* and arsenic hypersensitive *E. coli* strain AW3110 (*Δars*), and *E. coli* strain BL21.

Assessment of tolerance to inorganic arsenic in transformed *E. coli* (*Δars*)

WaarsM transformed *E. coli* (*Δars*) cells were inoculated into 5 ml of LB medium and incubated at 37°C overnight at 220 rpm. Late exponential phase cells were diluted in the medium at a concentration of 1×10^7 cells/ml and each culture was further diluted 10, 100, and 1000-fold. An equal amount of each suspension was spotted on LB medium containing 50 mg/L kanamycin with 0.1 mM isopropyl β D-thiogalactoside (IPTG) and known concentrations of AsV and AsIII, and was incubated at 37°C overnight and viability of cells was observed.

Assessment of arsenic tolerance and arsenate reductase activity of *WaarsM* in engineered *S. cerevisiae*

S. cerevisiae (Δ*acr2*) cells transformed with pYES2-*WaarsM* and pYES2 were inoculated into 5 ml of SC medium with 2% glucose at 30°C overnight at 220 rpm. Late exponential phase cells were diluted in the medium at a concentration of 1×10^7 cells/ml and each culture was further diluted 10, 100, and 1000-fold. An equal amount of each suspension was inoculated in SC medium with 2% galactose and indicated concentrations of AsV and AsIII in the presence of BSO and incubated at 30°C. The viability of cells was observed after 48 hours. *In vitro* arsenate reductase activity was assayed using the coupled enzymatic reaction described by Shi *et al.* (1999)[20] (See ESI).

Purification of WaarsM enzyme and methyltransferase activity analysis

The *WaarsM* cDNA was cloned into the pET28b(+) vector for protein expression in *E. coli* strain BL21 (DE3) and purified using Ni-NTA column (See ESI). The arsenic methyltransferase activity of the purified enzyme was determined using SAM Methyltransferase Assay kit (Calbiochem; Darmstadt, Germany) according to manufacturer instructions. The reaction mixture contained SAM methyltransferase assay buffer, adenosylhomocysteine, methyltransferase enzyme mix, SAM methyltransferase assay buffer additive, SAM colorimetric mix, 0.5 µg of purified WaarsM protein and 0.2 mM AsV and 0.1 mM AsIII as a substrate in a final volume of 100 µl. In methylation reaction, WaarsM generates S-adenosylhomocysteine, which rapidly gets converted to S-ribosyl homocysteine resulting in the end products: urate and hydrogen peroxide. The rate of production of hydrogen peroxide was measured by using a colorimetric spectrophotometer (PerkinElmer, USA).

The kinetic properties of WaarsM (0.1 µg) in a total volume of 100 µl were determined using different concentrations of AsV (0.2 to 1 mM) and AsIII (0.1 to 0.5 mM) as substrates. The K_M and V_{max} were calculated from Lineweaver-Burk plots using the computer program Graph Pad Prism, Version 1.0 (Graph Pad Software, San Diego, Calif., USA)[21].

In vitro arsenic methylation and reaction mechanism of WaarsM

In vitro arsenic methylation assays with purified WaarsM was performed in a buffer consisting of 50 mM phosphate buffer, pH 7.4, 8 mM reduced GSH and 0.3 mM AdoMet chloride using 200 ppb AsV, 100 ppb AsIII and 100 ppb MMA(V) as substrates. The methylated products were detected by HPLC-ICP-MS.

Arsenic methylation and volatilization by *S. cerevisiae* expressing *WaarsM*

S. cerevisiae strain (Δ*acr2*) bearing both pYES2-*WaarsM* and pYES2 cells were grown overnight at 30°C in 10 mL SC medium containing 2% galactose as an inducer and 20 ppm AsV and 2 ppm AsIII with gentle shaking at 30°C. After 24 hours, the arsenic species in the media was analyzed by HPLC-ICP-MS. For trapping of volatile arsenicals, the nitrocellulose membrane was placed in cap of vials impregnated with 0.15 ml of 6% H_2O_2 to oxidize TMA(III) to TMAO. The membranes were digested with 0.2 ml of 70% HNO_3 at 90°C for 20 min and analyzed by inductively coupled plasma Mass Spectrometry (ICP-MS).

Ability of arsenic bioremediation by *S. cerevisiae* expressing *WaarsM*

Rice seeds (Usar 3) were sterilized in 10% (v/v) H_2O_2 solution for 15 min, washed thoroughly with deionised water and then germinated in moist filter paper. After 10 days of germination, uniform rice seedlings were selected and transplanted to the nutrient solution. *S. cerevisiae* strains bearing *WaarsM* were pre-cultured in flasks containing 200 ml of YPD medium on a shaker at 30°C for 48 hours. Cells were collected from the culture by centrifugation (5 min; 3000 × *g*) and then suspended in fresh 1:1 (v/v) mixture (pH 5.5) nutrient solution containing 25 µM AsIII and 250 µM AsV. The nutrient solution was replaced with suspension containing transformed *S. cerevisiae* and cultured for 3 days at 28°C with a 16 h photoperiod. After the culture period, the plants growth was monitored.

Arsenic Speciation Analysis

Arsenic speciation was analyzed by HPLC-ICP-MS (PerkinElmer series) as described in ESI, Table S3. An Anion Exchange, Hamilton PRP-X100 (4.1 mm i.d. x 250 mm, 10 µm) was used. The mobile phase consisted of 20 mM ammonium bicarbonate (pH 8.5) and 20 mM

ammonium sulfate (pH 7.0). The mobile phase was pumped through the column at a flow rate of 1.0 mL min^{-1}.

Statistical analysis

Each experiment was carried with three replicates and repeated at least thrice. Data are presented as the average of the mean ± SE. One way ANOVA and two way ANOVA were used to determine significance as appropriate using Graph Pad Prism, Version 1.0 (Graph Pad Software, San Diego, Calif., USA), and the treatment mean values were compared at $P \leq 0.05$.

Results

Identification and cloning of arsenic methyltransferase gene from *W. aurantiaca*

Based on previous studies we identified an arsenic methyltransferase gene in *W. aurantiaca*. qRT-PCR analysis of this arsenic methyltransferase gene confirmed high mRNA accumulation after 9th day of AsV treatment (ESI, Fig. S1). To deduce CDS sequence, arsenic methyltransferase domain was amplified with the help of degenerate primers, designed from conserved regions of arsenic methyltransferase gene of different organisms. Partial sequence of this domain was used to design primers for 3' RACE and 5' RACE whereby 612 bp and 699 bp RACE products were obtained, respectively. Full length sequence of *WaarsM* comprised putative open reading frame of 876 bp with 51 bp 5' and 36 bp 3' UTR regions, including polyA tail, encoding the polypeptide of 291 amino acid residues. This Full length sequence was submitted to NCBI GenBank database (GenBank Accession no. KP165533.1).

In silico analysis showed a typical arsenic methyltransferase

The pI and molecular weight of the WaarsM protein is 4.94 and 30.7 kDa, respectively, predicted by the software pI/Mw tool at http://www.expasy.org[22]. Alignment score showed that WaarsM has a typical structure of arsenic methyltransferase, with highly conserved SAM

binding motifs (motif I- VIDLGSGAG, motif II- ADCIISNC and motif III- LLKRGGRVAI)[23] , and three cysteine residues at position 54, 143, 195, which form three coordinate surface sites for arsenic binding and methylation reaction[24] (Fig. 1a). Phylogenetic tree displayed high sequence homology among all arsenic methyltransferase and showed an evolutionary relationship closest to *Trichophyton rubrum* (XP_003237817.1) (Fig. 1b).

Expression of *WaarsM* enhances tolerance in arsenic hypersensitive *E. coli* (*Δars*)

The expression of *WaarsM* showed complementation in *E. coli* strain lacking *ars* operon and did not contain any *arsM* gene when cultured in arsenic containing medium (Fig. 2). The growth of *E. coli* (Δ*ars*) strain harboring the empty vector was impaired in the presence of 3 mM AsV and 0.6 mM AsIII. In contrast, *E. coli* (Δ*ars*) transformed with the *WaarsM* was able to grow in the presence of AsV at a concentration of up to 4 mM and up to 0.8 mM AsIII concentration.

Engineered *S. cerevisiae* (Δ*acr2*) showed arsenic resistance in GSH dependent manner

The heterologous expression of *WaarsM* showed enhanced tolerance in *S. cerevisiae* (*Δacr2*) on exposure to various concentrations of AsIII and AsV as compared to empty vector (pYES2). Tolerance in *S. cerevisiae* (Δ*acr2*) strain at a different concentration of AsV suggested that *WaarsM* gene also had additional ability to reduce AsV into AsIII. The tolerance to arsenic lowers after inhibition of GSH biosynthesis demonstrated that *WaarsM* requires endogenous reductant, preferably GSH for methyltransferase activity (Fig. 3).

Methyltransferase activity and kinetics of WaarsM

Immunoblot analysis of recombinant WaarsM using anti-his antibody confirmed the expression of WaarsM protein of 30.7 kDa (Fig. 4a, b). Methyltransferase activity was monitored

colorimetrically, which was evident by an increase in absorbance at 510 nm, confirming methyltransferase activity of WaarsM when incubated with AsV and AsIII in the presence of SAM (Fig. 4c, d). The WaarsM had a relatively lower K_M for AsIII than AsV, suggesting WaarsM can catalyze AsIII methylation more efficiently than AsV (ESI Table S4, Fig. 4e, f). Arsenate reductase activity of purified WaarsM was also assayed and revealed that the activity of purified WaarsM was about 0.4135 nM of NADPH oxidized per minute by 1 unit of WaarsM (ESI, Fig. S2).

Purified WaarsM exhibits *in vitro* methylation and volatilization

In vitro methylation reaction showed that methylation of arsenic can be initiated from both AsV and AsIII. When AsV (200 ppb) was used as a substrate in presence of GSH, after 3 hours of reaction volatile arsenicals (27.9 ppb) were detected with AsIII (13.4 ppb) and AsV (155.9 ppb). While, in absence of GSH, only AsIII (23 ppb) and AsV (173.9 ppb) were detected in the reaction suggesting that WaarsM was also involved in AsV reduction (Fig. 5a, c). When AsIII (100 ppb) was used as a substrate, all steps of methylation reaction took place in presence of GSH *ie.* DMA (29.3 ppb) and volatile arsenicals (29.3 ppb) were detected after 3 hours of reaction while in the absence of GSH only AsIII were detected after HPlC-ICP-MS (Fig. 5b, d). For further confirmation, MMA(V) was also used as initial substrate in the presence of GSH for methylation reaction, which showed that within 1 hour of reaction, total conversion of MMA(V) into DMA(V) takes place. These results gave the evidence that the second step of methylation was very fast (ESI, Fig. S3) and TMA, as a final product, was formed but not detected in reaction mixture due to its volatile nature.

WaarsM detoxifies arsenic by biomethylation and bio volatilization in _S. cerevisiae_

S. cerevisiae (Δ_acr2_) cells transformed with _WaarsM_ gene and grown in arsenic containing medium revealed that, after exposure to 2 ppm AsIII, approximately 0.64 ppm, DMA (V) and 0.58 ppm volatile arsenic were detected, whereas on exposure to 20 ppm AsV detected approximately 2.2 ppm volatile arsenic with an unknown peak that did not correspond to any of the standards and calculated by mass balance. No methylated species were observed in the cells bearing the vector, plasmid pYES2 (Fig. 6). Moreover, it was found that the amount of volatile arsenic increased with the exposure time (8, 16, 24, 32 h). It was also noted that arsenic volatilization increased slowly during the first 16 h, whereas at the latter period of the incubation, the production of volatile arsenic increased rapidly, particularly, at 32 h exposure (ESI, Table S5).

WaarsM transformed _S. cerevisiae_ act as bioremediation agent

Rice seedlings grown in medium containing AsV and AsIII with engineered yeast cells showed enhanced tolerance and less arsenic accumulation in shoot and leaves. Rice seedlings co-cultured with _WaarsM_ engineered yeast also showed better growth as determined by fresh weight than plants grown with yeast containing empty vector (Fig. 7).

Discussion

Our study suggested that _WaarsM_ gene is primarily liable for arsenic detoxification and extreme arsenic tolerance in _W. aurantiaca_ by methylation and volatilization, and the characterization of _WaarsM_ provides a potential approach for bioremediation.

The tolerance in arsenic-hypersensitive _E. coli_ (Δ_ars_) and _S. cerevisiae_ (Δ_acr2_) strongly suggests its role in arsenic detoxification. The _E. coli Ars_ operon encodes a series of proteins

involved in AsV reduction and AsIII extrusion[25], whereas *Acr2* gene in *S. cerevisiae* has an important role in arsenic tolerance and encodes protein for AsV reduction[26].

Arsenic tolerance in *S. cerevisiae* (Δ*acr2*) cells carrying *WaarsM* under limited intracellular GSH concentrations displayed that arsenic methylation activity of *WaarsM* gene was GSH dependent, even as other endogenous reductants also assisted the arsenic methylation in the same conditions. This result is also endorsed by previous studies of Waters and co-workers (2004) wherein endogenous reductants were shown to alleviate the methyltransferase activity of recombinant rat cyt19 arsenic methyltransferase[27].

In vitro methylation reaction initiated with both valence forms of inorganic arsenic suggested its broad function in arsenic detoxification. Arsenic methylation by WaarsM involves two types of reactions, thiol exchange and oxidation methylation reaction[8]. When AsV was used as substrate, either it interacted with cysteine residue of WaarsM and got reduced to AsIII or reduced by GSH non-enzymatically to form As $(GS)_3$ complex leading to further methylation reaction[28-30]. In the presence of GSH, first and second round of methylation reaction takes place. Second round of methylation reaction required GSH or other endogenous reductants to reduce cysteine which was oxidized during the first round of methylation[27]. Also the involvement of GSH as a reductant supporting the catalytic activity of WaarsM indicated an interaction between arsenic and a thiol reductant (e.g., GSH). GSH could be required to reduce disulfide bonds between any two of the nine cysteine residues in WaarsM. Maintenance of the correct ratio of reduced to oxidized cysteine residues may be crucial for the catalytic function of WaarsM. Alternatively, GSH could be the source of reducing equivalents used to reduce intermediates containing pentavalent arsenic to trivalent before methylation by WaarsM, as only single step reaction occurred in the absence of GSH. When AsIII was used as a substrate, all steps of

methylation reaction took place only in presence of GSH, suggesting that $As(GS)_3$ complex was the preferred substrate for further methylation reaction, which was also supported by the findings of Hayakawa and co-workers (2005). The peak of MMA was not detected probably because the methylation reaction of MMA to DMA was very fast. For further confirmation, MMA(V) was also used as initial substrate and found that within 1 hour of reaction, total conversion of MMA(V) takes place which suggesting that second step of methylation (MMA to DMA) is very fast. Catalytic ability of AsV reduction by WaarsM was also proved by the reaction without GSH, where AsV reduction takes place, but further reaction did not proceed due to absence of recycled cysteine residues, oxidized by AsV.

The ability of *WaarsM* in bioremediation was tested by the co-culture experiment which provides very strong evidence that the plants grown on arsenic-contaminated soil were able to tolerate at higher concentration and with low arsenic accumulation. The product of methylation by the genetically modified yeast cells is volatile arsenic with DMA (V) which is less toxic than inorganic arsenic. DMA(V) can also up taken by aquaporin channel OsLsi1[31-32],which may also compete with AsIII (much more toxic than DMA). In contrast, MMA(V) and DMA(V) are not form complex with thiols, which is very mobile and more efficiently translocated from roots to the shoot in rice and minimizes the arsenic toxicity in the plant.

Arsenic methyltransferase genes are widespread among fungi, which suggest that fungi also had a more sophisticated impact on global arsenic cycling mediated by the production of methylated arsenicals. Until this study no work has been carried out on characterization and heterologous expression of arsenic methyltransferase from soil fungi in *S. cerevisiae*. Earlier studies mainly focused on the use GE bacterium for arsenic bioremediation, which may also be pathogenic for humans and can cause health risk[33]. Earlier it was also suggested that *S. cerevisiae* can be used as

bio-fertilizers and bio-control agents for crops[15-16]. In this study we also compare biovolatilization of GE yeast cells with GE bacterial cells (ESI, Fig. S4). The low arsenic volatilization by GE bacteria might be due to codon bias of *WaarsM* resulting in post-transcriptional processing. Therefore, establishment of engineered *S. cerevisiae* on plant roots can possibly be a better and safer approach for arsenic bioremediation of contaminated soil[13].

Conclusion

In the present study, an arsenic methyltransferase gene was identified in fungus isolated from arsenic contaminated soil. Overexpression of *WaarsM* enhanced resistance in arsenic sensitive strains of *E. coli* and *S. cerevisiae* confirmed a role of *WaarsM* gene in arsenic detoxification. Purified WaarsM enzyme showed methyltransferase activity from both oxidation form of arsenic including AsV reductase activity. *In vitro* and *in vivo* arsenic methyltransferase activity of WaarsM displayed a potential role in arsenic detoxification via biomethylation and biovolatilization. Thus, our study reveals new insights in arsenic metabolism by *WaarsM* and provides a potential approach for bioremediation process by using highly efficient genetically engineered yeast that can also support plant growth and bio-control of harmful indigenous microbes.

Acknowledgments

The authors acknowledge the Director CSIR-NBRI for extending the facilities to carry out this work, and the Department of Biotechnology (DBT), and the Council of Scientific and Industrial Research (CSIR), Govt of India, for the funding through the Projects GAP-256225 and BSC-0107, respectively. S.V. and P. K.V. thankfully acknowledge Council of Scientific and Industrial Research (CSIR), India for senior research fellowship and A.K.M. for UGC-Senior

research fellowship. Authors have no conflict of interest. Authors are grateful to Dr. K. N. Nair, Sr. Principal Scientist, CSIR-NBRI for his help to correct the language.

References

1. P. N. Williams, M. Lei, G. Sun, Q. Huang, Y. Lu, C. Deacon, A. A. Meharg and Y. G. Zhu, *Environ Sci Technol*, 2009, **43**, 637-642.
2. S. M. I. Huq, J. C. Joardar, S. Parvin, R. Correll and R. Naidu, *J Health Popul Nutr*, 2006, **24**, 305-316.
3. K. M. Campbell and D. K. Nordstrom, *Reviews in Mineralogy and Geochemistry*, 2014, **79**, 185-216.
4. R. Bentley and T. G. Chasteen, *Microbiol Mol Biol Rev*, 2002, **66**, 250-271.
5. M. Vahter, *Toxicology*, 2001, **164**, 17-17.
6. M. Vahter, *Toxicology*, 2002, **181**, 211-217.
7. M. A. Rahman and C. Hassler, *Aquat Toxicol*, 2014, **146**, 212-219.
8. D. J. Thomas, J. X. Li, S. B. Waters, W. B. Xing, B. M. Adair, Z. Drobna, V. Devesa and M. Styblo, *Exp Biol Med*, 2007, **232**, 3-13.
9. W. R. Cullen, *Chem Res Toxicol*, 2014, **27**, 457-461.
10. W. R. Cullen, *J Environ Monit*, 2005, **7**, 11-15.
11. R. A. Wuana and F. E. Okieimen, *ISRN Ecology*, 2011, **2011**, 20.
12. P. K. Srivastava, A. Vaish, S. Dwivedi, D. Chakrabarty, N. Singh and R. D. Tripathi, *Sci Total Environ*, 2011, **409**, 2430-2442.
13. J. Chen, G. X. Sun, X. X. Wang, V. de Lorenzo, B. P. Rosen and Y. G. Zhu, *Environmental Science & Technology*, 2014, **48**, 10337-10344.
14. S. Liu, F. Zhang, J. Chen and G. X. Sun, *J Environ Sci-China*, 2011, **23**, 1544-1550.
15. S. Mukherjee and S. K. Sen, *J Sci Food Agr*, 2015, **95**, 1491-1499.
16. L. V. Ignatova, Y. V. Brazhnikova, R. Z. Berzhanova and T. D. Mukasheva, *Microbiol Res*, 2015, **175**, 78-83.
17. K.-o. Amprayn, M. T. Rose, M. Kecskés, L. Pereg, H. T. Nguyen and I. R. Kennedy, *Applied Soil Ecology*, 2012, **61**, 295-299.
18. C. Linhart and R. Shamir, *Bioinformatics*, 2002, **18 Suppl 1**, S172-181.
19. R. C. Edgar, *Nucleic Acids Res*, 2004, **32**, 1792-1797.
20. J. Shi, A. Vlamis-Gardikas, F. Aslund, A. Holmgren and B. P. Rosen, *J Biol Chem*, 1999, **274**, 36039-36042.
21. H. J. Motulsky, *Journal*, 2007, **www.graphpad.com**.
22. E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel and A. Bairoch, *Nucleic Acids Res*, 2003, **31**, 3784-3788.
23. R. M. Kagan and S. Clarke, *Arch Biochem Biophys*, 1994, **310**, 417-427.
24. A. A. Ajees, K. Marapakala, C. Packianathan, B. Sankaran and B. P. Rosen, *Biochemistry*, 2012, **51**, 5476-5485.
25. A. Carlin, W. Shi, S. Dey and B. P. Rosen, *J Bacteriol*, 1995, **177**, 981-986.
26. M. Ghosh, J. Shen and B. P. Rosen, *P Natl Acad Sci USA*, 1999, **96**, 5001-5006.

27. S. B. Waters, V. Devesa, L. M. Del Razo, M. Styblo and D. J. Thomas, *Chem Res Toxicol*, 2004, **17**, 404-409.
28. M. Delnomdedieu, M. M. Basti, J. D. Otvos and D. J. Thomas, *Chem Biol Interact*, 1994, **90**, 139-155.
29. T. Hayakawa, Y. Kobayashi, X. Cui and S. Hirano, *Arch Toxicol*, 2005, **79**, 183-191.
30. N. Scott, K. M. Hatlelid, N. E. MacKenzie and D. E. Carter, *Chem Res Toxicol*, 1993, **6**, 102-106.
31. R. Y. Li, Y. Ago, W. J. Liu, N. Mitani, J. Feldmann, S. P. McGrath, J. F. Ma and F. J. Zhao, *Plant Physiol*, 2009, **150**, 2071-2080.
32. X. Wang, B. Peng, C. Tan, L. Ma and B. Rathinasabapathi, *Environ Sci Pollut Res Int*, 2015, **22**, 5742-5750.
33. B. S. Thomas, K. Okamoto, M. J. Bankowski and T. B. Seto, *Infect Dis Clin Pract (Baltim Md)*, 2013, **21**, 147-213.

Figure Legends

Fig. 1 Pairwise alignment of amino acids and phylogeny of arsenic methyltransferases sequences. (a) Arsenic methyltransferase protein sequences from *Coprinopsis cinerea okayama* (XP_001832959.2), *Aspergillus fumigatesAf293* (XP_753155.1), *Talaromyces stipitatus ATCC 10500* (XP_002487876.1), *Trichophyton rubrum CBS 118892* (XP_003237817.1), *Fusarium fujikuroi IMI 58289* (CCT67817.1), *Oikopleura dioica* (XP_001913342.1), *Saccoglossus kowalevskii* (XP_006823910.1), *Salmo salar* (NP_001139863.1), *Danio rerio* (NP_001034928.1), *Ictalurus punctatus* (NP_001187373.1), *Mus musculus* (EDL42017.1), *Rattus norvegicus* (NP_543166.1), *Oryctolagus cuniculus* (XP_008268651.1), *Bos Taurus* (NP_001030195.1), *Homo sapiens* (NP_065733.2), *Xenopus tropicalis* (NP_001135714.1), *Callorhinchus milii* (XP_007882989.1), *Cyanidioschyzon sp* (ACN39190.1), *Acanthamoeba castellanii str. Neff* (XP_004346011.1) were obtained from NCBI protein database and using MUSCLE software. Residues highlighted by solid line above the amino acid alignment shows SAM binding motifs conserved in all arsenic methyltransferase proteins. Conserved cysteines residues are indicated by solid black triangles. (b) Phylogenetic analysis of WaarsM protein with arsenic methyltransferase proteins from various organisms. The phylogenetic tree was constructed using via the Neighbor-Joining (NJ) method using MEGA 6.0.1.

Fig. 2 Expression of *WaarsM* enhances resistance to arsenic hypersensitive *E. coli* AW3110 (*Δars*). The 10-fold serial dilutions (up to down) of liquid cultures of equal OD_{600} of *E. coli* (*Δars*) cells transformed with pET28b(+) and pET28b(+)-*WaarsM* were spotted on LB medium with or without AsV and AsIII. Each experiment was carried out at least three times.

Fig. 3 Heterologous expression of *WaarsM* increases arsenic tolerance in *S. cerevisiae* (*Δacr2*) *strain*. The 10 fold serial dilutions (up to down) of the liquid cultures of equal OD_{600} of *S. cerevisiae* (*Δacr2*) strain expressing pYES2-*WaarsM* and pYES2 were spotted on SC medium with various concentration of arsenic in presence and absence of 2 mM BSO. Growth was monitored after 24 hours at 30°C. Each experiment was carried out at least three times.

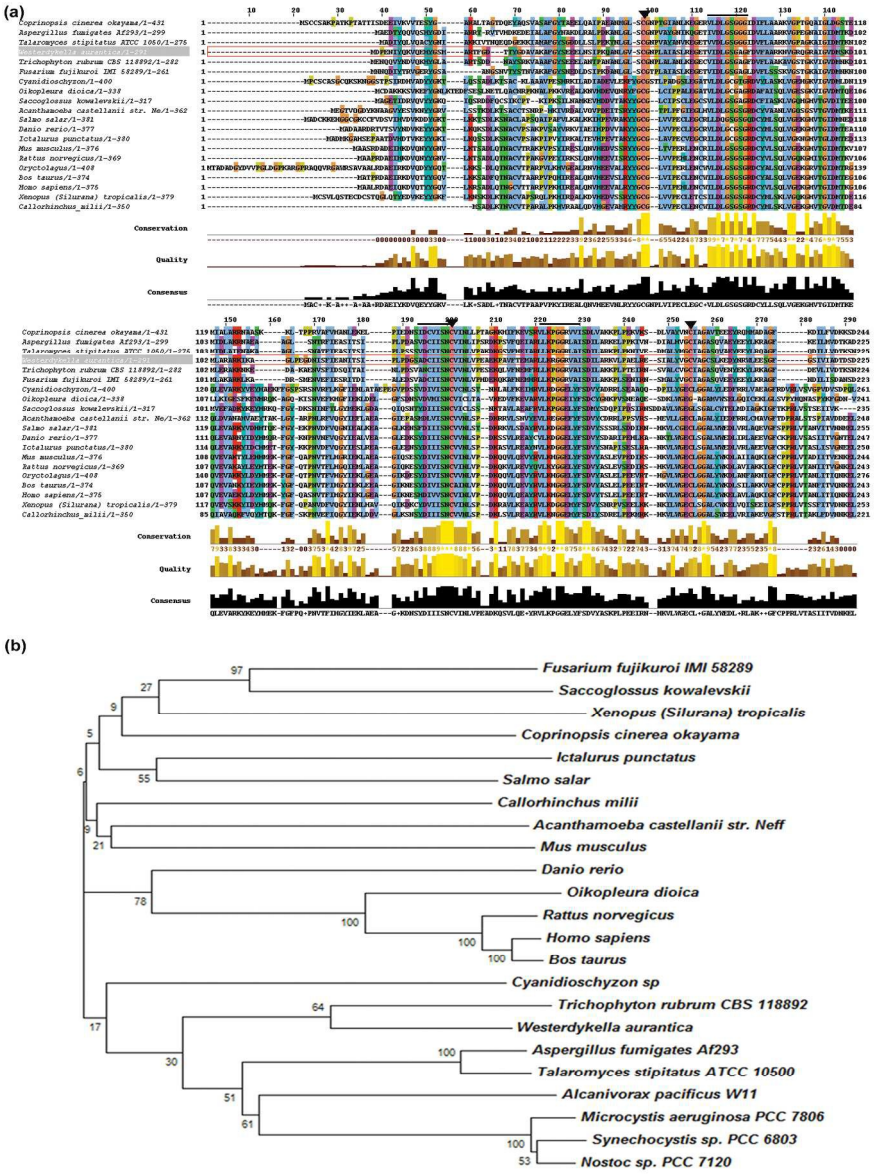
Fig. 4 Enzyme activity and kinetics of purified WaarsM. (a) SDS-PAGE of purified protein. (b) Immunobloting done by anti his antibody. Immunoblot analysis confirms the expression of desired sized protein (UI = Uninduced sample, SF = Soluble fraction, ISF = Insoluble fraction, P = Purified protein). (c) Enzyme activity of purified WaarsM when 0.1 mM AsIII used as the substrate. (d) Enzyme activity of purified WaarsM when 0.2 mM AsV used as the substrate. (e) Double reciprocal plots of WaarsM activity versus AsIII concentration. (f) Double reciprocal plots of WaarsM activity versus AsV concentration. The K_M and V_{max} were calculated from Lineweaver-Burk plots. Data are reported as mean ± SE for three independent experiments and three technical replicates.

Fig. 5 Purified WaarsM catalyzes arsenic methylation. (a) Arsenic species were analyzed after 3 hours of incubation with purified WaarsM when AsV used as a substrate. Curve 1, standard; curve 2, 200 ppb AsV; curve 3, purified WaarsM in the presence of 200 ppb AsV and 5 mM GSH after 3 h at 37°C; curve 4, purified WaarsM in the absence of 5 mM GSH after 3 h at 37°C. (b) Arsenic species were analyzed after 3 hours of incubation with purified WaarsM when AsIII used as the substrate. Curve 1, standard; curve 2, 100 ppb AsIII; curve 3, purified WaarsM in presence of 100 ppb AsIII and 5 mM GSH after 3 h at 37°C; curve 4, purified WaarsM in absence of 5 mM GSH after 3 h at 37°C. (c and d) The products of methylation were observed after 3 hours of reaction, AsV, AsIII and DMA(V) by HPLC-ICPMS and volatile arsenicals were

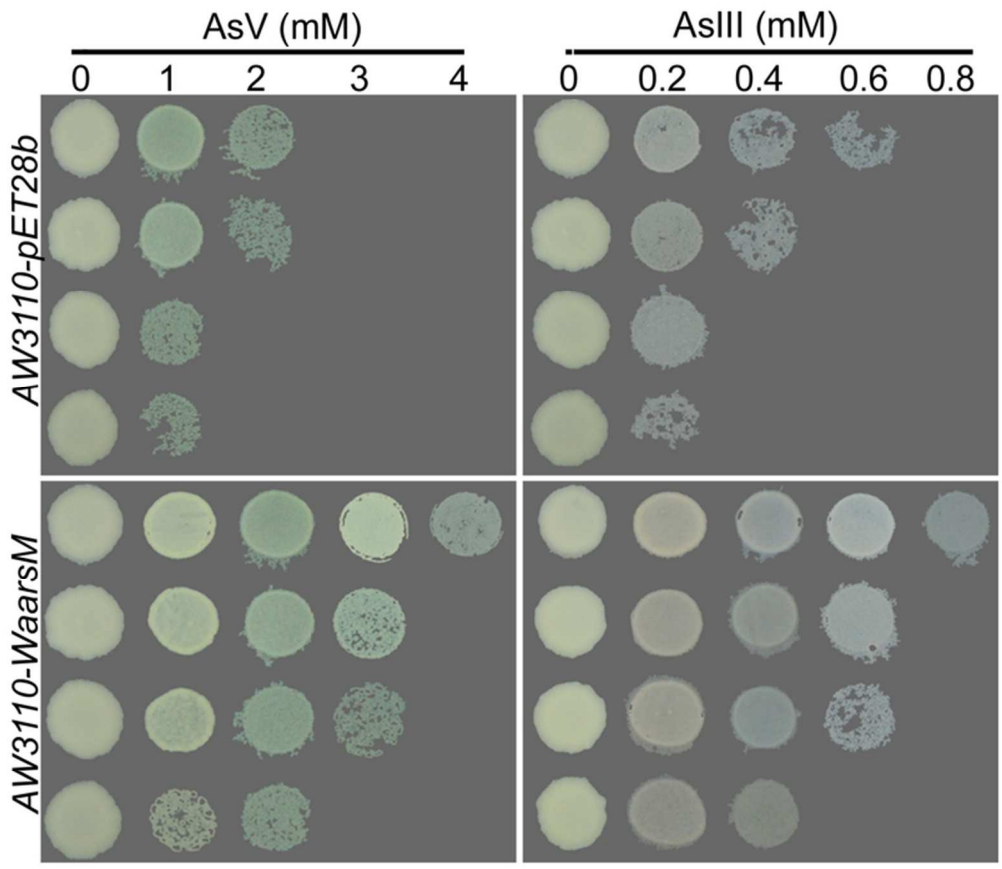
trapped on H_2O_2 impregnated filters and determined by ICP-MS. Data are reported as mean ± SE for three independent experiments and two technical replicates.

Fig. 6 *In vivo* methylation of AsIII and AsV. (a) *S. cerevisiae* (*Δacr2*) cells transformed with pYES2 and pYES2-*WaarsM* gene were grown on medium at 30°C containing 2 ppm AsIII and 20 ppm AsV, arsenic species in medium were analyzed after 24 hours by HPLC-ICPMS. Curve 1, standard; curve 2, cells transformed with pYES2 in presence of AsIII; curve 3, cells transformed with pYES2-*WaarsM* in presence of AsIII; curve 4, cells transformed with pYES2 in presence of AsV; curve 5, cells transformed with pYES2-*WaarsM* in presence of AsV. The unknown peak did not correspond to any of the standards. (b and c) The products of methylation by *WaarsM* after 24 hours were examined. Data are reported as mean ± SE for three independent experiments and two technical replicates.

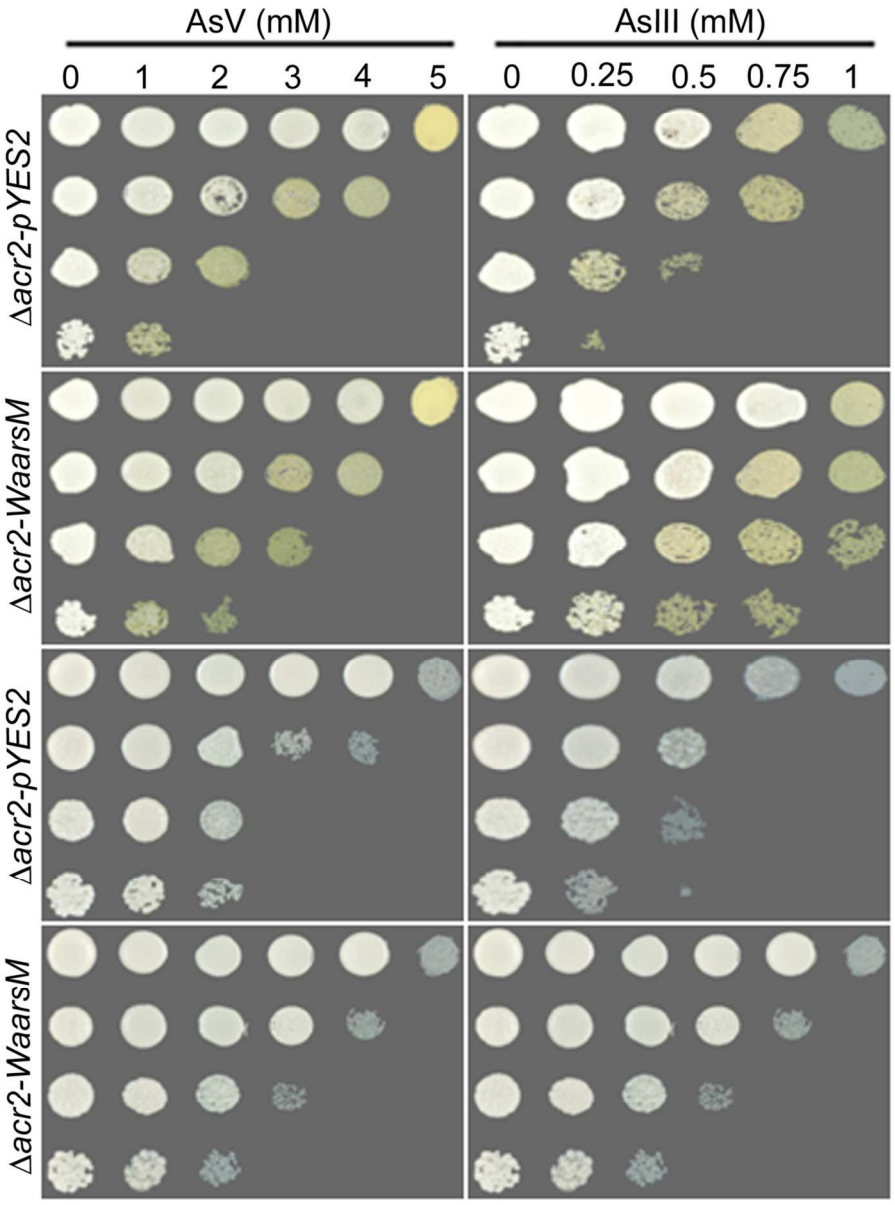
Fig. 7 Rice seedlings co-cultured with engineered yeast cells showed enhanced tolerance under arsenic stress. (a) Rice seedlings were grown with *S. cerevisiae* (empty vector), (b) Rice seedlings grown with *S. cerevisiae* (*WaarsM*) in presence of indicated concentration of AsV and AsIII. (c) Fresh weight of co-cultured rice seedlings in presence of indicated concentration of AsV and AsIII. Arsenic accumulation in co-cultured rice seedlings in (d) AsV stress and (e) AsIII stress. Data are reported as mean ± SE for three independent experiments.



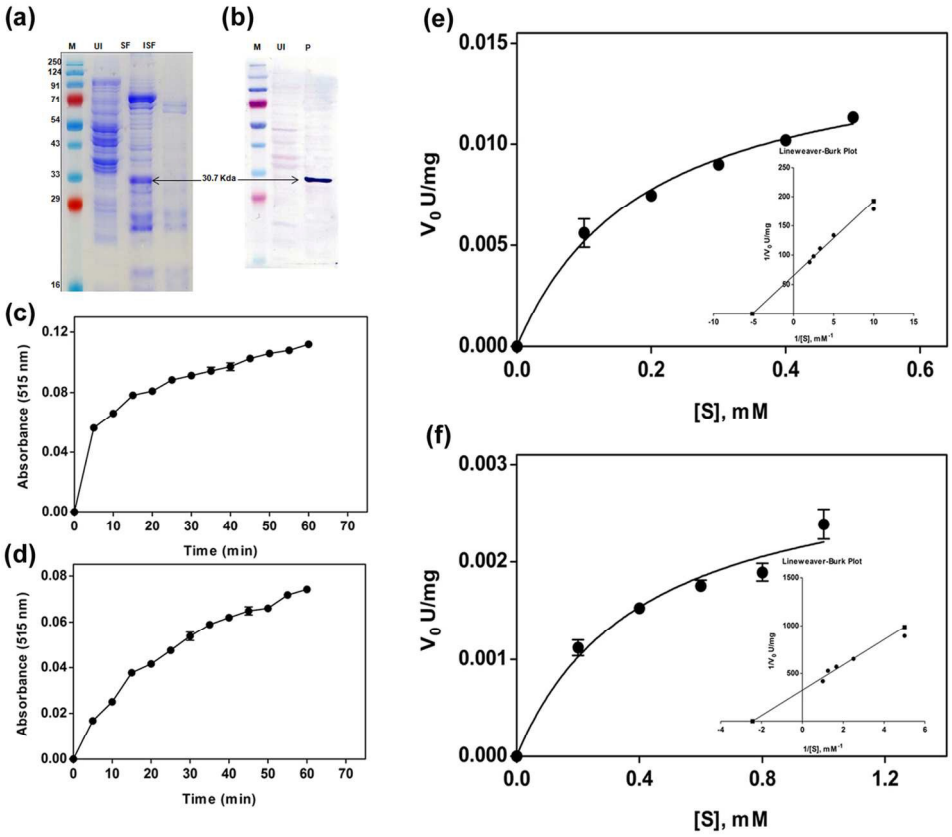
233x317mm (300 x 300 DPI)



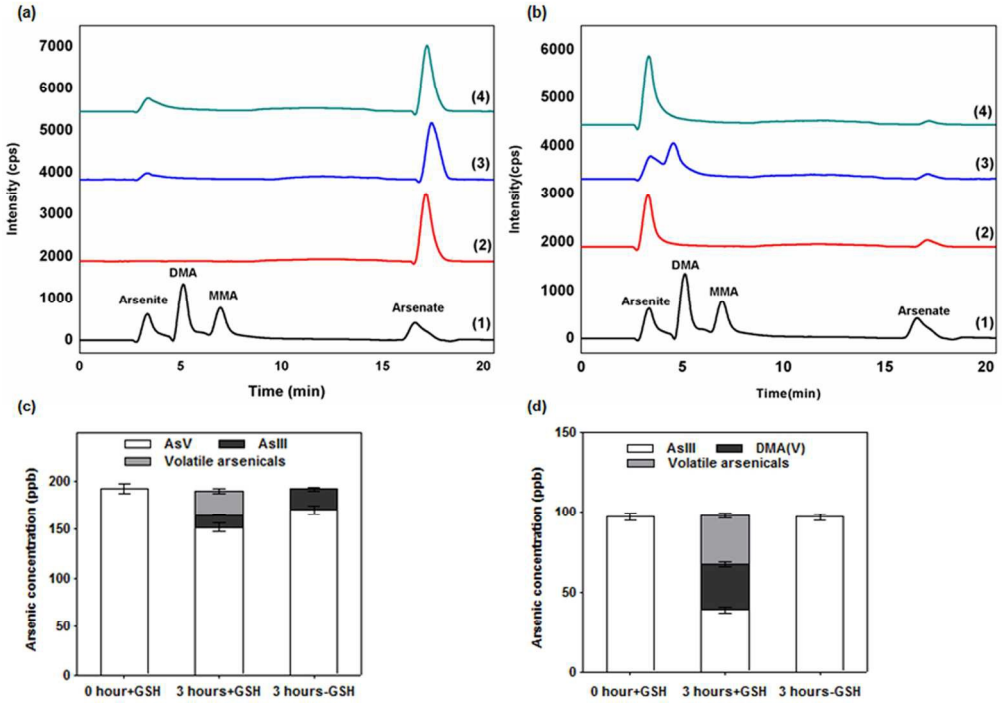
67x58mm (300 x 300 DPI)



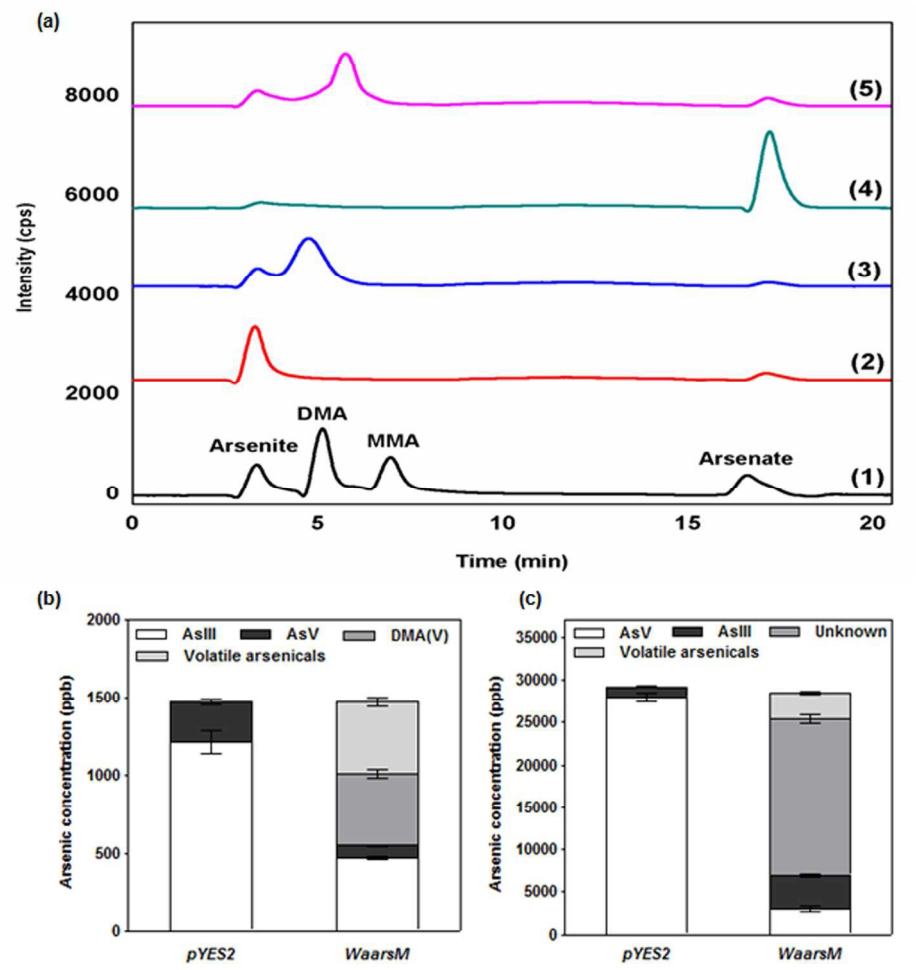
103x137mm (300 x 300 DPI)



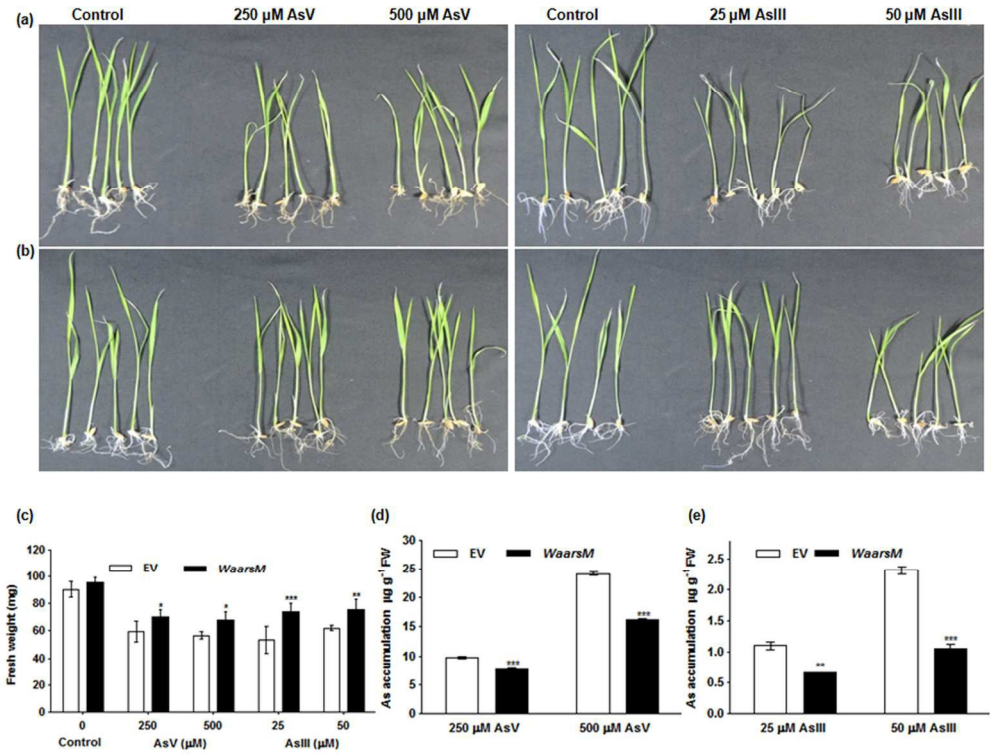
70x59mm (600 x 600 DPI)



119x89mm (300 x 300 DPI)



180x190mm (300 x 300 DPI)



130x99mm (300 x 300 DPI)