Glutathione-S-transferase from the arsenic hyperaccumulator fern *Pteris vittata* can confer increased arsenate resistance in *Escherichia coli*

Aftab A. Khan[1,4], Danielle R. Ellis[2,5], Xinyuan Huang[2,6], Gareth J. Norton[3], Andrew A. Meharg[3,7], David E. Salt[2,8], and Laszlo N. Csonka[1]*

Address: [1]Department of Biological Sciences, Purdue University, West Lafayette IN 47907 United States of America, [2]Departement of Horticulture and Landscape Architecture, Purdue University, West Lafayette IN 47907 United States of America, [3]Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, AB24 3UU, United Kingdom; Current addresses: [4]Laboratory of Renewable Resources Engineering, Purdue University, West Lafayette IN 47907 United States of America, [5]Biology Department, West Virginia University, Morgantown WV 26506-6057, [6]College of Resources and Environmental Sciences, Nanjing Agricultural University, Nanjing 210095 China, [7]School of Biological Sciences, Queen's University Belfast, Belfast, BT7 1NN, Northern Ireland, United Kingdom, [8]Division of Plant and Crop Sciences, School of Biosciences, University of Nottingham, Sutton Bonington Campus, Sutton Bonington, LE12 5RD, United Kingdom

Short title: GSTs from arsenate hyperaccumulator *P. vittata*

*Corresponding author

email addresses of corresponding author Laszlo N. Csonka csonka@purdue.edu

1

2 **Abstract**

1   Although arsenic is generally a toxic compound, there are a number of ferns in the genus *Pteris*

2   that can tolerate large concentrations of this metalloid. In order to probe the mechanisms of arsenic

3   hyperaccumulation, we expressed a *Pteris vittata* cDNA library in an *Escherichia coli ΔarsC*

4   (arsenate reductase) mutant. We obtained three independent clones that conferred increased

5   arsenate resistance on this host. DNA sequence analysis indicated that these clones specify

6   proteins that have a high sequence similarity to the phi class of glutathione-S-transferases (GSTs)

7   of higher plants. Detoxification of arsenate by the *P. vittata* GSTs in *E. coli* was abrogated by a

8   *gshA* mutation, which blocks the synthesis of glutathione, and by a *gor* mutation, which inactivates

9   glutathione reductase. Direct measurements of the speciation of arsenic in culture media of the *E.*

10   *coli* strains expressing the *P. vittata* GSTs indicated that these proteins facilitate the reduction of

11   arsenate. Our observations suggest that the detoxification of arsenate by the *P. vittata* GSTs

12   involves reduction of As(V) to As(III) by glutathione or a related sulfhydro compound.

19

Introduction

Several members of the fern genus *Pteris* can accumulate high levels of arsenic [1,2]. The first arsenic hyperaccumulator discovered was *Pteris vittata* (Chinese brake fern), which can contain arsenic at levels in excess of 20 mg/g in its fronds [1,3]. Such high level accumulation, and hence, tolerance of arsenic is remarkable in view of the toxicity of this element [4]. The exceptionally high arsenic tolerance of the hyperaccumulating ferns not only makes them good candidates for use in biological remediation of arsenic-contaminated environments, but it also makes them interesting systems to unravel the basic mechanisms of resistance to this widespread environmental pollutant.

Paradoxically, detoxification of arsenate involves reduction to arsenite [5,6], which has potentially higher human and ecological toxicity than arsenate [7,8]. Arsenite produced from arsenate is known to be removed from the cytoplasm by various efflux mechanisms. In *Escherichia coli*, it is pumped out by an ATP-dependent export system, and in yeast it is eliminated by a combination of efflux into the medium by a plasma membrane-bound carrier and transfer into the vacuole as a glutathione adduct [5,6]. In *Arabidopsis thaliana*, an arsenic sensitive non-hyperaccumulating plant, arsenite is transferred to the vacuole, as a phytochelatin adduct via two paralogous ABCC-type transporters (*AtABCC1* and *AtABCC2*) [9]. Three different families of arsenate reductases have been identified, represented by the ArsC protein of *E. coli* [10], the ArsC protein of *Staphylococcus aureus* [11], and the Acr2 protein of yeast [12]. The *E. coli* and yeast arsenate reductases use reduced glutathione (GSH) as the direct electron donor for the reduction of arsenate, whereas the *S. aureus* enzyme employs thioredoxin. Ellis et al. [13] cloned the *P. vittata PvACR2* arsenate reductase gene from a cDNA library of arsenate-treated gametophytes. This enzyme shows 25% amino acid sequence identity / 47% similarity to the yeast ACR2 protein and likewise uses GSH as substrate. Like the yeast enzyme, the *P. vittata* ACR2 lacks phosphatase activity [13]. Arsenate reductase has been shown to be induced 7-fold by arsenate in the roots of *P. vittata* [14]. The reduction product of arsenate is accumulated as free arsenite in the fronds of

1   this fern probably in the vacuoles [3,13,15]. Arsenate reductases similar to the *P. vittata* ACR2

2   protein have been cloned from rice [16] and the arsenate-tolerant grass *Holcus lanatus* [4], and

3   orthologous sequences have been identified in the genomes of other plants and fungi that are not

4   arsenic hyperaccumulators [16], though all these enzymes also have phosphatase activity.

5   Phytochelatins are cysteine-rich peptides derived from GSH that can bind to arsenite and

6   several reactive heavy metal cations [17]. *A. thaliana* mutants unable to synthesize phytochelatins

7   show enhanced sensitivity to arsenic [18], and overproduction of these compounds has been shown

8   to increase arsenate resistance [4,19,20]. Phytochelatins have been detected in *P. vittata* and in

9   the related species *Pteris cretica*, but unlike in non-hyperaccumulator plants, only a small portion

10   of arsenic is complexed with phytochelatins in these hyperaccumulators. Phytochelatin production

11   in these arsenic hyperaccumulators is limited [17,21], suggesting that arsenic-phytochelatin

12   complexes do not form a major storage form for arsenic, although they still may play an important

13   role in hyperaccumulation, as indicated by the occurrence of highly localized thiolate-coordination

14   of arsenic in close proximity to the vein and mid-veins in the fronds of the *P. vittata* sporophyte

15   [17,21].

16   Two other proteins that could play a role in arsenic tolerance or hyerpaccumulation in *P. vittata*

17   have been identified by selection for clones from cDNA libraries that increased the arsenate

18   tolerance of an arsenate reductase mutant (*ΔarsC*) of *E. coli*. One gene obtained by this approach

19   was found to specify a triosephosphate isomerase [22]. In addition to its expected enzymatic

20   activity, it was suggested that this enzyme promotes the reduction of arsenate by an unknown

21   mechanism. The second enzyme that was identified in this manner was a glutaredoxin [23], which

22   was able to confer increased resistance to both arsenate and arsenite by a mechanism that needs to

23   be determined.

24   In order to identify other proteins that could contribute to the arsenic resistance of *P. vittata*,

25   we carried out a similar procedure to isolate clones from a *P. vittata* cDNA library that elevated

26   the arsenate resistance of a *ΔarsC* mutant *E. coli*. Nucleotide sequence analysis revealed that the

4

1 cDNA clones in these derivatives encode proteins that have high sequence similarity to

2 glutathione-S-transferases (GSTs).

Materials and Methods

Bacterial Strains and Growth Conditions

5 The *E. coli* K12 strain W3110 (IN[*rrnD-rrnE*]) and its arsenate reductase mutant derivative

6 WC3110 (*ΔarsC* IN[*rrnD-rrnE*]) [12] were obtained from B. Rosen (Wayne State University), the

7 γ-glutamyl-cysteine synthetase mutant JTG10 (*gshA20::*Tn*10*kan *thr-1 araC14 leuB6*[Am] *Δ(gpt-*

8 *proA)62 lacY1 tsx-33 glnV44 galK2 hisG4 rfbC1 mgl-5 kdgK51 xylA5 mtl-1 argE3 thi-1 λ⁻ rpsL31*)

9 [24] was obtained from the *E. coli* Genetic Stock Center, and the glutathione reductase mutant

10 BW31342 (*Δgor::cat*) and the arsenate reductase mutant BW39212 (*ΔarsC::kan*) were obtained

11 from B. Wanner (Purdue University). The arsenate reductase-glutathione synthetase double

12 mutant strain KC1680 was constructed by P1 transduction of the *gshA20*::Tn*10*kan from strain

13 JTG10 into WC3110. The arsenate reductase-glutathione reductase double mutant strain KC1852

14 was constructed by transduction of the *ΔarsC::kan* and *Δgor::cat* insertions elements BW31342

15 and BW39212 into WC3110. Strains KC1858 (*Δgor::cat ΔarsC::kan* IN[*rrnD-rrnE*] /

16 pTriplEx2), KC1860 (*ΔarsC::kan Δgor::cat* IN[*rrnD-rrnE*] / p*PvGST35*), and KC1862

17 (*ΔarsC::kan Δgor::cat* IN[*rrnD-rrnE*] / p*PvGST41*) were made by electroporating the indicated

18 plasmids into KC1852.

19 Bacterial cultures were grown with aeration at 37° in the rich medium LB [25] or K medium,

20 which has a phosphate concentration of 1 mM [26]. When used, ampicillin (Ap), chloramphenciol

21 (Cm), and kanamycin sulfate (Km) were at 100, 25, and 75 mg l^{-1}, respectively. Solid media

22 contained 20 g l^{-1} Difco agar. $Na_2AsO_4 \cdot 7H_2O$ and $NaAsO_2$ were purchased from J. T. Baker and

23 Sigma, respectively.

24 Arsenate resistance was assessed qualitatively on solid media by the radial streaking method

25 [27]: Whatman GF/A filter disks (21 mm diameter) were placed in the middle of K medium plates

1   containing 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG; added to induce the expression of

2   the cDNA genes from the *lac* promoter on the cloning vector), and 30 μl of 2 M Na_2HAsO_4

3   solution was placed on the disks and air-dried. Strains were streaked with sterile toothpicks

4   outward from the disk, and their As^R was evaluated by the size of the zone of growth-inhibition

5   next to the arsenate-containing filter disk compared to that seen with WC3110 (*ΔarsC*) carrying

6   the empty cloning vector pTriplEx2. For a more quantitative determination of arsenate and

7   arsenite resistance in liquid culture, we used the procedure of Mukhopadhyay et al. [12] with slight

8   modifications: Cultures of strains carrying the *P. vittata* GST clones were grown overnight in LB

9   + Ap and were subcultured at 1:100 dilution into K medium + 1 mM IPTG, plus different

10   concentrations of Na_2HAsO_4 or $NaAsO_2$ and grown with aeration. After 24.0 h of incubation at

11   37°, the cell density of the cultures was measured as light scattering in a Shimadzu UV-1700

12   spectrophotometer at 600 nm (OD_{600}), with appropriate dilutions to correct for multiple scattering

13   at $OD_{600} \geq 1.0$ (where $OD_{600} = 1$ corresponds to ~1.5 x 10^9 cells/ml). For each arsenate or arsenite

14   concentration tested, the measurements were performed with three independent cultures.

Pteris vittata cDNA library construction and screening

16   RNA was purified from *P. vittata* gametophytes grown for 6 weeks in liquid culture [28]

17   containing 1 mM KH_2AsO_4. A cDNA library was synthesized using BD Biosciences Clonetech

18   SMART® cDNA library Construction kit (User Manual PT3000-1 [PR15738]) and cloned between

19   the SfiIA and *Sfi*IB sites of plasmid pTriplEx2 (Clonetech). The ligation mixture was

20   electroporated into *E. coli* strain DH10B [29] and transformants selected on LB Ap plates. Plasmid

21   DNA was isolated from a pool of ~$7x10^6$ colonies and electroporated into WC3110 (*ΔarsC*). The

22   cells were allowed to recover in 1 ml LB medium for one hour at 37^oC with aeration. Aliquots of

23   100 μl of the transformation mixtures were inoculated on K medium plates containing 6mM

24   $Na_2AsO_4 \cdot 7H_2O$ and 1 mM IPTG. After two days, the plates were examined for the appearance of

25   As^R colonies. Spontaneous As^R mutants arose in the untransformed control cells at a frequency of

26   ~2 per 10^7 cells; the frequency of such derivatives was generally 2- to 10-fold higher in the cultures

1   transformed with the library. For plates that contained fewer than ~50 As^R transformed colonies,

2   the strains were single colony isolated on K medium plates containing 6 mM $Na_2AsO_4 \cdot 7H_2O$, and

3   then tested for Ap resistance (Ap^R) on LB plates, confirming the presence of the cloning vector (as

4   opposed to a spontaneous mutation to As^R by the host). Plates that contained more than 50 As^R

5   colonies were replica-plated to LB Ap plates, and colonies that were stably Ap^R were single colony

6   isolated on K medium plates containing 6 mM arsenate. Plasmids were isolated from the

7   transformants that were confirmed to be both As^R and Ap^R and were retransformed into WC3110,

8   selecting Ap^R. The derivatives obtained in the latter step were tested for As^R by radial streaking.

9   Plasmids were isolated from the As^R strains and the sequences of the cDNA inserts were

10   determined at Purdue Genomics Core Facility. The sequences were analyzed with the blastn and

11   blastp algorithms (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The sequences of the three *P. vittata*

12   ST cDNA clones p*PvGST35*, p*PvGST41*, and p*PvGST57* that we obtained were entered into

13   GenBank under accessions numbers EU259318, EU259319, and EU259320, respectively.

14   Speciation of arsenic in cell cultures.

15   Strains KC1749 (pTriplEx2), KC1761 (p*PvGST35*) and KC1763 (p*PGSTv41*) were grown LB

16   + Ap, 0.03 ml were inoculated into 3 ml of K medium containing 10 mM glucose + 0.1 mM

17   Na_2HAsO_4 + Ap + 1 mM IPTG and incubated with aeration for 24 h. At the end of the growth,

18   the cell density was determined by light scattering at 600 nm (OD_{600}). The cells from 1.5 ml

19   samples were sedimented by a microfuge, 0.5 ml of the supernatant culture media were mixed with

20   0.5 ml of a buffer containing 13.36 mM NH_4NO_3, 13.36 mM $NH_4H_2PO_4$, and 10 mM EDTA (pH

21   6.2) and flash-frozen in dry ice ethanol. EDTA was used in the extraction buffer as it has been

22   shown to preserve the inorganic arsenic species [30]. The As(III) and As(V) content of the samples

23   from 5 independent cultures for each strain were determined as described by the method of

24   Williams et al. [31].

25   **Results**

Isolation of cDNA clones from *P. vittata* that confer arsenate-resistance in *E. coli*

Plasmids p*PvGST35*, p*PvGST41*, and p*PvGST57* were isolated by introducing a *P. vittata* cDNA expression library into the *ΔarsC* (arsenate reductase) mutant *E. coli* WC3110 and selecting derivatives that could grow in K medium containing 6 mM Na_2AsO_4. The arsenate resistance (As^R) phenotype of these transformants is illustrated in Figure 1: The three plasmids carrying the *P. vittata* cDNA clones conferred increased As^R on this host strain compared to the control carrying the empty vector pTriplex2. The phenotype conferred by these three plasmids was evaluated in a more quantitative form in Figure 2, which shows that in liquid medium, the plasmids conferred very similar increases in As^R above that seen with the control strain carrying the empty cloning vector. WC3110 derivatives expressing the *P. vittata* GST clones had an exponential doubling time of 1.6 h in K medium +10 mM Na_2AsO_4, whereas the empty vector control had a doubling time of >12 h (data not shown).

The cDNA clones conferring arsenate resistance encode GST-family proteins

DNA sequence analysis revealed that the three plasmids conferring As^R contain a 217-codon open reading frame whose deduced protein product shows high similarity to GSTs from higher plants. Plasmids p*PvGST*35 and p*PvGST*57 are identical in their nucleotide sequences in the central protein-coding region (Figure 3); the gene contained in these two plasmids was designated *PVGST*1. These two plasmids exhibited slight sequence differences at the 5' and 3' ends of the inserts, which could have been due to the fact that they were from different mRNAs, or differences could have been introduced during cloning. In any case, the fact that the plasmids are different indicates that they are independent isolates. The insert in p*PGSTv*41 shows a 13% nucleotide sequence difference from the inserts in p*PvGST35* and *pPvGST57*, resulting in 23-amino acid difference in the predicted open reading frames (Figure 3). Consequently, the open reading frame in p*PvGST*41, which is designated *PVGST*2, specifies a GST that is different from *PVGST*1. The expression vector pTriplEx2 contains translation start signals for 5' portions of the *ompA* and *lacZ* genes in different reading frames, followed by a run of T residues that provide a site for

8

transcriptional frameshifting. Consequently, cloned inserts may be translated in all three reading frames as N-terminal fusion proteins with OmpA or LacZ peptides. DNA sequencing revealed that the inserts in p*PvGST35* and p*PvGST57* contain a TAG termination codon 24 bases upstream of the ATG initiation codon for the GST, indicating that these two gene products are not made as fusion proteins. However, there is no analogous in-frame termination codon upstream of the GST sequence in p*PvGST41*, and so the latter may be translated as a fusion product with an N-terminal peptide from LacZ or OpmA.

GSTs of higher plants have been assigned into four classes based on their amino acid sequences: phi, tau, theta, and zeta; representatives of the former two classes are also found in bacteria and members of the latter two classes are also present in animals [32]. Figure 4A shows a multiple alignment of the two *P. vittata* genes against four of the most closely related enzymes from other plant species and one representative of each of the four classes of GSTs from *A. thaliana*. This analysis indicates that the *P. vittata* enzymes cluster together and with other members of the phi-class of plant GSTs (Figure 4B). One of the diagnostic features of the phi class of GSTs is that their GSH-binding site contains the amino acid triplet ESR [32], which is present in the *P. vittata* clones (Figure 4A).

The *P. vittata* GSTs do not show significant similarity (blastp Expect value > 0.019) to the three types of arsenate reductases represented by the *E. coli*, *S. aureus*, and the yeast/*P. vittata* enzymes. Genes encoding highly similar phi class GST can be found in the *Myxococcales* bacteria *Sorangium cellulosum* 'So ce 56' and *Myxococcus xanthus* DK 1622 (blastp similarity to *PVGST*1 Expect value = 5e-43 and 6e-28, respectively) and in over 100 other bacteria with blastp Expect values between 1e-18 and 1e-13. The function of these bacterial enzymes has not yet been determined.

Nucleotide sequence features of the *P. vittata* GST clones

An interesting feature of the three *P. vittata* GST clones is that they contain 10 repeats of the dinucleotide GA immediately upstream and overlapping the ATG start (Figure 3). This motif is

9

1   not present in the most closely related plant GSTs genes, including the one from the moss

2   *Physcomitrella patens*. However, a cDNA glutaredoxin clone from *P. vittata* [23] also contains

3   12 GA repeats seven nucleotides upstream of the translation start site (see sequence under

4   GenBank accession number EF052272). Whether these repeats are associated with the expression

5   of arsenic tolerance-related genes in *P. vittata* or whether they have some other function in gene

6   expression in this fern remains to be explored. There are precedents for strings of GA repeats

7   upstream of genes in other eukaryotic systems. In *Drosophila melanogaster*, the so-called GAGA

8   elements have been recognized as the binding sites of a chromatin remodeling factor that increases

9   transcription from the *hsp70* promoter [33 and references therein]. A GAGA element has been

10   found between the TATA box and the transcription start sites of the *Gsa1* gene of *Glycine max*,

11   which has been suggested to be involved in the transcriptional control of the gene. Repeats of GA,

12   or its complement CT, have also been found downstream of viral promoters, and it has been

13   suggested that they play a positive role in transcriptional activation [34,35].

14   ## Phenotypic characterization of the *P. vittata* GST clones in *E. coli*

15   We tested whether the *P. vittata* GST clones can confer increased As^R in an arsenate reductase

16   proficient ($arsC^+$) host (Figure 5). Our results reproduce the finding of Mukhopadhyay et al. [12]

17   that loss of the *E. coli* arsenate reductase greatly increases sensitivity to arsenate (compare

18   $arsC^+$/empty vector with $\Delta arsC$/empty vector). Although the *P. vittata* clones substantially elevate

19   As^R in the $\Delta arsC$ mutant, they resulted only in a marginal, statistically insignificant increase in the

20   already high As^R of the $arsC^+$ strain. These results indicate that although the *P. vittata* GSTs can

21   functionally substitute for the *E. coli* arsenate reductase in the $\Delta arsC$ mutant, they do not provide

22   improved resistance against this toxic chemical above the level conferred by the endogenous

23   enzyme.

24   The sequence similarity of our clones to known GSTs suggests that the *P. vittata* enzymes

25   probably use glutathione as a substrate. To test this, we determined whether blocking glutathione

26   synthesis with a *gshA* (γ-glutamyl-cysteine synthetase) mutation would diminish the arsenate

10

1   resistance conferred by these plasmids. The *gshA ΔarsC* double mutant carrying the empty cloning

2   vector was considerably more sensitive to arsenate in the absence of GSH than the *ΔarsC gshA$^+$*

3   counterpart (Figure 6A). More importantly, the *gshA* mutation cancelled the ability of the *P. vittata*

4   clones to confer arsenate resistance in the absence of glutathione. Supplementation with GSH,

5   which can be taken up by *E. coli* [36], restored the ability of the clones to confer As^R in the *gshA*

6   mutant background (Figure 6B). These results suggest that the *P. vittata* GSTs use GSH as

7   substrate for the detoxification of arsenate.

8   Finally, we examined whether the *P. vittata* GSTs could confer resistance to arsenite and found

9   that that the clones were unable to do so either in the *ΔarsC* mutant (Figure 7) or in the *arsC$^+$*

10   background (data not shown), indicating that the *P. vittata* GSTs can detoxify arsenate but not

11   arsenite.

12   We carried out GST activity assays of the *E. coli* strains carrying the *P. vittata* clones using 1-

13   chloro-2,4-dinitrobenzene (CDNB) as substrate, which has been used to assay phi class GSTs from

14   plants [32,37]. However, we were not able to detect statistically significant GST activity in the *E.*

15   *coli* strains carrying the *P. vittata* clones (data not shown), indicating that the *P. vittata* GSTs are

16   unusual among the phi class enzymes in that they have no detectable activity with CDNB as

17   substrate.

18

1   The *P. vittata* GSTs promote the reduction of arsenate in *E. coli*

2   In order to probe the mechanism by which the *P. vittata* GST clones confer increased arsenate

3   resistance, we carried out two experiments to determine whether this phenotype involves the

4   reduction of this toxic compound. In one of these experiments, we tested the effect of a glutathione

5   reductase (*gor*) mutation on the ability of the *P. vittata* GSTs to impart increased arsenate

6   resistance. The result was that the ability of the GST plasmids p*PvGST35* and p*PvGST41* to confer

7   increased arsenate resistance was compromised by the *Δgor::cat* mutation (Figure 8). This result

8   suggests that arsenate detoxification involves the oxidation of GSH, and glutathione reductase is

9   required to regenerate this reduced sulfhydro compound.

10   We also tested whether the *P. vittata* GSTs increases the formation of As(III) from arsenate.

11   For this experiment, we grew derivatives of strain WC3110 (*ΔarsC*) carrying the empty cloning

12   vector or plasmids p*PvGST35* and p*PvGST41* in medium containing an initial 1 mM Na_2HAsO_4,

13   and after 24 h of growth, we quantified the proportions of arsenic as As(III) and As(V) in the

14   culture medium. The final biomass density of the cultures was $OD_{600} = 0.96 \pm 0.05$ for the strain

15   carrying the empty cloning vector, and $OD_{600} = 1.71 \pm 0.10$ and 1.82 ± 0.15 for the strains carrying

16   p*PvGST35* and p*PvGST41*, respectively, confirming the ability of the GST plasmids to confer

17   increased arsenate resistance. The cells were sedimented by centrifugation, and the amounts of

18   arsenic in the medium as As(III) and As(V) were determined as described in Materials and

19   Methods. The results are shown in Table 1. For strain KC1749, which carried the empty cloning

20   vector, approximately 30% of the input arsenic was converted to As(III), with the balance

21   remaining as As(V). It is not clear whether the reduction of arsenate in this culture was due to an

22   enzymatic process or to some spontaneous reaction with GSH or some other sulfhydro compound.

23   However, in the cultures of the strains carrying p*PvGST35* and p*PvGST41*, the proportion of

24   arsenic that was recovered as As(III) increased to approximately 60% and 69%. This result

25   indicates that the *P. vittata* GSTs promoted the formation of an As(III) species, suggesting that the

26   increased arsenate resistance conferred by these enzymes involves the reduction of arsenate.

12

Discussion

In this work, we isolated three independent clones from a *P. vittata* cDNA library that imparted increased arsenate resistance to an *arsC* mutant *E. coli*. Sequence analysis indicated that these clones encode proteins that have high sequence similarity to the phi family of higher plant GSTs. GSTs have diverse functions related to the detoxification of harmful substances [32]. Some have glutathione dependent peroxidase or isomerase activities [38,39]. Others are involved in detoxification of xenobiotics by glutathione conjugation. It has been suggested that some GSTs function as ligands or carrier proteins for a variety of molecules [40]. They have also been implicated in signal transduction pathways [41] and in responses to a wide range of stresses [32].

Our phenotypic analysis in a *ΔarsC* mutant E. *coli* indicated that the *P. vittata* GSTs require GSH for arsenate resistance. The ArsC, Acr2p, and PvACR2 arsenate reductases of *E. coli*, yeast, and *P. vittata* use GSH as electron donor and produce oxidized glutathione (GSSG) [12,13,42]. Arsenate reduction in *E. coli* and yeast requires the participation of a glutaredoxin, which regenerates GSH from GSSG. However, the human monomethylarsonic acid reductase, which is a member of the omega family of mammalian GSTs [43], can carry out the GSH-dependent reduction of methylarsonate to methylarsenite without the direct involvement of a glutaredoxin [44]. It is possible that the *P. vittata* GSTs may be able to catalyze the analogous reduction of arsenate to arsenite.

Another type of mammalian GSTs, which belong to the omega subfamily, can detoxify arsenite by forming an arsenite-$(GSH)_3$ complex that is excreted from the cells [45,46]. Our observation that the *P. vittata* GSTs do not provide increased resistance to arsenite suggests that they do not operate by this mechanism. It is possible that the *P. vittata* GSTs might facilitate the formation of an arsenate-GSH adduct that could be expelled from *E. coli*. However, the latter mechanism would necessitate the stoichiometric consumption of at least one molecule of GSH for one molecule of arsenate that is detoxified. In view of the fact that our GST cDNA clones could provide resistance to at least 10 mM arsenate in a medium that contains only 10 mM glucose as carbon source, we

13

1   feel that this mechanism is unlikely. The requirement for glutathione reductase also supports the

2   conclusion that GSH has a catalytic rather than a stoichiometric role.

3   Arsenate resistance in the hyperaccumulator *P. vittata* is thought to involve reduction of As(V)

4   to As(III) and transport into the vacuoles in fronds and gametophytes. An arsenate reductase has

5   been cloned from *P. vittata* [13], and it has been recently reported that a triosephosphate isomerase

6   [22] and glutaredoxin [23] from this organism can increase the arsenic resistance in *E. coli*. Our

7   observations that the *P. vittata* GSTs promote the GSH-dependent conversion of As(V) to As(III)

8   in the *arsC* mutant *E. coli* suggest that these proteins participate in an arsenate reduction pathway,

9   which also involves the bacterial glutathione reductase. However, we do not have any direct

10   evidence that the *P. vittata* GSTs by themselves have arsenate reductase activity and cannot rule

11   out other possible mechanisms, such as that these enzymes might detoxify arsenate by promoting

12   the formation of a GSH complex. It will be necessary to purify these enzymes from *E coli* and

13   establish whether they have one of the above enzymatic activities in vitro. Finally, although the

14   fact that the GSTs conferred increased arsenate tolerance in *E. coli* makes it attractive to

15   hypothesize that these proteins are likely to contribute to the arsenic tolerance or

16   hyperaccumulation in *P. vittata*, it will be necessary to confirm the roles of these enzymes in

17   arsenic metabolism in the ferns.

18   **Acknowledgments**

19   We thank A. Murphy for helpful discussions, and N. Csonka for help with the photography and

20   image processing.

21   **Author Contributions**

22   Conceived and guided the project: DES LNC. Isolated *P. Vittata* GST cDNA clones in *E. coli*

23   and carried out initial phenotypic characterization: AAK. Constructed the *P. vittata* cDNA

1   expression library: DRE. Carried out the arsenic speciation and related data analysis: GJN and

2   AAM. Performed the GST assays: XH. Wrote the manuscript: LNC. All authors read and

3   approved the final manuscript.

4

References

1. Ma LQ, Komar KM, Tu C, Zhang WH, Cai Y, et al. (2001) A fern that hyperaccumulates arsenic: a hardy, versatile, fast-growing plant helps to remove arsenic from contaminated soils. Nature 409: 579.

2. Visoottiviseth P, Francesconi K, Sridokchan W (2002) The potential of Thai indigenous plant species for the phytoremediation of arsenic contaminated land. Environ Pollut 118: 453-461.

3. Wang J, Zhao FJ, Meharg AA, Raab A, Feldmann J, et al. (2002) Mechanisms of arsenic hyperaccumulation in *Pteris vittata*. Uptake kinetics, interactions with phosphate, and arsenic speciation. Plant Physiol 130: 1552-1561.

4. Bleeker PM, Hakvoort HW, Bliek M, Souer E, Schat H (2006) Enhanced arsenate reduction by a CDC25-like tyrosine phosphatase explains increased phytochelatin accumulation in arsenate-tolerant *Holcus lanatus*. Plant J 45: 917-929.

5. Rosen BP (2002) Transport and detoxification systems for transition metals, heavy metals and metalloids in eukaryotic and prokaryotic microbes. Comp Biochem and Physiol A 133: 689-693.

6. Vahter M (2002) Mechanisms of arsenic biotransformation. Toxicology 181-182: 211-217.

7. Bertolero F, Pozzi G, Sabbioni E, Saffiotti U (1987) Cellular uptake and metabolic reduction of pentavalent to trivalent arsenic as determinants of cytotoxicity and morphological transformation. Carcinogenesis 8: 803-808.

8. Carter DE, Aposhian HV, Gandolfi AJ (2003) The metabolism of inorganic arsenic oxides, gallium arsenide, and arsine: a toxicochemical review. Toxicol Appl Pharm 193: 309-334.

9. Song WY, Park J, Mendoza-Cózatl DG, Suter-Grotemeyer M, Shim D, et al. (2010) Arsenic tolerance in *Arabidopsis* is mediated by two ABCC-type phytochelatin transporters. Proc Natl Acad Sci U S A 107: 21187-21192.

10. Oden KL, Gladysheva TB, Rosen BP (1994) Arsenate reduction mediated by the plasmid-encoded ArsC protein is coupled to glutathione. Molec Microbiol 12: 301-306.

16

11. Ji G, Silver S (1992) Reduction of arsenate to arsenite by the ArsC protein of the arsenic resistance operon of *Staphylococcus aureus* plasmid pI258. Proc Natl Acad Sci U S A 89: 9474-9478.

12. Mukhopadhyay R, Shi J, Rosen BP (2000) Purification and characterization of Acr2p, the *Saccharomyces cerevisiae* arsenate reductase. J Biol Chem 275: 21149 – 21157.

13. Ellis DR, Gumaelius L, Indriolo E, Pickering IJ, Banks JA, et al. (2006) A novel arsenate reductase from the arsenic hyperaccumulating fern *Pteris vittata*. Plant Physiol 141: 1544-1554.

14. Duan G-L, Zhu Y-G, Tong Y-P, Cai C, R K (2005) Characterization of arsenate reductase in the extract of roots and fronds of Chinese brake fern, an arsenic hyperaccumulator. Plant Physiol 138: 461-469.

15. Pickering IJ, Gumaelius L, Harris HH, Prince RC, Hirsch G, et al. (2006) Localizing the biochemical transformations of arsenate in a hyperaccumulating fern. Environ Sci Technol 40: 5010-5014.

16. Duan G-L, Zhou Y, Tong Y-P, Mukhopadhyay R, Rosen BP, et al. (2007) A CDC25 homologue from rice functions as an arsenate reductase. New Phytol 174: 311-321.

17. Raab A, Feldmann J, Meharg AA (2004) The nature of arsenic-phytochelatin complexes in *Holcus lanatus* and *Pteris cretica*. Plant Physiol 134: 1113-1122.

18. Ha SB, Smith AP, Howden R, Dietrich WM, Bugg S, et al. (1999) Phytochelatin synthase genes from Arabidopsis and the yeast *Schizosaccharomyces pombe*. Plant Cell 11: 1153-1164.

19. Li Y, Dankher OP, Carreira L, Smith AP, Meagher RB (2006) The shoot-specific expression of ◎-glutamylcysteine synthetase directs the long-distance transport of thiol-peptides to roots conferring tolerance to mercury and arsenic. Plant Physiol 141: 288-298.

20. Li Y, Dhankher OP, L. C, Lee D, Chen A, et al. (2004) Overexpression of phytochelatin synthase in *Arabidopsis* leads to enhanced arsenic tolerance and cadmium hypersensitivity. Plant Cell Physiol 45: 1787-1797.

17

21. Zhang W, Cai Y, Downum KR, Ma LQ (2004) Thiol synthesis and arsenic hyperaccumulation in *Pteris vittata* (Chinese brake fern). Environ Pollut 131: 337-345.

22. Rathinasabapathi B, Wu S, Sundaram S, Rivoal J, Srivastava M, et al. (2006) Arsenic resistance in *Pteris vittata* L.: identification of a cytosolic triosephosphate isomerase based on cDNA expression cloning in *Escherichia coli*. Plant Mol Biol 62: 845-857.

23. Sundaram S, Rathinasabapathi B, Ma LQ, Rosen BP (2008) An arsenate-activated glutaredoxin from the arsenic hyperaccumulator fern *Pteris vittata* L. regulates intracellular arsenite. J Biol Chem 283: 6095-6101.

24. Greenberg JT, Demple B (1986) Glutathione in *Escherichia coli* is dispensable for resistance to H_2O_2 and gamma radiation. J Bacteriol 168: 1026 – 1029.

25. Davis RW, Botstein D, Roth JR (1980) Advanced Bacterial Genetics. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

26. Druger-Liotta J, Prange VJ, Overdier DG, Csonka LN (1986) Selection of mutations that alter the osmotic control of transcription of the *Salmonella typhimurium proU* operon. J Bacteriol 169: 2449-2459.

27. Ratzkin B, Roth J (1978) Cluster of genes controlling proline degradation in Salmonella typhimurium. J Bacteriol 133: 737-754.

28. Gumaelius L, Lahner B, Salt DE, Banks JA (2004) Arsenic hyperaccumulation in gametophytes of *Pteris vittata*: a new model system for analysis of arsenic hyperaccumulation. Plant Physiol 136: 3198- 3208.

29. Grant SG, Jessee J, Bloom FR, Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia col*i methylation-restriction mutants. Proc Natl Acad Sci U S A 87: 4645-4649.

30. Bednar AJ, Garbarino JR, Ranville JF, Wildeman TR (2002) Preserving the distribution of inorganic arsenic species in groundwater and acid mine drainage samples. Environ Sci Technol 36: 2213-2218.

31. Williams PN, Price AH, Raab A, Hossain SA, Feldmann J, et al. (2005) Variation in arsenic speciation and concentration in paddy rice related to dietary exposure. Environ Sci Technol 139: 5531-5540.

32. Wagner U, Edwards R, P. DD, Mauch F (2002) Probing the diversity of the Arabidopsis glutathione S-transferase gene family. Plant Molec Biol 49: 515-532

33. Wang YV, Tang H, Gilmour DS (2005) Identification in vivo of different rate-limiting steps associated with transcriptional activators in the presence and absence of a GAGA element. Molecand Cell Biol 25: 3543-3552.

34. He X, Fütterer J, Hohn T (2002) Contribution of downstream promoter elements to transcriptional regulation of the rice tungro bacilliform virus promoter. Nucleic Acids Res 30: 497-506.

35. Pauli S, Rothnie HM, Chen G, He X, Hohn T (2004) The cauliflower mosaic virus 35S promoter extends into the transcribed region. J Virol 78: 12120-12128.

36. Parry J, Clark DP (2002) Identification of a CysB-regulated gene involved in glutathione transport in *Escherichia coli*. FEMS Microbiol Lett 209: 81-85.

37. Habig WH, Jacoby WB (1981) Assays for differentiation of glutathione-S-transferases. In: Jacoby WB, editor. Methods in Enzymology. New York: Academic Press, Inc. pp. 398-405.

38. Dixon DP, Cole DJ, Edwards R (2000) Characterisation of a zeta class glutathione transferase from *Arabidopsis thaliana* with a putative role in tyrosine catabolism. Arch Biochem Biophys 384: 407-412.

39. Nutricati E, Miceli A, Blando F, De Bellis L (2006) Characterization of two *Arabidopsis thaliana* glutathione S-transferases. Plant Cell Rep 15: 997-1005.

40. Mueller LA, Goodman, C.D., Silady RA, Walbot V (2000) AN9, a petunia glutathione-S-transferase required for anthocyanin sequestration, is a flavonoid-binding protein. Plant Physiol 123: 1561-1570.

19

41. Chen I-C, Huang I-C, Liu M-J, Wang Z-G, S.-S. C, et al. (2007) Glutathione S-transferase interacting with far-red insensitive 219 is involved in phytochrome A-mediated signaling in Arabidopsis. Plant Physioly 143: 1189-1202.

42. Shi J, Vlamis-Gardikas A, Aslund F, Holmgren A, Rosen BP (1999) Reactivity of glutaredoxins 1, 2, and 3 from *Escherichia coli* shows that glutaredoxin 2 is the primary hydrogen donor to ArsC-catalyzed arsenate reduction. J Biol Chem 274: 36039-36042.

43. Zakharyan RA, Sampayo-Reyes A, Healy SM, Tsaprailis G, Board PG, et al. (2001) Human monomethylarsonic acid (MMA^V) reductase is a member of the glutathione-S-transferase superfamily. Chem Res Toxicol 14: 1051-1057.

44. Zakharyan RA, Aposhian HV (1999) Enzymatic reduction of arsenic compounds in mammalian systems: the rate-limiting enzyme of rabbit liver arsenic biotransformation is MMA^V reductase. Chem Res Toxicol 12: 1278-1283.

45. 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. J Biol Chem 279: 32700-32708.

46. Liu J, Chen H, Miller DS, Saavedra JE, Keefer LK, et al. (2001) Overexpression of glutathione S-transferase II and multidrug resistance transport proteins is associated with acquired tolerance to inorganic arsenic. Mol Pharmacol 60: 302-309.

47. Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16: 10881-10890.

1   **FIGURE LEGENDS**

2   **Figure 1. Qualitative demonstration of the arsenate resistance conferred by the *P. vittata***

3   **GST clones in *E. coli*.** Derivatives of strain WC3110 (*ΔarsC*) carrying the *P. vittata* clones

4   *PVGST35*, *PVGST41*, and *Pv57* were tested for arsenate resistance on solid media by radial

5   streaking, as described in Materials and Methods. The filter disk placed on the Petri dish on left

6   was spotted with 30 of μl H_2O and the one on the right was spotted with 30 μl of 1.8 M Na_2HAsO_4.

7   **Figure 2. Arsenate resistance conferred by the *P. vittata* GST clones on *E. coli* in liquid**

8   **cultures.** Resistance to the indicated concentrations of Na_2HAsO_4 was determined in liquid

9   cultures, as described in Materials and Methods. Strains used were derivatives of WC3110

10   (*ΔarsC*) carrying pTriplEx2 (empty vector), pTriplEx2 containing *P. vittata* GST1 (p*PvGST35* and

11   p*PvGST57*), or *P. vittata* GST2 (p*PvGST41*). The results are the averages ± standard deviation

12   (indicated by error bars) of data obtained with 3 independent cultures of each strain.

13   **Figure 3. Sequences of the *P. vittata* GST cDNA clones.** The nucleotide and predicted amino

14   acid sequences of the *pPvGST35*, *pPvGST57*, and *pPvGST41* cDNA clones are shown with lower

15   case and upper case letters, respectively. Nucleotide sequences that are common to p*PvGST35*

16   and p*PvGST57* are indicated as "p*PvGST35/57*"; where the two clones differ, the nucleotide

17   sequences are offset above (for p*PvGST57*) and below (for p*PvGST35*) the common sequences.

18   The predicted amino acid sequences of PVGST1 and PVGST2 are shown by capital letters above

19   and below the nucleotide sequences for p*PvGST35/57* and p*PvGST41*, respectively. The

20   nucleotide and amino acid sequences that are common to p*PvGST35* and p*PvGST41* are shown in

21   **black** letters, and sequences that are different between the two clones are shown in **red**. The GA

22   repeats before the translation start sites are underlined.

1   **Figure 4. Sequence relationship of the *P. vittata* GSTs with orthologs from other plants.**

2   **Panel A:** Alignment of the amino acid sequences of *PVGST*1 and *PVGST*2 with the four plant

3   GSTs that showed the highest similarity to *PVGST*1 (blastp Expect values between 1e-60 and

4   2e-62), and with one representative of each of the four classes of GSTs from *A. thaliana*. Multiple

5   sequence   alignment   was   performed   with   the   MultAlin   algorithm

6   (http://multalin.toulouse.inra.fr/multalin) [47]. The amino acids that are in boxes are conserved in

7   most GSTs and have been shown to constitute part of the catalytic site, and the amino acids shown

8   in **bold** letters have been found to be conserved in all *A. thaliana* GSTs [32]. The underlined ESR

9   sequence is diagnostic of the phi-class of GSTs and constitutes part of the GSH-binding site [32].

10   The characters "*", ":", and "." below the comparisons indicate fully conserved, strongly conserved

11   and weakly conserved residues, respectively, as defined in the instructions for the CLUTALW

12   algorithm (http://align.genome.jp/).

13   The sequences used were [accession numbers in brackets]: *P. patens*, moss *Physcomitrella*

14   *patens subsp. patens* hypothetical protein GST N family [XP_001764210.1]; *P. sitchensis*, Sitka

15   spruce *Picea sitchensis* hypothetical protein GST C family [ABK23015.1]); *C. sinensis*, navel

16   orange *Citrus sinensis* cultivar Tarocco glutathione S-transferase N family [ABA42223]; *C.*

17   *argyrophylla*, conifer yin shan *Cathaya argyrophylla* phi class glutathione transferase

18   [ACL37469.1]; AtGSTF2, *A. thaliana* phi-class glutathione S-transferase F2 [NM_116486.2];

19   ATGSTT1, *A. thaliana* theta-class glutathione S-transferase T1 [NM_123486.3]; AtGSTZ1, *A.*

20   *thaliana* zeta-class glutathione S-transferase Z1 [NM_201671.1]; AtGSTU5, *A. thaliana* tau-class

21   glutathione S-transferase U5 [NM_128499.3].

22   **Panel B:** The dendogram was generated from the multiple alignment in Panel with the MultAlin

23   algorithm.

24   **Figure 5. Arsenate resistance is not enhanced by *P. vittata* GST cDNAs in wild-type *E. coli*.**

25   Resistance was determined in liquid cultures as described in Materials and Methods in the arsenate

26   reductase proficient ($arsC^+$) W3110 strain carrying pTriplEx2 (empty vector), pTriplEx2

1   containing *P. vittata* GST1 (p*PvGST35*), or *P. vittata* GST2 (p*PvGST41*). Also included are data

2   for WC3110 (*ΔarsC*) carrying the empty vector pTriplEx2. The results are the averages ± standard

3   deviation (indicated by error bars) of data obtained with 3 independent cultures of each strain.

4   **Figure 6. Glutathione is required to confer arsenate resistance by the *P. vittata* GST cDNAs.**

5   **Panel A:** Arsenate resistance was determined in a *gshA20 ΔarsC* double mutant in liquid cultures

6   in the absence of reduced glutathione, as described in Materials and Methods. The strains carried

7   pTriplEx2 (empty vector), p*PvGST35*, and p*PvGST41*. Also included is a control with strain

8   WC3110 (*gshA⁺ ΔarsC*) carrying pTriplEx2 (empty vector). The results are the averages ±

9   standard deviation (indicated by error bars) of data obtained with 3 independent cultures of each

10   strain. **Panel B**: Arsenate resistance of the *gshA20 ΔarsC* strains carrying the indicated plasmids

11   was determined on solid media, as described in the Materials and Methods. The Petri dish on the

12   left contained K medium and the one on the right contained K medium augmented with 1 mM

13   GSH, and the filter disks on both plates were spotted with 30 μl of 1.8 M Na_2AsO_4.

14   **Figure 7. The *P. vittata* GST clones do not increase resistance to arsenite.** Resistance to the

15   indicated concentrations of $NaAsO_2$ was determined in liquid cultures, as described in Materials

16   and Methods. Strains tested were derivatives of WC3110 (*ΔarsC*) carrying pTriplEx2 (empty

17   vector), pTriplEx2 containing *P. vittata* GST1 (p*PvGST35*), or *P. vittata* GST2 (p*PvGST41*). The

18   results are the averages ± standard deviation (indicated by error bars) of data obtained with 3

19   independent cultures of each strain.

20   **Figure 8. Glutathione reductase is required for the ability of the *P. vittata* GST clones to**

21   **confer arsenate resistance.** Resistance to the indicated concentrations of Na_2HAsO_4 was

22   determined in liquid cultures, as described in Materials and Methods. Strains tested were

23   derivatives of KC1852 (*ΔarsC::kan Δgor::cat*) carrying pTriplEx2 (empty vector), pTriplEx2

24   containing *P. vittata* GST1 (p*PvGST35*), or *P. vittata* GST2 (p*PvGST41*). The results are the

1   averages ± standard deviation (indicated by error bars) of data obtained with 3 independent cultures

2   of each strain.

3

1   **Table 1**. Conversion of As(V) to As(III) by strains synthesizing the *P. vittata* GSTs

Strain	As(III) (μmol / ml culture)	As(V) (μmol / ml culture)	As(III) % of total As
KC1749 (empty vector control)	0.24 ± 0.04	0.55 ± 0.09	30.3 ± 0.5
KC1761 (p*PvGST35*)	0.51 ± 0.08	0.35 ± 0.08	59.6 ± 3.3
KC1763 (p*PvGST41*)	0.57 ± 0.07	0.26 ± 0.04	68.7 ± 0.7

2

3   The strains were grown in medium containing 1.0 μmol / ml Na_2HAsO_4 and the speciation

4 of arsenic as As(III) and As(V) was determined after 24 h of growth, as described in Materials and

5 Methods. Results are averages ± standard deviation of data obtained with 5 independent cultures
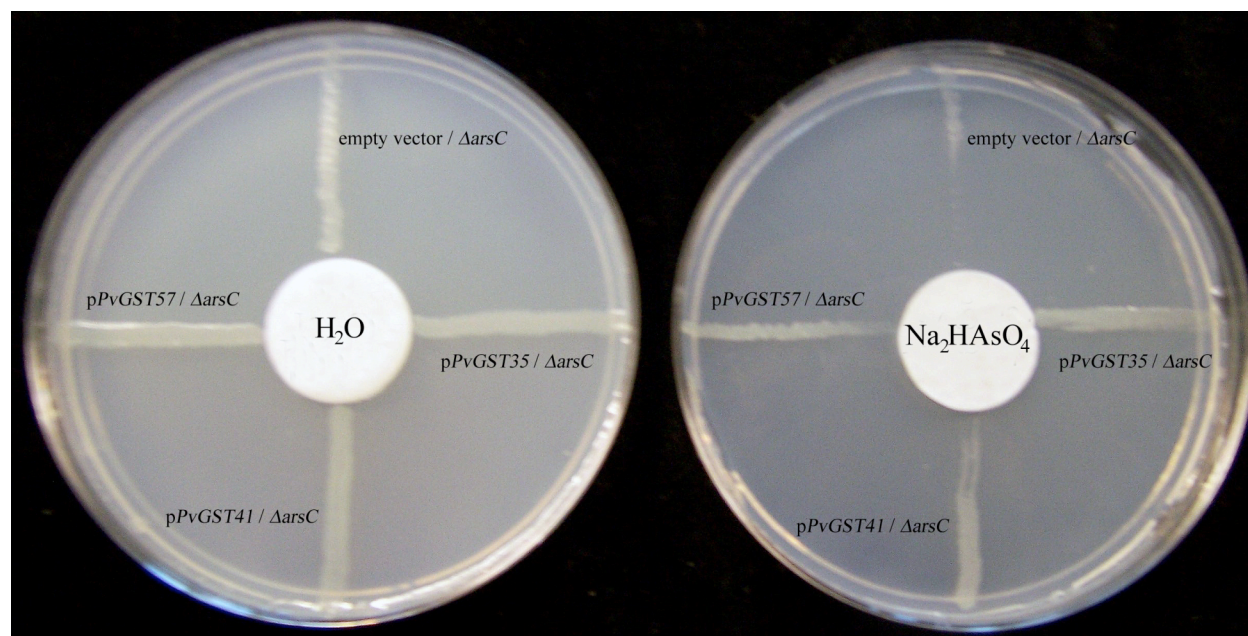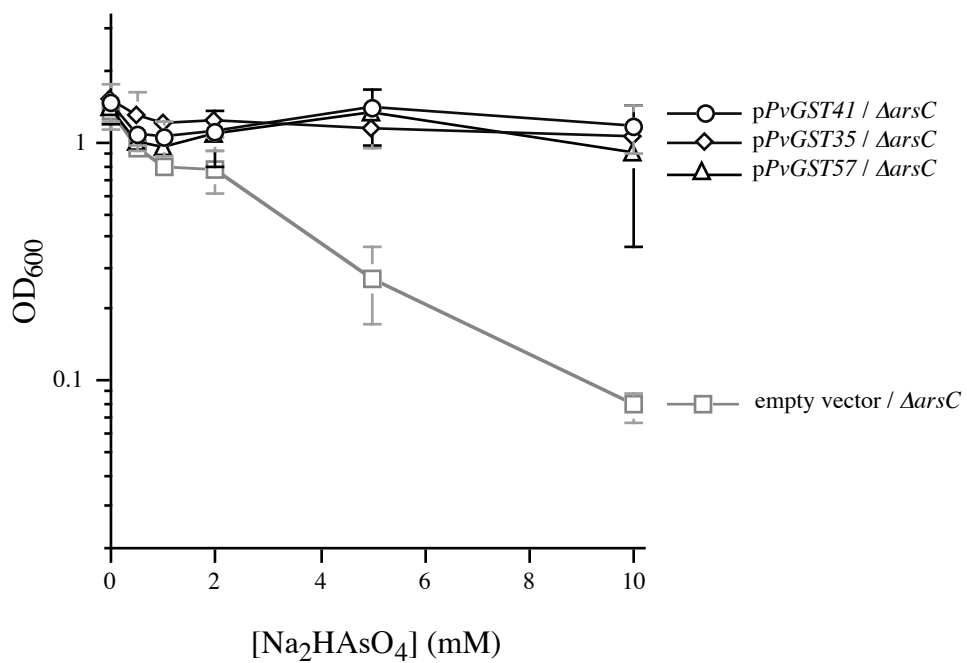
6 of each strain.

25

Figure 1.

Figure 2.

Figure 3

```
pPvGSH57     ------------------------------------g\
pPvGST35/57                                      gcgggccaatagagt-agactct---gagagag
pPvGST35     ------------------------------gacat/
pPvGST41     ggacgtgcagaaacacaagcgcgcgcgcacgcacacgcacacgcacgcactcacactgtagcgagagag

PvGST1                   M   T   I   T   V   H   G   Q   A   T   S   S   C   T   S   R   V   L   T
pPvGST35/57  agagagagagagagatgacgatcacagtgcacgggcaggcaacctccagctgcacctcacgcgtgcttacc
pPvGST41     agagagagagagagatgacgatcacagtacacgggcaggcaacctccagctgcacccaacgcgtgcttacc
PvGST2                   M   T   I   T   V   H   G   Q   A   T   S   S   C   T   Q   R   V   L   T

PvGST1         T   L   F   E   K   D   V   S   D   F   Q   L   L   H   V   D   L   S   T   G   A   H   K
pPvGST35/57  accctcttcgagaaggacgtctctgatttccagcttctgcacgtcgatctctccaccggcgcccacaag
pPvGST41     accctcttcgagaaggacgtctctgatttccagcttctgcacgtcgatctctctactggcgctcacaag
PvGST2         T   L   F   E   K   D   V   S   D   F   Q   L   L   H   V   D   L   S   T   G   A   H   K

PvGST1         Q   P   D   Y   L   A   L   Q   P   F   G   V   I   P   V   V   Q   D   G   D   L   T   I
pPvGST35/57  cagcccgactacctcgccctccagcctttggcgtcataccagttgttcaagatggagaccttaccatc
pPvGST41     cagcctgaatatctcgccctgcagccctttggcgtcataccagtcgttcaagatggggaccttaccatc
PvGST2         Q   P   E   Y   L   A   L   Q   P   F   G   V   I   P   V   V   Q   D   G   D   L   T   I

PvGST1         F   E   S   R   A   I   I   R   Y   L   A   L   K   Y   E   D   Q   G   S   P   L   Y   G
pPvGST35/57  ttcgaatcaagagctataataagatatcttgctctcaagtatgaggaccaaggtagcccattgtatggg
pPvGST41     ttcgaatcaagagctataataagatacttggctctcaaatatgaggaccaaggaagcccattgtatggg
PvGST2         F   E   S   R   A   I   I   R   Y   L   A   L   K   Y   E   D   Q   G   S   P   L   Y   G

PvGST1         R   T   L   E   E   R   A   L   V   E   Q   W   L   E   V   E   S   Q   N   F   H   V   A
pPvGST35/57  agaacattagaggaaagggctctagtggagcagtggctagaagtggaatcacaaaacttccacgtggca
pPvGST41     agaacattagaggaaagggctaatgtggagcaatggttggaagtagagtcgcagaatttccacgtggca
PvGST2         R   T   L   E   E   R   A   N   V   E   Q   W   L   E   V   E   S   Q   N   F   H   V   A

PvGST1         A   S   A   L   V   Y   Q   L   S   S   R   A   K   K   G   L   P   P   D   A   D   V   V
pPvGST35/57  gcatcagcattagtctaccagctatcatcacgagctaaaaaaggattgcctcctgatgcagatgtggtt
pPvGST41     gcatcggcaatcgtcaaccaactatcatctcgtgcaaaaaaagggctcccccctgaagagggtgtggtt
PvGST2         A   S   A   I   V   N   Q   L   S   S   R   A   K   K   G   L   P   P   E   E   G   V   V

PvGST1         K   A   N   I   E   K   L   E   A   T   L   D   I   Y   E   K   R   L   S   K   S   S   F
pPvGST35/57  aaggcaaacattgaaaagcttgaagccacccttagatatatatgagaagcgcttgtcaaaaaagctccttc
pPvGST41     aagacaaacatggaaaagcttgaggccatcttagatgtgtatgagaagcgcttgtccgagagctccttc
PvGST2         K   T   N   M   E   K   L   E   A   I   L   D   V   Y   E   K   R   L   S   E   S   S   F

PvGST1         L   A   G   D   F   F   S   L   A   D   L   S   H   L   P   R   T   K   S   L   M   K   S
pPvGST35/57  ttagcagggggatttctttagcttggcagacctctctcaccttcctcgcacaaagtcgctcatgaagtct
pPvGST41     ttagcaggagacttctttagcttggctgacctctcccaccteccctcgtaccaagtccettgtaaagtct
PvGST2         L   A   G   D   F   F   S   L   A   D   L   S   H   L   P   R   T   K   S   L   V   K   S

PvGST1         A   N   L   T   H   L   I   T   A   R   E   H   V   H   A   W   W   Q   R   I   S   T   R
pPvGST3/575  gccaacctgactcacctaattaccgctcgcgagcatgtccatgcatggtggcaacgtatctccactcgc
pPvGST41     tgcaaattgcctcaccttattacctctcgcaagcacgtccatgaatggtggcaacgcatctctactcgt
PvGST2         C   K   L   P   H   L   I   T   S   R   K   H   V   H   E   W   W   Q   R   I   S   T   R

PvGST1         P   S   W   V   K   V   T   E   M   A   T   P   S   K   X
pPvGST35/57  ccttcatgggggttaaggtcacagaaatggctactccttcaaaatgattttgagacctcatgttgagtgaaa
pPvGST41     ccttcttggatcaaggtcattgaaatggctgccccttcaaaatgatgtaagacctcacattt-gcaaaa
PvGST2         P   S   W   I   K   V   I   E   M   A   A   P   S   K   X

pPvGSH57                             /tc\     /tgtc\ /tgg\    /ggtgtttgtagttt
pPvGST35/57  cattatagaataataaaagaaattatcagtc    aaaa       a       aaa
pPvGSH35                                  \aa/    \aaaa/ \aaa/   \aaaaaaaaaaaaa
pPvGST41     catggtggcata-cacaactcctcaaaagtc gc aaaa aaaa a aaa aaa aaaaaaaaaaaaaaa

pPvGSH57     tttttttttttaataccagaacagggcatgaattgtat
```

```
PvGST1       ------MTITVHGQATSSCTSRVLTTLFEKDVSDFQLLHVDLSTGAHKQ-------PDYLALQPFG-V
PvGST2       ------MTITVHGQATSSCTQRVLTTLFEKDVSDFQLLHVDLSTGAHKQ-------PEYLALQPFG-V
V. vinifera  ------MVVKVYG-PDFASAKRVLVCLIEKEV-EFETLPIDIIKGQNKD-------PEFLKLQPFG-V
G. max       -----------YG-PTYGSPKRVLVCLIEKEI-EFETVHVDLFKGENKE-------PEFLKLQPFG-S
C. sinensis  ------MVVKVYGSVKAACPQRVLACLLEKGV-EFEIVQVDLDEGEHKR-------PEFLLRQPFG-Q
P. hybrida   ------MVVKVHGSAMAACPQRVMVCLIELGV-DFELIHVDLDSLEQKK-------PEFLVLQPFG-Q
AtGSTF2      -----MAGIKVFGHPASIATRRVLIALHEKNL-DFELVHVELKDGEHKK-------EPFLSRNPFG-Q
AtGSTT1      -----MMKLKVYADRMSQ-PSRAVIIFCKVNGIQFDEVLISLAKRQQLS-------PEFKDINPLG-K
AtGSTZ1      MANSGEEKLKLYSYWRSSCAHRVRIALALKGL-DYEYIPVNLLKGDQFDSVYRFDLQDFKKINPMG-T
AtGSTU5      --MAEKEEVKLLGIWASPFSRRVEMALKLKGI-PYEYVEEILENKS--------PLLLALNPIHKK
                             .    .  *.    :          ::  :     :              :*:

PvGST1       IPVVQDGDLTIFESRAIIRYLALKYEDQGSPLYG---RTLEERALVEQWLEV ESQNFHVAASALVYQL
PvGST2       IPVVQDGDLTIFESRAIIRYLALKYEDQGSPLYG---RTLEERANVEQWLEV ESQNFHVAASAIVNQL
V. vinifera  VPVIQDGDYTLFESRAIMRYYAEKYKSQGTDLLG---KTIEERGVVEQWLEV EAQSYHPAIDNLVIEI
G. max       LPVIQDGDYTLYESRAIIRYLAEKYKDQGTDLLG---KTIEEKGLVEQWLEV EAHNFHPPLYNLVINV
C. sinensis  VPVIEDGDFKLFESRAIIRYYAAKYANQGPNLLG---NTLEEKALVDQWLEV EAHNFNDLAFNLVLQL
P. hybrida   VPVIEDGDFRLFESRAIIRYYAAKYEVKGSKLTG---TTLEEKALVDQWLEV ESNNYNDLVYNMVLQL
AtGSTF2      VPAFEDGDLKLFESRAITQYIAHRYENQGTNLLQTDSKNISQYAIMAIGMQV EDHQFDPVASKLAFEQ
AtGSTT1      VPAIVDGRLKLFESHAILIYLSSAFPSVADHWYP---NDLSKRAKIHSVLDW HHTNLRRGAAGYVLNS
AtGSTZ1      VPALVDGDVVINDSFAIIMYLDEKYPE--PPLLP---RDLHKRAVNYQAMSI VLSGIQPHQNLAVIRY
AtGSTU5      VPVLVHNGKTILESHVILEYIDETWPQ--NPILP---QDPYERSKARFFAKL VDEQIMNVGFISMARA
               :*..  ..     :  :*  .*    *       :              :   .       .                    .

PvGST1       SSRAKKGLPPDADVVKANIEKLEATLDIYEKRLSKS---SFLAGDFFSLADLSHLPRTKSLMKSAN--
PvGST2       SSRAKKGLPPEEGVVKTNMEKLEAILDVYEKRLSES---SFLAGDFFSLADLSHLPRTKSLVKSCK--
V. vinifera  LFGRKRGIPPDAKVIEESEKKLAKVLDIYEERLSKS---KYLAGDFFSLADLSHLPFTKYLADMG---
G. max       LFAPLTGAPSDQKVIEESDKKIEKVLDVYEERLSKS---KYLAGDFFSLADLSHLPFGHYLVNQTG--
C. sinensis  VILPRMGQRSDTALVHNLEQKLEAVLNIYEQRLSKS---NYLAGDSFTLADLSHLPALRYLMNEAG--
P. hybrida   LVFPKMGQTSDLTLVTKCANKLENVFDIYEQRLSKS---KYLAGEFFSLADLSHLPSLRFLMNEGG--
AtGSTF2      IFKSIYGLTTDEAVVAEEEAKLAKVLDVYEARLKEF---KYLAGETFTLTDLHHIPAIQYLLGTP---
AtGSTT1      VLGPALGLPLNPKAAAEAEQLLTKSLSTLETFWLKGNAKFLLGSNQPSIADLSLVCELMQLQVLDDKD
AtGSTZ1      IEEKINVEEKTAWVNNAITKGFTALEKLLVNCAGKH-----ATGDEIYLADLFLAPQIHGAINRFQIN
AtGSTU5      DEKGREVLAEQVRELIMYLEKELVGKDYFGGKTVGFLDFVAGSLIPFCLERGWEGIGLEVITEEKFPE
                          .                                   :

PvGST1       LTHLITAREHVHAWWQ--RISTRPSWVKVTEMATPSK--------------------
PvGST2       LPHLITSRKHVHEWWQ--RISTRPSWIKVIEMAAPSK--------------------
V. vinifera  KMYLIEERKHVKAWWD--DISNRPSWKKVFSSRWPLLE--------------------
G. max       RGNLVRDRKHVSAWWD--DISNRPAWQKVLQLYKYPV--------------------
C. sinensis  MAHLVTQRKHVNAWWD--KISSRPAWKKLASLAH-----------------------
P. hybrida   FSHLVTKRKCLHEWYL--DISSRDSWKKVLDLMMKKISEIEAVSIPAKEEAKV-----
AtGSTF2      TKKLFTERPRVNEWVA--EITKRPASEKVQ---------------------------
AtGSTT1      RLRLLSTHKKVEQWIENTKKATMPHFDETHEILFKVKEGFQKRREMGTLSKPGLQSKI
AtGSTZ1      MEPYPTLAKCYESYNELPAFQNALPEKQPDAPSSTI--------------------
AtGSTU5      FKRWVRNLEKVEIVKDCVPPREEHVEHMNYMAERVRSS--------------------
```

Figure 4A.
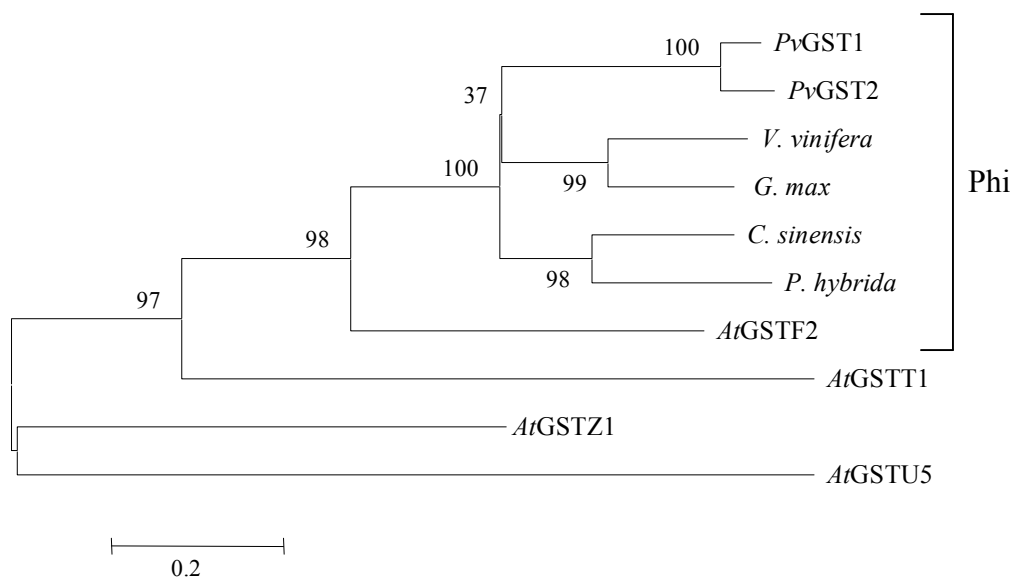
Figure 4 B.
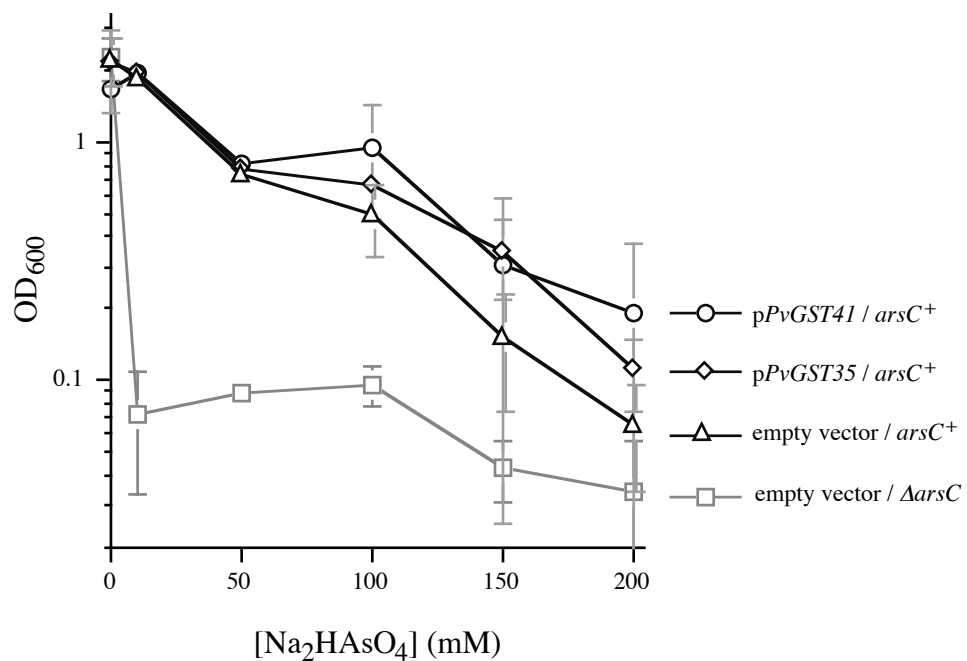
Figure 5.
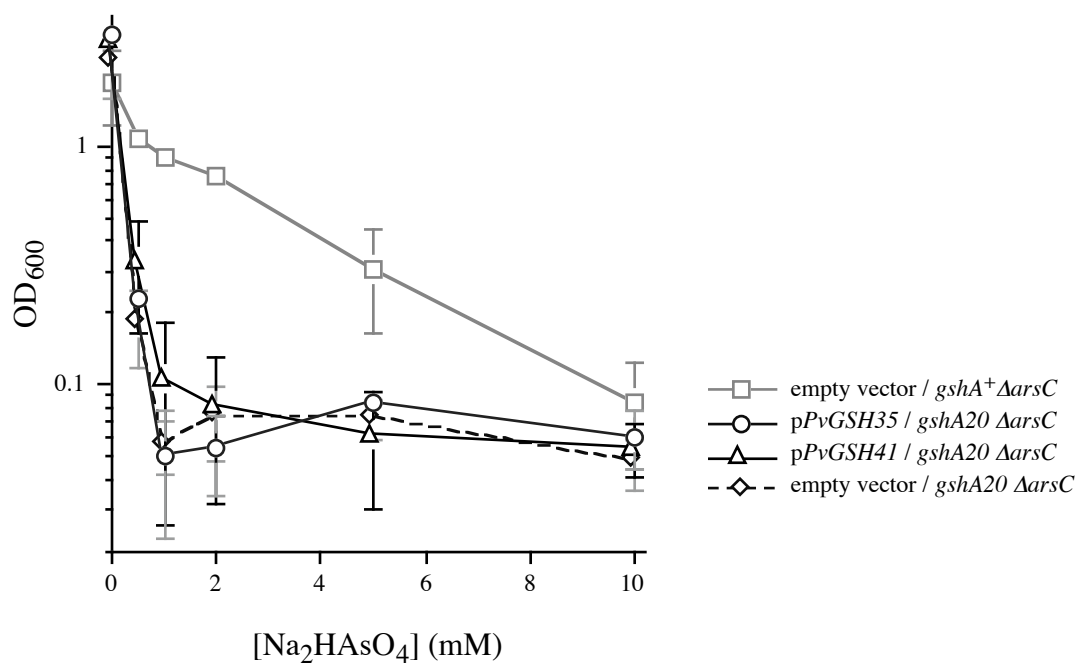
Figure 6A.
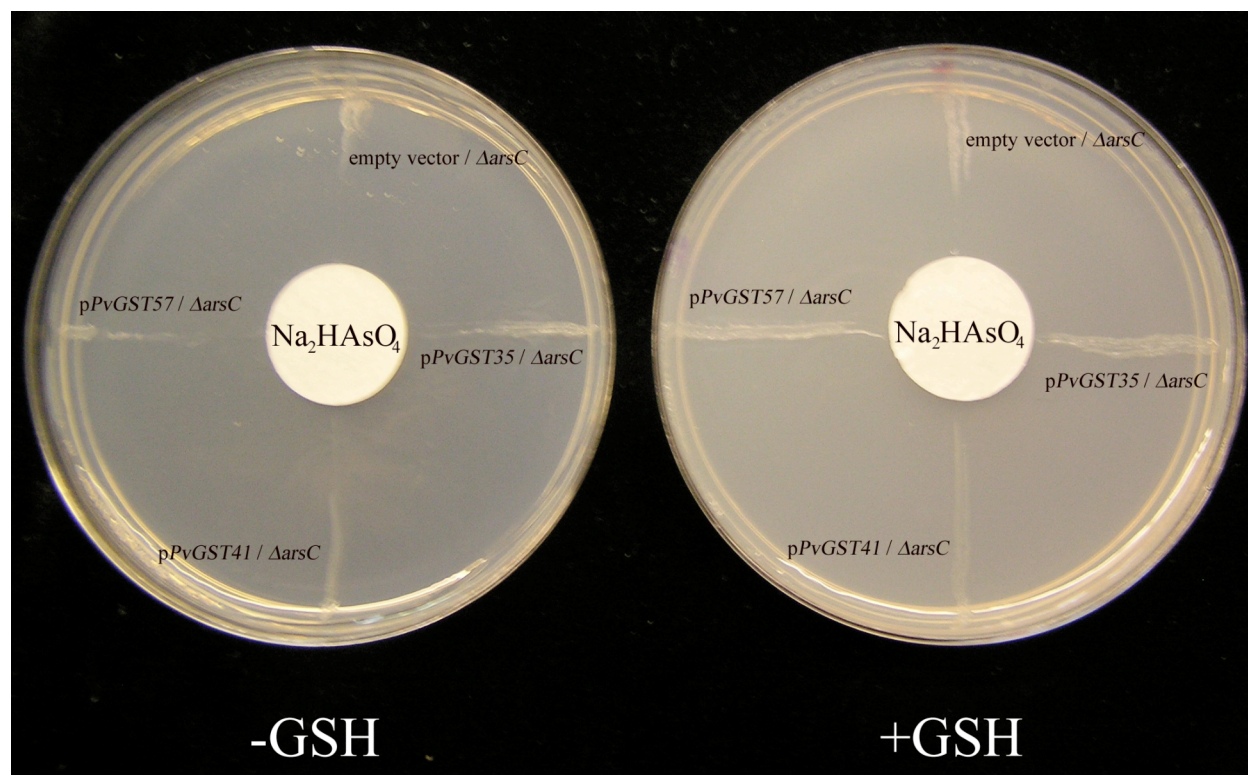
Figure 6B.

Figure 7.

Figure 8.