Appl Microbiol Biotechnol (2009) 83:715–725
DOI 10.1007/s00253-009-1929-4

Novel gene clusters involved in arsenite oxidation and resistance in two arsenite oxidizers: *Achromobacter* sp. SY8 and *Pseudomonas* sp. TS44

Lin Cai · Christopher Rensing · Xiangyang Li · Gejiao Wang

Received: 6 January 2009 / Revised: 21 February 2009 / Accepted: 23 February 2009 / Published online: 13 March 2009

Abstract This study describes three gene clusters involved in arsenic redox transformation of two arsenite oxidizers: *Achromobacter* sp. SY8 and *Pseudomonas* sp. TS44. A 17.5-kb sequence containing the arsenite oxidase (*aox*) gene cluster (*aoxX-aoxS-aoxR* and *aoxA-aoxB-aoxC-aoxD*) was isolated from SY8 using a fosmid library approach. Similarly, a 14.6-kb sequence including the *aox* cluster (*arsD-arsA-aoxA-aoxB*) and the arsenic resistance (*ars*) gene cluster (*arsC1-arsR-arsC2-ACR3-arsH*-dual specificity phosphatase (DSP)-glyceraldehyde-3-phosphate dehydrogenase (GAPDH)-major facilitator superfamily (MFS)) was obtained from TS44 by inverse polymerase chain reaction (PCR). According to reverse transcription (RT) PCR experiments, SY8 *aoxXSR* and *aoxABCD* transcribed as two different transcripts in opposite directions, and TS44 *aox* and *ars* clusters transcribed as a single transcript in their respective cluster. All of these genes were found to be upregulated by the addition of arsenite [As(III)], arsenate [As(V)], and antimonite [Sb(III)], except that TS44 *arsC1-arsR* appeared to be expressed constitutively. The SY8 *aox* cluster was predicted to be regulated by a two-component signal transduction system and a potential regulatory model was proposed. The TS44 *aox* cluster is unusual since it contains structural genes only and *arsDA* in its upstream. The TS44 *ars* cluster includes several genes previously identified not associated with arsenic resistance or transformation. This study showed novel structures and arrangements of arsenic gene clusters associated with bacterial As(III) oxidation and As(V) reduction.

Keywords Arsenic · Arsenite oxidizer · *aox* cluster · *ars* cluster

L. Cai · X. Li · G. Wang (✉)
State Key Laboratory of Agricultural Microbiology,
College of Life Science and Technology,
Huazhong Agricultural University,
Wuhan 430070, China
e-mail: gejiaow@yahoo.com.cn

C. Rensing
Department of Soil, Water and Environmental Science,
The University of Arizona,
Tucson, AZ 85721, USA

Introduction

Arsenic is recognized as one of the most toxic oxyanion in the natural environment, which caused severe contamination of soil–water systems and subsequent endemic arsenicosis in many countries, especially in Bangladesh, India, and China. Many microorganisms have evolved different arsenic detoxification pathways to cope with the widespread distribution of the poisonous arsenic (Rosen 2002). Four distinct microbial arsenic resistance mechanisms have previously been described: (1) As(III) oxidation, (2) cytoplasmic As(V) reduction and As(III) extrusion, (3) respiratory As(V) reduction, and (4) As(III) methylation (Qin et al. 2006; Silver and Phung 2005). These mechanisms confer arsenic resistance in microorganisms that play an important role in the transformation and geological cycle of arsenic.

An ever increasing number of As(III)-oxidizing bacteria have been detected and studied including *Alcaligenes faecalis* (Philips and Taylor 1976), *Agrobacterium tumefaciens* 5A (Kashyap et al. 2006), *Thiomonas* sp. 3As (Duquesne et al. 2008), *Herminiimonas arsenicoxydans* ULPAs1 (Muller et al. 2007), and *Thermus* sp. HR13 (Gihring et al. 2001), etc. These strains can oxidize As(III) to the less toxic As(V) by As(III) oxidase. It is noteworthy to point out that two types of arsenite oxidizers exist in the environment. One type is able to use As(III) as the sole

electron donor and grow as a chemolithoautotroph (Duquesne et al. 2008; Santini and vanden Hoven 2004). The other one is chemoorganoheterotrophic and employs As(III) oxidation only as a detoxification mechanism.

In *A. tumefaciens* 5A, the As(III) oxidase structural genes *aoxA* and *aoxB* encodes the small Rieske subunit and the large molybdopterin subunit, respectively. Expression of *aoxA* and *aoxB* is under control of the upstream regulatory genes *aoxS* and *aoxR* which encode histidine kinase (HK) and response regulator (RR) of the two-component signal transduction system, respectively. In addition to *aoxA* and *aoxB*, the downstream genes *aoxC* and *aoxD* encode cytochrome c and an enzyme involved in molybdopterin biosynthesis, respectively. These genes are cotranscribed as a single operon (Kashyap et al. 2006). However, in *H. arsenicoxydans* ULPAs1, the upstream regulatory genes *aoxS* and *aoxR* are transcribed in the reverse direction together with an additional gene whose function has not been identified (termed *aoxX* here; Muller et al. 2007). Research on *A. tumefaciens* 5A suggested that As(III) oxidation was regulated by a two-component signal transduction system and quorum sensing was also involved (Kashyap et al. 2006). However, the detailed regulatory mechanism of As(III) oxidation is still unclear.

Unlike the *aox* operon, the *ars* operon is well studied and the regulatory mechanism is better understood. A typical *ars* operon contains either three (*arsRBC*) or five (*arsRDABC*) genes that generally transcribe as a single unit (Rosen 1999). ArsR is a repressor that binds the promoter region and regulates the *ars* operon. ArsB is a membrane-located transport protein that can pump As(III) out of cells using the proton-motive-force. ArsC was shown to be a cytoplasmic As(V) reductase. ArsA is an As(III)-activated ATPase (Zhou et al. 2000). Furthermore, ArsA and ArsB can form an ArsA/ArsB complex that functions as a detoxification pump thereby strongly enhancing As(III) efflux ability (Dey and Rosen 1995). ArsD regulates the *ars* operon as a secondary and weak repressor (Chen and Rosen 1997; Wu and Rosen 1993). Recent research indicates that ArsD can function as an As(III) metallochaperone that transfers As(III) to ArsA and increases its affinity to ArsA (Lin et al. 2006). Microbial genome sequencing showed that another As(III) transport protein (termed Acr3p) was much more widespread than previously anticipated. Current genomic survey suggests that *ACR3* is predominant over *arsB* (Achour et al. 2007; Cai et al. 2009).

Several groups have reported that the *aox* and *ars* operon can be induced by both As(III) and Sb(III) (Kashyap et al. 2006; Lehr et al. 2007; Sato and Kobayashi 1998). Interestingly, the phenomenon of As(III) oxidation and As (V) reduction may happen simultaneously in 5A since disruption of the *aox* operon not only caused the loss of oxidizing ability but As(V) reduction could also be readily detected (Kashyap et al. 2006). This implies that activities of As(III) oxidation and As(V) reduction in 5A may have certain connections.

The impacts of microbial arsenic redox transformation have been reported to influence the geocycle of environmental arsenic (Mukhopadhyay et al. 2002; Oremland et al. 2004). A better understanding of arsenic species and their detoxification mechanisms provides the basis of arsenic bioremediation. The main purpose of this study was to identify three gene clusters responsible for As(III) oxidation and As(V) reduction in two novel arsenite oxidizers.

Materials and methods

Bacterial strains and media

The bacterial strains and plasmids used in this study were listed in Table 1. *E. coli* strains were grown in Luria–Bertani (LB) medium (Sambrook and Russell 2001). As (III)-oxidizing bacteria SY8 and TS44 were cultured in LB medium or chemically defined medium (CDM; Weeger et al. 1999) as required. The working concentrations of chloramphenicol, ampicillin, IPTG, and X-Gal were 12.5, 100, 100, and 50 μg ml^{-1}, respectively.

Table 1 Bacterial strains and plasmids used in this study

Strain or plasmid	Description	Source
Strains		
Achromobacter sp. SY8	Wild-type, As(III)-oxidizing phenotype	(Cai et al. 2009)
Pseudomonas sp. TS44	Wild-type, As(III)-oxidizing phenotype	(Cai et al. 2009)
Escherichia coli EPI300™	Fosmid cloning host	EPICENTER
E. coli JM109	Subcloning or TA cloning host	Promega
Plasmids		
pCC1FOS™	Cm^r, Fosmid cloning vector	EPICENTER
pGEM-4Z	Ap^r, subcloning vector	Promega
pGEM-T	Ap^r, PCR TA cloning vector	Promega

As(III)-oxidizing efficiency analysis

As(V) can react with molybdate to form a complex and then be reduced by ascorbic acid to produce blue color under conditions of certain acidity and temperature while As(III) cannot under the same conditions. The blue complex has an absorbance peak at 838 nm and can be measured by a colorimetric method (Lenoble et al. 2003; Zhou 1990).

A single colony was picked and inoculated in 100 ml CDM broth with 800 μM $NaAsO_2$ and then shaken at 160 rpm at 28°C. A total of 2 ml culture was taken each hour and used to determine cell OD values and concentrations of As(V) by a BECKMAN DU800 UV/Vis spectrophotometer; 1.5 ml sample was used for the cell OD_{600} determination using pure CDM as control. The remaining 0.5 ml sample was centrifuged, and 0.3 ml of the resulting supernatant was added to a mixture of 4 ml ddH_2O, 0.4 ml 50% H_2SO_4 (*w/v*), 0.4 ml of 3% Na_3MoO_4 (*w/v*), 0.2 ml 2% ascorbic acid (*w/v*), and kept immersed in a 90°C water bath for 20 min. The samples were then cooled to room temperature, and ddH_2O was added to a final volume of 10 ml. A_{838} was determined using the described mixture without 0.3 ml supernatant as control. A standard curve correlating A_{838} and concentrations of As (V) was made beforehand, and A_{838} could be converted into concentrations of As(V).

Isolation of arsenic gene clusters in SY8 and TS44

A genomic library was constructed and used to isolate *aox* cluster and nearby sequence of SY8 by the following steps: (1) constructing a fosmid library using the CopyControl™ Fosmid Library Production Kit according to the manufacturer's instruction (http://www.EpiBio.com); (2) screening positive clones using PCR by primers screenF and screenR since *aoxB* fragment (~530 bp) was obtained as the probe by degenerate PCR (Inskeep et al. 2007); (3) digesting the positive clones with *Eco*RI, *Bam*HI, *Pst*I, and *Hin*dIII and subcloning into pGEM-4Z using an appropriate restriction endonuclease; (4) sequencing the subclones and assembling; (5) analyzing and, if necessary, designing primers according to the gained sequence and performing PCR paired with the fosmid vector primer (pCC1FOS F or pCC1FOS R) for further amplification and sequencing to obtain the whole *aox* cluster; (6) designing primers for PCR using total genomic DNA as template and sequencing a second time to confirm the validity.

TS44 *aox* and *ars* clusters were isolated using the inverse PCR method. The following steps were employed: (1) designing inverse primers based on the known *aoxB* fragment (~530 bp, tUp1 and tDn1) and *ACR3* fragment (~750 bp, tUp3 and tDn3) obtained from another degenerate PCR (Achour et al. 2007); (2) preparing pure genomic DNA and determining its concentration (NanoDrop ND-1000 Spectrophotometer); (3) digesting the DNA (~5 μg in 100 μl volume), respectively, with *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I, *Nco*I, *Pae*I, *Pst*I, *Sac*I, *Sal*I, and *Xba*I overnight; (4) detecting the digestion products by running agarose gel; (5) purifying the fully digested DNA by phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation (add 50 μl 3M NaAc (pH7.0), 350 μl ddH_2O, and 500 μl phenol/chloroform/isoamyl alcohol (25/24/1, *v/v/v*); mix horizontally and centrifuge for 5 min at top speed; aspirate the supernatant and amend 2.5 volumes of ethanol; mix gently and centrifuge for 15 min at top speed; wash the precipitated DNA twice with 70% cold ethanol; air-dry and dissolve the DNA in 20 μl ddH_2O; measure the DNA concentration); (6) self-ligating the digested DNA (~500 ng in 500 μl) at 16°C overnight; (7) purifying the self-ligated DNA as described in step 5; (8) performing long distance inverse PCR in a volume of 1 μl self-ligated DNA, 1 ng μl^{-1} of each primer, 200 μM dNTPs, 5 μl 10× PCR buffer (Mg^{2+} plus), 2.5 U TaKaRa LA Taq DNA polymerase and ddH_2O to 50 μl (initial 5 min denaturation at 94°C; 30 cycles of 1 min at 94°C, 1 min at 55°C, 5 min at 72°C; and final 5 min extension at 72°C); (9) purifying the products and TA cloning for sequencing; (10) analyzing and, if necessary, doing an additional inverse PCR to get the whole sequence; (11) resequencing as SY8 did to assure the validity.

Gene annotation and bioinformatic analyses

ContigExpress soft was used for sequence assembly. Gene annotation was performed by NCBI ORF Finder and BlastX (http://www.ncbi.nlm.nih.gov/). Protein families and domains were analyzed by Pfam (http://pfam.sanger.ac.uk/) and PROSITE (http://www.expasy.org/prosite/). Prediction of transmembrane helices in proteins was performed by TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/). MyDomains-Image Creator was employed to generate custom domain figures (http://www.expasy.org/tools/mydomains/). Sequence alignment was analyzed by ClustalW (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_clustalw.html).

Expression analyses of arsenic gene clusters

RT-PCR was used to assess cotranscription and induction of the arsenic gene clusters identified in this study. Primers and sequences used to analyze transcription of the gene clusters are shown in Fig. 2 and listed in Table 2. In SY8, primer pairs of sP1 and sP2, sP3 and sP4, sP5 and sP6, and sP7 and sP8 were designed to monitor the transcription of *aoxR-aoxS* (415 bp), *aoxS-aoxX* (213 bp), *aoxA-aoxB* (146 bp), and *aoxB-aoxC-aoxD* (660 bp), respectively. In

Table 2 Primers used in this study

Primer	Sequence
screenF	5′-TAAATGGCCCGAGCAAAG-3′
screenR	5′-CCGTGGTCGAAACAGGAG-3′
sUp1	5′-CGCTCACTTTCTTCCGCAAAG-3′
sUp2	5′-TGAACTCGGCGTGTCTCAAG-3′
sUp3	5′-CGAGGACCCCAATACCAGT-3′
pCC1FOS F	5′-GGATGTGCTGCAAGGCGATTAAGTTGG-3′
pCC1FOS R	5′-CTCGTATGTTGTGTGGAATTGTGAGC-3′
6kbF	5′-ACCGGCAATACTTCCTCT-3′
3kbR	5′-GTCGGGCATTCGTCGTAG-3′
sP1	5′-AATGTCGCTGATCACCGC-3′
sP2	5′-AGCTAACGGTGGAAGTCG-3′
sP3	5′-GTGTGTCTCGAAGCATGC-3′
sP4	5′-ATGTACGACGACATCCGC-3′
sP5	5′-TGAGGGACTGATCTACGG-3′
sP6	5′-TACCCACAGCCTACGATG-3′
sP7	5′-TTCGACGCACGATCAGCT-3′
sP8	5′-TCTGGGCATGCAGTACGT-3′
tUp1	5′-CCCTTATTTACCTCGCACTCC-3′
tDn1	5′-ATTCTGTTCAATCTTTTCGACC-3′
tUp2	5′-GGACGATGAGCCCTGCTT-3′
tDn2	5′-AGAGCCCCGTTCAACTGG-3′
tUp3	5′-AGACCTGTCATGTACTCCG-3′
tDn3	5′-TCACCATCGTCGCGATGTT-3′
tP1	5′-AATGCTGTTCGGGACGTG-3′
tP2	5′-TTCGATGCTGGGTCAGTG-3′
tP3	5′-TAGGCCTGAATCGTCTGC-3′
tP4	5′-ATCTGCCTCAGATCGTGC-3′
tP5	5′-TGGTGATCTGCATTGGGG-3′
tP6	5′-TTCGAAGACCCAGTCAGC-3′
tP7	5′-GGTGAGTTCGATGATGCG-3′
tP8	5′-AAGGATGCTTCGGCAGCT-3′
tP9	5′-GCATAGCTGCACTTGAGG-3′
tP10	5′-ATCGAGATTCGGGAGCAG-3′
tP11	5′-TGATCGCTATTCCGAGCG-3′
tP12	5′-AGGATCGTTGATCCGGAC-3′
tP13	5′-GTCTCCATCGATTACCGC-3′
tP14	5′-ACTTCCACAGACAGGGTC-3′

TS44, primer pairs of tP1 and tP2, tP4 and tP5, tP3 and tP5, tP6 and tP7, tP6 and tP8, tP9 and tP10, tP11 and tP12, and tP13 and tP14 were used to detect the transcription of *arsD-arsA* (187 bp), *aoxA-aoxB* (312 bp), *arsA-aoxA-aoxB* (1,140 bp), *arsC1-arsR* (379 bp), *arsC1-arsR-arsC2* (557 bp), *arsC2-ACR3-arsH* (1,119 bp), *arsH*-DSP-GAPDH (710 bp), and GAPDH-MFS (209 bp), respectively.

To detect genes expression, SY8 and TS44 cells were cultured at 28 °C in LB medium containing 200 μM As(III) until logarithmic phase. To monitor genes induction, SY8 and TS44 were inoculated in LB medium in PA bottles and grown at 28 °C for about a week, then induced with 200 μM As(III), 200 μM As(V), 200 μM Sb(III), and no addition as control for 3 h, respectively.

Total RNA was isolated using the Invitrogen Trizol reagent as described in the manufacturer's instruction (http://www.invitrogen.com) treated with DNase I and then purified again using the Trizol reagent. First strand cDNA synthesis used TaKaRa PrimeScript™ 1st Strand cDNA Synthesis Kit following the manufacturer's instructions (http://www.takara.com.cn), followed by cDNA amplification by PCR using the corresponding primer pairs and the first strand cDNA as template (initial 5 min denaturation at 94 °C; 32 cycles of 1 min at 94 °C, 1 min at 55 °C, 0.5–1.5 min at 72 °C; and final 5 min extension at 72 °C).

Deposit of strains and nucleotide sequences

Achromobacter sp. SY8 and *Pseudomonas* sp. TS44 were deposited in China Center for Type Culture Collection (CCTCC, http://www.cctcc.org/). Their accession numbers are: M207048 for SY8 and AB209010 for TS44. The two nucleotide sequences isolated in this study are available in the NCBI GenBank database (EF523515 for SY8 *aox* cluster, EU311944 for *aox* and *ars* clusters of TS44).

Results

Identification of As(III)-oxidizing bacteria SY8 and TS44

SY8 and TS44 (identified as *Achromobacter* and *Pseudomonas*) were isolated from soils with intermediate and high levels of arsenic contamination, respectively (Cai et al. 2009). The MICs of As(III) were 13 mM (SY8) and 23 mM (TS44). Both SY8 and TS44 could not gain energy from As (III) oxidation; rather, they appeared to function as a detoxification mechanism (data not shown). Both of the two strains were Gram-negative bacteria based on Gram stain and 16S rDNA identification.

As(III) oxidation and cell growth curves of SY8 and TS44 were monitored by spectrophotometry (Fig. 1a, b). The average oxidizing velocity between SY8 (52.9 μM h^{-1}) and TS44 (59.1 μM h^{-1}) was almost equivalent, but the efficiency of SY8 (721.1 μM h^{-1} OD^{-1}) was much higher than TS44 (172.6 μM h^{-1} OD^{-1}) since SY8 was present at very low cell density during the oxidizing process. According to the curves, As(III) oxidation in TS44 did not occur until the OD_{600} reached 0.2 whereas in SY8 had already been finished at an OD_{600} of about 0.1. Hence, As (III) oxidation in TS44 depended on certain cell density, but in SY8, it did not (Fig. 1a, b).

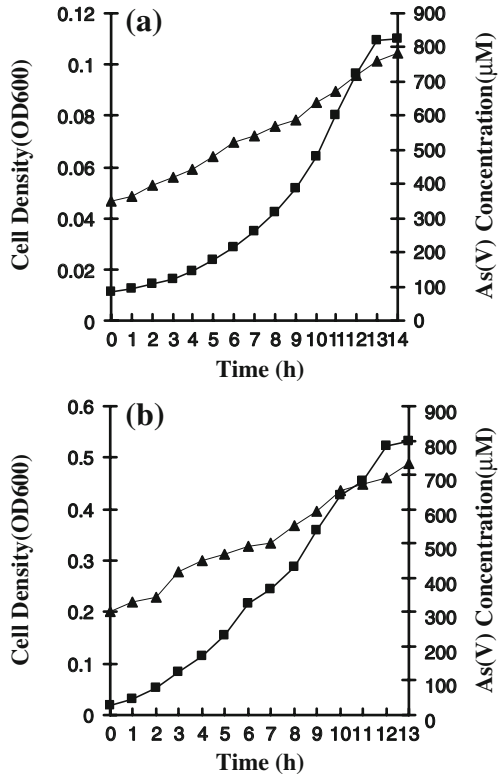

Fig. 1 As(III) oxidation and growth curve of SY8 (**a**) and TS44 (**b**). *filled triangle*, cell density (OD_{600}); *filled square*, As(V) concentration (μM)

Assembly of arsenic gene clusters in SY8 and TS44

A fosmid library with about 10^5 clones and an average size of 40-kb insert was successfully constructed for SY8. Seven positive clones containing *aoxB* were generated after screening 1,248 clones. Restriction maps were made for the positive clones and two fragments (~3.7, ~6.6 kb) digested by *Hin*dIII were chosen for further subcloning and sequencing. Primers 6kbF and 3kbR were designed to ascertain the ligation and relative position between the two fragments (Fig. 2a). Subsequent BLAST analyses showed that *aox* cluster upstream genes were not contained in the initial two fragments. Based on the known downstream sequence, three primers sUp1, sUp2, and sUp3 were designed, respectively (Fig. 2a). Pairing with the fosmid vector primer (pCC1FOS F or pCC1FOS R), ~2.3, ~3.1, and ~1.8 kb fragments were amplified and sequenced, respectively. Finally, a total of 17.5-kb sequence (GenBank, EF523515) was assembled together (Fig. 2a).

An inverse PCR method was used for functional sequence isolation in TS44. Two fragments of ~4.9 kb (*Pae*I) and ~5.4 kb (*Sal*I) were amplified using primers tUp1 and tDn1, and an extended ~1.7-kb (*Nco*I) fragment was obtained using primers tUp2 and tDn2 (Fig. 2b). Using primers tUp3 and tDn3, a ~6.0-kb (*Sac*I) fragment was amplified (Fig. 2b). For further assembly, we found *aox* cluster and *ars* cluster overlapped, and finally, a total of 14.6-kb sequence (GenBank, EU311944) was obtained (Fig. 2b).

Gene annotation and bioinformatic analyses

Two functional sequences were analyzed by NCBI ORF finder and BlastX, and precise physical maps were drawn (Fig. 2a, b). Annotation and analyses for each gene were listed in Table 3. Genes shared in SY8 and TS44 whose deduced amino acids displayed 75%/87% (*marR*), 64%/79% (*trxB*), 55%/69% (*aoxA*), 64%/79% (*aoxB*), and 21%/34% (*arsR*) in identity/similarity. Although the deduced ArsRs displayed low identity and similarity, both of them were analyzed to have the ArsR-type HTH domain.

Comparative physical maps of *aox* clusters and *ars* clusters were made for SY8, TS44, and other identified species using GenBank available data (Fig. 2c, d). SY8 contained a typical *aox* cluster whose regulatory *aoxXSR* cluster structure was identical to *A. faecalis* NCIB 8687 and *H. arsenicoxydans* ULPAs1. However, the downstream structural gene cluster *aoxABCD* had the same arrangement as in *A. tumefaciens* 5A (Fig. 2c). The TS44 *aox* cluster was atypical for lack of upstream regulatory genes and downstream-related genes in the immediate vicinity of the cluster but *arsDA* located in the upstream of the structural genes. The upstream structure of TS44 *ars* cluster (*arsC1-arsR-arsC2-ACR3-arsH*) was similar to *P. putida* W619, and downstream gene composition and order (*arsH*-DSP-GAPDH-MFS) was identical to *P. mendocina* ymp (Fig. 2d). However, it is important to note that no *aoxAB* homologs were found in both genomes of *P. putida* W619 (GenBank, NC_010501) and *P. mendocina* ymp (GenBank, NC_009439) after analysis of their complete genomic data.

As predicted by gene annotation, the SY8 *aox* cluster is regulated by a two-component signal transduction system including histidine kinase sensor AoxS and response regulator AoxR. Multiple sequence alignment of AoxS and AoxR homologues revealed H, N, G1, G2 boxes of histidine kinase and the phosphorylation site of response regulator (Fig. 3a; Stock et al. 2000). Bioinformatic analyses based on the protein primary structure, domains and conserved regions were performed and conformed to the typical two-component systems well (Fig. 4a). One potential signal transduction pathway between AoxS and AoxR was proposed (Fig. 4a).

Evidence of the arsenic gene clusters

According to multiple sequence alignment of As(III) oxidase, the twin arginine in TAT (twin arginine translocation)

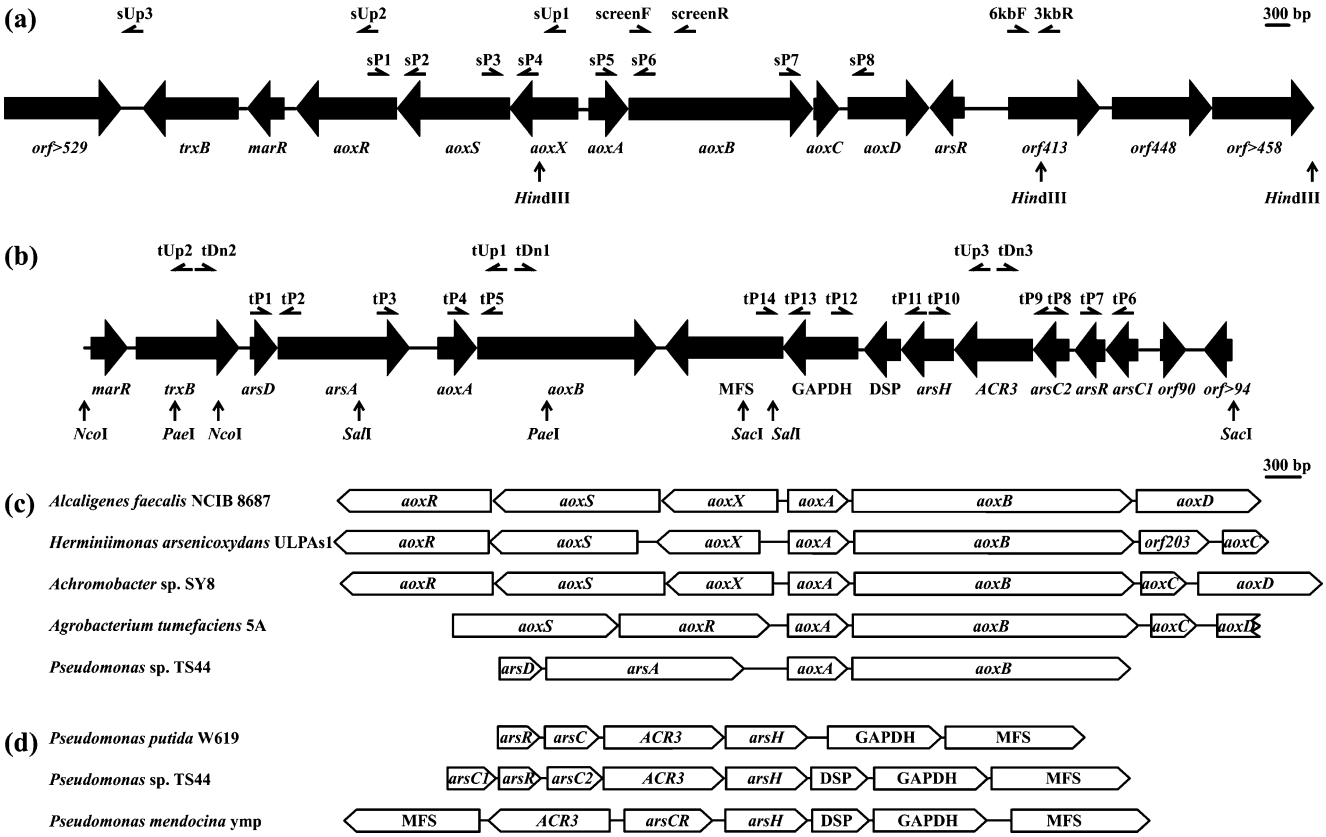

Fig. 2 Physical maps of arsenic gene clusters of SY8 (**a**; EF523515) and TS44 (**b**; EU311944). Comparative physical maps of *aox* clusters (**c**) and *ars* clusters (**d**). Primers *sP1* to *sP8*, *tP1* to *tP14* were used for transcriptional study, and the others were involved in sequence isolation. The amplification directions of all primers were shown by *horizontal arrows*. Important restriction endonuclease sites related to (sub)cloning were indicated by *vertical arrows*. Representative gene clusters (accession number *in parentheses*) are from: *A. faecalis* NCIB 8687 (AY297781); *H. arsenicoxydans* ULPAs1 (NC_009138); *A. tumefaciens* 5A (DQ151549); *P. putida* W619 (NC_010501); and *P. mendocina* ymp (NC_009439)

signal peptide and also the [2Fe–2S] cluster binding motif (C-X-H-X_{15}-C-X_2-H) of AoxA and the [3Fe–4S] cluster binding motif (C-X_2-C-X_3-C-X_{70}-S) as well as the As(III) binding residues (His, Glu, Arg, and His) of AoxB were displayed well (Ellis et al. 2001; Silver and Phung 2005; Fig. 3b). These domains and residues showed high evolutionary conservation with the five widely studied arsenite oxidizers. Furthermore, comparison of *aox* clusters demonstrated that the genes and their arrangements were highly similar with the well identified *aox* clusters (Fig. 2c).

Transcriptional analyses identified four distinct arsenic operons. Gene coexpression in the *aox* and *ars* clusters was detected using RT-PCR by primers end to end method. In SY8, *aoxXSR* and *aoxABCD* were cotranscribed in opposite directions and belonged to two operons (Fig. 5a). In TS44, *arsD-arsA-aoxA-aoxB* and *arsC1-arsR-arsC2-ACR3-arsH*-DSP-GAPDH-MFS were cotranscribed respectively and belonged to two distinct operons (Fig. 5b). Induction of gene expression was monitored in these arsenic gene clusters (Fig. 6) and indicated that all genes could be upregulated by the addition of As(III), As(V), and Sb(III) except that TS44 *arsC1-arsR* appeared to be expressed constitutively (Fig. 6b).

Discussion

SY8 displayed a much higher As(III) oxidizing efficiency than TS44 during the oxidation. There are a number of possibilities to explain this phenomenon; however, our data suggest the structure and arrangement of the respective *aox* clusters may have an influence on the gene expression level. The SY8 *aox* cluster includes a two-component signal transduction system for sensing As(III) and regulating expression in high efficiency, while the TS44 *aox* cluster contains the structural genes only.

Two-component signal transduction systems exist in a wide array of species and regulate a multitude of microbial operons efficiently (Stock et al. 2000). SY8 contains a typical *aox* cluster which includes a two-component regulatory gene cluster (*aoxXSR*) in one direction and a structural gene cluster (*aoxABCD*) in the opposite direction.

Table 3 Gene analysis of SY8 and TS44 based on BlastX

Gene	Length	Putative protein	Identity (%)	Similarity (%)	Reference species (accession number)
SY8					
trxB	425 aa	Thioredoxin reductase	56	73	*Alcanivorax borkumensis* SK2 (YP_693563)
marR	163 aa	Transcriptional regulator, MarR family	69	78	*Alcanivorax borkumensis* SK2 (YP_693562)
aoxR	449 aa	Two-component signal transduction regulator	58	76	*Alcaligenes faecalis* NCIB8687 (AAQ19842)
aoxS	501 aa	Two-component histidine kinase sensor	57	73	*Alcaligenes faecalis* NCIB8687 (AAQ19841)
aoxX	312 aa	Oxyanion binding protein	56	71	*Alcaligenes faecalis* NCIB8687 (AAQ19840)
aoxA	176 aa	Arsenite oxidase Rieske subunit	68	80	*Alcaligenes faecalis* NCIB8687 (AAQ19839)
aoxB	827 aa	Arsenite oxidase Mo-pterin subunit	82	90	*Alcaligenes faecalis* NCIB8687 (AAQ19838)
aoxC	109 aa	Cytochrome c, class I	53	67	*Ralstonia solanacearum* GMI1000 (NP_521379)
aoxD	364 aa	Molybdenum cofactor biosynthesis protein A	57	72	*Alcaligenes faecalis* NCIB8687 (AAQ19837)
arsR	151 aa	Arsenical resistance operon repressor	51	65	*Nitrobacter hamburgensis* X14 (YP_571841)
TS44					
marR	160 aa	Transcriptional regulator, MarR family	65	77	*Alcanivorax borkumensis* SK2 (YP_693562)
trxB	458 aa	Thioredoxin reductase	53	68	*Alcanivorax borkumensis* SK2 (YP_693563)
arsD	120 aa	Arsenical resistance operon repressor	71	83	*Pseudomonas stutzeri* A1501 (YP_001170787)
arsA	585 aa	Arsenite-activated ATPase	77	84	*Pseudomonas stutzeri* A1501 (YP_001170786)
aoxA	173 aa	Arsenite oxidase Rieske subunit	54	71	*Alcaligenes faecalis* NCIB8687 (AAQ19839)
aoxB	825 aa	Arsenite oxidase Mo-pterin subunit	64	80	*Alcaligenes faecalis* NCIB8687 (AAQ19838)
MFS	408 aa	Permease of the major facilitator superfamily	89	95	*Pseudomonas mendocina* ymp (YP_001187019)
GAPDH	334 aa	Glyceraldehyde-3-phosphate dehydrogenase	90	95	*Pseudomonas mendocina* ymp (YP_001187018)
DSP	163 aa	Dual specificity protein phosphatase	73	80	*Pseudomonas mendocina* ymp (YP_001187017)
arsH	238 aa	Arsenical resistance protein	86	92	*Pseudomonas mendocina* ymp (YP_001187016)
ACR3	353 aa	Arsenite transport protein	88	94	*Pseudomonas putida* W619 (YP_001751993)
arsC2	160 aa	Arsenate reductase	73	81	*Pseudomonas putida* W619 (YP_001751992)
arsR	118 aa	Arsenical resistance operon repressor	73	77	*Pseudomonas putida* W619 (YP_001751991)
arsC1	137 aa	Arsenate reductase	69	86	*Pseudomonas mendocina* ymp (YP_001187015)

These genes and their products constitute a complete signal transduction pathway and As(III)-oxidizing capability (Fig. 4b). The functions of AoxS, AoxR, AoxA, AoxB, AoxC, and AoxD could be analyzed with relative confidence using freely available bioinformatic tools. However, the *aoxX* gene is often present in *aox* clusters but its function has not been determined so far (Fig. 2c). AoxX is a potential periplasmic component, and homologs have been predicted to be involved in oxyanion transport and metabolism based on BlastX analysis. Moreover, the TAT signal peptide was detected in AoxX by protein domain analysis further suggesting a periplasmic location. AoxX probably participates in the regulation of As(III) oxidation because *aoxX* was cotranscibed with *aoxSR* and induced by As(III). This immediate response to periplasmic As(III) could also be responsible for the higher As(III)-oxidizing efficiency of SY8.

Cysteine sulfhydryl groups are generally thought to be responsible for the interaction between As(III) and proteins, however, serine, histidine, glutamic acid, and arginine residues also have the potential ability to bind As(III) (Ellis et al. 2001; Zhou et al. 2000). Based on protein domain analysis, the AoxS periplasmic loop of SY8 is thought to be an As(III) sensing region, but no cysteine residue was detected there. This is not surprising since the periplasm is often an oxidizing environment. At this point, it is not known which residues bind As(III) or other inducing oxyanions. Another possibility would be a pathway dependent on the periplasmic binding protein. AoxX is a potential As(III) binding protein that could first interact with As(III) and then transfer AoxX with bound oxyanion or As(III) to the projected signal sensing domain in the periplasmic loop of AoxS (Fig. 4b). However, further experiments will be needed to clarify the real function of AoxX.

The TS44 *aox* cluster was atypical due to the absence of upstream regulatory genes and downstream related genes, but *arsDA* were predicted to be present as part of the structural genes and subsequent experiments demonstrated

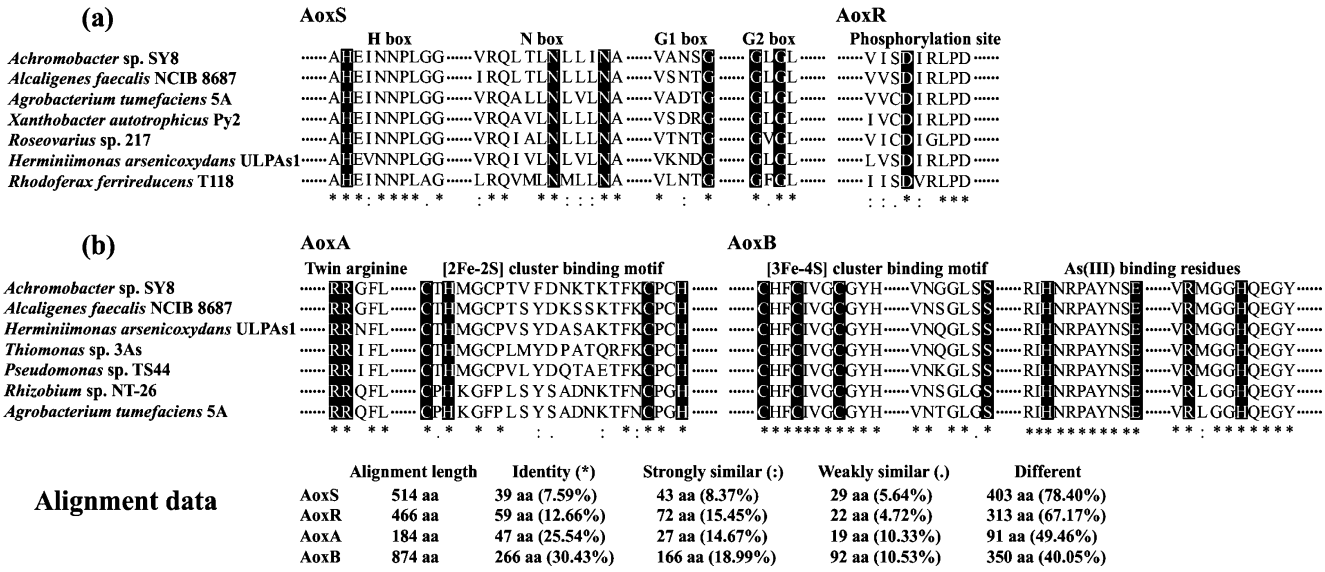

Fig. 3 Multiple sequence alignments of AoxSR homologs (**a**) and AoxAB homologs (**b**). Representative homologs (accession numbers *in parentheses*) are form: *A. faecalis* NCIB 8687 (AY297781); *H. arsenicoxydans* ULPAs1 (NC_009138); *T.* sp. 3As (AM502288); *R.* sp. NT-26 (AY345225); *A. tumefaciens* 5A (DQ151549); *X. autotrophicus* Py2 (NC_009720); *R.* sp. 217 (NZ_AAMV00000000); and *R. ferrireducens* T118 (NC_007908). Strains NCIB 8687, ULPAs1, 3As, NT26, and 5A had been well studied as arsenite-oxidizers. No *aoxSR* were identified in the upstream of *aoxAB* from strains TS44, 3As, and NT-26. However, whole genome sequence of strains Py2, 217, and T118 displayed putative *aox* genes and similar *aox* cluster arrangements, so their *aoxSR* products were added for AoxSR alignment analyses

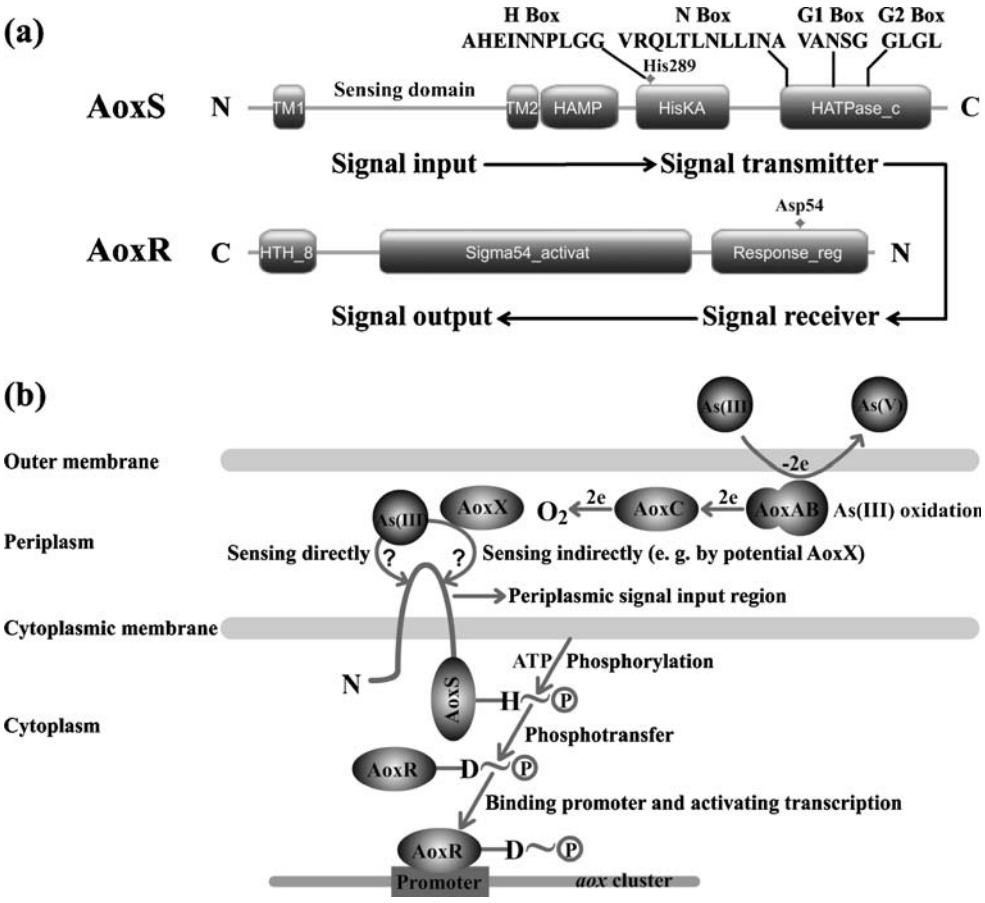

Fig. 4 Prediction of domains and potential signal transduction pathway of SY8 AoxS and AoxR (**a**) and proposed regulatory model of two-component signal transduction system of SY8 *aox* cluster (**b**)

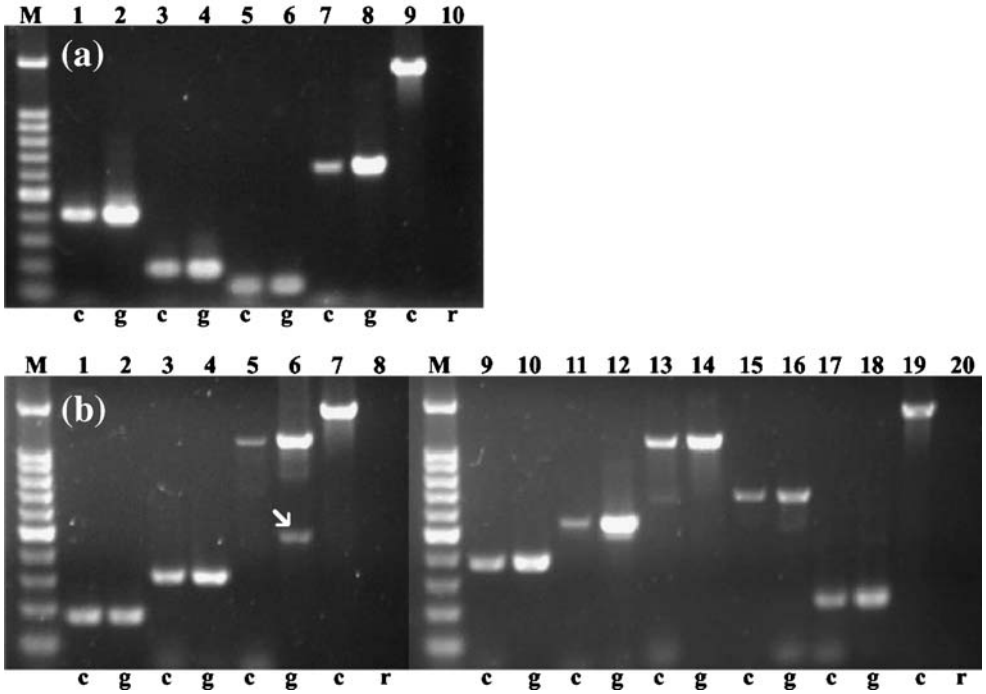

Fig. 5 Detection of cotranscriptional gene clusters by RT-PCR of SY8 (**a**) and TS44 (**b**). *M* 100-bp ladder (100, 200, 300, 400, 500, 600, 700, 800, 900, 1,000, and 1,500 bp); *c*, *g*, and *r* represented PCR templates of first strand cDNA, genomic DNA, and total RNA, respectively. The *arrow* indicated the unspecific band. Controls were monitored by 16S rDNA (27F and 1492R); for (**a**), *lanes 1*, *2* (sP1 and sP2, *aoxR-aoxS*), *lanes 3*, *4* (sP3 and sP4, *aoxS-aoxX*), *lanes 5*, *6* (sP5 and sP6, *aoxA-aoxB*), *lanes 7*, *8* (sP7 and sP8, *aoxB-aoxC-aoxD*), and *lanes 9*, *10* (control); for (**b**), *lanes 1*, *2* (tP1 and tP2, *arsD-arsA*), *lanes 3*, *4* (tP4 and tP5, *aoxA-aoxB*), *lanes 5*, *6* (tP3 and tP5, *arsA-aoxA-aoxB*), *lanes 7*, *8* (control), *lanes 9*, *10* (tP6 and tP7, *arsC1-arsR*), *lanes 11*, *12* (tP6 and tP8, *arsC1-arsR-arsC2*), *lanes 13*, *14* (tP9 and tP10, *arsC2-ACR3-arsH*), *lanes 15*, *16* (tP11 and tP12, *arsH*-DSP-GAPDH), *lanes 17*, *18* (tP13 and tP14, GAPDH-MFS), *lanes 19*, *20* (control)

that they were cotranscribed with *aoxAB*. This is an interesting phenomenon since *arsDA* was generally present in *ars* operons (typically in *arsRDABC*) to enhance the ability of As(III) extrusion via the ArsAB pump (Dey and Rosen 1995). Whether *arsDA* have any relationship with the *aox* cluster or only function with the *ars* operon needs further investigation. There are examples of the presence of *arsDA* alone without an adjacent *arsB*, e.g., in *Halobacterium* sp. strain NRC-1 (Wang et al. 2004); however, TS44 could also have an *arsB* gene on the chromosome away from *arsDA*. Immediately downstream of the *aox* cluster was the *ars* cluster transcribed in the opposite direction. Based on RT-PCR experiments, a total of eight genes in this cluster were transcribed as a unit (Fig. 5b). Although interesting, it is hard to explain why the last three genes (DSP, GAPDH, and MFS) were expressed as part of the *ars* cluster. DSP has both Ser-/Thr- and Tyr-specific protein phosphatase activity that is able to dephosphorylate serine, threonine, and tyrosine residues in the phosphoprotein. DSP is a key regulatory component in signal transduction pathways by controlling the phosphorylation state of serine, threonine, and tyrosine residues (Pils and Schultz 2004). GAPDH is an important enzyme in the sixth step of glycolysis, which can catalyze the oxidation and phosphorylation of glyceraldehyde-3-phosphate to glycerate 1,3-bisphosphate reversibly. Recent publications revealed the nonmetabolic function of GAPDH that could participate in transcriptional regulation (Hara et al. 2005; Zheng et al. 2003). Since phosphate and arsenate have many similar properties, DSP and GAPDH might be linked to arsenate detoxification in a yet unidentified way. MFS represents the largest group of secondary membrane transporters and typically contains 12 transmembrane α helices (Huang et al. 2003). This family of transporter occurs ubiquitously in all classifications of organisms and can carry out multiple functions (e.g., sugars uptake, drugs efflux, Krebs cycle metabolites, organophosphate/phosphate exchangers, and so on; Pao et al. 1998). This implies that MFS transporter is widely distributed and diverse in function. Hence, it might be associated with arsenic transport because its phosphoric exchange function has been reported. However, similar genes and organization of *ars* clusters were also found in the genome of *P. putida* W619 and *P. mendocina* ymp (Fig. 2d).

Another interesting phenomenon was that although the *aox* and *ars* clusters in TS44 were neighbors, this strain

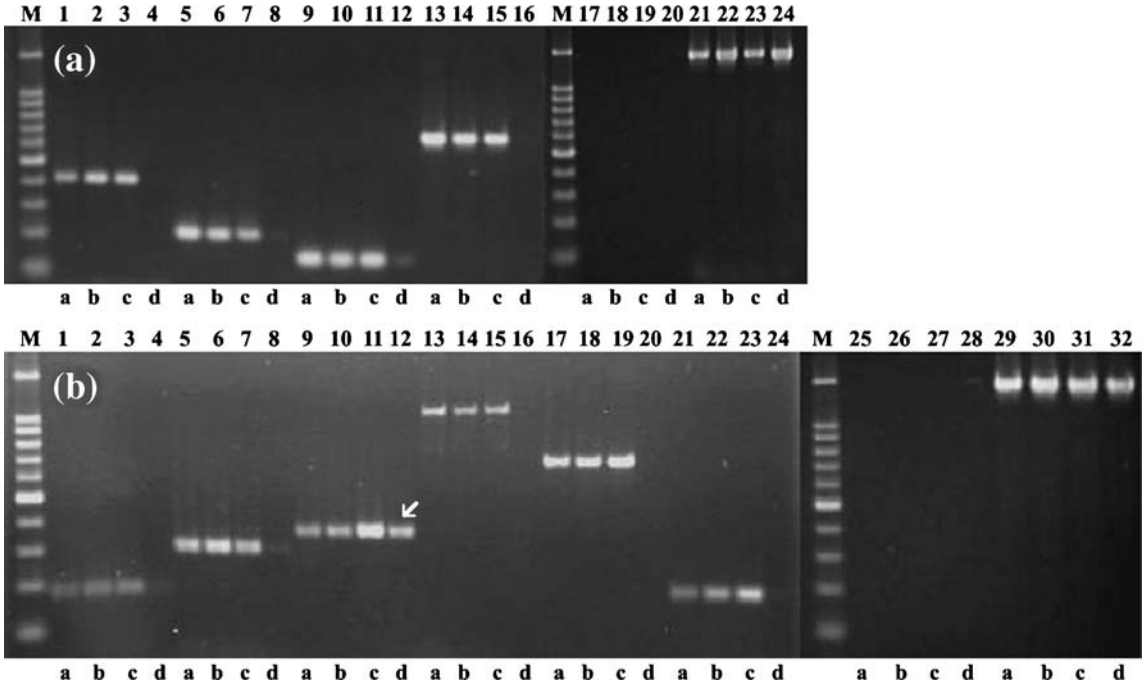

Fig. 6 Induction of arsenic gene clusters by RT-PCR for SY8 (**a**) and TS44 (**b**). *M* and *controls* were identical to Fig. 5. The inducers were indicated by *a*, As(III); *b*, As(V); *c*, Sb(III); and *d*, no addition. The band of constitutive expression was shown by the *arrow*. For (**a**), *lanes 1–4* (sP1 and sP2, *aoxR-aoxS*), *lanes 5–8* (sP3 and sP4, *aoxS-aoxX*), *lanes 9–12* (sP5 and sP6, *aoxA-aoxB*), *lanes 13–16* (sP7 and sP8, *aoxB-aoxC-aoxD*), *lanes 17–24* (control). For (**b**), *lanes 1–4* (tP1 and tP2, *arsD-arsA*), *lanes 5–8* (tP4 and tP5, *aoxA-aoxB*), *lanes 9–12* (tP6 and tP7, *arsC1-arsR*), *lanes 13–16* (tP9 and tP10, *arsC2-ACR3-arsH*), *lanes 17–20* (tP11 and tP12, *arsH*-DSP-GAPDH), *lanes 21–24* (tP13 and tP14, GAPDH-MFS), *lanes 25–32* (control)

exhibited an As(III)-oxidizing phenotype. RT-PCR experiments indicated that the genes in both clusters could be expressed by arsenic induction. One possible explanation is that arsenic redox transformation happened simultaneously and the velocity of As(III) oxidation was much higher than the velocity of As(V) reduction. There should be a dynamic equilibrium between the two processes; however, how the two processes in a single strain are dealt with is still unknown and needs further work. It is important to point out that As(V)-reducing phenotype could be observed only after the loss of As(III)-oxidizing ability in *A. tumefaciens* 5A (Kashyap et al. 2006).

In this study, we found that arsenic clusters could be induced by As(III), As(V), or Sb(III). It was generally thought that they could only be induced by As(III) or Sb (III) but not As(V). However, a model proposed by a recent study could explain this observation well (Li and Krumholz 2007). As(V) entered the cell and subsequently was reduced to As(III) by the constitutively expressed ArsC. The reduced As(III) then acted as an inducer to activate transcription. In this study, we detected constitutive *arsC1* of TS44 in support of this model. The constitutive *arsC* of SY8 could occur in the genome away from the *aox* cluster. If regulated in this way, the arsenic clusters appeared to be induced by As(V) due to the reduction to As(III) by the constitutively expressed ArsC. The TS44 *ars* cluster was shown to be transcribed as a single unit; however, *arsC1-arsR* were constitutively expressed whereas expression of *arsC2-ACR3-arsH*-DSP-GAPDH-MFS could be induced by As(III), As(V), and Sb(III). It is strange how this *ars* cluster regulates differently. One reasonable explanation might be that two different transcripts occur, one being *arsC1-arsR* (constitutive expression) and the other being *arsC1-arsR-arsC2-ACR3-arsH*-DSP-GAPDH-MFS (inducible expression). Such phenomenon has already been demonstrated by Mateos and coworkers (Ordonez et al. 2005).

In both genomes of SY8 and TS44, another interesting phenomenon was the presence of *marR* and *trxB* genes in upstream of the respective *aox* cluster (Fig. 2a, b). In the beginning, we thought they might be linked to the *aox* cluster. However, subsequent analysis showed that they could not be induced by As(III) (data not shown).

In conclusion, we isolated two functional sequences associated with arsenic redox transformation from two novel arsenite oxidizers. Although gene deletions were not carried out, analyses based on reliable bioinformatic tools and subsequent gene expression experiments were able to show that (1) the described three arsenic gene clusters were involved in As(III) oxidation and As(V) reduction and (2)

these clusters were complex and had interesting and novel arrangements.

Acknowledgments This work was supported by the National Natural Science Foundation of China (30570058), the PhD Supervisor Fund (20060504027), and the Retuning Oversea Scientist Fund of the Ministry of Education of China.

References

Achour AR, Bauda P, Billard P (2007) Diversity of arsenite transporter genes from arsenic-resistant soil bacteria. Res Microbiol 158:128–137

Cai L, Liu G, Rensing C, Wang G (2009) Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. BMC Microbiol 9:4

Chen Y, Rosen BP (1997) Metalloregulatory properties of the ArsD repressor. J Biol Chem 272:14257–14262

Dey S, Rosen BP (1995) Dual mode of energy coupling by the oxyanion-translocating ArsB protein. J Bacteriol 177:385–389

Duquesne K, Lieutaud A, Ratouchniak J, Muller D, Lett MC, Bonnefoy V (2008) Arsenite oxidation by a chemoautotrophic moderately acidophilic *Thiomonas* sp.: from the strain isolation to the gene study. Environmental microbiology 10: 228–237

Ellis PJ, Conrads T, Hille R, Kuhn P (2001) Crystal structure of the 100 kDa arsenite oxidase from Alcaligenes faecalis in two crystal forms at 1.64 angstrom and 2.03 angstrom. Structure 9:125–132

Gihring TM, Druschel GK, McCleskey RB, Hamers RJ, Banfield JF (2001) Rapid arsenite oxidation by *Thermus aquaticus* and *Thermus thermophilus*: field and laboratory investigations. Environ Sci & Technol 35:3857–3862

Hara MR, Agrawal N, Kim SF, Cascio MB, Fujimuro M, Ozeki Y, Takahashi M, Cheah JH, Tankou SK, Hester LD, Ferris CD, Hayward SD, Snyder SH, Sawa A (2005) S-nitrosylated GAPDH initiates apoptotic cell death by nuclear translocation following Siah1 binding. Nat Cell Biol 7:665–674

Huang Y, Lemieux MJ, Song J, Auer M, Wang DN (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. Science 301:616–620

Inskeep WP, Macur RE, Hamamura N, Warelow TP, Ward SA, Santini JM (2007) Detection, diversity and expression of aerobic bacterial arsenite oxidase genes. Environmental microbiology 9:934–943

Kashyap DR, Botero LM, Franck WL, Hassett DJ, McDermott TR (2006) Complex regulation of arsenite oxidation in *Agrobacterium tumefaciens*. J Bacteriol 188:1081–1088

Lehr CR, Kashyap DR, McDermott TR (2007) New insights into microbial oxidation of antimony and arsenic. Appl Environ Microbiol 73:2386–2389

Lenoble V, Deluchat V, Serpaud B, Bollinger JC (2003) Arsenite oxidation and arsenate determination by the molybdene blue method. Talanta 61:267–276

Li X, Krumholz LR (2007) Regulation of arsenate resistance in *Desulfovibrio desulfuricans* G20 by an *arsRBCC* operon and an *arsC* gene. J Bacteriol 189:3705–3711

Lin YF, Walmsley AR, Rosen BP (2006) An arsenic metallochaperone for an arsenic detoxification pump. Proc Natl Acad Sci U S A 103:15617–15622

Mukhopadhyay R, Rosen BP, Phung LT, Silver S (2002) Microbial arsenic: from geocycles to genes and enzymes. FEMS Microbiol Rev 26:311–325

Muller D, Medigue C, Koechler S, Barbe V, Barakat M, Talla E, Bonnefoy V, Krin E, Arsene-Ploetze F, Carapito C, Chandler M, Cournoyer B, Cruveiller S, Dossat C, Duval S, Heymann M, Leize E, Lieutaud A, Lievremont D, Makita Y, Mangenot S, Nitschke W, Ortet P, Perdrial N, Schoepp B, Siguier P, Simeonova DD, Rouy Z, Segurens B, Turlin E, Vallenet D, Van Dorsselaer A, Weiss S, Weissenbach J, Lett MC, Danchin A, Bertin PN (2007) A tale of two oxidation states: bacterial colonization of arsenic-rich environments. PLoS Genet 3:518–530

Ordonez E, Letek M, Valbuena N, Gil JA, Mateos LM (2005) Analysis of genes involved in arsenic resistance in *Corynebacterium glutamicum* ATCC 13032. Appl Environ Microbiol 71: 6206–6215

Oremland RS, Stolz JF, Hollibaugh JT (2004) The microbial arsenic cycle in Mono Lake, California. FEMS Microbiol Ecol 48:15–27

Pao SS, Paulsen IT, Saier MH Jr (1998) Major facilitator superfamily. Microbiol Mol Biol Rev 62:1–34

Philips SE, Taylor ML (1976) Oxidation of arsenite to arsenate by *Alcaligenes faecalis*. Appl Environ Microbiol 32:392–399

Pils B, Schultz J (2004) Evolution of the multifunctional protein tyrosine phosphatase family. Mol Biol Evol 21:625–631

Qin J, Rosen BP, Zhang Y, Wang G, Franke S, Rensing C (2006) Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. Proc Natl Acad Sci U S A 103:2075–2080

Rosen BP (1999) Families of arsenic transporters. Trends Microbiol 7:207–212

Rosen BP (2002) Biochemistry of arsenic detoxification. FEBS Lett 529:86–92

Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor: NY

Santini JM, vanden Hoven RN (2004) Molybdenum-containing arsenite oxidase of the chemolithoautotrophic arsenite oxidizer NT-26. J Bacteriol 186:1614–1619

Sato T, Kobayashi Y (1998) The *ars* operon in the skin element of *Bacillus subtilis* confers resistance to arsenate and arsenite. J Bacteriol 180:1655–1661

Silver S, Phung LT (2005) Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. Appl Environ Microbiol 71:599–608

Stock AM, Robinson VL, Goudreau PN (2000) Two-component signal transduction. Annu Rev Biochem 69:183–215

Wang G, Kennedy SP, Fasiludeen S, Rensing C, DasSarma S (2004) Arsenic resistance in *Halobacterium* sp. strain NRC-1 examined by using an improved gene knockout system. J Bacteriol 186: 3187–3194

Weeger W, Lievremont D, Perret M, Lagarde F, Hubert JC, Leroy M, Lett MC (1999) Oxidation of arsenite to arsenate by a bacterium isolated from an aquatic environment. BioMetals 12:141–149

Wu J, Rosen BP (1993) The *arsD* gene encodes a second trans-acting regulatory protein of the plasmid-encoded arsenical resistance operon. Mol Microbiol 8:615–623

Zheng L, Roeder RG, Luo Y (2003) S phase activation of the histone H2B promoter by OCA-S, a coactivator complex that contains GAPDH as a key component. Cell 114:255–266

Zhou Y (1990) Arsenite and arsenate determination by the molybdene blue method. Environmental Protection Science (Chinese) 16:45–47

Zhou T, Radaev S, Rosen BP, Gatti DL (2000) Structure of the ArsA ATPase: the catalytic subunit of a heavy metal resistance pump. Embo J 19:4838–4845