Structural and functional genomics of plasmid pSinA of *Sinorhizobium* sp. M14 encoding genes for the arsenite oxidation and arsenic resistance

Lukasz Drewniak[1*], Lukasz Dziewit[2], Martyna Ciezkowska[1], Jan Gawor[3], Robert Gromadka[3] and Aleksandra Sklodowska[1]

[1] Laboratory of Environmental Pollution Analysis, Faculty of Biology, University of Warsaw, Miecznikowa 1, 02-096 Warsaw, Poland

[2] Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, Miecznikowa 1, 02-096 Warsaw, Poland

[3] DNA Sequencing Laboratory, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawińskiego 5a, 02-106 Warsaw, Poland

[*] corresponding author address:

Laboratory of Environmental Pollution Analysis, Faculty of Biology, University of Warsaw, Miecznikowa 1, 02-096 Warsaw, Poland

Tel./Fax +48 22 55 41 219

ldrewniak@biol.uw.edu.pl

Running title: self-transferable arsenic metabolism and resistance plasmid

Abstract

Plasmid pSinA of *Sinorhizobium* sp. M14 (*Alphaproteobacteria*) is the first described, natural, self-transferable plasmid harboring a complete set of genes for oxidation of arsenite. Removal of this plasmid from cells of the host strain caused the loss of resistance to arsenic and heavy metals (Cd, Co, Zn and Hg) and abolished the ability to grow on minimal salt medium supplemented with sodium arsenite as the sole energy source. Plasmid pSinA was introduced into other representatives of *Alphaproteobacteria* which resulted in acquisition of new abilities concerning arsenic resistance and oxidation, as well as heavy metals resistance. Microcosm experiments revealed that plasmid pSinA can also be transferred *via* conjugation into other indigenous bacteria from microbial community of As-contaminated soils, including representatives of *Alpha-* and *Gammaproteobacteria*. Analysis of "natural" transconjugants showed that pSinA is functional (expresses arsenite oxidase) and is stably maintained in their cells after approximately 60 generations of growth under nonselective conditions. This work clearly demonstrates that pSinA is a self-transferable, broad-host-range plasmid, which plays an important role in horizontal transfer of arsenic metabolism genes.

Keywords: arsenite oxidation, arsenic resistance, plasmid pSinA, horizontal gene transfer

1. Introduction

Arsenic is one of the most abundant pollutants in the world. According to the World Health Organization (WHO), arsenic contamination of groundwater has been detected in 70 countries and probably occurs in many others. Dissemination of arsenic pollution in the environment is a result of anthropogenic activity, as well as natural, biogeochemical phenomena.

Arsenic oxidizing microorganisms are amongst the key players involved in arsenic mobilization/immobilization processes and thus in the biogeochemical cycle of this toxic element. They can transform soluble arsenite to less toxic form – arsenate, as well as catalyze oxidation of arsenic-bearing minerals like arsenopyrite (Rhine et al., 2008) or As (III) surface-bound to mineral sulfides (Duquesne et al., 2003). The process of arsenic oxidation catalyzed by bacteria is much faster than chemical oxidation (in the presence of O_2) and has been observed in many environments (e.g. Osborne et al., 2010). Several pure cultures of bacterial strains (Rhine et al., 2006), as well as microbial consortia (Battaglia-Brunet et al., 2002) able to transform As (III) to As(V) have been isolated. For some of them, biogeochemical function in the environment has been described (Connon et al., 2008; Drewniak et al., 2010) and/or a practical application in bioremediation has been proposed (e.g. Simeonova et al., 2005).

Microbial arsenite oxidation processes are carried out either for detoxification or energy production. The transformation of As (III) to As (V) ions is catalyzed by the arsenite oxidase. This enzyme is a heterodimeric protein composed of a small subunit containing a Rieske [2Fe-2S] cluster and a large molybdopterin subunit, harboring a [3Fe-4S] cluster (Ellis et al., 2001). According to the previous studies and unified nomenclature proposed by Lett et al. (2012), these subunits are encoded by *aioB* (small Rieske subunit, previously known as *aoxA/aroB/asoB*) and *aioA* (large molybdopyterin subunit, previously known as *aoxB/aroA/asoB*) genes. Regulatory genes: *aioS* (previously *aoxS*) and *aioR* (previously *aoxR*), as well as cytochrome C gene (*cytC)* were identified in a majority of arsenite oxidizers.

Moreover, many arsenite oxidizing bacteria harbor additional genes involved in arsenic resistance (*ars* genes). In some of them, as in the *Alcaligenes faecalis* NCBI8687 genome, *ars* genes are located in close proximity to *aio* genes, within the "arsenic gene island" (Silver and Phung, 2005).

Recent metagenomic studies (e.g. Quemeneur et al., 2010, Drewniak et al., 2012) showed that *aioA*-like genes, are widely distributed within microbial communities of arsenic-contaminated environments. However, little is known about environmental dissemination and horizontal transfer of arsenite oxidase genes. In contrast to the arsenic resistance genes, almost all found and documented arsenite oxidase genes are located within chromosomes. To date, the *aio*-like loci have been found only within three plasmids: megaplasmid pSI07 of *Ralstonia solanacearum* (Remenant et al., 2010), plasmid pPB12 of *Nitrobacter hamburgiensis* (Starkenburg et al., 2008), and plasmid pTT27 of *Thermus thermophilus* HB8 (AP008227). Information concerning above-mentioned plasmids and their functionality in the context of arsenite oxidation, as well as horizontal transfer of arsenic metabolism genes, is rather residual. Unfortunately, the application of arsenic-metabolizing microorganisms in bioremediation processes have so far been confined to the laboratory environment and *ex situ* technologies. *In situ* bioremediation of contaminated regions is problematic because of the rapid decrease in the number of the introduced organisms in their "new" environment. This phenomenon occurs due to the physical and chemical conditions of such environments, which are usually inappropriate for organisms used in bioremediation, and competition between the introduced organisms and indigenous microbial communities. To overcome these problems, attempts have been made to use naturally occurring multi-resistant bacteria with a broad tolerance to diverse physical and chemical conditions. An alternative approach is to use genetically modified indigenous microorganisms.

Different strategies can be used to construct genetically engineered microorganisms, including the introduction of new genes (i) directly into the bacterial chromosome by genetic manipulation, (ii) on man-made vectors, or (iii) by the transfer of naturally occurring plasmids. Of these methods, natural plasmid conjugation appears to be the most applicable for the creation of strains for use in bioremediation processes. The enhanced bioaugmentation obtained by the transfer of a conjugal plasmid from an introduced donor strain to indigenous microbial recipients has been successfully used in the remediation of soil environments contaminated with toxic organic compounds, such as 2,4-dichlorophenoxyacetic acid (Quan *et al*., 2010).

In this study, we have performed genomic and functional characterization of an approximately 109-kb plasmid pSinA of *Sinorhizobium* sp. M14 [a strain isolated from the ancient Zloty Stok gold mine (Poland), highly polluted by arsenic], which carries genetic information responsible for arsenite oxidation and heavy metal resistance. We demonstrate that pSinA is a suitable candidate for the construction of novel arsenic bioremediation tools, since it can be used both in the production of genetically modified strains and in bioaugmentation.

2. Materials and methods

2.1. Bacterial strains, plasmids and culture conditions.

Sinorhizobium sp. M14 (Drewniak et al., 2008) was the host strain of the analyzed plasmid pSinA. Other bacterial strains and plasmids used in this study are listed in Table S1. All strains were routinely grown in lysogeny broth (LB) medium (Sambrook and Russell, 2001) at 22ºC (*Sinorhizobium* spp., *Aeromonas* spp., *Pseudomonas* spp., *Serratia* spp., *Brevundimonas* spp.), 30ºC (*Paracoccus* spp. and *Agrobacterium* spp.) or 37ºC (*Escherichia coli* TG1). Where necessary, the LB medium was supplemented with (i) antibiotics in the following concentrations: kanamycin, 50 μg/ml; rifampicin 50 μg/ml; tetracycline, 20 μg/ml;

or (ii) heavy metals and metalloids in the following concentrations: As(III), 0.1-25 mM; As(V), 5-250 mM; Hg(II), 0.01-2.5 mM; Cd(II), 0.1-2.5 mM; Co(II), 0.1-2.5 mM, Zn(II), 0.1-2.5 mM. For chemolithoautotrophic growth, bacterial strains were cultivated at 22ºC in mineral salt minimal (MSM) medium containing 5 mM sodium arsenite as electron donor and sodium bicarbonate as the only source of carbon (Drewniak et al., 2008).

2.2. Chemical analysis

Arsenite transformation in culture supernatants was determined following filtration through sterile 0.22-μm pore size Durapore filters (Millipore). Arsenic species were separated by high-performance liquid chromatography on a reversed-phase polymeric resin (Waters IC-Pak™ Anion HC column, 150 by 4.6 mm, pore size 10 μm) and quantified by graphite furnace atomic absorption spectrometry (GFAAS) (AA Solaar M6 Spectrometer, TJA Solutions, UK). Total arsenic concentration was quantified by GFAAS using ASA standards (Merck, Darmstadt, Germany) prepared in 3% HNO_3.

2.3. Plasmid pSinA DNA isolation.

Plasmid pSinA DNA was isolated from 200 ml of *Sinorhizobium* M14 culture using a modified alkaline lysis method (Sambrook and Russell, 2001) with the following changes. The cell pellet was resuspended in 20 ml of 10 mM EDTA (pH 8.0) prior to lysis. Residual proteins in the sample were removed by additional extractions with chloroform-isoamyl alcohol (24:1) solution. Finally, the isolated DNA was precipitated using 2.5 M NaCl/20% PEG8000 solution, the pellet was washed with 70% ethanol and resuspended after drying in 0.1 ml of 10 mM Tris-Cl (pH 8.0).

2.4. Shotgun sequencing and assembly.

Plasmid pSinA was sequenced using a shotgun approach with a GS FLX Titanium (454) pyrosequencer. Approximately 5 ug of plasmid DNA was used for shotgun library construction following the manufacturer's instructions. A library sample was sequenced on a

small region of a Picotiter plate. Sequence assembly was performed using Newbler *de novo* assembler software (Roche). Contig alignment was performed using Seqman software from the Lasergene package (DNAStar). A single gap was bridged by PCR and confirmed by capillary sequencing.

2.5. DNA sequence analysis and annotation

Annotation of the complete pSinA nucleotide sequence (submitted to the GenBank database under accession number JF809815) was carried out manually using Artemis software and BLAST programs. Searches for IS elements were performed using the IS Finder database (http://www-is.biotoul.fr/is.html).

2.6. DNA manipulations

Megaplasmid visualization was achieved by in-gel lysis and DNA electrophoresis according to the procedure of Hynes and McGregor (1990). Common DNA manipulation methods were performed as described by Sambrook and Russell (2001). DNA probes for Southern hybridization were labeled with digoxigenin (Roche). Hybridization and visualization of bound probes were carried out as recommended by the manufacturer.

2.7. PCR amplification

The primer pairs aioAF/aioAR, repAF/repAR, traI/traIR and 12F/12R (Table S2) were used for amplification of the *aioA*, *repA, traI* and *orf12* genes, respectively. Amplification was performed in the standard PCR mixtures containing oligonucleotide primers, appropriate template DNAs, deoxynucleoside triphosphates, and *Pfu* polymerase (Qiagen) with the supplied buffer. A Mastercycler EpigradientS (Eppendorf) was used to perform the PCR thermocycle. Amplified DNA fragments were analyzed by electrophoresis on 0.8 or 2.0% agarose gels and were used as probes for Southern hybridization.

2.8. Introduction of plasmid DNA into bacterial cells

Transformation of *E. coli* TG1 was performed according to the method of Kushner (1978). The triparental mating [the *E. coli* TG1 donor strain carrying the mobilizable vector, *E. coli* TG1 carrying the helper plasmid pRK2013, and a suitable recipient strain (*A. tumefaciens* LBA288, *P. alcaliphilus* JCM7364R, *Brevundimonas* sp. OS16R, *Aeromonas* sp. W54R, *S. grimesii* OS9R, *Pseudomonas* sp. OS8R or *P. marcusii* OS22R)] was performed as previously described by Bartosik et al. (2001).

2.9. Microcosm (plasmid pSinA) release experiment

Arsenic contaminated soil was obtained from the soil subsurface layer (about 20 cm below the surface) from the area of "Zloty Potok" and "Potok Trujaca" (streams passing through the Zloty Stok gold mine, SW Poland). The experiments were conducted in a microcosm scale, using 100 g soil samples, by inoculation with *Sinorhizobium* sp. M14. Non-sterile soil samples, without enrichment, were used as the control samples. All tests were performed in triplicate. Putative recipients of plasmid pSinA, referred to as transconjugants, were collected at the beginning, and every 15 days of 60-day experiments.

2.10. Selection of presumptive transconjugants

Presumptive transconjugants were selected according to the strategy of arsenite oxidizers isolation, previously described by Drewniak et al., 2008. Isolation of these transconjugants was performed in aerobic conditions on MSM medium supplemented with 5 mM sodium arsenite and after 7 days of incubation at 22°C test with 0.1 M $AgNO_3$ solution was used. The reaction between $AgNO_3$ and As(III) or As(V) results in the formation of a coloured precipitate. A brownish precipitate reveals the presence of arsenate in the medium (colonies expressing arsenite oxidase), while a yellow precipitate shows the presence of arsenite.

2.11. Identification of transconjugants carrying plasmid pSinA

For identification of transconjugants carrying plasmid pSinA, multiplex-PCR and Southern hybridization were used. The primer pairs aioAF/aioAR, repAF/repAR, traI/traIR and

12F/12R (Tab. S2) were used for amplification of the *aioA*, *repA, traI* and *orf12* genes, respectively. Amplification by PCR was performed with a Mastercycler EpigradientS (Eppendorf) using *Pfu* polymerase (Qiagen) with the supplied buffer. Amplified DNA fragments were analyzed by electrophoresis on 0.8% agarose gels and were used as probes for Southern hybridization. DNA probes for Southern hybridization were labeled with digoxigenin (Roche). Hybridization and visualization of bound probes were carried out as recommended by the manufacturer. The plasmids carried by transconjugants were identified by screening colonies using a rapid alkaline extraction procedure (Birnboim and Doly, 1979) and agarose gel electrophoresis. The final step of identification of transconjugants was the analysis of 16S rRNA genes. Partial 16S rDNA sequences (~1.4 kbp) were amplified from each isolate (transconjugant) using primers 27f and 1492r (Lane, 1991), then sequenced, and the bacteria were classified according to the similarity of their 16S rDNA sequences to those deposited in the GenBank database. The unique strains representing the soil ecosystem (Soil I or Soil II) are listed in Table S3.

3. Results

3.1. Plasmid pSinA is responsible for arsenite oxidation

In a previous study, we described the As(III)-oxidizing psychrotolerant *Sinorhizobium* sp. M14 (Drewniak et al., 2008). This strain was isolated from the arsenic-rich bottom sediments of an ancient gold mine in Lower Silesia, SW Poland. Physiological characteristic showed that *Sinorhizobium* sp. M14 can grow chemolithoautotrophically using sodium arsenite or arsenopyrite as a source of energy, and carbon dioxide/bicarbonate as the sole carbon source. In addition, M14 is able to grow heterotrophically in extremely high concentrations of arsenate (up to 250 mM) and arsenite (up to 20 mM) at temperatures ranging from 10°C to 37°C and at pH values between 4 and 9. *Sinorhizobium* sp. M14 is also resistant to many heavy metals, including: Cd, Co, Cr, Cu, Fe, Mn, Ni, Se, V, Zn and (Drewniak et al., 2008).

These abilities make this strain an ideal candidate for use in the removal of arsenic from contaminated environments. Since some of these features are likely to be encoded by mobile genetic elements, that may have been acquired during horizontal gene transfer, we decided to investigate this possibility.

To identify plasmids of M14 strain we performed an alkaline lysis and megaplasmids visualization. We found that the strain carries two plasmids, pSinA (approx. 109 kb) and pSinB (approx. 300 kb). In this paper we present a complex structural and functional analysis of pSinA plasmid.

To determine the role of pSinA encoded modules in *Sinorhizobium* sp. M14 metabolism, we have constructed plasmid's miniderivatives, which were then used to cure the host cells of the native plasmid due to the incompatibility phenomenon. For the construction of pSinA miniderivatives different EcoRI and BamHI restriction fragments of the plasmid were cloned into the mobilizable, *E. coli*-specific (unable to replicate in *Alphaproteobacteria*) vector pABW1 (Bartosik et al., 1997). These constructs were transferred *via* triparental conjugation into *A. tumefaciens* LBA288 (Hooykaas et al., 1980), a bacterium closely related to *Sinorhizobium* sp. Two different-sized replicons were obtained: (i) pMSA5 – containing a 3.5-kb EcoRI restriction fragment of pSinA, and (ii) pMSA10 – containing a 4.5-kb BamHI restriction fragment of pSinA. Sequencing showed that pMSA10 carries: (i) a single replication initiation protein gene (*repC*) (ORF3) without other components of the typical *repABC* operon (Bartosik et al., 1998), (ii) a gene potentially involved in recombinational resolution of plasmid multimers (*mrs*; ORF4) and (iii) two additional genes: *phd* (ORF5) and *pemK* (ORF6), which probably create a putative hybrid *phd-pemK* addiction system. Despite the absence of the *repAB* genes, pMSA10 was still able to replicate in various representatives of *Alphaproteobacteria*, including: *A. tumefaciens* LBA288, *Brevundimonas* sp. OS16R, *P. marcusii* OS22R and *Paracoccus alcaliphilus* JCM7364R (Bartosik et al., 2002). The plasmid

was very stably maintained in the host cells, since after approximately 60 generations of growth under nonselective conditions it was present in 99% of the cells. Such high stability is probably a consequence of the presence of the toxin-antitoxin system (*phd-pemK*) within this plasmid.

The second minireplicon (pMSA5) has an unusual structure since no similarity to known replication systems was found. Plasmid pMSA5 carries an incomplete ORF69, which codes for a predicted protein similar (86% identity) to a putative helicase of *R. leguminosarum* bv. *viciae* 3841 (Young et al., 2006), and ORF70 which encodes a predicted product with similarity (71% identity) to partitioning protein ParB of *A. rhizogenes* (Mankin et al., 2007). Plasmid host range studies surprisingly revealed that this minireplicon is able to replicate in both *Alphaproteobacteria* (*Brevundimonas* sp. OS16R, *P. marcusii* OS22R, *P. alcaliphilus* JCM7364R) and *Gammaproteobacteria* (*Pseudomonas* sp. OS8R, *S. grmiesii* OS9R, and *Aeromonas* sp. W54R). Thus, we hypothesize that this replication system can serve as an additional replication region, like those found in other *repABC*-type plasmids such as pSmeSM11a (Stiens et al., 2006) and pRmeGR4a (Burgos et al., 1996). It is also probable that pMSA5 does not encode a functional replication initiation protein, but contains only an origin of replication (*oriV*) that requires the activity of additional gene products encoded within the bacterial chromosome to function. Nevertheless, based on these preliminary studies, we conclude that pSinA is a stable, broad-host-range plasmid, which actively participates in horizontal gene transfer and this makes it potentially extremely valuable for use in future bioremediation technologies, especially in bioaugmentation.

Using pMSA10 as an incompatibility determinant we were able to remove plasmid pSinA from M14 cells, which resulted in obtaining a plasmid-less strain M14SL. The loss of this plasmid completely abolished the ability to chemoautotrophic growth with arsenite as the energy source (Fig. 1), which indicates a crucial role of pSinA encoded genetic modules in

arsenite oxidation. Growth analysis of the wild-type strain (M14) and the mutant lacking pSinA (M14SL) showed that the loss of this plasmid from *Sinorhizobium* sp. M14 cells causes not only disturbances in the ability to oxidize As(III), but also leads to deficiencies in arsenic resistance. Strain M14SL tolerated significantly lower concentrations of toxic arsenite (1 mM) and arsenate (10 mM), than the wild-type strain (20 mM and 250 mM, respectively). Those findings confirmed that plasmid pSinA is engaged in the arsenite oxidation and arsenic resistance and encouraged us to perform a genomic analysis of plasmid pSinA.

3.2. The structure of plasmid pSinA

To identify the genetic determinants responsible for arsenite oxidation, supercoiled plasmid pSinA DNA was isolated from *Sinorhizobium* sp. M14 with a scaled-up, modified alkaline-lysis procedure and used for shotgun library construction. Sequencing of the constructed library allowed us to obtain a circular DNA-molecule of 108,938 bp with a mean G+C content of 59.5%. The genetic map of plasmid pSinA is shown in Fig. 2A and a summary of the functional predictions for pSinA coding sequences (CDSs) is presented in Table S4 in the supplementary material.

Plasmid pSinA contains 103 CDSs (accounting for 89% of the plasmid) and has a mosaic structure with single genes grouped in gene clusters and modules responsible for different functions. The plasmid backbone (36 ORFs, which encodes putative proteins responsible for replication, stabilization and conjugative transfer) and phenotypic modules (47 ORFs, which are clustered in two main modules responsible for arsenic metabolism functions and heavy metal resistance), are highlighted on the outer circle of the pSinA genetic map in Fig. 2A and described below.

3.2.1. The plasmid backbone

The conserved backbone contains three modules: (i) REP+STA, comprising 6 ORFs, which encode proteins responsible for replication and stabilization; (ii) REP+STA2, comprising two

additional ORFs - encoding proteins responsible for replication and stabilization and (iii) TRA, comprising 29 ORFs, which encode proteins involved in conjugative transfer and mating-pair formation.

Genomic analysis of plasmid pSinA confirmed the results obtained by functional studies of pSinA miniderivatives and showed that pSinA harbors two maintenance modules with genes that may be involved in replication and stabilization. The pSinA REP-STA module consists a complete component of the hybrid replication-stabilization system *repABC* and previously mentioned two other stabilization systems responsible for (i) recombinational resolution of plasmid multimers (*mrs*; ORF4) and (ii) postsegregational elimination of plasmid-less cells from the bacterial population (*phd-pemK;* ORF5 and ORF6) (Fig. 2B). The pSinA *repABC* system encodes the partitioning proteins RepA (ParA family ATPase; ORF1), RepB (ParB-like nuclease domain; ORF2), and replication initiation protein RepC (ORF3). The *repA* and *repB* genes encode proteins with significant similarity to the RepA and RepB polypeptides of pRi1724 of *Agrobacterium rhizogenes* (93% and 71%, respectively), whereas the *repC* product is similar (73% amino-acid identity) to the replication protein encoded on plasmid p42a of *Rhizobium etli* CFN 42. Sequence analysis showed that the second maintenance module (REP-STA2) contains only two genes, previously identified during minireplicons analysis.

The transfer region of pSinA consists of 29 genes and its structure (Fig. 2B) is similar to that of the conjugative transfer and mating-pair formation region of many plasmids, e.g. pSmeSM11b of *S. meliloti* strain SM11 (Stiens et al., 2007). The predicted origin of conjugal transfer (*oriT*) of pSinA is placed between the *traC* and *traA* genes, and its location and nucleotide sequence (5'-CCAAGGGCGCAATTATACGTCGCTGACGCGACGTT-3') are conserved among plasmids belonging to the MOB_{Q2} family (Garcillán-Barcia et al., 2009). Phylogenetic analysis of *traA* genes and *repC* genes (Fig. 1S) confirmed the common origin

of the pSinA plasmid and repABC replicons, representing the MOB_{Q2} family. However, it should be noted, that the remaining part of the plasmid (with phenotypic modules) has a completely different origin, and is a mixture of gene modules derived from various ancestors (see G+C contents in Fig. 2A and phylogenetic analysis in Fig. 1S).

3.2.2. Arsenic metabolism and resistance gene island

Bioinformatics analysis of plasmid pSinA genome allowed for identification of a 24.8-kb region that contains 27 ORFs, encoding putative proteins involved in arsenite oxidation (AIO module), arsenic/phosphorus transport (PST/PHN module) and arsenic resistance (ARS module) (Fig. 2C).

***The arsenite oxidation gene cluster* (*AIO*)** is comprised of seven genes in the same transcriptional orientation: *aioX* (oxyanion binding protein), *aioS* (signal transduction), *aioR* (response regulator), *aioAB* (structural genes encoding two subunits of arsenite oxidase), *cyt*-c2 (c-type cytochrome) and *moeA* (molybdopterin biosynthesis gene). The predicted amino-acid sequences of particular proteins encoded by the AIO cluster exhibit 91-100% identity in the amino-acid sequence to the gene products of arsenite oxidation modules of the facultatively chemolithoautotrophic *Rhizobium* sp. NT26 (Santini and van den Hoven, 2004), and the heterotrophic strains *Agrobacterium tumefaciens* strain 5A (Kashyap et al., 2006) and *Ochrobactrum tritici* SCII24 (Branco et al., 2009). Furthermore, the genetic organization of the plasmid pSinA AIO module (*aioSRABcytCmoeA*) is identical to that of the arsenite oxidase genes of *Rhizobium* sp. NT26, *A. tumefaciens* strain 5A and *O. tritici* SCII24. The same synteny, but only for five genes, *aioSRABcytC*, is also found in the arsenite oxidase module of *Achromobacter* sp. SY8 strain (Cai et al., 2009). Homologs of the *aioSR* (two-component signal-transduction system) and *aioAB* (arsenite oxidase) genes are also present in the *aio/aox* loci of *Herminiimonas arsenicoxydans* (Muller et al., 2007) and *Alcaligenes*

faecalis (Silver and Phung 2005a), but the regulatory genes (*aioSR/aoxSR*) in these hosts are separated from the structural genes (*aioAB*) by an additional ORF.

The phosphate and phosphonate transport region (9.7 kb), contains 12 ORFs encoding putative proteins sharing substantial similarities with (i) a phosphate-specific ABC transporter (*pst*), (ii) a phosphonate uptake ABC transporter (*phn*), and (iii) a phosphate/phosphonate regulon transcriptional regulatory protein (*pho*). The *pst* locus (ORF30-ORF33) of pSinA comprises genes similar to *pstS,-C,-A,-B* found in the "arsenic gene island" of *A. faecalis* (70-75% identity at the amino acid level) (Silver and Phung, 2005a). This type of system usually consists of four components: PstS – a periplasmic phosphate binding protein, PstC and PstA – integral membrane proteins, and PstB – an ATPase. PST modules encode high-affinity phosphate transporters, but some bacteria may also use these for the transport of arsenate into bacterial cells as oxyanions (Mukhopadhyay et al., 2002, Silver and Phung, 2005a, Paez-Espino et al., 2009). In light of this information, we assume that the PST module of plasmid pSinA also encodes proteins which can be used for this purpose, and it will be a subject of further study.

The PHN module is composed of *phnD*, *phnC* and two *phnE* genes, which encode products with homology to proteins required for the transport and catabolism of phosphonates (Kononova and Nesmeyanova, 2002): a periplasmic binding protein (PhnD), hydrophilic ATP-binding protein (PhnC), and two integral membrane proteins (PhnE1 and PhnE2). The systems, potentially responsible for phosphate (*pst*) and phosphonate (*phn*) transport, are separated by three ORFs (ORF34-ORF36) that encode putative proteins involved in transcriptional regulation. ORF34 and ORF35 encode phosphate regulon transcriptional regulatory proteins PhoU and PhoB, respectively (49 and 40% amino-acid identity with *Pseudovibrio* sp. JE062; GenBank accession nos. **ZP_05085295** and **ZP_05085350**), while ORF36 encodes a bifunctional protein with an N-terminal transcriptional regulator (ArsR)

domain and a C-terminal arsenate reductase (ArsC) domain (47% amino-acid identity with the ArsRC protein encoded by plasmid pPB12 of *N. hamburgensis* X14; Starkenburg et al., 2008).

The arsenic resistance gene cluster is a 6.1-kb region that contains eight ORFs, including genes that encode an arsenite efflux transporter (ArsB), and tyrosine (ArsCa) and glutaredoxin (ArsCb) arsenate reductases. ArsB permease and ArsC reductase are the key enzymes in the arsenic detoxification systems of most *Archaea* and *Bacteria* (Silver and Phung, 2005b). Genes encoding other types of reductases were also identified within the pSinA *ars* locus: *arsH*, that encodes an NADPH-dependent FMN reductase and *trkA*, that encodes a major facilitator superfamily protein. All of the structural genes involved in arsenic resistance are flanked by two transcriptional regulatory genes (*arsR*). Homologs of the pSinA arsenic resistance genes have been found in various arsenite oxidizing bacteria, but the gene order is variable. Moreover, detailed phylogenetic analysis showed that arsenic resistance genes (*ars*) have a different evolutive origin than arsenite oxidation genes (*aio*), and are closely related to genes derived from bacteria that are unable to oxidize arsenite (Fig. 1S).

3.2.3. *Heavy metal resistance gene module*

In addition to genes involved in arsenic metabolism, plasmid pSinA contains a 7.5-kb region that might be involved in heavy metal resistance (ORF54-ORF62). It encodes (i) a predicted efflux pump for cadmium, zinc and cobalt (CzcD**)**, and (ii) putative proteins involved in the transport and detoxification of mercury compounds (MerRTPA). These two gene clusters are separated by an additional *nhaA* gene, which encodes a pH-dependent sodium/proton antiporter. Homologs of this gene have been identified in many bacterial genomes, and they have been shown to be involved in adaptation to high salinity and alkaline pH (Padan et al., 2004). Thus, the presence of *nhaA* gene within the pSinA plasmid structure may partly explain the adaptation of *Sinorhizobium* sp. M14 to high salinity (up to 20mg/L of NaCl) and

alkaline pH (up to 10), which was observed in our previous works (Drewniak et al., 2008, Drewniak, 2009).

The *czc* (Cd/Zn/Co resistance) locus of pSinA comprises three ORFs (54-56), and the gene product of ORF55 is most probably responsible for transport of heavy metals and detoxification. The putative polypeptide encoded by ORF55 is in 71% identical to the cation efflux protein of *Methylobacterium nodulans* ORS 2060 (GenBank accession no. YP_002497120), which belongs to the cation diffusion facilitator (CDF) protein family. Chemiosmotic efflux pumps belonging to the CDF family are polypeptides spanning a single membrane, and the best-known member is the CzcD Cd^{2+} and Zn^{2+} efflux system of *Ralstonia metallidurans* (Nies, 2003). ORF55 is flanked by two genes (ORF54 and ORF56) which encode hypothetical proteins. Analysis with InterProScan (EMBL-EBI) showed that both putative proteins contain predicted transmembrane regions, which suggests that they could be permeases involved in the transport of heavy metals.

The mercury resistance genes in plasmid pSinA are identical to those within the *mer* operon of *Ochrobactrum anthropi* ATCC 49188 (GenBank accession no. CP000759). The gene product of *merR* from Gram-negative and Gram-positive bacteria functions as an activator protein for the regulation of mercury (Hg^{2+}) resistance. MerT is an inner membrane uptake protein and MerP is a small protein that binds Hg^{2+} in the periplasm, whereas MerA is a reductase enzyme involved in reduction of Hg^{2+} to Hg^{0} (Silver and Phung, 2005b).

To investigate whether plasmid pSinA mediates heavymetal resistance, the growth of *Sinorhizobium* sp. M14 and its plasmid-less derivative (M14SL) in the presence of Cd, Co, Zn and Hg salts were compared. MIC analysis showed that removal of plasmid pSinA from M14 cells leads to deficiencies in heavy-metal resistance. Strain M14SL (pSinA-less) tolerated significantly lower concentrations of Cd (II) (0.1 mM), Co(II) (0.2 mM), Zn(II) (0.1 mM) and

Hg(II) (0.01 mM) comparing with the wild-type strain (Cd, 0.5 mM; Co, 0.9 mM; Zn, 0.9 mM; Hg, 0.25 mM).

3.3. Functionality of the pSinA TRA module

In many bioaugmentation field experiments, a rapid decline in the inoculum size or massive cell death have frequently been observed (El Fantroussi and Agathos 2005). The use of bacterial strains that contain conjugal or mobilizable plasmids carrying the desired genes (e.g. catabolic or metal resistance genes) may help to solve these problems. The enhanced bioaugmentation caused by the transfer of such a plasmid from an introduced donor strain to indigenous microbial recipients has been demonstrated in the remediation of soil contaminated with toxic organic compounds, such as dichlorophenoxyacetic acid (2,4-D) (Top *et al*., 1998). Before using plasmids in bioaugmentation methods or constructing novel strains useful in bioremediation, it is necessary to examine the functionality of the transfer genes.

To confirm the functionality of the predicted TRA module of pSinA and to check its self-transfer ability, di- and triparental conjugation analyses with *Sinorhizobium* sp. M14 as a donor strain were performed. For both types of conjugation, two representatives of *Alphaproteobacteria,* closely related to *Sinorhizobium* sp., (*A. tumefaciens* LBA288 and *P. alcaliphilus* JCM7364R) and unrelated strains (*Aeromonas* spp., *Pseudomonas* spp., *Serratia spp*.,) which are unable to catalyze arsenite oxidation and are sensitive to As(III) (1mM of sodium arsenite inhibits the growth of both strains), were used as the recipients. Both (bi- and triparental) conjugation experiments yielded arsenite resistant transconjugant strains carrying plasmid pSinA with only *Alphaproteobacteria* (*A. tumefaciens* and *P. alcaliphilus*) representatives as recipient strains (the presence of the plasmid in tranconjugants was confirmed by electrophoresis and Southern hybridization – data not shown). This result

clearly showed that pSinA is a self-transferable plasmid that can be introduced into *Alphaprotebacteria* cells by conjugation even without a helper strain.

Two transconjugants, *A. tumefaciens* D10 and *P. alcaliphilus* C10, obtained by biparental conjugation were subjected to further analysis. An arsenite oxidation assay revealed that both strains had acquired the ability to oxidize arsenite. They grew aerobically in MSM with arsenite as the electron donor, and carbon dioxide (CO_2) or bicarbonate (HCO_3^-) as the carbon source (Fig. 3 ABC). Moreover, the two strains showed higher tolerance to arsenite than their parental, wild-type strains. *A. tumefaciens* LBA288 tolerated arsenite concentration levels up to 0.5mM, whereas its derivative with plasmid pSinA (*A. tumefaciens* D10) was able to grow in arsenite concentrations up to 15mM. In turn, *P. alcaliphilus* JCM7364R tolerated arsenite up to 0.8mM, whereas its derivative with plasmid pSinA (*P. alcaliphilus* C10) was able to grow in arsenite concentrations up to 20 mM. However, the introduction of pSinA into *Agrobacterium* sp. and *Paracoccus* sp. cells did not enhance their tolerance to arsenate or heavy metals. The wild-type strains (without pSinA) were able to grow efficiently in the presence of high concentrations of As(V) (250 mM), Cd (II) (2.5 mM), Co (II) (2.5 mM) and Zn(II) (5 mM), since most of them probably possess their own resistance mechanisms.

In order to explore the potential use of pSinA in the construction of new strains able to oxidize arsenite and useful in arsenic bioremediation, its stability in host cells was also tested. No plasmid-less cells of *A. tumefaciens* D10 and *P. alcaliphilus* C10 were detected after approximately 60 generations of growth under nonselective conditions (without arsenite selection). This result demonstrated the stable maintenance of pSinA and suggested that it could be employed in the laboratory-based construction of new strains useful in arsenic bioremediation as well as in plasmid-mediated bioaugmentation of microbial communities in arsenic polluted soils and sediments.

3.4. pSinA is a stably maintained, broad host range, self-transferable plasmid – conclusions from the microcosm release experiment

Plasmid pSinA transfer ability was also confirmed in the microcosm release experiment with the use of two independent microbial communities of naturally As-contaminated soils and *Sinorhizobium* sp. M14 as a donor strain. Transconjugants (carrying plasmid pSinA) were detected in both tested soil samples, and the number of transconjugants (after 60 days of incubation) reached 25-40% of arsenic oxidizing isolates (Fig. 4). The frequency of pSinA plasmid transfer differed between analyzed soil communities and doubtlessly it was connected with biogeochemical character of the selected soils. The number of natural transconjugants carrying the pSinA plasmid was relatively stable between days 30 and 60 (20-25% of total arsenite oxidizing bacteria) in the soil sample I (ZP) collected from the Zloty Potok, whereas increased progressively between days 30 and 60 (from 10% to 40% of total arsenite oxidizing bacteria) in the soil sample II (PT) collected from the Potok Trująca. Interestingly, diversity analysis of a pool of transconjugants revealed that transfer of plasmid pSinA was not restricted to the strains closely related to *Sinorhiozbium* sp. M14, as could be suggested by the results obtained from the traditional conjugation experiments (see chapter 3.3). Among recipient strains were representatives of phylogenetically distant taxonomic groups, including *Alphaproteobacteria* and *Gammaproteobacteria* (Table S3), which confirmed that plasmid pSinA is a broad host range replicon. Furthermore, analysis of the obtained transconjugants showed that pSinA is functional (expressed arsenite oxidase) and is stably maintained in their cells after approximately 60 generations of growth under nonselective conditions.

This clearly demonstrates that plasmid pSinA is a natural, self-transferable and broad host range vector that could be employed in the laboratory-based construction of novel strains useful in arsenic bioremediation (able to oxidize arsenite), as well as in plasmid-mediated bioaugmentation of microbial communities in arsenic polluted soils and sediments.

4. Discussion

Plasmids are used in both, basic research (e.g. cloning vectors) and biotechnology, e.g. bioremediation. There have been numerous applications of plasmids in bioremediation of contaminated environments (Bathe and Hausner, 2010), but only a few proposals for the use of plasmids carrying "arsenic genes". Among the first was that of Rawlings and co-workers (1984), who constructed and patented arsenic resistance vectors for *Acidithiobacillus ferroxidans.* The introduction of these plasmids into recipient cells enhanced their arsenic resistance, which is extremely important for biomining microbes, especially bacteria used for arsenopyrite oxidation. Most recently, the construction of a genetically engineered microorganism with a high tolerance to arsenite and strong arsenite oxidative ability was described by Yang and colleagues (2010). Their method relies on the introduction of a broad-host-range vector (pBBR) containing two *aioAB* structural genes, coding for arsenite oxidase (derived from *Thermus thermophilus* HB8), into the host cells. This procedure can increase the capacity for arsenite oxidation, but is limited to strains that are already capable of oxidizing As(III). Moreover, in addition to the limitations of the hosts, the vectors constructed in both of the aforementioned studies carry antibiotic resistance genes, which is undesirable due to the risk of their uncontrolled spread in the environment.

The natural, self-transferable plasmid pSinA described here represents an alternative to laboratory-constructed vectors. This is the first described plasmid that harbors an "arsenic gene island" with all the genes necessary for arsenite oxidation, and arsenic resistance. Removal of plasmid pSinA from M14 cells leads to deficiencies in arsenic [As(III) and As(V)] resistance and completely abolishes the ability to grow on minimal salt medium supplemented with sodium arsenite as the energy source.

We have suggested earlier (Drewniak et al. 2008) that *Sinorhizobium* isolate M14 has two mechanisms of adaptation to the high concentrations of arsenic occurring in the mine. It

can reduce arsenate to arsenite by means of the *Ars* system (detoxification only) and is able to simultaneously oxidize arsenite in a respiratory process which generates energy for growth. Calculated energy balance between arsenite oxidation (respiratory process) and reverse process of As(V) reduction in the detoxification process showed that almost 73% of the energy gained from the respiratory process in the absence of organic source of carbon may be consumed in the detoxification reaction.

Our experiments did not directly show which genes are responsible for arsenite oxidation, but functional analysis of *A. tumefaciens* and *P. alcaliphilus* strains carrying plasmid pSinA clearly demonstrated that the entire plasmid is necessary for this process. Both constructed strains acquired the ability to oxidize arsenite and their tolerance to As(III) was enhanced following the introduction of pSinA.

It is worth to mention that both constructed strains, in contrast to *Sinorhizobium* sp. M14 (the pSinA parent strain), completely oxidize arsenite and did not accumulate arsenite within their cell. In our previous work (Drewniak et al., 2008) we showed that the oxidation of arsenite to arsenate mediated by *Sinorhizobium* sp. M14, the pSinA parent strain, did not yield oxidizing equivalents at a ratio of 1:1 and the total arsenic level did not remain constant throughout the experiment. We showed that arsenic was accumulated in cells in form of electron dense granules.

Thus, these strains are an example of how the potential of plasmid pSinA may be exploited for the modification of indigenous strains, originally unable to oxidize arsenic.

Plasmid pSinA may also be applied in plasmid-mediated bioaugmentation, because it carries REP/STA and TRA modules, which enable its conjugal transfer without a helper strain and permits its stable maintenance and functioning in various host cells. It should be noted that both minireplicons as well as the entire plasmid are characterized by broad host range and are stably maintained in various hosts, which is extremely important in the context of the

use in infinitive latitude bioremediation. The use of bacterial strains that contain natural conjugal or mobilizable plasmid carrying the desired genes (e.g. catabolic or metal resistance genes) may help to solve survival problems of strains introduced into the environments. It is commonly known that most of the laboratory constructed (or studied) strains are not able to survive under natural, environmental conditions. In many field experiments, a rapid decline in the inoculum size or massive cell death have frequently been observed (El Fantroussi and Agathos 2005), and one of the best ways to avoid such problems is the application of the plasmid-mediated bioaugmenation approach.

The enhanced bioaugmentation caused by the transfer of conjugal plasmid from an introduced donor strain to indigenous microbial recipients has been demonstrated in the remediation of soil contaminated with toxic organic compounds, such as dichlorophenoxyacetic acid (2,4-D) (Top et al., 1998). Here, we showed that plasmid pSinA can also be used in bioaugmentation efforts of arsenic contaminated sites. In the microcosm experiment, we observed a successful lateral transfer of plasmid pSinA and arsenite oxidation genes from *Sinorhizobium* sp. M14 into microorganisms (from microbial communities) of naturally As-contaminated soils. In parallel with the passage of time and reduced viability of the M14 strain, the number of transconjugants carrying the plasmid pSinA increased. These results are very promising and important in the context of the bioremediation of arsenic contaminated regions. Most probably, we will be able to use M14 strain carrying pSinA in every arsenic contaminated environment, since in each of them there are indigenous microbes, that, in case of a rapid decline of inoculum (M14) size, could become the hosts for plasmid pSinA and could support the arsenite oxidation process.

Acknowledgements

This work was supported by the Polish Ministry of Science and Higher Education in the form of grant N N302 083639. We gratefully acknowledge Prof. D. Bartosik for providing plasmids and bacterial strains.

Competing interests

The authors declare that they have no competing interests

Authors' contribution

LDR carried out most of experiments (removal of plasmid pSinA, functional analysis of wild type and plasmid-less strains, annotation of the plasmid, constructions and functional analysis of minireplicons, removal of the plasmid, phenotypic analysis of selected strains, microcosm experiment), wrote the manuscript and was the head of the project; **LDZ** - supported the plasmid pSinA annotation and was involved in the manuscript preparation; **MC** - performed the construction and functional analysis of novel strains able to oxidize arsenite. **JG** and **RG** conducted plasmid DNA isolation, sequencing and assembling the plasmid. **AS** - is the head of the group and was involved in consultation and article preparation. All authors read and approved the final manuscript.

References

Bartosik, D., Bialkowska, A., Baj, J., Wlodarczyk M., 1997. Construction of mobilizable cloning vectors derived from pBGS18 and their application for analysis of replicator region of a pTAV202 mini-derivative of *Paracoccus versutus* pTAV1 plasmid. Acta Microbiol. Pol. 46, 387–392.

Bartosik, D., Baj, J., Wlodarczyk, M., 1998. Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. Microbiology 144, 3149–3157.

Bartosik, D., Baj, J., Piechucka, E., Waker, E., and Wlodarczyk, M., 2002. Comparative characterization of *repABC*-type replicons of *Paracoccus pantotrophus* composite plasmids. Plasmid 48, 130–141.

Bartosik, D., Szymanik, M., Wysocka, E., 2001. Identification of the partitioning site within the repABC-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. J. Bacteriol. 183, 6234–6243.

Bathe, S., and Hausner, M., 2010. Plasmid-mediated bioaugmentation of wastewater microbial communities in a laboratory-scale bioreactor. Methods Mol. Biol. 599, 185–200.

Battaglia-Brunet, F., Dictor, M.C., Garrido, F., Crouzet, C., Morin, D., Dekeyser, K., Clarens, M. and Baranger, P., 2002. An arsenic (III)-oxidizing bacterial population: selection, characterization, and performance in reactors. J. Appl. Microbiol. 93, 1–12.

Birnboim, H.C., and Doly, J., 1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* 7, 1513–1523.

Branco, R., Francisco, R., Chung, A.P., Morais, P.V., 2009. Identification of an aox system that requires cytochrome c in the highly arsenic-resistant bacterium *Ochrobactrum tritici* SCII24. Appl. Environ. Microbiol. 75, 5141–5147.

Burgos, P.A., Velázquez, E., and Toro, N., 1996. Identification and distribution of plasmid-type A replicator region in rhizobia. Mol. Plant. Microbe. Interact. 9, 843–849.

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.

Carver, T., Thomson, N., Bleasby, A., Berriman, M., Parkhill, J., 2009. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 25, 119–120.

Connon, S.A., Koski, A.K., Neal, A.L., Wood, S.A., Magnuson, T.S., 2008. Ecophysiology and geochemistry of microbial arsenic oxidation within a high arsenic, circumneutral hot spring system of the Alvord Desert. FEMS Microbiol. Ecol. 64, 117-128.

Drewniak, L., Matlakowska, R., Sklodowska, A., 2008. Arsenite and arsenate metabolism of *Sinorhizobium* sp. M14 living in the extreme environment of the Zloty Stok gold mine. Geomicrobiol. J. 25, 363–370.

Drewniak, L., 2009. Characterization of arsenic bacteria isolated from Zloty Stok gold mine. PhD thesis. University of Warsaw, Laboratory of Environmental Pollution Analysis, Warsaw.

Drewniak, L., Matlakowska, R., Rewerski, B., Sklodowska, A., 2010. Arsenic release from gold mine rocks mediated by the activity of indigenous bacteria. Hydrometallurgy 104, 437-442.

Drewniak, L., Maryan, N., Lewandowski, W., Sklodowska, A., 2012. The contribution of microbial mats to the arsenic geochemistry of an ancient gold mine. Environ. Pollut. 162, 190-201

Duquesne, K.; Lebrun, S.; Casiot, C.; Bruneel, O.; Personne, J. C.; Leblanc, M.; Elbaz-Poulichet, F.; Morin, G.; Bonnefoy, V., 2003. Immobilization of arsenite and ferric iron by *Acidithiobacillus ferrooxidans* and its relevance to acid mine drainage. Appl. Environ. Microbiol. 69, 6165–6173.

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

El Fantroussi, S., and Agathos, S.N., 2005. Is bioaugmentation a feasible strategy for pollutant removal and site remediation? Curr. Opin. Microbiol. 8, 268–275.

Garcillán-Barcia, M.P., Francia, M.V., de la Cruz, F., 2009. The diversity of conjugative relaxases and its application in plasmid classification. FEMS Microbiol. Rev. 33, 657–687.

Hooykaas, P.J.J., den Dulk-Ras, H., and Schilperoort, R.A., 1980. Molecular mechanism of Ti plasmid mobilization by R plasmids: isolation of Ti plasmids with transposon-insertions in *Agrobacterium tumefaciens*. Plasmid 4, 64–75.

Hynes, M. F., and McGregor, N.F., 1990. Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. Mol. Microbiol. 4, 567–574.

Kashyap, D.R., Botero, L.M., Franck, W.L., Hassett, D.J., McDermott, T.R., 2006. Complex regulation of arsenite oxidation in *Agrobacterium tumefaciens*. J. Bacteriol. 188, 1081–1088.

Kononova, S.V., and Nesmeyanova, M.A., 2002. Phosphonates and their degradation by microorganisms. Biochemistry 67, 184–195.

Kushner, S.R., 1978) An improved method for transformation of *E. coli* with ColE1 derived plasmids, p. 17–23. In H. B. Boyer and S. Nicosia (ed.), Genetic engineering. Elsevier/North-Holland, Amsterdam, Netherlands.

Lane, D.J., 1991. 16S/23S rRNA sequencing. In: Stackebrandt, E., Goodfellow, M. (Eds.), Nucleic Acid Techniques in Bacterial Systematics. Wiley, New York, pp. 115-175.

Lett, M.C., Muller, D., Lièvremont, D., Silver, S., Santini J. 2012. Unified nomenclature for genes involved in prokaryotic aerobic arsenite oxidation. J. Bacteriol. 194, 207-208.

Mankin, S.L., Hill, D.S., Olhoft, P.M., Toren, E., Wenck, A.R., Nea, L., et al., 2007. Disarming and sequencing of *Agrobacterium rhizogenes* strain K599 (NCPPB2659) plasmid pRi2659. In Vitro Cell Dev. Biol. - Plant 43, 521–535.

Mukhopadhyay, R., Rosen, B.P., Phung, L.T., and 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., et al., 2007. A tale of two oxidation states: bacterial colonization of arsenic-rich environments. PLoS Genet. 3, E53

Nies, D.H., 2003) Efflux-mediated heavy metal resistance in prokaryotes. FEMS Microbiol. Rev. 27, 313–339

Osborne, T.H, Jamieson, H.E., Hudson-Edwards, K.A., Nordstrom, D.K., Walker, S.R., Ward, S.A., Santini J.M., 2010. Microbial oxidation of arsenite in a subarctic environment: diversity of arsenite oxidase genes and identification of a psychrotolerant arsenite oxidiser. BMC Microbiology 10, 205

Padan, E., Tzubery, T., Herz, K., Kozachkov, L., Rimon, A., Galili, L., 2004. NhaA of *Escherichia coli*, as a model of a pH-regulated Na^+/H^+ antiporter. Biochim. Biophys. Acta 1658, 2–13.

Paez-Espino, D., Tamames, J., de Lorenzo, V., Canovas, D., 2009. Microbial responses to environmental arsenic. Biometals 22, 117-130.

Quemeneur, M., Cébron, A., Billard, P., Battaglia-Brunet, F., Garrido, F., Leyval, C., Joulian, C., 2010. Population structure and abundance of arsenite-oxidizing bacteria along an arsenic pollution gradient in waters of the upper isle River Basin, France. Appl. Environ. Microbiol. 76, 4566-4570.

Rawlings, D.E., Pretorius, I.M., and Woods, D.R., 1984. Construction of arsenic-resistant *Thiobacillus ferrooxidans* recombinant plasmids and the expression of autotrophic plasmid genes in a heterotrophic cell-free system. J. Bacteriol. 1, 129–133.

Remenant, B., Coupat-Goutaland, B., Guidot, A., Cellier, G., Wicker, E., Allen, C., et al., 2010. Genomes of three tomato pathogens within the *Ralstonia solanacearum* species complex reveal significant evolutionary divergence. BMC Genomics 11, 379.

Rhine, E.D., Onesios, K.M., Serfes, M.E., Reinfelder, J.R., Young, L.Y., 2008. Arsenic transformation and mobilization from minerals by the arsenite oxidizing strain WAO. Environ. Sci. Techol. 42, 1423-1429.

Rhine, E.D., Phelps, C.D., Young, L.Y., 2006. Anaerobic arsenite oxidation by novel denitrifying isolates. Environ. Micriobiol. 8, 899-908

Sambrook, J., and Russell, D.W., 2001. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Santini, J.M., and van den Hoven, R.N., 2004. Molybdenum-containing arsenite oxidase of the chemolithoautotrophic arsenite oxidizer NT-26. J. Bacteriol. 186, 1614–1619.

Silver, S., and Phung, L.T., 2005a. Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. Appl. Environ. Microbiol. 71, 599–608.

Silver, S., and Phung, L.T., 2005b. A bacterial view of the periodic table: genes and proteins for toxic inorganic ions. J. Ind. Microbiol. Biotechnol. 32, 587–605.

Simeonova, D.D., Micheva, K., Muller, D.A., Lagarde, F., Lett, M.C., Groudeva, V.I., Lièvremont, D., 2005. Arsenite oxidation in batch reactors with alginate-immobilized ULPAs1 strain. Biotechnol. Bioeng. 91, 441–446.

Starkenburg, S.R., Larimer, F.W., Stein, L.Y., Klotz, M.G., Chain, P.S., Sayavedra-Soto, L.A., et al., 2008. Complete genome sequence of *Nitrobacter hamburgensis* X14 and comparative genomic analysis of species within the genus *Nitrobacter*. Appl. Environ. Microbiol. 74, 2852–2863.

Stiens, M., Schneiker, S., Keller, M., Kuhn, S., Pühler, A., Schlüter, A., 2006. Sequence analysis of the 144-kilobase accessory plasmid pSmeSM11a, isolated from a dominant *Sinorhizobium meliloti* strain identified during a long-term field release experiment. Appl. Environ. Microbiol. 72, 3662–3672.

Stiens, M., Schneiker, S., Keller, M., Pühler, A., Schlüter, A., 2007. Sequence analysis of the 181-kb accessory plasmid pSmeSM11b, isolated from a dominant *Sinorhizobium meliloti* strain identified during a long-term field release experiment. FEMS Microbiol. Lett. 271, 297–309.

Top, E.M., Van Daele, P., De Saeyer, N., Forney L.J., 1998. Enhancement of 2,4-dichlorophenoxyacetic acid (2,4-D) degradation in soil by dissemination of catabolic plasmids. Antonie Leeuwenhoek 73, 87–94.

World Health Organization. 2003.Arsenic in drinking water. http://www.who.int/mediacentre/factsheets/fs210/en/index.html

Yang, C., Xu, L., Yan, L., Xu, Y., 2010. Construction of a genetically engineered microorganism with high tolerance to arsenite and strong arsenite oxidative ability. J. Environ. Sci. Health A Tox. Hazard. Subst. Environ. Eng. 45, 732–737.

Young, J.P., Crossman, L.C., Johnston, A.W., Thomson, N.R., Ghazoui, Z.F., Hull, K.H et al. 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biol 7, R34.

Fig 1.

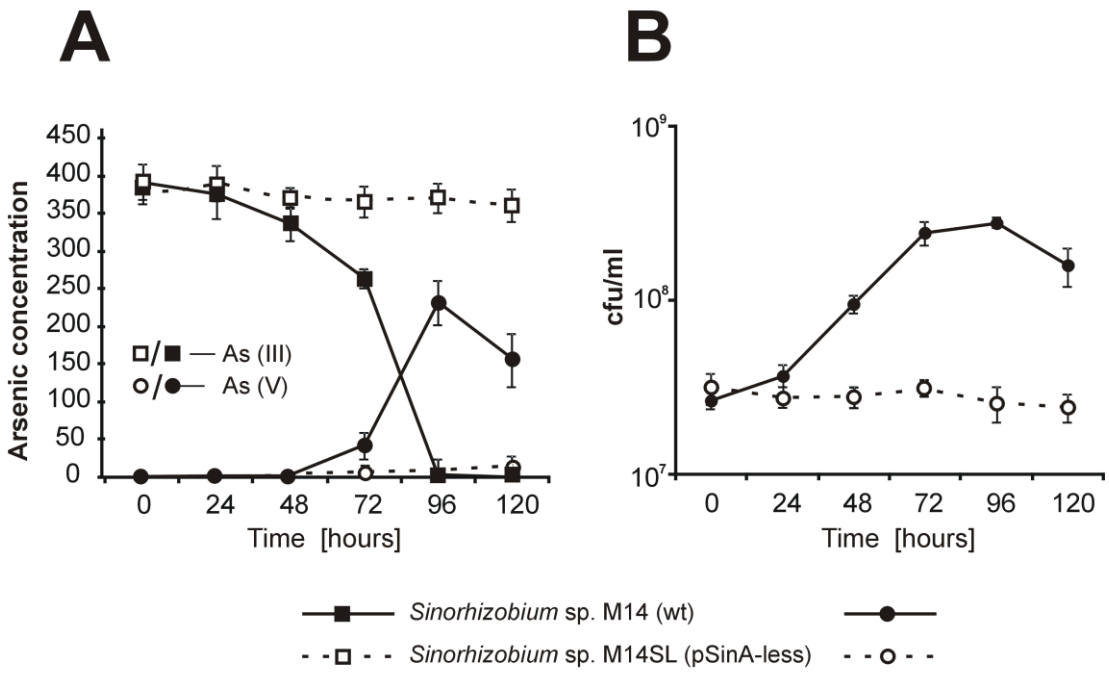

Figure 1. Arsenite oxidation and growth of *Sinorhizobium* sp. M14 (wild type) and *Sinorhizobium* sp. M14SL (pSinA-less) strains in minimal medium containing 5 mM (375 ppm) sodium arsenite. (**A**) arsenite and arsenate concentration and (**B**) colony forming units per milliliter (cfu/ml) were measured every 24 hours for 5 days.

Fig 2.

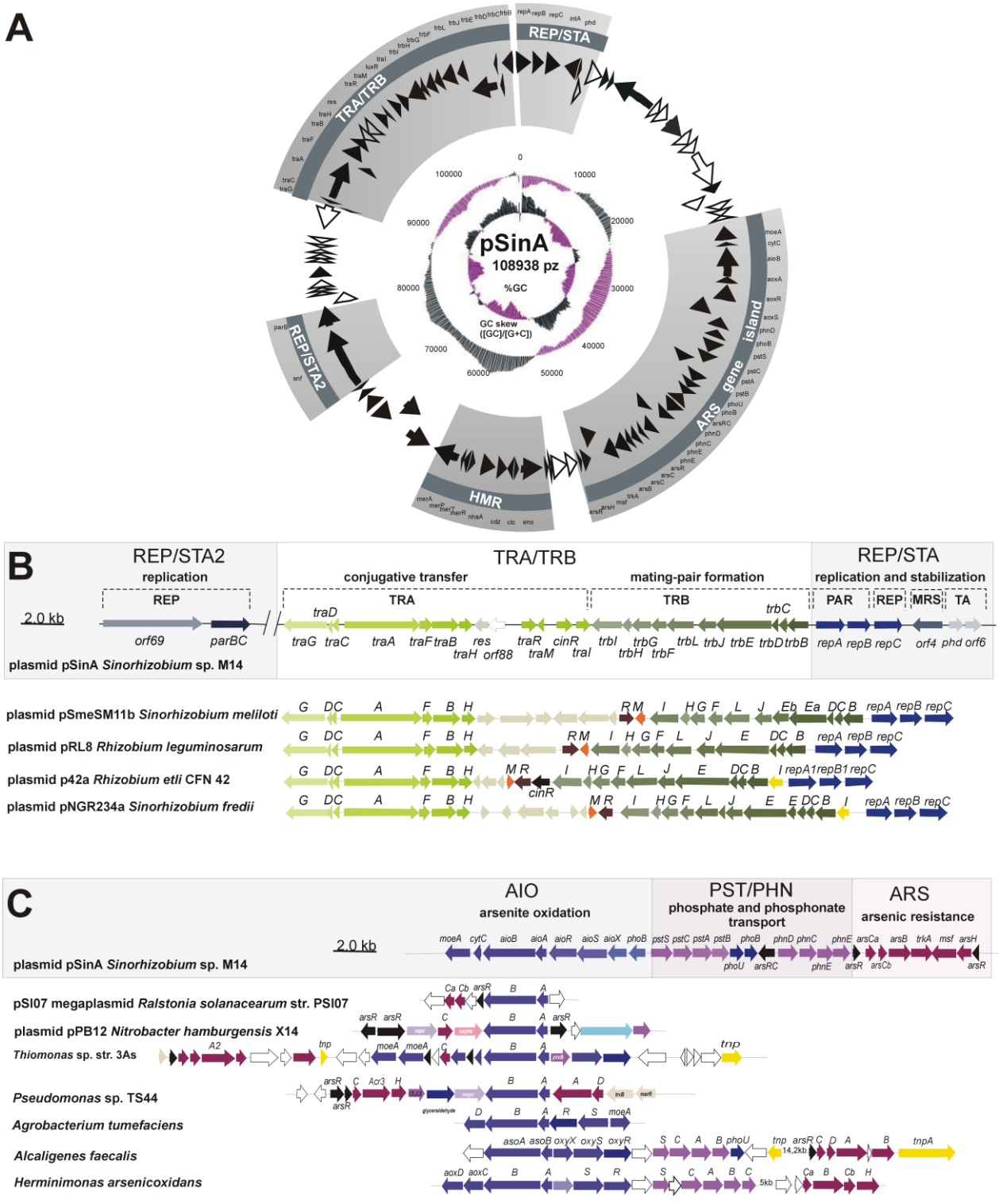

Figure 2. A - Genetic map and organization of plasmid pSinA of *Sinorhizobium* sp. M14.

Arrows indicate the location and orientation of coding regions that are numbered clockwise from 1 to 103. The relative G+C content and the GC skew [(GC)/(G+C)] are shown on the

two inner circles, where a G+C content of >50% is indicated in black and a G+C content of <50% is indicated in violet. G/C plots were generated using DNA Plotter (Carver et al., 2009). The middle circle gives the scale of the plasmid in base pairs (bps). The first nucleotide of the *repA*-coding region was chosen as the start point for numbering the pSinA nucleotide sequence. Plasmid modules are shown as gray blocks: (i) REP/STA and REP/STA2 – regions involved in plasmid replication and stabilization; (ii) ARS genomic island – region encoding arsenic metabolism genes; (iii) HMR – a heavy metal resistance region; and (iv) TRA/TRB – region involved in conjugative transfer. **B** – **organization of the plasmid backbone**. Arrows indicate the location and orientation of coding regions of the plasmid backbone modules: REP/STA (replication and stabilization), TRA/TRB (conjugation) and REP/STA2 (replication). Dashed brackets indicate the systems located within the backbone modules: TA (toxin-antitoxin stabilization system), MRS (multimers resolution system), REP (replication system), PAR (partitioning system), TRB (mating-pair formation system), TRA (conjugative transfer system). **C – organization of arsenic genomic island.** Arrows indicate the location and orientation of coding regions of: AIO (arsenite oxidation), PST/PHN (arsenic/phosphorus transport) and ARS (arsenic resistance genes) modules.

Fig 3.

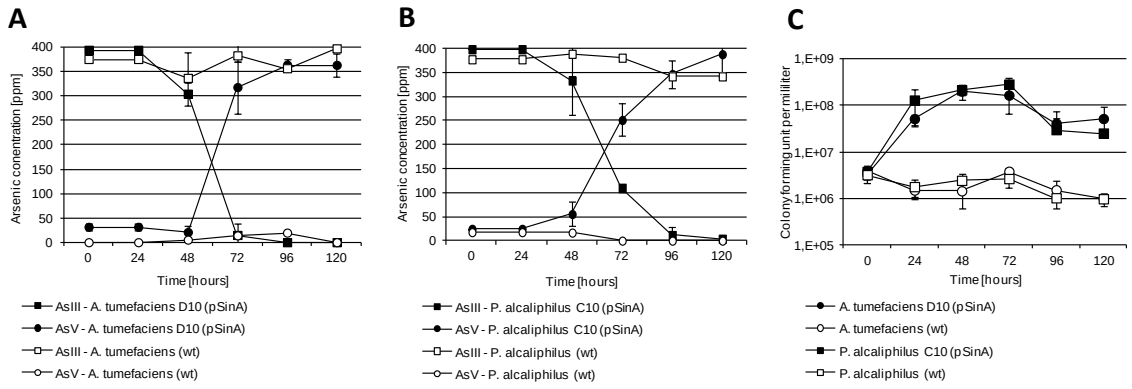

Figure 3. Comparison of arsenite metabolism and tolerance to As (III) of wild-type *A. tumefaciens* LBA288 and *P. alcaliphilus* JCM7364R, and their derivatives carrying plasmid pSinA. To examine arsenite oxidation, cultures were grown in minimal salt medium containing 5 mM (375 ppm) sodium arsenite at 22°C. Arsenite and arsenate concentration in culture supernatants of (**A**) *A. tumefaciens* LBA288 and D10, (**B**) *P. alcaliphilus* JCM7364R and C10, as well as (**C**) colony forming units per milliliter (cfu/ml) were measured every 24 hours for 5 days.

Fig 4.

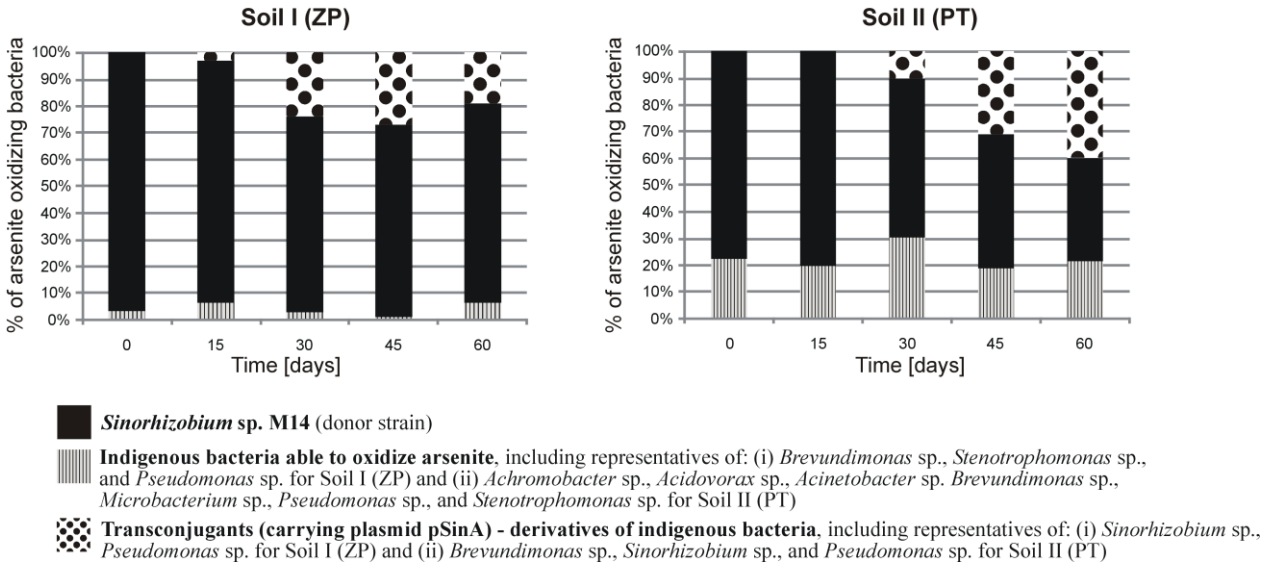

Figure 4. Transfer frequencies of plasmid pSinA from *Sinorhizobium* sp. M14 to As-contaminated soil bacteria. Two independent microbial communities from soil samples collected from different locations (ZP -Zloty Potok and PT - Potok Trujaca) near the ancient arsenic/gold mine in Zloty Stok (SW Poland) were used in bioaugmentation experiments. Soil samples were enriched with *Sinorhizobium* sp. M14 (donor of plasmid pSinA) and were monitored during two-month period. The analysis refers to the cultivable fraction of the community and relied on an estimation of the number of transconjugants carrying plasmid pSinA. Isolation of presumptive transconjugants was performed in aerobic conditions on MSM medium supplemented with 5mM sodium arsenite. After 7 days of incubation at 22°C the test with 0.1M AgNO3 solution was used. For identification of transconjugants carrying plasmid pSinA multiplex-PCR and Southern hybridization were used. Detailed information about identified indigenous bacteria and transconjugants able to oxidize arsenite are presented in Table S3. One hundred randomly selected colonies grown on the selective medium had been considered for statistical analysis.