

Research article

Open Access

## Phylogenetic analysis of bacterial and archaeal *arsC* gene sequences suggests an ancient, common origin for arsenate reductase

Colin R Jackson\* and Sandra L Dugas

Address: Department of Biological Sciences, SLU 10736, Southeastern Louisiana University, Hammond, LA 70402, USA

Email: Colin R Jackson\* - [cjackson@selu.edu](mailto:cjackson@selu.edu); Sandra L Dugas - [sdugas@selu.edu](mailto:sdugas@selu.edu)

\* Corresponding author

Published: 23 July 2003

Received: 20 May 2003

*BMC Evolutionary Biology* 2003, **3**:18

Accepted: 23 July 2003

This article is available from: <http://www.biomedcentral.com/1471-2148/3/18>

© 2003 Jackson and Dugas; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

### Abstract

**Background:** The *ars* gene system provides arsenic resistance for a variety of microorganisms and can be chromosomal or plasmid-borne. The *arsC* gene, which codes for an arsenate reductase is essential for arsenate resistance and transforms arsenate into arsenite, which is extruded from the cell. A survey of GenBank shows that *arsC* appears to be phylogenetically widespread both in organisms with known arsenic resistance and those organisms that have been sequenced as part of whole genome projects.

**Results:** Phylogenetic analysis of aligned *arsC* sequences shows broad similarities to the established 16S rRNA phylogeny, with separation of bacterial, archaeal, and subsequently eukaryotic *arsC* genes. However, inconsistencies between *arsC* and 16S rRNA are apparent for some taxa. Cyanobacteria and some of the  $\gamma$ -Proteobacteria appear to possess *arsC* genes that are similar to those of Low GC Gram-positive Bacteria, and other isolated taxa possess *arsC* genes that would not be expected based on known evolutionary relationships. There is no clear separation of plasmid-borne and chromosomal *arsC* genes, although a number of the Enterobacteriales ( $\gamma$ -Proteobacteria) possess similar plasmid-encoded *arsC* sequences.

**Conclusion:** The overall phylogeny of the arsenate reductases suggests a single, early origin of the *arsC* gene and subsequent sequence divergence to give the distinct *arsC* classes that exist today. Discrepancies between 16S rRNA and *arsC* phylogenies support the role of horizontal gene transfer (HGT) in the evolution of arsenate reductases, with a number of instances of HGT early in bacterial *arsC* evolution. Plasmid-borne *arsC* genes are not monophyletic suggesting multiple cases of chromosomal-plasmid exchange and subsequent HGT. Overall, *arsC* phylogeny is complex and is likely the result of a number of evolutionary mechanisms.

### Background

Arsenic is a toxic element that is found in both natural environments such as geothermal springs, and in sites contaminated by a number of industries. Inorganic arsenic exists primarily in two valence states: arsenite ( $\text{As}_{\text{III}}$ ) and arsenate ( $\text{As}_{\text{V}}$ , or the arsenate ion  $\text{AsO}_4^{3-}$ ). Both forms are toxic to microorganisms, with arsenite disrupt-

ing enzyme function, and arsenate behaving as a phosphate analog and interfering with phosphate uptake and utilization [1]. Microorganisms have evolved a variety of mechanisms for coping with arsenic toxicity, including minimizing the amount of arsenic that enters the cell (e.g. through increased specificity of phosphate uptake, [2]), arsenite oxidation through the activity of arsenite oxidase

[3,4], or peroxidation reactions with membrane lipids [5,6]. Other microorganisms utilize arsenic in metabolism, either as a terminal electron acceptor in dissimilatory arsenate respiration [7–10] or as an electron donor in chemoautotrophic arsenite oxidation [11,12]. However, the most well-characterized microbial arsenic detoxification pathway involves the *ars* operon [2,13].

The *ars* operon consists of a group of genes coding for a transmembrane pump and an arsenate reductase (*arsC*). The operon includes a regulatory gene (*arsR*) and a gene coding for an arsenite-specific pump (*arsB*) as well as *arsC* [13]. Arsenite is pumped directly out of the cell by the *arsB* protein; however arsenate must first be reduced to arsenite by the cytoplasmic arsenate reductase coded from *arsC*. Some bacteria also possess other *ars* genes: *arsA* produces an arsenite-stimulated ATPase [14] that results in more efficient arsenite extrusion; *arsD* encodes for a regulatory protein that controls the upper level of *ars* expression [15]; *arsH* has been identified but has an uncertain function [16]. The *ars* operon was initially recognized in plasmids of *Staphylococcus aureus* and *S. xylosus* [17,18] but has subsequently been found in other microorganisms (e.g. *Escherichia coli* [19,20], *Acidiphilum multivorum* [21], *Bacillus subtilis* [22], *Pseudomonas aeruginosa* [23]). The genes can be plasmid-borne or chromosomal, and genome-sequencing projects have identified putative *ars* genes in both Bacteria and Archaea that have not been specifically characterized in terms of arsenic resistance. An analogous genetic system (*Arr* or *ACR*) has also been described in the eukaryote *Saccharomyces cerevisiae* [24]. Thus the *ars* operon, or similar arsenic resistance systems, appears to be relatively widespread throughout microorganisms.

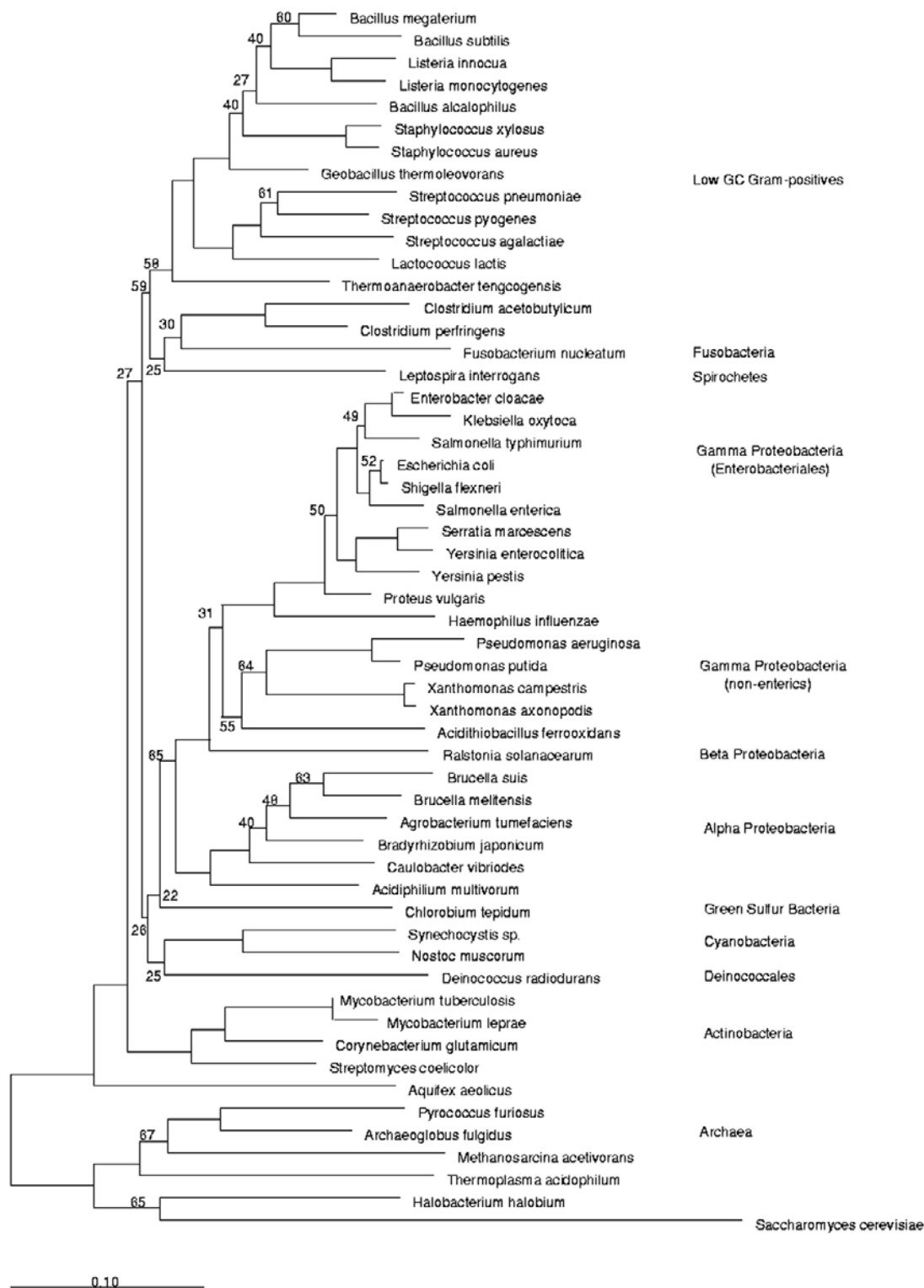
The *arsC* gene is of particular interest in that its product, the soluble enzyme arsenate reductase, catalyzes the reduction of arsenate to arsenite. Arsenate is the thermodynamically favorable form of arsenic under aerobic conditions [25,26], so it is likely to be the most common form of arsenic in many environments. Thus, the presence and expression of *arsC* is likely to be required for microorganisms inhabiting such areas. Furthermore, arsenite is generally more labile and toxic than arsenate [26], so that expression of *arsC*, and the ensuing reduction of arsenate to arsenite, might increase the toxicity of arsenic in the environment. Despite the potential importance of the *arsC* gene at both a physiological and environmental level, a thorough study of the phylogenetic distribution of different *arsC* genes has not been performed.

A simple phylogenetic tree showing the relationships between the arsenate reductases of seven Bacteria (those that had been confirmed to possess the gene at that time) was presented as part of a study that identified the arsenic resistance genes of *Thiobacillus ferrooxidans* (now *Acidithio-*

*bacillus ferrooxidans*) [27]. Saltikov and Olson [28] used probes/primers based on the *E. coli ars* operon to detect *ars* genes in natural environments, and presented a phylogenetic analysis of these genes, however their analysis was largely confined to enteric bacteria (those that were similar enough to *E. coli* to be detected) and they emphasized the *arsB* gene rather than *arsC*. As part of a review of microbial arsenic transformations, we recently determined a preliminary *arsC* phylogeny based on 19 sequences [29]. These three studies reported phylogenies that suggested a common origin for the *ars* genes examined [27–29]. An alternative view suggested by Mukhopadhyay et al. [30] is that three distinct classes of arsenate reductases exist, which have developed similar mechanisms and reaction centers through convergent evolution. Two of these classes are bacterial (one typified by the *arsC* found in enteric bacteria such as *E. coli*; the other typified by the *arsC* found in *Staphylococcus* plasmids and other Gram-positive Bacteria); the third is the *Arr2* gene of *S. cerevisiae*. Phylogenies of these gene families based on 18 *arsC* sequences were recently reported [30]. However, as well as microorganisms that have been shown to have demonstrable arsenate resistance, whole genome analyses of various microorganisms have revealed an increasing number of open reading frames (ORFs) that have homology to *arsC*. Inclusion of these putative *arsC* genes in phylogenetic analyses might resolve the basic question: Did arsenate resistance systems evolve from a common origin or develop convergently in multiple taxa? In this study we attempt to answer this question through a phylogenetic analysis of both confirmed *arsC* genes and those inferred from whole genome analyses. We compare this phylogeny to one derived from 16S rRNA, and suggest possible implications for the evolutionary history of arsenic resistance in microorganisms.

## Results and Discussion

The 16S rRNA based phylogenetic tree (Figure 1) generally matched the accepted 16S rRNA model with clear separation of the two prokaryotic domains and the subsequent divergence of Eukarya (represented by *Saccharomyces cerevisiae*) from Archaea [31]. Minor inconsistencies such as *Ralstonia solanacearum* ( $\beta$ -Proteobacteria) forming a deep branch within the  $\gamma$ -Proteobacteria, or *Fusobacterium nucleatum* (Fusobacteria) and *Leptospira interrogans* (Spirochetes) grouping loosely with the Gram-positive Bacteria are likely the result of the limited data set used (they were the only representatives of their respective phyla that had listed *arsC* genes) and do not challenge established relationships. Otherwise the major divisions and subdivisions of the Bacteria (e.g. the various Proteobacteria, Low GC Gram-positives, Actinobacteria) formed expected patterns. An interesting finding was the grouping (albeit weak) of *Nostoc muscorum* and *Synechocystis* sp. with *Deinococcus radiodurans* to form a high level clade (Figure 1).



**Figure 1**  
**Phylogenetic tree based on 16S rRNA gene sequences.** The tree was constructed by neighbor joining methods using 1326 informative positions. Numbers represent percentages of 1000 bootstraps and are only shown for bootstrap values <80%. The names of the major divisions of prokaryotes represented are shown to the right of corresponding clades.

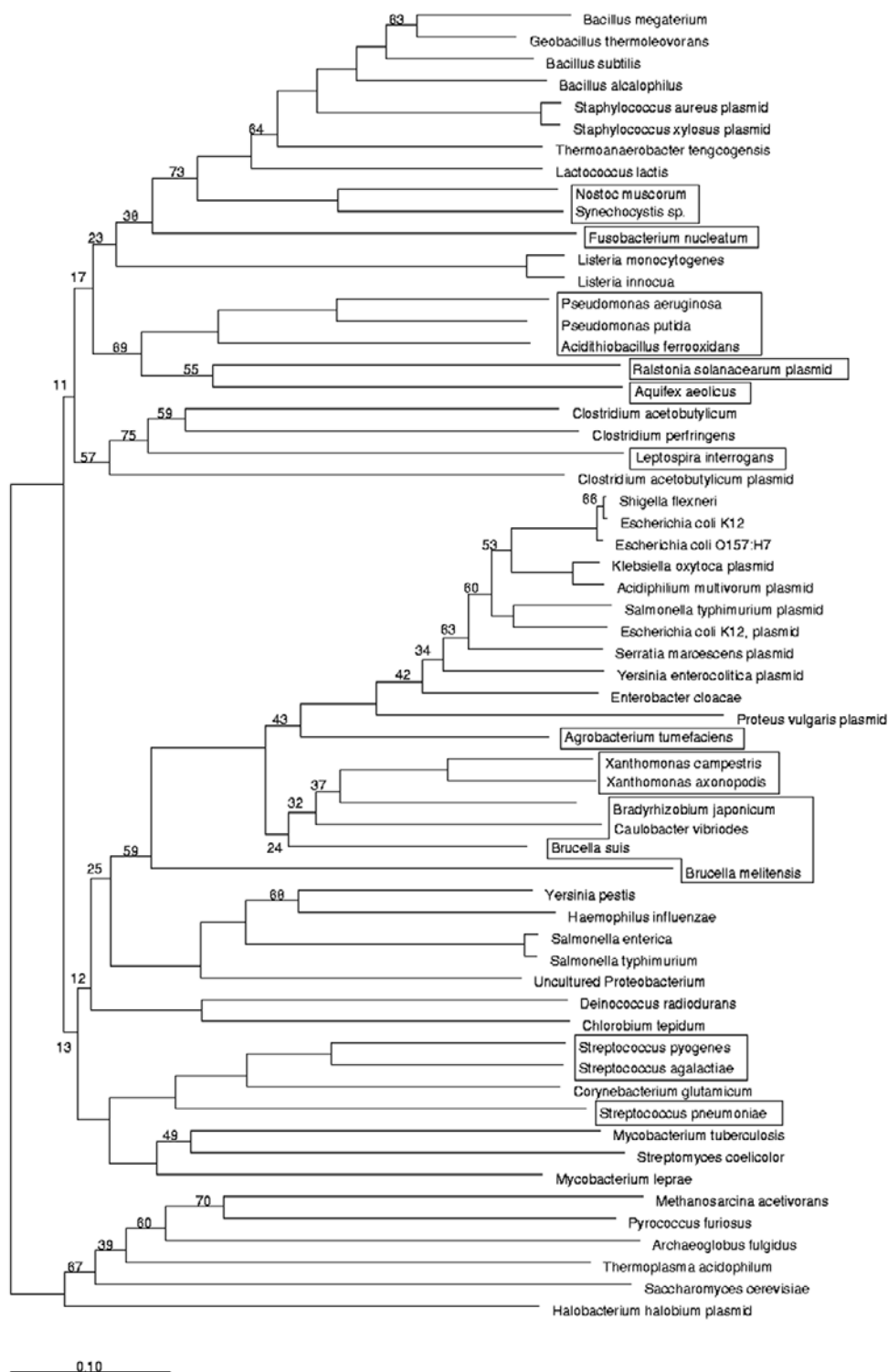
This is similar to findings from genome trees that suggest a close relationship between the Cyanobacteria and Deinococcales [32], although our 16S rRNA tree does not suggest a relationship between these taxa and the Actinobacteria. However, the Actinobacteria (High GC Gram-positives) did clearly separate from the Low GC Gram-positive Bacteria (Figure 1).

The *arsC* phylogenies obtained by evolutionary distance (ED) analysis (Figure 2) and maximum parsimony (MP; Figure 3) showed a number of broad similarities to the 16S rRNA tree. There was clear separation of Archaea and Bacteria (with the sole eukaryote, *S. cerevisiae*, grouping towards Archaea), and also general grouping of the different bacterial divisions; all the archaeal *arsC* genes were from the Euryarchaeota so it was not possible to examine division level phylogeny for Archaea. Both treeing methods support the existence of at least three major classes of *arsC* genes, corresponding to the Archaea/Eukarya, the Enterobacteriales (enteric  $\gamma$ -Proteobacteria) and  $\alpha$ -Proteobacteria, and the Low GC Gram-positive Bacteria. At a basic level, these broad groupings of *arsC* correspond to the three distinct classes of arsenate reductases that others have observed [30], and these three groups had low sequence similarity to each other (less than 33% similarity between the Bacteria and Archaea/Eukarya, and 48% similarity between the two major bacterial groups). However, the analysis reported here includes more diverse *arsC* sequences and suggests that other deeply branching types of *arsC* exist. For example, the *arsC* sequences from major divisions of Bacteria such as the Green Sulfur Bacteria (represented by *Chlorobium tepidum*) and the Deinococcales (represented by *D. radiodurans*) are loosely associated with either the Enterobacteriales/ $\alpha$ -Proteobacteria (Figure 2) or Low GC Gram-positive Bacteria (Figure 3) depending upon the analysis used, but in either case diverge to form their own deep branches, suggesting that they possess distinct arsenate reductases. Thus, the suggestion that the three previously reported classes of arsenate reductases developed through convergent evolution [30] seems flawed: not only does *arsC* phylogeny show broad parallels to the accepted 16S rRNA phylogeny, but deep bacterial divisions appear to possess the distinct arsenate reductases that would be expected from divergence from a common origin. Given the alternate hypotheses of (1) Common origin followed by sequence divergence, or (2) Independent origin of multiple *arsC* sequence types with broadly similar reaction centers and general mechanisms, based on the phylogenies observed and the general rarity of convergent evolution at a sequence level [33,34] we suggest that the former is more likely.

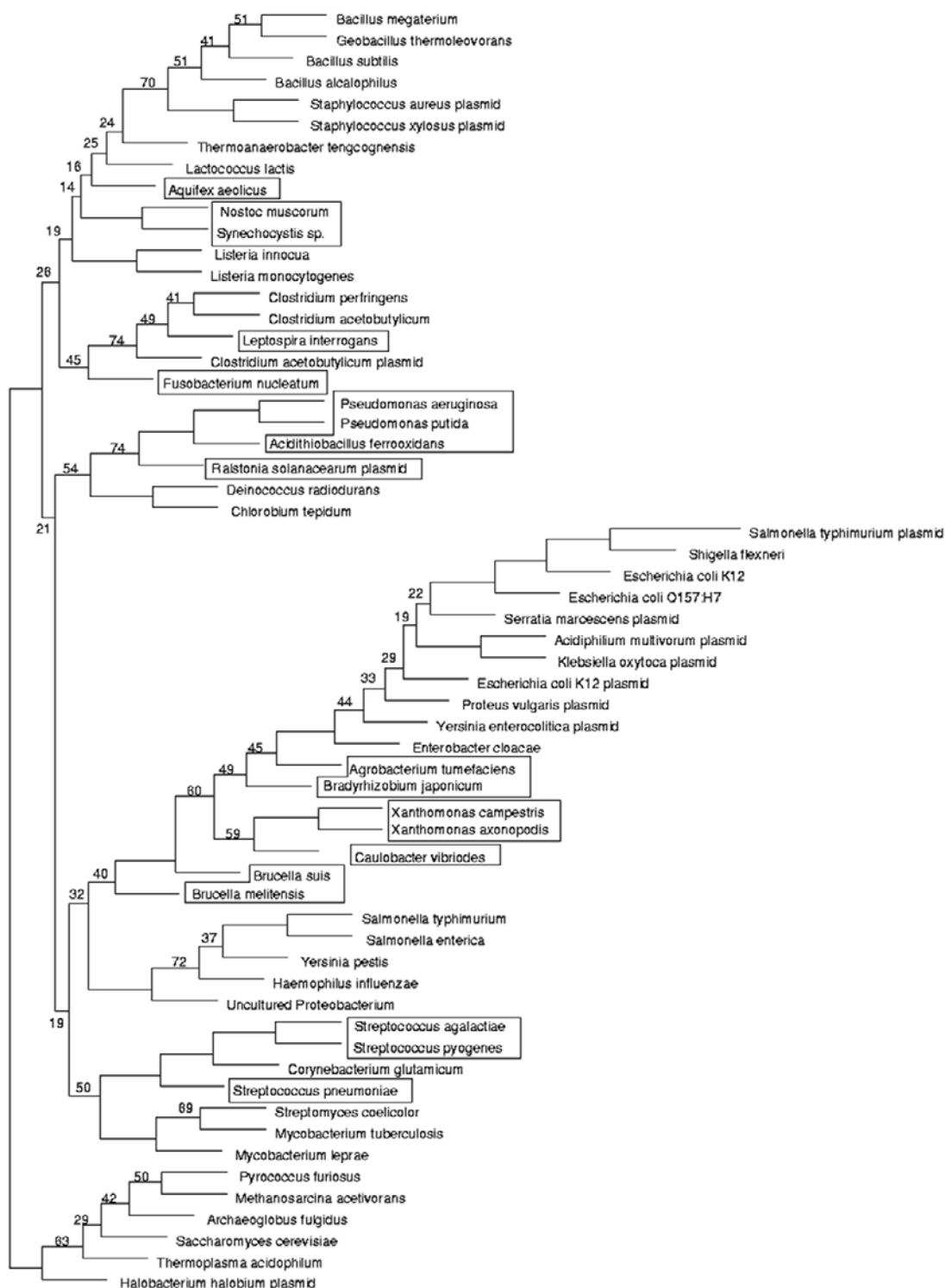
A number of taxa do show *arsC* phylogenies that are inconsistent with the established 16S rRNA evolutionary tree, perhaps the most dramatic being the separation of

some of the non-enteric  $\gamma$ -Proteobacteria (*Pseudomonas putida*, *P. aeruginosa*, and *A. ferrooxidans*) from the Enterobacteriales. The non-enteric  $\gamma$ -Proteobacteria actually appear to be paraphyletic with regards to *arsC*, in that the two *Xanthomonas* species group with the  $\alpha$ -Proteobacteria, whereas the others form a distinct group with the sole representative of the  $\beta$ -Proteobacteria (*R. solanacearum*). The phylogenetic affiliation of the latter clade is uncertain: ED analysis placed it as a deep branch in the Low GC Gram positive Bacteria group (Figure 2), whereas MP placed it loosely with the large Enterobacteriales/ $\alpha$ -Proteobacteria group (Figure 3). Either placement had poor bootstrap support, and given the relatively low sequence homology of this clade to either of the other major bacterial *arsC* groups (50% similarity to the Low GC Gram positives, 40% similarity to the Enterobacteriales/ $\alpha$ -Proteobacteria group) we suspect that this group might represent a fourth distinct group of arsenate reductases, making its placement in a binary tree difficult. A second distinct difference between the *arsC* and 16S rRNA trees is the tight grouping of the Cyanobacteria with the *arsC* sequences from Low GC Gram-positive Bacteria. Other discrepancies are placement of the three *Streptococcus* sequences (Low GC Gram-positive Bacteria) with the Actinobacteria and the placement of *Aquifex aeolicus* (basal in the 16S rRNA-defined bacterial clade) and *F. nucleatum* within the Low GC Gram-positives (Figures 2 and 3). The latter two sequence types were the only *arsC* sequences reported for the Aquificales and Fusobacteria, respectively, and as with some of the 16S rRNA tree branches, might represent the limited data set rather than true evolutionary relationships. However, *A. aeolicus* is known to have exchanged numerous genes with other microorganisms [35], and the phylogeny of its arsenate reductase gene might represent another example of this.

Regardless of the particular phylogenetic methods used, the *arsC* trees clearly show a number of non-orthologous *arsC* genes, assuming that the 16S rRNA tree reflects true phylogeny. One explanation for these inconsistencies could be the existence of *arsC* paralogs that have diverged after *arsC* gene duplication events. While this might be the case for those organisms with *arsC* sequences that were not closely related to other sequences (e.g. *D. radiodurans* or *C. tepidum*), it is an unreasonable explanation for the similar *arsC* sequences (e.g. those of the two Cyanobacteria) that group within unexpected larger clades. Rather than multiple gene duplication events or convergent evolution, this suggests horizontal gene transfer (HGT) of *arsC* genes. Recent studies suggest that HGT events may have been much more widespread during prokaryotic evolution than had been previously thought, with genetic exchange even occurring between Bacteria and Archaea [35–37]. While we see no evidence of cross-domain transfer of *arsC* genes, a number of HGT events could explain



**Figure 2**  
**Evolutionary distance tree based on *arsC* gene sequences.** The tree was constructed by neighbor joining methods using 408 informative positions. Numbers represent percentages of 1000 bootstraps and are only shown for bootstrap values <80%. Plasmid-borne *arsC* genes are indicated by "plasmid" following the organism name. Names in boxes represent branches that are inconsistent with the 16S rRNA tree.



**Figure 3**  
**Maximum parsimony tree based on *arsC* gene sequences.** The tree was constructed using the same 408 informative positions used in the ED analysis. Numbers represent percentages of 1000 bootstraps and are only shown for bootstrap values <80%. Plasmid-borne *arsC* genes are indicated by "plasmid" following the organism name. Names in boxes represent branches that are inconsistent with the 16S rRNA tree.

some of the discrepancies between 16S rRNA-based and *arsC*-based phylogenies. HGT of *arsC* from an ancestral low GC Gram-positive bacterium to an ancestral cyanobacterium seems likely, and a similar HGT event from ancestral Actinobacteria to the Streptococci would explain the unexpected phylogeny of the *Streptococcus arsC* sequences. Similar gene transfer events from the low GC Gram-positives to *F. nucleatum* could also have occurred, but with an *arsC* sequence only available from a single member of the Fusobacteria it is difficult to draw conclusions. The paraphyletic *arsC* sequences in the non-enteric  $\gamma$ -Proteobacteria likewise may have arisen from HGT events, although the difficulty in resolving the affiliation of some of these sequences makes speculation about their evolution premature. Indeed, the exact placement of many of the non-orthologous *arsC* genes was difficult and poor bootstrap support for some nodes suggests that refinement of *arsC* phylogeny as more sequences become available should be an ongoing process. Regardless, a number of bacteria possess *arsC* genes that are inconsistent with established phylogenies, and HGT events are one possible explanation. It should also be noted that some of the non-orthologous sequences represent putative *arsC* genes identified from ORF's, rather than confirmed arsenate reductases. These ORF's show homology to identified *arsC* genes, but may not necessarily code for a functional enzyme. Thus, some of the non-orthologous sequences might represent elevated genotypic variation if the organism has no selective pressure to maintain a functional *arsC* gene.

We were hoping that patterns in the phylogeny of plasmid-borne *arsC* genes might indicate whether HGT of *arsC* genes has occurred (or is still occurring) in recent times. Specifically, if the plasmid-borne arsenate reductases are similar, this would suggest a common origin for these genes, which must ultimately be chromosomal. However, the plasmid-borne *arsC* genes appear to be paraphyletic and clearly separate into very different *arsC* types (Figures 2 and 3). The two major groups of plasmid-borne genes are those of the Staphylococci (which were the first *arsC* genes to be recognized [17,18]) and those of the Enterobacteriales. Both of these groups of plasmid-borne genes are similar to the chromosomal genes found in related organisms. The only exception is the *arsC* found on the pKW301 plasmid of *A. multivorum*, which shows high sequence similarity to the plasmid-borne *arsC* sequences found in the Enterobacteriales, strongly suggesting relatively recent plasmid transfer between a member of this group and *A. multivorum*. Indeed, the *ars* operon found in *A. multivorum* is expressed and confers arsenate resistance when transferred to *E. coli* [21]. Three other types of plasmid-borne *arsC* genes are found in *R. solanacearum*, *Clostridium acetobutylicum*, and *Halobacterium halobium*, demonstrating that plasmid-borne *arsC*

genes are phylogenetically widespread and suggesting multiple incidences of chromosomal-plasmid transfer. The *arsC* sequence of *C. acetobutylicum* is typical of the clostridial *arsC* genes, and also shows some similarity to the arsenate reductase sequence in *L. interrogans* (a Spirochete). Thus, it is possible that ancestral clostridial plasmids may have transferred *ars* genes to the Spirochetes. As well as *C. acetobutylicum*, two other bacteria were represented by both chromosomal and plasmid-borne *arsC* sequences. The two *arsC* genes of *E. coli* K12 are similar, but they are equally similar to the plasmid-borne genes of other enteric bacteria. The plasmid-borne *arsC* gene from *Salmonella typhimurium*, shows little similarity to its chromosomal equivalent, and appears to be related to the other *arsC* genes found in plasmids in the Enterobacteriales. The high similarity (at least 70%) between all the enteric plasmids suggests a common source, which from this analysis appears to be very similar to the *E. coli* chromosomal *arsC* gene. However, it is clear that within the enterics, relationships based on *arsC* do not parallel relationships based on 16S rRNA, and significant genetic exchange within this group (possibly inside a host organism) has likely occurred.

## Conclusions

Despite the suggestion of HGT and duplication events, the *arsC* phylogeny suggests that arsenate reductase is an evolutionarily old enzyme, a hypothesis that was recently suggested for arsenite oxidase [38]. This is the first phylogenetic analysis of an arsenic redox active gene to include all three domains and the Archaea and Bacteria were clearly separated, with *S. cerevisiae* subsequently branching from Archaea. Thus, the *arsC* gene, and by extension, arsenic resistance mechanisms, must have been present in early organisms, either in the last universal ancestor or after the divergence of the two prokaryotic domains (i.e. in an ancestral bacterium or archaean). In the latter case, early HGT event(s) must have transferred arsenate resistance to the other domain, followed by subsequent divergence to the phylogeny seen today. Such gene exchange between Archaea and primitive Bacteria has certainly taken place [36] and makes absolute statements about the initial origin of genes difficult to make. Regardless, it seems likely that arsenate resistance developed early in the evolution of microorganisms. Recent ideas on the origin of life suggest that early cellular structures were abiotic iron-sulfur formations at submarine hydrothermal vents in the Hadean ocean, in which basic biochemical and metabolic processes could have developed before the formation of cell membranes [39,40]. As well as iron and sulfur, hydrothermal areas are often characterized by high levels of arsenic [41] and may even contain arsenic redox active microbial communities [42]. Given the integral role of phosphate in early (and current) cellular metabolism [39], an ability to reduce the levels of arse-

nate (a phosphate analog) might have been an essential biochemical process. There would have been selective pressure to develop an arsenate reductase that could change this arsenate to arsenite, and this early evolution of an *arsC*-like ancestor coupled with sequence divergence and HGT would result in the diversity and widespread distribution of *arsC* genes seen in microorganisms today.

## Methods

### Sequence acquisition and alignment

Sequences of *arsC* genes were obtained from GenBank [43] and were characterized into two broad types: confirmed *arsC* genes (i.e. those sequences that were identified in studies that explicitly tested arsenate reduction and/or resistance in that given organism), and putative *arsC* genes (i.e. open reading frames (ORFs) obtained from genome projects that showed homology to known *arsC* genes). Genes were also noted as being either chromosomal or plasmid-borne. A total of 60 sequences were analyzed consisting of 54 bacterial, 5 archaeal, and the *Arr2* gene of *Saccharomyces cerevisiae*. Accession numbers for these sequences are given in Table 1 [see Additional file 1]. The *Arr2* gene (or ACR2 [24]) has been identified as coding for a protein necessary for arsenate resistance in yeast, and appears to have some homology to prokaryotic *arsC* genes. Sequences were imported into the ARB software package (distributed by W. Ludwig and O. Strunk, Technical University of Munich, Germany; <http://www.arb-home.de/>) running on a SuSE Linux 6.2 platform.

In ARB, amino acid sequences were derived from nucleotide sequences using the Translate DNA to Protein function and the standard genetic code. Because some *arsC* genes were incomplete, all three possible reading frames were examined to ensure correct translation. Amino acid sequences of the confirmed bacterial arsenate reductases (12 total) were initially aligned automatically using the Fast Aligner function of ARB Edit 4.1, and then manually adjusted to ensure that secondary structure considerations of the arsenate reductase protein were met. The secondary structure of *E. coli* arsenate reductase (17978 in the NCBI Molecular Modeling Database, [44]) was visualized using Cn3D 3.0 and One-D Viewer 1.0 (available from <http://www.ncbi.nih.gov>) and used as a model for arsenate reductase secondary structure. Nucleotide sequences of these confirmed *arsC* genes were aligned according to the amino acid alignment, and the putative *arsC* ORFs subsequently aligned to these confirmed sequences. Multiple alignments of all *arsC* genes used in this study are presented as an additional ARB file [see Additional file 2]. Aligned 16S rRNA sequences of all organisms used in the *arsC* analysis (with the exception of an *arsC* sequence obtained from an unknown Proteobacterium) were obtained from the Ribosomal Database Project [45] and

imported into ARB. Because three *arsC* genes have been reported for different strains/plasmids of *E. coli*, and two *arsC* genes have been identified for both *Salmonella typhimurium* and *Clostridium acetobutylicum* (plasmid and chromosomal for each; Table 1 [see Additional file 1]) only one 16S rRNA sequence was used for each to avoid potential biases and inaccurate attractions due to sequence identity.

### Phylogenetic analyses

The aligned 16S rRNA sequences (55) were used to construct a baseline phylogenetic tree to serve as a comparison for *arsC* sequences. 1326 informative positions (corresponding to *E. coli* positions 140–1475) were compared using evolutionary distance (ED) methods. An ED tree was constructed in ARB for the 56 16S rRNA sequences using neighbor joining [46] with 1000 bootstraps. Distance and maximum parsimony (MP) methods were used to construct *arsC* trees using aligned nucleotide sequences. Both methods used all 60 aligned *arsC* nucleotide sequences. The ED tree was constructed in ARB using neighbor joining following a simple Jukes and Cantor model with the Olsen correction. Because visual examination of the aligned sequences revealed a number of gaps suggesting insertion/deletion events within different taxa, gaps were treated as a fifth base for construction. Thus, ED phylogeny was based on a total of 408 informative positions with 1000 bootstraps performed as part of the procedure. MP phylogeny used the same informative positions and number of bootstraps, and was performed using the PHYLIP package included with ARB.

### Authors' contributions

CRJ carried out the phylogenetic analyses, participated in sequence alignment and drafted the manuscript. SLD obtained the sequence information and participated in sequence alignment. Both authors read and approved the final manuscript.



## Additional material

### Additional File 1

Excel file showing the sources of the sequences used in both *arsC* and 16S rRNA phylogenetic analyses. Each *arsC* sequence was designated as either plasmid or chromosomal, and as either confirmed (based on known arsenic resistance) or putative (from whole genome sequencing).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-3-18-S1.xls>]

### Additional File 2

ARB data file containing multiple alignments of all *arsC* and 16S rRNA sequences used in this manuscript. File requires the ARB package for Linux or Solaris systems, freely available from <http://www.arb-home.de>

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-3-18-S2.arb>]

## Acknowledgements

Funding for this research was provided by award LEQSF(2002-05)-RD-A-25 from the Louisiana Board of Regents Support Fund, Research and Development Program, to CRJ. Rick Miller, Evelyn F. Jackson and two anonymous reviewers provided useful comments on an earlier version of this manuscript.

## References

- Tamaki S and Frankenberger WT Jr: **Environmental biochemistry of arsenic** *Rev Environ Cont Toxicol* 1992, **124**:79-110.
- Cervantes C, Ji G, Ramirez JL and Silver S: **Resistance to arsenic compounds in microorganisms** *FEMS Microbiol Rev* 1994, **15**:355-367.
- Anderson GL, Williams J and Hille R: **The purification and characterization of arsenite oxidase from *Alcaligenes faecalis*, a molybdenum-containing hydroxylase** *J Biol Chem* 1992, **267**:23674-23682.
- Muller D, Lievreumont D, Simeonova DD, Hubert J-C and Lett M-C: **Arsenite oxidase *aox* genes from a metal-resistant  $\beta$ -Proteobacterium** *J Bacteriol* 2003, **185**:135-141.
- Abdrashitova SA, Abdullina GG and Ilyaletdinov AN: **Role of arsenites in lipid peroxidation in *Pseudomonas putida* cells oxidizing arsenite** *Mikrobiologiya* 1986, **55**:212-216.
- Abdrashitova SA, Abdullina GG, Mynbaeva BN and Ilyaletdinov AN: **Oxidation of iron and manganese by arsenic-oxidizing bacteria** *Mikrobiologiya* 1990, **59**:85-89.
- Ahmann D, Roberts AL, Krumholz LR and Morel FMM: **Microbe grows by reducing arsenic** *Nature* 1994, **371**:750.
- Newman DK, Ahmann D and Morel FMM: **A brief review of microbial arsenate respiration** *Geomicrobiol* 1998, **15**:255-268.
- Stolz JF and Oremland RS: **Bacterial respiration of arsenic and selenium** *FEMS Microbiol Rev* 1999, **23**:615-627.
- Huber R, Sacher M, Vollmann A, Huber H and Rose D: **Respiration of arsenate and selenate by hyperthermophilic archaea** *System Appl Microbiol* 2000, **23**:305-314.
- Ilyaletdinov AN and Abdrashitova SA: **Autotrophic oxidation of arsenic by a culture of *Pseudomonas arsenitoxidans*** *Mikrobiologiya* 1981, **50**:197-204.
- Santini JM, Sly LI, Schnagl RD and Macy JM: **A new chemolithoautotrophic arsenite-oxidizing bacterium isolated from a gold mine: phylogenetic, physiological, and preliminary biochemical studies** *Appl Environ Microbiol* 2000, **66**:92-97.
- Ji G, Silver S, Garber EAE, Ohtake H, Cervantes C and Corbisier P: **Bacterial molecular genetics and enzymatic transformations of arsenate, arsenite, and chromate** In *Biohydrometallurgical Technologies* Edited by: Torma AE, Apel ML, Briery CL. Warrendale, PA: The Minerals, Metals & Materials Society; 1993:529-539.
- Kaur P and Rosen BP: **Plasmid-encoded resistance to arsenic and antimony** *Plasmid* 1992, **27**:29-40.
- Wu J and Rosen BP: **The *arsD* gene encodes a second trans-acting regulatory protein of the plasmid-encoded arsenical resistance operon** *Mol Microbiol* 1993, **8**:615-623.
- Neyt CN, Iriarte M, Thi VH and Cornelis GR: **Virulence and arsenic resistance in *Yersinia*** *J Bacteriol* 1997, **179**:612-619.
- Ji G and Silver S: **Regulation and expression of the arsenic resistance operon from *Staphylococcus aureus* plasmid pI258** *J Bacteriol* 1992, **174**:3684-3694.
- Rosenstein R, Peschel A, Wieland B and Gotz F: **Expression and regulation of the antimonite, arsenite, and arsenate resistance operon of the *Staphylococcus xylosus* plasmid pSX267** *J Bacteriol* 1992, **174**:3676-3683.
- Carlin A, Shi W, Dey S and Rosen BP: **The *ars* operon of *Escherichia coli* confers arsenical and antimicrobial resistance** *J Bacteriol* 1995, **177**:981-986.
- Diorio C, Cai J, Marmor J, Shinder R and DuBrow MS: **An *Escherichia coli* chromosomal *ars* operon homolog is functional in arsenic detoxification and is conserved in Gram-negative bacteria** *J Bacteriol* 1995, **177**:2050-2056.
- Suzuki K, Wakao N, Kimura T, Sakka K and Ohmiya K: **Expression and regulation of the arsenic resistance operon of *Acidiphilum multivorum* AIU 301 plasmid pKW301 in *Escherichia coli*** *Appl Environ Microbiol* 1998, **64**:411-418.
- Sato T and Kobayashi Y: **The *ars* operon in the skin element of *Bacillus subtilis* confers resistance to arsenate and arsenite** *J Bacteriol* 1998, **180**:1655-1661.
- Cai J, Salmon K and DuBrow MS: **A chromosomal *ars* operon homologue of *Pseudomonas aeruginosa* confers increased resistance to arsenic and antimony in *Escherichia coli*** *Microbiology* 1998, **144**:2705-2713.
- Bobrowicz P, Wysocki R, Owsianik G, Goffeau A and Ulaszewski S: **Isolation of three contiguous genes, *ACR1*, *ACR2* and *ACR3*, involved in resistance to arsenic compounds in the yeast *Saccharomyces cerevisiae*** *Yeast* 1997, **13**:819-828.
- Cullen WR and Reimer KJ: **Arsenic speciation in the environment** *Rev* 1989, **89**:713-764.
- Ferguson JF and Gavis J: **A review of the arsenic cycle in natural waters** *Water Res* 1972, **6**:1259-1274.
- Butcher BG, Deane SM and Rawlings DE: **The chromosomal arsenic resistance genes of *Thiobacillus ferrooxidans* have an unusual arrangement and confer increased arsenic and antimony resistance to *Escherichia coli*** *Appl Environ Microbiol* 2000, **66**:1826-1833.
- Saltikov CW and Olson BH: **Homology of *Escherichia coli* R773 *arsA*, *arsB*, and *arsC* genes in arsenic-resistant bacteria isolated from raw sewage and arsenic-enriched creek waters** *Appl Environ Microbiol* 2002, **68**:280-288.
- Jackson CR, Jackson EF, Dugas SL, Gamble K and Williams SE: **Microbial transformations of arsenite and arsenate in natural environments** *Recent Res Devel Microbiol* 2003.
- Mukhopadhyay R, Rosen BP, Phung LT and Silver S: **Microbial arsenic: from geocycles to genes and enzymes** *FEMS Microbiol Rev* 2002, **26**:311-325.
- Woese CR: **Bacterial evolution** *Microbiol Rev* 1987, **51**:221-271.
- Wolf YI, Rogozin IB, Grishin NV, Tatusov RL and Koonin EV: **Genome trees constructed using five different approaches suggest new major bacterial clades** *BMC Evolutionary Biology* 2001, **1**:8.
- Doolittle RF: **Convergent evolution: The need to be explicit** *Trends Biochem Sci* 1994, **19**:15-18.
- Nei M and Kumar S: **Molecular Evolution and Phylogenetics** New York: Oxford University Press; 2000.
- Aravind L, Tatusov RL, Wolf YI, Walker DR and Koonin EV: **Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles** *Trends Genet* 1998, **14**:442-444.
- Brown JR and Doolittle WF: **Archaea and the prokaryote-to-eukaryote transition** *Microbiol Mol Biol Rev* 1997, **61**:456-502.
- Mirkin BG, Fenner TI, Galperin MY and Koonin EV: **Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes** *BMC Evolutionary Biology* 2003, **3**:2.

38. LeBrun E, Brugna M, Baymann F, Muller D, Lièvreumont D, Lett M-C and Nitschke W: **Arsenite oxidase, and ancient bioenergetic enzyme** *Mol Biol Evol* 2003, **20**:686-693.
39. Russell MJ and Hall AJ: **The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front** *J Geol Soc Lond* 1997, **154**:377-402.
40. Martin W and Russell MJ: **On the origin of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells** *Phil Trans R Soc Lond* 2003, **358**:59-85.
41. Langner HW, Jackson CR, McDermott TR and Inskeep WP: **Rapid oxidation of arsenite in a hot spring ecosystem in Yellowstone National Park** *Environ Sci Technol* 2001, **35**:3302-3309.
42. Jackson CR, Langner HW, Donahoe-Christiansen J, Inskeep WP and McDermott TR: **Molecular analysis of microbial community structure in an arsenite-oxidizing acidic thermal spring** *Environ Microbiol* 2001, **3**:532-542.
43. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA and Wheeler DL: **GenBank** *Nucleic Acids Res* 2000, **28**:15-18.
44. Martin P, DeMel S, Shi J, Gladysheva T, Gatti DL, Rosen BP and Edwards BF: **Insights into the structure, solvation, and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme** *Structure* 2001, **9**:1071-1081.
45. Maidak BL, Cole JR, Lilburn TG, Parker CT Jr, Saxman PR, Stradwick JM, Garrity GM, Li B, Olsen GJ, Pramanik S, Schmidt TM and Tiedje JM: **The RDP (Ribosomal Database Project) continues** *Nucleic Acids Res* 2000, **28**:173-174.
46. Saitou N and Nei M: **The neighbor-joining method: A new method for reconstructing phylogenetic trees** *Mol Biol Evol* 1987, **4**:406-425.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

