Research article

Diversity of Betaproteobacteria revealed by novel primers suggests their role in arsenic cycling

Anirban Chakraborty a,1, Chanchal K. DasGupta a, Punyasloke Bhadury b, *

a Department of Life Science and Biotechnology, Jadavpur University, Kolkata, 700032, West Bengal, India
b Integrative Taxonomy and Microbial Ecology Research Group, Department of Biological Sciences and Centre for Climate and Environmental Studies, Indian Institute of Science Education and Research Kolkata, Mohanpur, Nadia, 741246, West Bengal, India

ARTICLE INFO

Keywords:
Environmental science
Environmental geochemistry
Hydrology
Groundwater
Microbiology
Bacteria
Microorganism
Primer
16S rRNA gene
Betaproteobacteria
Aquifer
Arsenic

ABSTRACT

High arsenic concentration in groundwater is a severe environmental problem affecting human health, particularly in countries of South and South-East Asia. The Bengal Delta Plain (BDP) distributed within India and Bangladesh is a major arsenic-affected region where groundwater is the primary source of drinking water. Previous studies have indicated that members of the bacterial class Betaproteobacteria constitute a major fraction of the microbial community in many of the aquifers within this region. Bacteria belonging to this class are known to be involved in redox cycling of arsenic as well as other metals such iron and manganese, thereby impacting arsenic mobilization and immobilization. While microbial diversity in arsenic-contaminated environments is generally assessed using universal 16S rRNA gene primers, targeted evaluation of Betaproteobacteria diversity remains poorly constrained. In this study, bacterial diversity was investigated in the groundwater from two shallow aquifers (West Bengal, India) based on 16S rRNA gene clone libraries and sequencing using a custom-designed pair of primers specific to Betaproteobacteria. Specificity of the primers was confirmed in silico as well as by the absence of PCR amplification of other bacterial classes. Four major families (Burkholderiaceae, Comamonadaceae, Gallionellaceae and Rhodocyclaceae) were detected among which members of Burkholderiaceae represented 59% and 71% of the total community in each aquifer. The four OTUs (operational taxonomic units; 97% sequence identity) within Burkholderiaceae were close phylogenetic relatives of bacteria within the genus Burkholderia known to solubilize phosphate minerals. Additionally, the OTUs belonging to Gallionellaceae were closely related to the members of the genera Gallionella and Sideroxydans, known to oxidize iron under microaerophilic conditions. These results suggest that members of Betaproteobacteria can potentially influence iron and phosphorus cycling which can influence biogeochemistry in arsenic-contaminated aquifers of the BDP.

1. Introduction

Elevated natural concentrations of arsenic (As) in groundwater is a major environmental health problem in many South and South-East Asian countries including India and Bangladesh, where consumption of As-contaminated water in the fertile Bengal Delta Plain (BDP) affects millions of people on a daily basis [1, 2]. In particular, the extent of contamination is greater in the shallow Holocene grey sand aquifers which supply majority of the drinking water in these regions, whereas groundwater in the deeper Pleistocene brown sand aquifers is largely known to be arsenic-free [3]. The primary source of arsenic in the BDP aquifers is considered to be geogenic, the most widely accepted explanation being the release of adsorbed and co-precipitated As from iron (Fe) and manganese (Mn) oxyhydroxides by microbially mediated reductive dissolution [4]. On the other hand, arsenic can be microbially immobilized by arsenic- and iron-oxidizing bacteria through direct enzymatic oxidation [5] or adsorption onto biogenic iron oxides [6]. A number of environmental parameters such as biodegradable organic matter [7], iron [8], phosphate [9] and other electron acceptors [10, 11] have been shown to influence microbial arsenic cycling as well as groundwater microbial community composition and functions [12]. Therefore, an in-depth understanding of the composition and distribution of indigenous microbial communities in BDP aquifers is crucial, particularly in the context of...
developing cost-effective long-term arsenic mitigation strategies and for supply of arsenic free potable water.

Over the years, several investigations of microbial communities in the BDP aquifers have shown diversity in microbial assemblages spatially (across aquifers) as well as temporarily (e.g. seasonal influences such as precipitation) within the same broad geographic location [13, 14, 15, 16, 17]. In 2009, an investigation of microbial communities in shallow and deep tube wells in Bangladesh revealed that bacteria belonging to the class Betaproteobacteria were clearly the dominant members, constituting up to ~84% of the entire community [13]. Another study by Sultana et al showed greater abundance of Betaproteobacteria in a shallow (21m) monitoring well compared to that in a deeper (85m) well [14]. Similar observations were reported by Hassan et al where abundance of Betaproteobacteria was observed as the most abundant class (~69% of the community) in a comparative examination of microbial communities from 24 groundwater samples in Bangladesh [17]. In addition to BDP aquifers, this bacterial class also has been shown to dominate the microbial communities in As-contaminated soils from China and the United Kingdom [18]. Taken together, these observations are indicative of an emerging pattern where members of Betaproteobacteria are major inhabitants of arsenic-contaminated subsurface in geographically distant regions. Comprising over 75 genera and 400 species, Betaproteobacteria is a large class within the phylum Proteobacteria [19]. In addition, members of Betaproteobacteria are frequently detected in subsurface environments in general and display an enormous range of metabolic diversity such as redox transformation of As [5] and Fe [20], denitrification [21] and biodegradation of recalcitrant organic compounds [22]. Investigations focusing on Betaproteobacteria diversity are therefore critical to better understand the potential of members of this class in arsenic cycling within contaminated aquifers.

In India, arsenic contamination in the BDP extends into the state of West Bengal, where the Nadia district is one of the severely impacted regions. The shallow grey sand aquifers in this region contain much higher As concentrations than the deeper brown sand aquifers [23]. Previous geochemical data has shown marked heterogeneity in the distribution of dissolved organic carbon in these aquifers, with greater input and preferential biodegradation of mature hydrocarbons as well as increased Fe and Mn concentrations in the shallow aquifers [24]. A recent investigation of bacterial community structures used clone libraries of the 16S rRNA gene and the arsenic oxidase large subunit aioA gene to demonstrate that Betaproteobacteria continued to dominate microbial communities in two shallow aquifers in the Nadia district [16]. A number of genera such as Hydrogenophaga and Albidiferax were prominent in the aioA clone library, which strongly suggested involvement of these microorganisms in As oxidation. On the other hand, the 16S rRNA gene clone library indicated presence of genera like Acidovorax and Dechloromonas, which are known to oxidize Fe(II) under anoxic conditions [25]. Interestingly, the above genera have also been detected in hydrocarbon-contaminated aquifers [26]. Relative abundance of the above bacterial groups was observed to be seasonally variable, likely due to monsoon-induced changes in groundwater.

The objective of the current study was to develop a community fingerprinting tool specific towards the class Betaproteobacteria. Taxon-specific primers have been shown to improve detection capability of less abundant organisms in targeted diversity analyses using community fingerprinting methods [27]. In the current study, a pair of robust degenerate Betaproteobacteria-specific 16S rRNA gene primers was custom designed. Additionally, groundwater from the two shallow aquifers described above was re-sampled and clone library analyses were conducted using the above new pair of primers. These analyses showed that the Betaproteobacteria communities in both wells were massively dominated by members of the family Burkholderiaceae, which was not previously detected in these wells. Additionally, the new primers were able to detect the presence of Sideroxydans-like organisms that are well known for microaerophilic Fe oxidation. These primers therefore can provide deeper understanding of the functioning of members of Betaproteobacteria in arsenic contaminated aquifers.

2. Materials and methods

2.1. Groundwater sampling and geochemical analyses

In November 2013, groundwater was collected from two As-rich grey sand shallow aquifers (Well 28 and Well 204) located in Karimpur Block II of Nadia district in the state of West Bengal, India (Figure 1). A detailed geological description of these aquifers and groundwater sampling methods has been provided elsewhere [16]. Briefly, fresh groundwater was collected from these two wells after pumping out roughly three times the well volumes including the stagnant water in the wells. Groundwater temperature, pH and total dissolved solids (TDS) were measured immediately after collection. In addition, two sets of water samples (50 mL) were sterile filtered (syringe filter, 0.45μm pore size) and stored in sterile falcon tubes for further geochemical analyses of major ions and metals. Colorimetric assays were subsequently used in the laboratory to measure dissolved nitrate [28], dissolved phosphate [29] as well as total Fe [30] and As [31].

Two liters of groundwater from each well were preserved with absolute molecular grade ethanol (2% final concentration) immediately after collection to limit microbial activity and prevent denaturation of nucleic acids. The samples were immediately transported to laboratory and passed through 0.22 μm pore size Sterivex filters (Millipore Sigma, Danvers, MA, USA) using a peristaltic pump for concentrating the biomass. The filters were stored at -20°C and subjected to environmental DNA extraction following a protocol described elsewhere [32].
2.2. Designing 16S rRNA gene primers specific to class Betaproteobacteria

Twenty seven near-full-length 16S rRNA gene sequences representing all known families of the class Betaproteobacteria were downloaded from nucleotide databases (GenBank/ENA/DDBJ), along with 9 additional sequences representative of other classes in the phylum Proteobacteria (Table 1). These 36 sequences were subsequently aligned in ClustalW [33] for identification of regions conserved only within the class Betaproteobacteria. Based on this alignment, a new set of forward and reverse primers Beta52F (5’ AAGTGCAAGCGGACRSRG 3’) and Beta1014R (5’ GTGYYGCGAAARGRGCYK 3’), respectively, were designed. The primers were subsequently examined using probeCheck web server [34] for self-complementation, low GC content and low Tm value to ensure sufficient thermal window for efficient annealing. The size of 16S rRNA gene amplicons generated using these primers were approximately 950bp.

2.3. PCR amplification of the 16S rRNA gene fragments

PCR amplification was conducted from environmental DNA using the newly designed primers. Each PCR reaction consisted of 0.5 μl of DNA Dream Taq polymerase (5 U/μl; Fermentas, Thermo Fisher Scientific, Waltham, MA, USA), 5.0 μl 10x Dream Taq buffer, 5.0 μl dNTPs (final concentration 0.2 mM), 5.0 μl MgCl₂ (final concentration 2.0 mM), 0.5 μl of each primer (final concentration 5 μM), 0.5 μl (~20 ng) DNA template, 0.5 μl Bovine Serum Albumin (1 mg/ml) and nuclease free water to make a final volume of 50 μl [16]. The PCR conditions for the Betaproteobacteria-specific primers were as follows: initial denaturation at 95 °C for 3 min, 35 cycles of 95 °C for 1 min, 46 °C for 45 s, 72 °C for 2 min, and final extension at 72 °C for 10 min. The optimum annealing temperature was selected based on a gradient PCR approach.

All PCR reactions were performed in triplicate and subsequently pooled together. The pooled PCR products were then purified using a gel purification kit (Qiagen, Hilden, Germany) as per manufacturer’s instructions.

2.4. Clone library construction, DNA sequencing and sequence processing

Purified PCR products were cloned using pGEM®-T Easy Vector system (Promega, Madison, WI, USA) following the manufacturer’s instructions. Plasmid DNA containing the inserts was sequenced in both orientations.

---

Table 1. The names, NCBI Accession Numbers and taxonomy of the 36 bacterial species (27 belonging to Betaproteobacteria) which were chosen for designing the pair of primers used in this study.

| Organism (Betaproteobacteria)                          | Accession Number# | Domain         | Phylum         | Class              | Order            | Family          | Genus          |
|--------------------------------------------------------|-------------------|----------------|----------------|--------------------|------------------|-----------------|----------------|
| Acidovorax delafieldii strain 2AN                      | HM625980          | Bacteria       | Proteobacteria | Betaproteobacteria | Burkholderiales  | Comamonadaceae  | Acidovorax     |
| Acidovorax ehrus strain TPYSY                          | CP01392           | Bacteria       | Proteobacteria | Betaproteobacteria | Burkholderiales  | Comamonadaceae  | Acidovorax     |
| Dechloromonas agilata strain CKB                       | NR_024884         | Bacteria       | Proteobacteria | Betaproteobacteria | Rhodocyclaceae   | Dechloromonas    |                |
| Dechloromonas aromatica strain RGB                      | NC_007298         | Bacteria       | Proteobacteria | Betaproteobacteria | Rhodocyclaceae   | Dechloromonas    |                |
| Azospira saliva strain PS                               | AF170348          | Bacteria       | Proteobacteria | Betaproteobacteria | Rhodocyclaceae   | Azospira         |                |
| Siderocysta luteovaricosa strain ES-1                  | NR_074731         | Bacteria       | Proteobacteria | Betaproteobacteria | Nitrosomonadales | Siderocysta      |                |
| Siderocysta luteovaricosa strain LD-1                  | DQ386859          | Bacteria       | Proteobacteria | Betaproteobacteria | Nitrosomonadales | Siderocysta      |                |
| Leptothrix discophora strain SS-1                      | NR_029516         | Bacteria       | Proteobacteria | Betaproteobacteria | Burkholderiales  | Leptothrix       |                |
| Ferritrophicum radicicola strain CCJ                   | DQ386263          | Bacteria       | Proteobacteria | Betaproteobacteria | Ferritrophicales | Ferritrophicum   |                |
| Pseudogulbenkiania sp. strain 2002                     | AY609199          | Bacteria       | Proteobacteria | Betaproteobacteria | Neisseriales     | Chromobacteriaceae | Pseudogulbenkiania |
| Comamonas denitrificans isolate SG15                  | HE681736          | Bacteria       | Proteobacteria | Betaproteobacteria | Burkholderiales  | Comamonas        |                |

Organism (Non-Betaproteobacteria)

| Dechlorosporum sp. strain M1                            | GQ262802          | Bacteria       | Proteobacteria | Alphaproteobacteria | Rhodospirillales | Dechlorosporum   |                |
| Maripropionibacterium ferrooxidans strain PV-1         | EF493243           | Bacteria       | Proteobacteria | Alphaproteobacteria | Maripropionidales | Maripropionidales |                |
| Geobacter metallicus strain GS-15                      | L07834            | Bacteria       | Proteobacteria | Deltaproteobacteria | Desulfuromonadales | Geobacter         |                |
| Ferrovibrio denitrificans strain Sp-1                  | GQ356620          | Bacteria       | Proteobacteria | Alphaproteobacteria | Rhodospirillales | Ferrovibrio      |                |
| Shewanella oneidensis strain MR-1                      | NR_074798         | Bacteria       | Proteobacteria | Gammaproteobacteria | Alteronomadales  | Shewanella       |                |
| Paracoccus ferrooxidans strain BDN-1                   | AJ954687          | Bacteria       | Proteobacteria | Alphaproteobacteria | Rhodobacterales  | Paracoccus       |                |
| Thermomonas brevis strain LMG 21746T                   | AJ519989          | Bacteria       | Proteobacteria | Gammaproteobacteria | Xanthomonadales  | Thermomonas      |                |
| Escherichia coli strain CJ707                          | AE014075          | Bacteria       | Proteobacteria | Gammaproteobacteria | Enterobacterales  | Escherichia      |                |
| Acinetobacter sp. strain X9-2                          | KC156905          | Bacteria       | Proteobacteria | Gammaproteobacteria | Pseudomonadales  | Moraxellaceae    | Acinetobacter   |

A. Chakraborty et al. Heliyon 6 (2020) e03089
Betaproteobacteria targeting genomic DNA extracted from three well boreholes of the BDP aquifers of groundwater in this well [24]. The sequences were further checked for Chimeras using Bellerophon [36] and chimeric sequences were excluded from downstream analyses. Following quality control, the generated sequences were clustered into operational taxonomic units (OTU) based on 97% sequence using Mothur [37]. An OTU table was generated and subsequently used for calculating the diversity and richness indices. Taxonomy was assigned to the representative OTU sequences using the SILVA database (version 123) within Mothur. A Fisher's exact test with Storey's FDR correction for multiple comparisons was applied using the STAMP software package version 2.1 [38] to determine differential abundance of the Betaproteobacteria OTUs between the two wells. Corrected p-values < 0.05 were considered significant.

2.5. Phylogenetic analysis

Representative sequences of the OTUs were first automatically aligned using the web-based SINA aligner [39] and imported into the ARB-SILVA database SSU Ref NR 123 [40] using the ARB software package [41]. A maximum likelihood (PhyML) tree was calculated with almost full-length 16S rRNA gene sequences (>1300 nt) of closely related positions by using a positional variability filter for bacteria. Using the ARB parsimony tool, the OTU sequences were subsequently added to this tree one at a time by using a 50% sequence conservation filter and positional variability filters covering the individual length of each representative OTU sequence, without changing the overall tree topology. Filled circles at the node of the branches indicate lineages with >80% bootstrap support (100 re-samplings). The scale bar represents 5% estimated substitutions. The tree was visualized using iTOL version 3 [42] and a bubble plot showing the relative sequence abundance of the OTUs was created using the R package ggplot2 [43].

2.6. Availability of sequences

All clone sequences generated in this study were submitted in GenBank at NCBI under Accession Numbers KY458643 – KY458731.

3. Results and discussion

3.1. Groundwater sampling and geochemical analyses

Groundwater pH, temperature, dissolved oxygen (DO), total dissolved solids (TDS), total arsenic (As), total iron (Fe), nitrate and phosphate in the two BDP wells (Hydrogenophaga aypica strain BDIP10, Hydrogenophaga bisansensis strain BDIP20 and Acidovorax faeicilis strain BDIP24) isolated previously in our laboratory from groundwater of our study site [12]. Successful amplification of a ca. 950 bp long fragment was observed for all of the above isolates. Pairwise alignment of the above amplicon sequences to the full-length 16S rRNA gene sequences of the same isolates obtained after amplification with universal primers revealed 100% sequence identity in all cases. Contrarily, amplification was not observed when genomic DNA from Escherichia coli (class - Gammaproteobacteria) and Bacillus subtilis (class - Bacilli), representing bacteria belonging to other classes, were used as template.

Two clone libraries (28W2013 and 204W2013) consisting of 44 and 45 clones respectively, were generated using the above primers in order to investigate diversity of Betaproteobacteria in Well 24 and Well 204. The diversity indices for both libraries have been summarized in Table 3. When all 89 clone sequences were grouped into OTUs based on a cut-off of 97% identity on nucleotide level, a total of 16 OTUs were produced (Figure 2). Taxonomic assignments placed the OTUs in four major families, Burkholderiaceae (5 OTUs), Comamonadaceae (3 OTUs), Gallionellaceae (5 OTUs) and Rhodocyclaceae (3 OTUs). Burkholderiaceae accounted for the majority of the clone sequences in both libraries, representing a total of 26 (59% relative abundance; 3 OTUs) and 32 (71% relative abundance; 5 OTUs) clones in libraries 28W2013 and 204W2013, respectively. Members of Gallionellaceae were relatively more abundant in Well 28 (29.5% relative abundance; 4 OTUs) than in Well 204 (8.9% relative abundance; 1 OTU). Conversely, clones belonging to Rhodocyclaceae were more prominent in Well 204 (15.6% relative abundance; 2 OTUs) than in Well 28 (6.8% relative abundance; 1 OTU). Comamonadaceae was the least represented family in both libraries with 4.4% (Well 28) and 4.5% (Well 204) relative abundances, respectively. The observed OTU richness was slightly higher in library 28W2013 (10 OTUs) than in 204W2013 (9 OTUs), consistent with the pattern of estimated OTU richness (Chao1 and ACE values; Table 3). Additionally, rarefaction analysis revealed a relatively greater under-

| Table 2. Summary of three year’s geochemical data comprised of groundwater pH, temperature, dissolved oxygen (DO), total dissolved solids (TDS), total arsenic (As), total iron (Fe), nitrate and phosphate in the two BDP wells. |
| Well 28 | Well 204 |
| Year | Y2013a | Y2011a | Y2010a | Y2013b | Y2011b | Y2010b |
| pH | 7.3 | 7.3 | 6.9 | 6.9 | 7.3 | 6.8 |
| Temperature (°C) | 24.3 | 22.3 | 24.7 | 25.6 | 23.4 | 24.4 |
| DO (mg/l) | 2.8 | n.m. | n.m. | 3.1 | n.m. | n.m. |
| TDS (mg/l) | 351 | 358 | 336 | 382 | 379 | 373 |
| Total As (µg/l) | 81 | 55 | 35 | 125 | 117 | 37 |
| Total Fe (mg/l) | 1.3 | 1.65 | 1 | 2.4 | 3.45 | 0.39 |
| NO3 (µg/l) | 0.5 | n.m. | n.m. | 5.14 | n.m. | n.m. |
| PO4 (µg/l) | 50 | n.m. | n.m. | 260 | n.m. | n.m. |

*This study; † previous two year's sampling data; n.m.: not measured.

| Table 3. Alpha diversity estimates of the Betaproteobacteria communities inhabiting the two BDP wells. |
| OTUs | Total | Well 28 | Well 204 |
| Clone sequences | 89 | 44 | 45 |
| Number of OTUsa | 16 | 10 | 9 |
| Chao1 richness estimate | 16.75 | 11 | 9 |
| ACE richness estimate | 17.66 | 12.33 | 9 |
| Shannon diversity index (H') | 2.34 | 1.87 | 2.07 |
| Evenness indexb | 0.84 | 0.81 | 0.94 |

a OTU clustering was performed based on 97% sequence identity. 
b Evenness was calculated by dividing Shannon index by Ln (OTUs).
saturation of the curve for library 28W2013, indicating that deeper sequencing effort is warranted to capture a greater extent of microbial diversity in Well 28 (Figure 3). Contrarily, higher Shannon index and sequence evenness values in library 204W2013 indicated greater diversity of Betaproteobacteria in Well 204. Our observations are consistent with data from previous years (2010 and 2011) where bacterial diversity in Well 204 was also observed to be higher [16]. Fisher’s exact test comparing the two communities revealed a non-uniform pattern in OTU distribution as the abundances of 7 top OTUs (except OTU 2) were found to be significantly different across the two wells (Figure 4).

The top four OTUs belonged to Burkholderia, a genus which was not detected in the previous investigation from these two wells when bacterial communities were analyzed using the full-length 16S rRNA and arsenite oxidase large subunit aioA genes as molecular markers [16]. The closest cultivated phylogenetic relative of OTUs 1 and 4 was Burkholderia ferrariae, which was isolated from a phosphate-containing iron ore (Figure 2) [44] and was able to generate dissolved phosphate from insoluble phosphate minerals. Such capability of mineral weathering for the purpose of nutrient acquisitioning is quite commonly observed among members of this genus. Burkholderia fungorum, which is a close
relative of OTU 3 in our study, was indeed demonstrated to release mineral-bound arsenic in the process of solubilizing phosphate from solid-phase minerals [45]. Such ancillary mobilization of arsenic due to nutrient limitation has been hypothesized to occur conditionally depending upon the availability of poorly weathered minerals as well as supply of organic carbon. Since the BDP aquifers meet all of these criteria, it is conceivable that bacteria belonging to *Burkholderia* could play key roles in mobilization of arsenic in groundwater. It has also been demonstrated that Burkholderia can enhance arsenic toxicity by reducing methyl-arsenate compounds to toxic methyl-arsenite [46]. Interestingly, *Burkholderia* was also among the most prominent genus in recent metagenomic surveys from other arsenic-contaminated regions located in India [47] and Bangladesh [48].

*Gallionellaceae* was the next major family observed in both communities, although the overall abundance of the OTUs belonging to *Gallionellaceae* was significantly higher in Well 28 (Figure 4). Three of the 5 *Gallionellaceae* OTUs were observed to be close relatives of *Sideroxydans lithotrophicus* strain LD-1, a bacterium well studied for its capability of oxidizing Fe(II) in microaerophilic conditions at neutral pH [49]. The availability of iron and low DO concentration in the BDP wells are suggestive of ideal conditions for such organisms to outcompete abiotic Fe(II) oxidation and produce biogenic hydrous ferric oxides which are known to immobilize dissolved arsenic as well as phosphate, a structural analog of arsenate. A recent molecular survey of 24 groundwater wells from Bangladesh revealed frequent occurrence of *Sideroxydans*- and *Gallionella*-like bacteria in BDP region [17]. Another study showed that the abundance of *Sideroxydans* populations increased during nitrate treatment of arsenic-rich groundwater from Cambodia [6]. Nitrate was found in both BDP wells (Table 2) and a number of genera known to contain nitrate-dependent Fe(II)-oxidizing bacteria (e.g. *Acidovorax* and *Dechloromonas*) have previously been observed in these wells. Moreover, recent cultured studies from BDP wells (204 and 28) have reported the presence of several strains of *Acidovorax*, many of which are potentially new species [12]. These results suggest that potential of *Gallionellaceae* in trapping arsenic indirectly through the production of biogenic iron oxides.

4. Conclusion

The findings of this study demonstrate previously undetected diversity of *Betaproteobacteria* in the BDP wells by using a novel group-specific set of primers. The relatively longer amplicon generated by these primers certainly provides greater taxonomic resolution compared to those produced by the primers that target specific variable regions of the 16S rRNA and are frequently used for generating short-read amplicon libraries using next-generation sequencing platforms. This primer set can also principally be used for designing nested PCR assays to generate *Betaproteobacteria*-specific short-read amplicon libraries by high-throughput sequencing platforms. The physiology of the dominant genera detected in this study suggests that they are potentially key players in iron oxidation and phosphate dissolution in these aquifers. Both of these microbially mediated processes immensely impact arsenic release and immobilization in the shallow aquifers of the BDP region. The results of this study suggest that in addition to their capability of oxidizing and reducing As compounds, members of *Betaproteobacteria* have the potential to impact arsenic cycling indirectly by altering the redox states and solubility of compounds that in turn affect the entrapment or release of As in the BDP aquifers.

Figure 3. Rarefaction curves of observed *Betaproteobacteria* OTUs for the 16S rRNA gene clone libraries from the two BDP wells.

Figure 4. Differences in *Betaproteobacteria* communities (at the OTU level) between Well 28 and Well 204.
Declarations

Author contribution statement

Anirban Chakraborty: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. Chanchal K. DasGupta: Contributed reagents, materials, analysis tools or data; Wrote the paper. Punyasloke Bhadury: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

This work was supported by the Dr. D. S. Kothari Postdoctoral Fellowship awarded to Anirban Chakraborty. Punyasloke Bhadury acknowledges FIRE and ARF grants of IISER Kolkata to undertake the study.

Competing interest statement

The authors declare no conflict of interest.

Additional information

All clone sequences generated in this study were submitted in GenBank at NCBI under accession numbers KY458643 – KY458731.

Acknowledgements

We acknowledge the Public Health Engineering Department, Government of West Bengal for providing useful background information. We thank Debaprasad Parai for assistance with fieldwork and the DNA Sequencing Facility of IISER Kolkata.

References

[1] K.M. McCarty, H.T. Hanh, K.W. Kim, Arsenic geochemistry and human health in South east Asia, Rev. Environ. Health 26 (1) (2011) 71–78.
[2] B. Nath, J.P. Moity, J.-S. Jean, G. Birch, S. Kar, H.-J. Yang, M.-K. Lee, R. Hazra, D. Chatterjee, Geochemical characterization of arsenic-affected alluvial aquifers of the bengal delta (West Bengal and Bangladesh) and chinain plains (SW Taiwan): implications for human health, Appl. Geochem. 26 (5) (2011) 705–713.
[3] A.K. Sharma, J.C. Tjell, J.J. Stoth, P.E. Holm, Review of arsenic contamination, exposure through water and food and low cost mitigation options for rural areas, Appl. Geochem. 41 (2014) 11–33.
[4] R.T. Nickson, J.M. McArthur, P. Ravenscroft, W.G. Burgess, K.M. Ahmed, Mechanism of arsenic release to groundwater, Bangladesh and West Bengal, Appl. Geochem. 15 (4) (2000) 403–413.
[5] R.S. Oremlund, J.F. Stolz, Arsenic, microbes and contaminated aquifers, Trends Microbiol. 13 (2) (2005) 45–49.
[6] E.O. Omorogie, R.M. Couture, P. Van Cappellen, C.L. Corkhill, J.M. Charnock, D.A. Polya, D. Vaughan, K. Vanbroekhoven, J.R. Lloyd, Arsenic bioremediation by biogenic iron oxides and sulfides, Appl. Environ. Microbiol. (2013).
[7] R.B. Neumann, K.N. Ashfaque, A.B.M. Badruzzaman, M. Ashraf Ali, J.K. Shoemaker, Mechanism of arsenic release to groundwater: Bangladesh and West Bengal, Appl. Geochem. 28 (3) (2011) 198–211.
[8] S. Ghosh, P. Sar, Identification and characterization of metabolic properties of bacterial populations recovered from arsenic contaminated ground water of North East India (Assam), Water Res. 47 (19) (2013) 6992–7005.
[9] D. Ghosh, P. Bhadury, J. Routh, Diversity of arsenic oxidizing bacterial communities in arsenic-rich deltaic aquifers in West Bengal, India, Front. Microbiol. 5 (2014).
[10] Z. Hassan, M. Sultana, B.M. van Breukelen, S.I. Khan, W.F. Roling, Diverse arsenic- and iron-cycling microbial communities in arsenic- contaminated aquifers used for drinking water in Bangladesh, FEMS Microbiol. Ecol. 91 (2015).
[11] Y. Gu, J.D. Van Nostrand, L. Wu, Z. He, Y. Qin, F.-J. Zhao, J. Zhou, Bacterial community and arsenic functional genes diversity in arsenic contaminated soils from different geographic locations, PLoS One 12 (5) (2017), e0176966.
[12] K. Katrak, P. De Vos, M. Gillis, J. Swings, P. Vandamme, E. Stackebrandt, Introduction to the probebacteria, in: M. Dworkin, et al. (Eds.), The Prokaryotes: Volume 5: Proteobacteria: Alpha and Beta Subclasses, Springer New York, New York, NY, 2006, pp. 3–37.
[13] S. Hiedrich, M. Schlimman, D.B. Johnson, The iron-oxidizing probebacteria, Microbiology 157 (6) (2011) 1551–1564.
[14] H. Lu, K. Chandran, D. Stensel, Microbial ecology of denitrification in biological wastewater treatment, Water Res. 64 (2014) 237–254.
[15] Z. Liszer, The ecology of anaerobic degraders of BETX hydrocarbons in aquifers, FEMS Microbiol. Ecol. 93 (1) (2017).
[16] A. Biwas, B. Nath, P. Bhattacharya, D. Halder, A.K. Kudu, U. Mandal, A. Mukherjee, D. Chatterjee, C.M. Morth, G. Jacks, Hydrogeochemical contrast between brown and grey aquifers in shallow depth of Bengal Basin: consequences for sustainable drinking water supply, Sci. Total Environ. 431 (2012) 402–412.
[17] D. Ghosh, J. Routh, P. Bhadury, Characterization and microbial utilization of dissolved organic iron fraction in arsenic impacted aquifers (India), J. Hydoll. 527 (2015) 221–233.
[18] K. Weber, L. Achenbach, J. Coates, Microorganisms pumping iron: anaerobic microbial iron oxidation and reduction, Nat. Rev. Microbiol. 4 (2006) 752–764.
[19] A. Tánczos, I. Szabo, K. Baka, S. Szoboslay, J. Kukolya, B. Kritzer, K. Máagtaligi, Investigation of catechol 2,3-dioxygenase and 16S rRNA gene diversity in hypoxic, petroleum hydrocarbon contaminated groundwater, Syst. Appl. Microbiol. 33 (7) (2010) 398–406.
[20] M. Muhling, J. Woolven-Allen, J.C. Murrell, I. Joint, Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities, ISME J. 2 (4) (2008) 379–392.
[21] D.A. Cataldo, M. Hanson, L.E. Schrader, V.L. Youngs, Rapid colorimetric method for measuring arsenic concentrations in groundwater, Anal. Chim. Acta 526 (2) (2004) 203–209.
[22] M.R. Liles, L.L. Williamson, J. Rodburrer, V. Torsvik, L.C. Parsley, R.M. Goodman, J. Handelman, Isolation and cloning of high-molecular-weight metagenomic DNA from soil microorganisms, Cold Spring Harb. Protoc. 2009 (8) (2009) pdb.prot5271.
[23] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, S.писанов, J. Gerken, T. Schweer, P. Yarza, J. Peplies, A.W. Ginhart, O. Gross, W. Ludwig, O. Strunk, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and clustal X version 2.0, Bioinformatics 23 (21) (2007) 2947–2948.
[24] A. Loy, R. Arnold, P. Tischler, T. Ratel, M. Wagner, M. Horn, probeCheck – a central resource for evaluating oligonucleotide probe coverage and specificity, Environ. Microbiol. 10 (10) (2008) 2894–2898.
[25] T. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp. Ser. 41 (1999) 95–98.
[26] T. Huber, G. Faulkner, P. Hugenholtz, Bellerophon: a program to detect chimeric sequences in multiple sequence alignments, Bioinformatics 20 (4) (2004) 2317–2318.
[27] P.D. Schloss, S.L. Westcott, T. Ryabin, J.R. Hall, M. Hartmann, E.B. Holscher, R.A. Lesniewski, B.B. Oakley, D.H. Parks, C.J. Robinson, J.W. Sahl, B. Stres, G.G. Hallinger, D.J. Van Horn, C.F. Weber, Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities, Appl. Environ. Microbiol. 75 (23) (2009) 7537–7541.
[28] D.H. Parks, G.W. Tyson, P. Hugenholtz, R.G. Beiko, STEMPATH: statistical analysis of taxonomic and functional profiles, Bioinformatics 30 (21) (2014) 3123–3124.
[29] E. Pruessa, J. Jeples, F.O. Glockner, SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes, Bioinformatics 28 (14) (2012) 1823–1829.
[30] C. Quast, E. Pruessa, F. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Jeples, F.O. Glockner, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools, Nucleic Acids Res. 41 (2013) D590–D596 (Database issue).
S. Grumann, S. Hermann, R. Jost, A. König, T. Liss, R. Lussmann, M. May, B. Nohoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, K.H. Schleifer, ARB: a software environment for sequence data, Nucleic Acids Res 32 (4) (2004) 1363–1371.

[42] I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees, Nucleic Acids Res. 44 (W1) (2016) W242–W245.

[43] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag, New York, 2009.

[44] A. Valverde, P. Delvanto, A. Peix, E. Velázquez, I. Santa-Regina, A. Ballester, C. Rodríguez-Barrueco, C. García-Balboa, J.M. Igual, Burkholderia ferrariae sp. nov., isolated from an iron ore in Brazil, Int. J. Syst. Evol. Microbiol. 56 (Pt 10) (2006) 2421–2425.

[45] B.J. Mailloux, E. Alexandrova, A.R. Keinowitz, K. Wovkulich, G.A. Freyer, M. Herron, J.F. Stolz, T.C. Kenna, T. Pichler, M.L. Polizzotto, H. Dong, M. Bishop, P.S.K. Knappett, Microbial mineral weathering for nutrient acquisition releases arsenic, Appl. Environ. Microbiol. 75 (8) (2009) 2558–2565.

[46] S.S. Pawitwar, S.M. Utturkar, S.D. Brown, M. Yoshinaga, B.P. Rosen, Draft genome sequence of Burkholderia sp. MR1, a methylarsenate-reducing bacterial isolate from Florida golf course soil, Genome Announc. 3 (3) (2015) e00608–e00615.

[47] S. Das, S.S. Bora, R.N.S. Yadav, M. Barooah, A metagenomic approach to decipher the indigenous microbial communities of arsenic contaminated groundwater of Assam, Genomics Data 12 (2017) 89–96.

[48] A.C. Layton, A. Chauhan, D.E. Williams, B. Mailloux, P.S. Knappett, A.S. Ferguson, L.D. McKay, M.J. Alam, K. Matin Ahmed, A. van Geen, G.S. Sayler, Metagenomes of microbial communities in arsenic- and pathogen-contaminated well and surface water from Bangladesh, Genome Announc. 2 (6) (2014).

[49] J.V. Weiss, J.A. Rentz, T. Plaia, S.C. Neubauer, M. Merril-Floyd, T. Lilburn, C. Bradburne, J.P. Megenical, D. Emerson, Characterization of neutrophilic Fe(II)-oxidizing bacteria isolated from the rhizosphere of wetland plants and description of Ferritrophicum radicicola gen. nov. sp. nov., and Sideroxydans paludicola sp. nov, Geomicrobiol. J. 24 (7) (2007) 599–597.