Data in Brief 33 (2020) 106610

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Data on the optimization of an archaea-specific probe-based qPCR assay

Malin Bomberg*, Hanna Miettinen

VTT Technical Research Centre of Finland Ltd., Tietotie 2, FIN-02150 Espoo, Finland

ARTICLE INFO

Article history:
Received 28 July 2020
Revised 26 November 2020
Accepted 27 November 2020
Available online 30 November 2020

Keywords:
Microorganisms in mine processing water
Microbial metabolism
High throughput sequencing
Bioinformatics
Bacteria

ABSTRACT

Estimation of archaeal numbers by use of fluorescent DNA binding dyes is challenging, because primers targeting the archaeal 16SrRNA genes readily also bind to bacterial 16S rRNA gene sequences, especially when the relative abundance of bacteria is greater than that of archaea. In order to increase specificity, we optimized a fluorescent probe-based assay using previously published archaeal primers and probe. The assay was tested on genomic DNA of pure bacterial and archaeal cultures and optimized using PCR amplicons of the archaeal pure cultures. The used bacterial strains showed slight amplification using the fluorescent dye assay, whereas all archaeal strains could be amplified with the archaea primers used. Due to differences in genome size and number of 16S rRNA gene copies between the tested archaeal strains, the amplification level varied greatly between the strains. Therefore, we also tested the amplification using PCR amplified fragments of the archaeal 16S rRNA genes. The tests with the archaeal 16S rRNA gene amplicons showed good amplification, although the amplification efficiency still varied between archaeal strains. The qPCR assay was used to estimate the archaeal numbers in process water of a multi-metal mine's metallurgical plant [1] and will be used in similar future microbiological analysis included in the H2020 ITERAMS project (Grant agreement# 730480).

DOI of original article: 10.1016/j.resmic.2020.07.001
* Corresponding author.
E-mail address: malin.bomberg@vtt.fi (M. Bomberg).
Social media: (M. Bomberg)

https://doi.org/10.1016/j.dib.2020.106610

Specifications Table

Subject	Applied Microbiology and Biotechnology
Specific subject area	The qPCR assay for estimating archaeal numbers in environmental samples was optimized for more specific detection of archaea. Archaeal 16S rRNA gene detection may be hampered by unspecific binding of archaeal primers to bacterial targets. This assay used archaeal specific primers in combination with an archaea specific probe equipped with a fluorescent label and greatly improved the specific detection of archaea simultaneously decreasing background noise.
Type of data	Tables with information on the DNA concentrations and purity of the used bacterial and archaeal strains (Table 1) and a comparison of the average Cp values of the bacterial and archaeal genomic DNA tested with bacterial and archaeal primers in a Sybr green based qPCR assay (Table 2) and comparison of the qPCR performance between the different tested conditions (Table 3). Charts presenting the average detection level between the different qPCR tests (Fig. 1) and the between run reproducibility of the used standard dilution series (Fig. 2). Supplementary data in an excel file containing method description, curated raw data from the tests, and information about the used qPCR standard.
How data were acquired	The data was produced with real-time quantitative PCR (qPCR) using fluorescent dye-based and fluorescently labeled probe-based assays. Instrument: LightCycler 480 (Roche)
Data format	Raw data in excel sheets Analyzed data in excel sheets
Parameters for data collection	Data was collected by testing the specificity of archaeal 16S rRNA gene targeting primers to DNA of archaeal pure cultured strains. The test was performed using both bacterial and archaeal strains.
Description of data collection	The data was obtained by running qPCR assays using pure cultured bacterial and archaeal strains (Table 1) at three different annealing temperatures (57 °C, 62 °C or 64 °C) using genomic DNA and a fluorescent dye-based qPCR kit. The assay was further optimized by using archaeal 16S rRNA gene amplicons as template for a qPCR assay employing a fluorescently labeled probe. Two figures are included in the text. Fig. 1 compares the level and specificity of the detection of archaeal 16S rRNA genes in the genomic DNA extracts of archaeal and bacterial strains using Sybr green detection and archaea-specific fluorescent probe (A) and the detection level and sensitivity of the probe-based qPCR when using PCR amplicons of archaeal 16S rRNA genes as template. Fig. 2 presents the between run reproducibility of the standard dilution series.
Data source location	Institution: VTT Technical Research Centre of Finland, Ltd City/Town/Region: Espoo Country: Finland
Data accessibility	The data is presented as figures within the article as well as a supplementary excel file with raw data and analyzed data.
Related research article	Malin Bomberg, Hanna Miettinen, Benjamin Musuku, Päivi Kinnunen, 2020. First insights to the microbial communities in the plant process water of the multi-metal Kevitsa mine, Research in Microbiology, https://doi.org/10.1016/j.resmic.2020.07.001.

Value of the Data

- The data describes the optimization of a probe-based archaea specific qPCR assay.
- The method can be used by any researcher that desires higher specificity to detect archaea in environmental samples.

- Using the optimized method for determining the size of the archaeal community in mixed-community environmental samples will decrease the background and false positive amplification caused by the bacteria present in the same samples.
- The added value of this approach is the specific targeting of the archaea.

1. Data Description

This data is part of a publication [1] and describes the optimization of a quantitative PCR (qPCR) assay for specific detection and enumeration of archaea from natural and process water samples. This optimization was needed in order to decrease background noise caused by unspecific binding of primers to bacterial templates. Specific primers and probe targeting the archaeal 16S rRNA gene were tested on DNA from bacterial and archaeal pure cultures (Table 1). Suitable annealing temperature conditions were tested (57 °C, 61 °C, 64 °C) for increased specificity, but also for testing the upper limits of detectability for different archaea. The raw data from these temperature tests with genomic DNA and the amplification efficiency test using archaeal 16S rRNA gene amplicon as template, obtained from the LightCycler480 used, as well as analyzed data is presented in a supplemental data file (SupplementaryData.xlsx). The outcome from the qPCR optimization with annealing temperature 57 °C are presented in Fig. 1A,B and the between runs reproducibility is presented in Fig. 2.

The supplementary data contains the following information per sheet;

1 Description of the method used for optimizing the qPCR assay using genomic DNA extracts of bacterial and archaeal strains (Method 1), and 16S rRNA gene amplicons from archaeal strains (Method 2).
2 The data from the qPCR run with genomic DNA of bacterial and archaeal strains with Method 1 and annealing temperature of 57 °C. Each sample is run in triplicate reactions from which the average detection values with standard deviation was calculated.
3 The data from the qPCR run with genomic DNA of bacterial and archaeal strains with Method 1 and annealing temperature of 61 °C. Each sample is run in triplicate reactions from which the average detection values with standard deviation was calculated.
4 The data from the qPCR run with genomic DNA of bacterial and archaeal strains with Method 1 and annealing temperature of 64 °C. Each sample is run in triplicate reactions from which the average detection values with standard deviation was calculated.

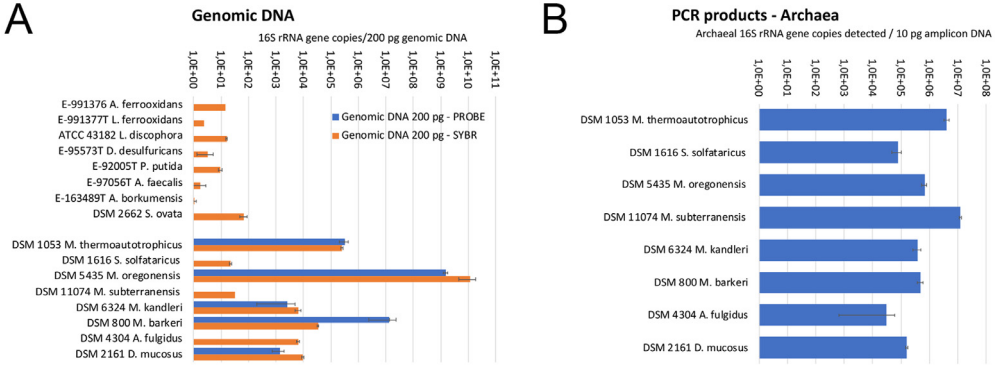

Fig. 1. The detection of amplification by qPCR from A) genomic DNA and B) 16S rRNA gene amplicon, using annealing temperature 57 °C. In A) the orange bars show the amplification detected using fluorescent dye (Sybr) and the blue bars show the detection using a probe specific to archaeal 16S rRNA genes from 200 pg genomic DNA. In B) the detection was tested using a PCR amplicon as template in order to have equal number of targets in all reactions. The DNA concentration/reaction in B was approximately 0.01 pg/reaction and the probe was used for detection.

Table 1
The bacterial and archaeal strains used as template for optimizing the archaeal qPCR assay. The genomic DNA concentration of all strains was set to a working concentration of 200 pg/µL. The qPCR efficiency was further tested using amplicons of the archaeal 16S rRNA genes. The targeted DNA concentrations of the prepared amplicons were 10 pg/µL, but ranged in general between 8 and 11 pg/µL, except for DSM 800 *M. barkerii*, for which the amplicon DNA concentration was lower. The qPCR was run using two different dilutions of the amplicons.

1Genomic DNA, strain	DNA extraction	DNA purity, Nanodrop-1000 (A260/A280)	DNA yield, Qbit, ng/mL	Working DNA concentration pg/µL
Bacteria				
E-991376 Acidithiobacillus ferrooxidans	NucleoSpin Microbial DNA (Macherey-Nagel)	1.25	696	200
E-991377T Leptospirillum ferrooxidans	NucleoSpin Microbial DNA (Macherey-Nagel)	5.84	360	200
ATCC 43182 Leptothrix discophora	NucleoSpin Microbial DNA (Macherey-Nagel)	2.63	23,800	200
E-95573T Desulfovibrio desulfuricans	NucleoSpin Microbial DNA (Macherey-Nagel)	2.16	38,600	200
E-92005T Pseudomonas putida	NucleoSpin Microbial DNA (Macherey-Nagel)	1.89	159,000	200
E-97056T Alcaligenes faecalis	NucleoSpin Microbial DNA (Macherey-Nagel)	4.76	3560	200
E-163489T Alcanivorax borkumensis	NucleoSpin Microbial DNA (Macherey-Nagel)	3.19	14,900	200
DSM 2662 Sporomusa ovata	NucleoSpin Microbial DNA (Macherey-Nagel)	3.1	278	200
Archaea				
DSM 1053 Methanothermobacter thermoautotrophicus	MoBio PowerSoil DNA extraction kit (MoBio)	1.72	85,400	200
DSM 1616 Saccharolobus solfataricus	MoBio PowerSoil DNA extraction kit (MoBio)	1.16	97,400	200
DSM 5435 Methanolobus oregonensis	MoBio PowerSoil DNA extraction kit (MoBio)	1.43	61,000	200
DSM 11074 Methanobacterium subterranensis	MoBio PowerSoil DNA extraction kit (MoBio)	1.44	279	200
DSM 6324 Methanopyrus kandleri	MoBio PowerSoil DNA extraction kit (MoBio)	2.27	3120	200
DSM 800 Methanosarcina barkeri	MoBio PowerSoil DNA extraction kit (MoBio)	2,44	6940	200
DSM 4304 Archaeoglobus fulgidus	MoBio PowerSoil DNA extraction kit (MoBio)	2.11	83,600	200
DSM 2161 Desulfurococcus mucosus	MoBio PowerSoil DNA extraction kit (MoBio)	2.94	65,200	200
		DNA purity, DeNovix DS-11 Nanodrop (A260/A280)	DeNovix DS-11 Nanodrop ng/µL	Working DNA concentration pg/µL
DSM 1053 Methanothermobacter thermoautotrophicus	NucleoSpin gel and PCR cleanup kit (Macherey-Nagel)	1.7	32.6	9.9
DSM 1616 Saccharolobus solfataricus	NucleoSpin gel and PCR cleanup kit (Macherey-Nagel)	1.7	41.4	8.2
DSM 5435 Methanolobus oregonensis	NucleoSpin gel and PCR cleanup kit (Macherey-Nagel)	1.3	63.4	8.9
DSM 11074 Methanobacterium subterranensis	NucleoSpin gel and PCR cleanup kit (Macherey-Nagel)	3.0	150.2	9.2
DSM 6324 Methanopyrus kandleri	NucleoSpin gel and PCR cleanup kit (Macherey-Nagel)	2.9	144.5	10.5
DSM 800 Methanosarcina barkeri	NucleoSpin gel and PCR cleanup kit (Macherey-Nagel)	2.5	122.6	11.2
DSM 4304 Archaeoglobus fulgidus	NucleoSpin gel and PCR cleanup kit (Macherey-Nagel)	1.1	54.2	8.7
DSM 2161 Desulfurococcus mucosus	NucleoSpin gel and PCR cleanup kit (Macherey-Nagel)	1.0	50.1	9.3

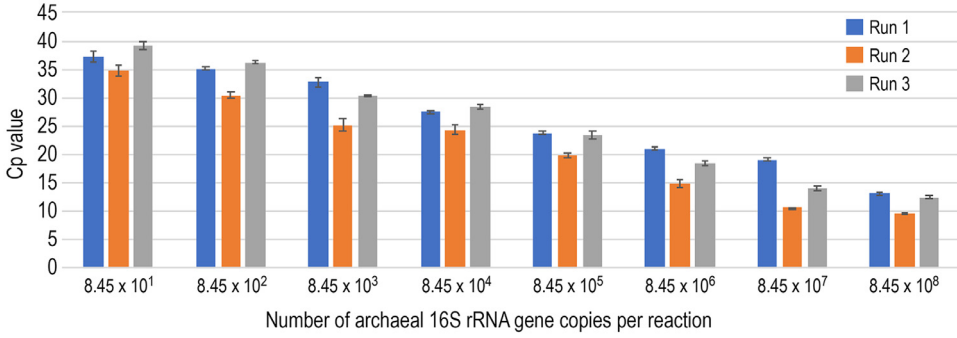

Fig. 2. Variation in the Cp values of the *H. salinarium* standard dilutions between runs. Each column shows the average Cp value calculated from three replicate reactions and the error bars show standard deviation.

5 Compiled comparison of the results of the genomic DNA tests with different annealing temperature detected with the Sybr green assay.
6 The data from the probe-based qPCR run with archaeal 16S rRNA gene amplicons tested with Method 2. Each sample is run in triplicate reactions from which the average detection values with standard deviation was calculated.
7 Comparison of Method 1 and Method 2, annealing temperature of 57 °C.
8 Data on the reproducibility between runs using standard dilutions containing $8.45 \times 10^{1\text{-}8}$ copies of *Halobacterium salinarum* 16S rRNA gene fragment in TOPO-TA plasmids.
9 Calculation of the qPCR standard concentration

2. Experimental Design, Materials and Methods

The primers used for specific detection of archaea have been shown to also target bacteria to some degree, e.g. [2]. For estimation of archaeal numbers with qPCR assays, even a low degree of unspecific amplification may have great implications for the outcome of the test and may greatly overestimate the number of archaea in any given environment. In order to reduce unspecific detection of archaea, a probe-based qPCR assay was developed. The v4v5 region of the archaeal 16S rRNA gene was targeted using primers A344F [3] and A744R, modified from [4], which produce an approximately 400-bp fragment of the archaeal 16S rRNA gene. The amplification was detected with a FAM-labeled probe, A516F [5]. The amplification level was compared to that of a tenfold dilution series of plasmids containing a fragment of the 16S rRNA gene of the archaeon *Halobacterium salinarum*^{T} strain VTT E-103154.

The specificity of the primer and probe combination was tested on pure culture DNA extracts from 8 bacterial and 9 archaeal strains (Table 1) in addition to the archaeal standard. DNA from the different strains was extracted from 1 mL of liquid culture. Bacterial DNA was extracted using the Macherey-Nagel Genomic DNA extraction kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany) and the archaeal DNA was extracted using the MoBio UltraClean DNA extraction kit (MoBio Laboratories, Carlsbad, CA, USA). DNA extraction reagent controls were included in the extraction procedure as contamination controls. No RNase or DNase treatment was performed. The concentration and purity (A260/A280) of the extracted DNA samples was examined using a Nanodrop ND-1000 or Nanodrop Denovix DS-11 spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA).

All genomic DNA extracts were diluted to 200 pg/µL in order to normalize the assay. The bacterial genomic DNA extracts were tested in a bacterial 16S rRNA gene targeted qPCR assay before use in order to test them for inhibition or contamination. In addition, both bacterial and archaeal genomic DNA extracts were tested in an archaeal 16S rRNA gene targeted qPCR assay

Table 2
The average Crossing point (Cp) values for the original genomic DNA extracts, negative DNA extraction kit controls and no template controls in the bacterial and archaeal 16S rRNA gene targeted Sybr green assays, respectively. Archaeal DNA extracts were not tested in the bacterial assay (na, not analysed). – indicates negative result.

	Bacterial 16S, 60 °C, Cp	Archaeal 16S, 57 °C, Cp
E-991376 Acidithiobacillus ferrooxidans	18.87	32.98
E-991377T Leptospirillum ferrooxidans	22.84	35.32
ATCC 43182 Leptothrix discophora	14.66	32.83
E-95573T Desulfovibrio desulfuricans	20.53	35.13
E-92005T Pseudomonas putida	10.58	33.52
E-97056T Alcaligenes faecalis	20.26	35.89
E-163489T Alcanivorax borkumensis	12.38	36.40
DSM 2662 Sporomusa ovata	18.70	30.93
No template control	28.08	–
Negative DNA extraction kit control	32.09	–
DSM 1053 Methanothermobacter thermoautotrophicus	na	20.02
DSM 1616 Saccharolobus solfataricus	na	32.41
DSM 5435 Methanolobus oregonensis	na	–
DSM 11074 Methanobacterium subterranensis	na	31.90
DSM 6324 Methanopyrus kandleri	na	24.83
DSM 800 Methanosarcina barkeri	na	22.65
DSM 4304 Archaeoglobus fulgidus	na	24.85
DSM 2161 Desulfurococcus mucosus	na	24.32

in order to test specificity of the archaeal primers to archaeal 16S rRNA genes. The assay was performed using the SensiFAST SYBR No-ROX 2 × master mix for LightCycler480 (Bioline, London, UK). Triplicate 10 µL reactions were prepared for each DNA sample. The reaction mixtures contained 800 nM of primers S-D-Bact-0341-b-S-17 and S-D-Bact-0785-a-A-21 [6] for the bacterial 16S rRNA gene assay and each bacterial primer, and primers A344F and A744R [3,4] for the archaeal assay. A total of 1 µL DNA, negative kit control or PCR grade water was added as template to each reaction. The qPCR reactions were prepared in white-walled 96-well plates with transparent film cover (4titude, Surrey, UK). The amplification programme consisted of an initial denaturation step at 95 °C for 15 min, followed by 35 amplification cycles of 10 s at 95 °C, 35 s at 60 °C for the bacterial assay and 57 °C for the archaeal assay, and 30 s at 72°C, and a final elongation step of 3 min at 72 °C and a melting curve analysis. The average Crossing points (Cp) for the cycle of detection of bacterial 16S rRNA genes in the bacterial assay was between 10.58 and 22.87 for the bacterial DNA extracts, which was always clearly earlier than the Cp for the no template control reactions or the DNA extraction kit controls (Table 2). In the archaea-targeting Sybr green assay, the bacterial DNA extracts were detected at Cp 30.93 and later, whereas most archaeal isolates were detected at Cp between 20.02 and 24.85. Strains DSM 1616 *Saccharolobus solfataricus,* DSM 5435 *Methanolobus oregonensis* and DSM 11074 *Methanobacterium subterranensis* were exceptions and were detected at Cp 31.90–32.41, or not at all. No signal was obtained from the no template control reactions or the DNA extraction kit controls in the archaeal Sybr green assay. Serial dilutions of the bacterial and archaeal genomic DNA extracts and archaeal PCR amplicons were prepared (dilutions 10^{-1} to 10^{-5}) in order to test for inhibition and limit of detection (LOD). No inhibition of the amplification reaction was seen, as determined by extrapolating the amplification results according to the dilution factor. However, low binding of the archaeal primers to the DSM 1616 *Saccharolobus solfataricus,* DSM 5435 *Methanolobus oregonensis* and DSM 11074 *Methanobacterium subterranensis* 16S rRNA genes in the genomic DNA samples was observed, but was not affected by the dilution of the template. In contrast, the annealing temperature had a greater effect on primer binding to these targets, indicating that either the assay conditions or the primers are unsuitable for detecting the genomes of these archaea. It should be noted, however, that similar low primer binding was not observed with the PCR amplicons of the 16S rRNA genes of these three strains.

The probe assay was also run in triplicate in 10 µL volumes using the SensiFAST Probe No-ROX kit (Bioline, London, UK). The reactions contained 500 nM of each primer, 200 nM probe,

Table 3
Comparison between the Sybr green (Sybr) and probe-based assays under different conditions. LOD indicated limit of detection. The limit of detection is based on the lowest standard concentration and the lowest detected number of gene copies in the dilution series of the samples.

Assay	Error	Efficiency	Y-intercept	Slope	LOD, gene copies per reaction	Cp for LOD	Cp for NTC
Archaea, 57 °C, Sybr	0.347	2.128	33.68	−3.049	20	32.41	–
Archaea, 57 °C, probe	0.336	2.09	39.45	−3.127	20	32.62	–
Archaea, 61 °C, probe	0.358	2.03	45.11	−3.259	380	36.86	–
Archaea, 63 °C, probe	0.976	1.993	44.98	3.338	3570	34.10	–

and 200 pg of template DNA. The amplification reaction consisted of enzyme activation at 95 °C for 3 min and 40 cycles of 10 s at 95 °C, 30 s at 57 °C, 62 °C or 64 °C, and 1 s at 72 °C. All qPCR runs were done using the LightCycler 480 instrument (Roche Diagnostics, Basel, Switzerland). Bacteria were not detected with any of the tested annealing temperatures. However, the amplification level obtained from the DNA extracts varied greatly between the archaeal strains. Thus, a fragment of each archaeal 16S rRNA gene was amplified with the 2× MyTaq™ Red Mix (Bioline, London, UK), using 20 pmol each of primer A109f and A934r in 25 µL reaction volumes on an Eppendorf MasterCycler Gradient (Eppendorf, Hamburg, Germany). The amplification programme consisted of an initial denaturation step at 98 °C for 3 min, followed by 35 cycles of 15 s denaturation at 95 °C, 15 s primer annealing at 50 °C, and 15 s elongation at 72 °C, and a final 30 s elongation step at 72 °C. The amplification products were checked for correct size with agarose gel electrophoresis on a 1% agarose in 1× TAE buffer, and run for 30 min on 150 V. The pre-stained (MidoriGreen, Nippon Genetics, Japan) agarose gel was imaged using a Gel-Doc (Bio-Rad, Hercules, CA, USA) imager and the correct sized bands were cut from the gel with sterile scalpels. The excised ca 800 bp PCR amplicon was purified from agarose using the NucleoSpin gel and PCR cleanup kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany) and the DNA concentration of each preparate was set to approximately 10 pg/µL, of which 1 µL of 10^{-3} dilutions (i.e. 0.01 pg) were used for the qPCR test. The amplification was tested with primers A344F and A744R and probe A516F as described above using only the 57 °C annealing temperature.

The plasmid standard dilution series containing the *H. salinarum* 16S rRNA gene fragment amplified and run in an agarose gel, cut from the gel and purified as described above, except using annealing temperature 50 °C. The purified amplicon was then cloned into the TOPO-TA vector (ThermoFisher Scientific, Waltham, MA, USA) using the heat-shock protocol by the manufacturer. The transformants were plated on Luria-Bretani agar plates supplemented with ampicillin (100 mg mL^{-1}) without blue-white screening. After over-night incubation at 37 °C 10 colonies were screened for correct size insert using colony-PCR with the same conditions and primers as previously. Clones containing the correct insert were regrown in 5 mL Luria-Bretani + ampicillin broth over-night, shaken at 150 rpm. After the growth, the bacterial mass was pelleted by centrifugation in 2-mL aliquots in sterile 2-mL microcentrifuge tubes at 13 000 rpm using an Eppendorf table-top microcentrifuge (Eppendorf, Hamburg, Germany). The supernatant was removed by decanting and the plasmid DNA was extracted using the NucleoSpin plasmid kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany) according to the manufacturer's protocol. The concentration of the eluted plasmid DNA was measured using the DeNovix DS-11 Nanodrop spectrophotometer. The concentration of archaeal 16S rRNA gene fragments in the plasmid preparates was calculated as described in the SupplementaryData.xlsx.

The data was analyzed using the LightCycler480 analysis software version 1.5 (Roche Diagnostics, Basel, Switzerland) using the absolute quantification with the Fit points method. The noise band was adjusted manually to above any background noise, if needed. A standard curve consisting of triple reactions of a standard dilution series of 7 tenfold dilutions ranging from 8.4×10^1 to 8.4×10^8 gene copies per reaction was calculated using the LightCycler480 software, which also calculates the error rate, efficiency of the amplification, slope of the standard curve and the Y-intercept of the amplification (Table 3).

The test with DNA from pure cultures showed that several of the bacterial strains used showed amplification with the archaea-specific primers when using the Sybr qPCR kit, although only at a very low degree (Fig. 1A). In contrast, 6 out of the 8 tested archaeal strains showed higher amplification than what was detected for the bacterial strains, and three strains showed only low amplification. When the probe was used, none of the bacteria were detected, and 5 of the 8 archaeal strains could be detected when using genomic DNA. The uneven detection from the genomic DNA of the archaeal strains may be caused by the primers and probe not exactly matching the targeted gene sequences, or the number of gene copies and the size of the genomes of the tested archaea differed greatly. A possibility is also that the DNA was unevenly degraded during storage. Due to these factors, we decided to additionally test the amplification and detection with the A516F probe using PCR amplicons of the archaeal 16S rRNA genes from the tested strains. The amplification levels indicated good detection with all tested archaeal strains (Fig. 1B), although the detection level between the strains varied with two orders of magnitude between the most readily detectable *M. subterranensis* and *M. thermoautotrophicus* and the least detectable *A. fulgidus*. This shows that different archaeal strains are unequally detectable with universal archaeal primers, which should be kept in mind when estimating microbial numbers using qPCR as well as when analyzing microbial communities using amplicon sequencing. Finally, the repeatability of the assays tested by always running all samples in triplicate reactions was high within a run, with a standard deviation of the average Cp value ranging between 0.04 and 0.6 in of the standard curve and 0.1–0.44 in the samples. However, the reactions with the genomic DNA of *M. subterranensis* and *M. thermoautotrophicus* and *A. fulgidus* showed higher variation with standard deviations of the average Cp of 1.3 to 2.6. Between assays the Cp values varied more (Fig. 2), depending on the run. The highest between run variations were seen in the most dilute samples. Due to the variation in the Cp values between runs, we recommend that a standard is included in all runs.

CRediT Author Statement

Malin Bomberg: conceptualization, methodology, validation, formal analysis, investigation, data curation, writing original draft, review and editing, visualization; **Hanna Miettinen:** validation, review and editing of the manuscript.

Declaration of Competing Interest

This work was funded by the European Union's Horizon 2020 Research and Innovation program under Grant Agreement no. 730480, ITERAMS project (Integrated mineral technologies for more sustainable raw material supply). The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Acknowledgments

Tuula Kuurila is thanked for invaluable assistance in the laboratory. This project has received funding from the European Union's Horizon 2020 Research and Innovation program under Grant Agreement no. 730480, ITERAMS project (integrated mineral technologies for more sustainable raw material supply).

Supplementary Materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.dib.2020.106610.

References

[1] M Bomberg, H Miettinen, B Musuku, P Kinnunen, First insights to the microbial communities in the plant process water of the multi-metal Kevitsa mine, Res. Microbiol. 171 (2020) 230–242.
[2] MA Fischer, S Güllert, SC Neulinger, WR Streit, RA Schmitz, Evaluation of 16S rRNA gene primer pairs for monitoring microbial community structures showed high reproducibility within and low comparability between datasets generated with multiple archaeal and bacterial primer pairs, Front Microbiol. 7 (2016) 1297.
[3] N Bano, S Ruffin, B Ransom, JT Hollibaugh, Phylogenetic composition of Arctic Ocean archaeal assemblages and comparison with Antarctic assemblages, Appl. Environ. Microbiol. 70 (2004) 781–789.
[4] SM Barns, RE Fundyga, MW Jeffries, NR Pace, Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment, Proc. Natl. Acad. Sci. 91 (1994) 1609–1613.
[5] KE Takai, K Horikoshi, Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes, Appl. Environ. Microbiol. 66 (2000) 5066–5072.
[6] DP Herlemann, M Labrenz, K Jürgens, S Bertilsson, JJ Waniek, AF Andersson, Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea, ISME J. 5 (2011) 1571–1579.