Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrated Assessment of DNA Extraction Protocols for Microbiome Characterization in Low-Biomass Inorganic Building Materials

Trung Duc Le
Technical University of Liberec
Klara Liskova
Technical University of Liberec
Petr Vecernik
Nuclear Research Institute Rez (Czechia)
Jakub Riha
Technical University of Liberec
Alena Sevcu
Technical University of Liberec
Veronika Hlavackova
veronika.hlavackova@tul.cz

Technical University of Liberec

Research Article

Keywords: DNA extraction, cementitious materials, low biomass microbiome, quantitative PCR, mock community, 16s rRNA gene sequencing

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11667277/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Reliable microbiome characterization of cementitious materials is increasingly important for understanding microbial interactions with engineered environments. Although geopolymers have gained attention as alternatives to conventional Portland cement, the influence of their distinct physicochemical characteristics on microbial DNA recovery and downstream community analysis remains poorly understood. This study compared a modified chloroform-based protocol and a commercial silica column-based method for DNA extraction from geopolymer and ordinary Portland cement matrices using a defined bacterial mock consortium across a biomass dilution series (10^4–10^9 cells mL^{-1}). Extraction performance was evaluated through DNA quantification, quantitative PCR, 16S rRNA gene amplicon sequencing, and scanning electron microscopy. DNA recovery differed substantially between the two materials, with geopolymer generally yielding higher DNA concentrations and more consistent molecular signals than conventional cement. Although chloroform extraction produced higher total DNA concentrations in geopolymer samples, the column-based method recovered greater quantities of amplifiable bacterial DNA and showed better correspondence with the expected dilution series. Sequencing revealed extraction-dependent differences in recovered microbial community composition, particularly at lower biomass levels, where chloroform extracts exhibited greater variability and increased representation of additional taxa. Microscopic observations further revealed material-dependent differences in microbial attachment and mineral coverage, suggesting potential constraints on DNA accessibility. These findings demonstrate that DNA yield alone is insufficient to assess extraction performance and highlight the importance of considering material-dependent biases for reliable microbial community characterization in cementitious environments, including emerging geopolymer materials.

Introduction

Cementitious materials are widely used in infrastructure and environmental engineering, including nuclear facilities and deep geological repositories for radioactive waste disposal. Although their highly alkaline nature initially limits microbial colonization, progressive environmental exposure and physicochemical alteration can create conditions favorable for microbial attachment, biofilm development, and metabolic activity (Bone et al., 2022; Cwalina, 2008; Gaylarde & Ortega-Morales, 2023). Microbial communities associated with these materials may contribute to mineral precipitation, chemical alteration, and changes in material properties, potentially influencing their long-term performance (Le et al., 2026; Ruiz-Fresneda et al., 2023; Turick & Berry, 2016). Consequently, understanding microbial colonization and community composition in cementitious environments is increasingly relevant to both microbial ecology and the assessment of engineered material durability.

Alongside conventional ordinary Portland cement (OPC), geopolymers (GPs) have attracted growing attention as alternative inorganic binders because of their chemical durability and potential applications in sustainable construction and radioactive waste management (Duxson et al., 2007; Provis et al., 2015). Unlike OPC, which primarily forms calcium-rich hydration products, geopolymers are produced through

alkali activation of aluminosilicate precursors, resulting in distinct chemical compositions and mineral surface characteristics (Pouhet & Cyr, 2014; Singh et al., 2015). These differences may influence microbial attachment, biofilm development, and interactions between microbial cells, extracellular DNA, and mineral surfaces. Previous investigations have demonstrated microbial interactions with geopolymer and conventional cement matrices, including biofilm formation and material alteration under microbial exposure (Drugă et al., 2018; Grengg et al., 2020). Recent study has investigated the incorporation of microbial communities into geopolymer materials, further illustrating the growing interest in their microbiological characteristics (Danilyan et al., 2025). However, despite increasing application of molecular approaches to these materials, the influence of their distinct physicochemical properties on microbial DNA recovery and subsequent community characterization remains insufficiently understood.

Molecular techniques, particularly quantitative PCR and 16S rRNA gene amplicon sequencing, have become important tools for investigating microbial communities associated with cementitious materials. These approaches enable the detection and characterization of microorganisms that may be difficult to recover through conventional cultivation, providing insights into microbial diversity and community composition (Kiledal et al., 2021, 2023). Nevertheless, reliable molecular characterization depends strongly on the efficiency and reproducibility of DNA extraction, especially in low-biomass environments. Cementitious matrices present several challenges, including high alkalinity, low microbial abundance, potential PCR inhibitors, and interactions between DNA and mineral surfaces. In calcium-rich cement, DNA may adsorb onto mineral phases through electrostatic interactions and cation bridging, whereas aluminosilicate-rich geopolymer matrices may exhibit different DNA–mineral interactions (Direito et al., 2012; Freeman et al., 2023; Romanowski et al., 1991). These processes can influence DNA recovery and amplifiability, potentially introducing biases into downstream microbial community analyses. Furthermore, as microbial biomass decreases, background DNA and laboratory contamination can contribute disproportionately to sequencing results, complicating the interpretation of low-abundance microbial taxa (Salter et al., 2014).

Different DNA extraction approaches have been applied to mineral-rich and cementitious materials, including organic solvent-based protocols and commercial silica-based extraction kits (Kiledal et al., 2021; Maresca et al., 2016). Organic solvent-based approaches typically incorporate chemical and enzymatic lysis, mineral-associated DNA release, and subsequent purification, while commercial kits provide standardized extraction procedures intended to improve reproducibility and remove inhibitory compounds. However, higher total DNA recovery does not necessarily correspond to greater quantities of amplifiable microbial DNA or more representative community profiles. Previous studies have highlighted the importance of extraction methodology in mineral-rich and low-biomass environments (Direito et al., 2012; Gómez-Acata et al., 2019), yet direct comparisons of extraction performance across chemically distinct cementitious matrices remain limited. In particular, the extent to which extraction methodology and matrix composition influence DNA recovery, amplification efficiency, and the apparent composition of microbial communities in geopolymer and conventional cement has not been adequately characterized.

Therefore, this study evaluated the performance of a modified chloroform-based extraction protocol and a commercial silica column-based method for microbial DNA recovery from two chemically distinct matrices: ordinary Portland cement and geopolymer. Using a defined bacterial mock consortium across a controlled biomass dilution series, we assessed extraction performance through DNA quantification, qPCR, and 16S rRNA gene amplicon sequencing. Scanning electron microscopy was additionally employed to examine microbial attachment and material-associated surface features that may influence microbial accessibility. By integrating these complementary approaches, the study aimed to determine how extraction methodology and material composition affect the reliability of microbial DNA detection and community reconstruction, providing methodological guidance for microbiome investigations in low-biomass cementitious environments.

Materials and Methods

Cementitious material preparation

The reference matrix (CEM) was prepared as a conventional Portland cement paste using CEM I 42.5 R (CEMEX, Prachovice, Czech Republic), mixed with demineralized water at a water-to-solids ratio of 0.45. The paste was cast into cubic and cylindrical molds depending on the intended test and cured under controlled laboratory conditions. The alternative cement–bentonite matrix (AFM) was prepared by mixing CEM I 42.5 R cement with Czech Ca–Mg bentonite and demineralized water containing a superplasticizer (Sika) at a water-to-solids ratio of 0.36. The homogeneous paste was cast into 5 × 5 × 5 cm cubic molds and aged under laboratory conditions before testing. The nano zero-valent iron matrix (NNM) was prepared by incorporating commercially available nanoscale zero-valent iron particles (Nanofer STAR, Nanoiron s.r.o.) into a Portland cement paste at concentrations ranging from 0.02 to 1.0 wt% (relative to cement mass). The mixture was cast into 5 × 5 × 5 cm cubic molds and cured for at least two months before experimentation. The geopolymer matrix (GP) consisted of an alkali-activated aluminosilicate binder prepared from ground granulated blast furnace slag, sodium silicate solution, demineralized water and proprietary aluminosilicate precursors supplied by CHEMCOMEX, a.s. company (the detail recipe of GP is protected by copyright). Geopolymer paste was cast into 5 × 5 × 5 cm cubic molds and cured under controlled laboratory conditions - approximately 98% relative humidity for 28 days. Prior to microbial experiments, cured matrices were mechanically crushed and homogenized into fine powder. For DNA extraction experiments, 2.5 g of crushed CEM or GP material was used for each experimental unit.

Biofilm observation experimental setup

Biofilm formation experiments were conducted using four types of matrix materials. Each material was incubated with the mock microbial consortium at a final cell concentration of 1×10^7 cells·mL^{-1}. Incubations were performed at 28°C under static conditions. Biofilm development was monitored at three time points: T0 (after one hour), T1 = 20 days, and T2 = 40 days using scanning electron microscopy to assess microbial attachment and biofilm formation on the material surfaces.

Scanning electron microscope

Prior to the SEM observations, samples were fixed following a previously described method (Mijnendonckx et al., 2022) with modifications. Briefly, samples were immersed in a fixation solution containing 3% glutaraldehyde and 0.132M cacodylate and maintained at 4°C for one hour. Then 0.1M cacodylate pH 7 was used twice to rinse the glutaraldehyde excess and left to dissipate the remaining buffer for 5 mins. The samples afterward were dehydrated through a graded ethanol series (30%, 50%, 70%, 90%, 95%, 100% v/v) with each concentration applied as a single rinse followed by a 10-minute standing period. A final rinse with 100% ethanol followed by a treatment of hexamethyldisilazane for 5 mins. Each sample was then air-dried under a fume hood for at least 1 hour and stored in a desiccator until SEM analysis. The samples were subsequently analyzed using a Zeiss ULTRA Plus field-emission scanning electron microscope (SEM; Zeiss, Germany). Images were captured at an accelerating voltage of 5 kV and a low probe current (approximately 15 pA) utilizing an InLens secondary electron detector, operated via SmartSEM software.

Preparation of the mock microbial consortium

The mock consortium consisted of three bacterial strains selected to represent microorganisms of different ecological relevance to cementitious materials: *Pseudomonas stutzeri*, an environmental isolate from our laboratory collection; *Bacillus cohnii*, a spore-forming alkalinity-tolerant bacterium; and *Staphylococcus epidermidis* CCM 4418, a common human skin commensal. Each strain was cultivated separately in Luria–Bertani medium under shaking conditions (120 rpm) at 37°C for 5 days. After incubation, 50 mL of each individual culture was combined to obtain a mixed suspension with a final volume of 150 mL.

Quantification of microbial cell counts

Absolute cell concentrations in the mock consortium were determined by flow cytometry using a NovoCyte system (Agilent Technologies, Santa Clara, CA, USA). Prior to measurement, 1 mL of the mixed consortium was diluted 100-fold in phosphate-buffered saline. From the diluted suspension, 250 µL was stained with propidium iodide and incubated for 30 min in the dark before analysis. Cell counts were determined based on event detection by gating. The measured event counts and analyzed volume were used to calculate the absolute cell concentration of the stock consortium, which was approximately 2.24 × 10^{9} cells mL^{-1}. This stock suspension was subsequently used to prepare serial dilutions for inoculation experiments and biofilm formation assays.

Sample processing and DNA extraction

Prior to DNA extraction, only CEM and GP samples were mechanically crushed under aseptic conditions. For method comparison experiments, 2.5 g of crushed CEM or GP material was inoculated with 1 mL of a serially diluted mock microbial consortium, covering a concentration range from 1 × 10^{9} to 1 × 10^{4} cells mL^{-1}. The same sample mass was used for both extraction methods to enable direct comparison

under identical experimental conditions. Although larger input masses are commonly recommended for commercial extraction kits, the suitability of these recommendations for highly alkaline cementitious materials remains unclear, as increasing sample mass may increase the release of alkaline pore solution, calcium ions, dissolved mineral constituents, and mineral-associated DNA, while also increasing the co-extraction of PCR inhibitors. Following 1h ambient temperature incubation, DNA was extracted using two different extraction approaches to compare their DNA recovery performance and suitability for downstream molecular analyses. Each extraction was performed in duplicate for both materials and for each method.

The first approach was a modified physicochemical lysis protocol adapted from Bruijn (denoted as chloroform) (Bruijn, 2011; Yeates et al., 1998). Briefly, 2.5 g of crushed material was mixed with 2 mL lysis buffer containing 0.1 M potassium phosphate (pH 5.8), 15 mM $MgSO_4$, 30 mM Tris–acetate, 5 mM EDTA, 25 mM KCl, and 1% (w/v) SDS. Mechanical disruption was enhanced by the addition of 2.5 g autoclaved glass beads of two size fractions (0.25–0.5 mm and 2.85–3.45 mm; Carl Roth GmbH + Co. KG, Karlsruhe, Germany). To improve cell dispersion and reduce adsorption to mineral surfaces, 2.5 mL EDTA/NaPP solution (0.025 M EDTA and 0.4% w/v sodium pyrophosphate, pH 8.9) was added. Enzymatic lysis was performed using lysozyme (15 mg g^{-1} material) and proteinase K (20 mg mL^{-1}), followed by repeated cycles of vortexing and incubation at 65°C. After centrifugation, the supernatant was subjected to chloroform extraction to remove proteins and cell debris. Nucleic acids were further purified by selective precipitation using sodium acetate and isopropanol, washed with 70% ethanol, and resuspended in TE buffer. Final cleanup and concentration of DNA were performed using the Zymo Research ZR-96 DNA Clean & Concentrator kit, according to the manufacturer's instructions.

In parallel, DNA was extracted from identical samples using the DNeasy PowerMax Soil Kit (denoted as kit) (QIAGEN), following the manufacturer's instructions after adding 2.5 mL EDTA/NaPP solution (0.025 M EDTA and 0.4% w/v sodium pyrophosphate, pH 8.9). Final cleanup and concentration of DNA were performed using the Zymo Research ZR-96 DNA Clean & Concentrator kit, according to the manufacturer's instructions. DNA concentrations obtained using both extraction methods were quantified using a Qubit fluorometer and used for downstream molecular analyses.

Molecular-genetic analyses

Bacterial abundance was assessed by quantitative PCR (qPCR) targeting the bacterial 16S rRNA gene using fluorescence DNA marker i.e. SYBR green, U16SRT-F primer- ACTCCTACGGGAGGCAGCAGT and U16SRT-R primer – ATTACCGCGGCTGCTGGC (Clifford et al., 2012). LightCycler® 480 (Roche, Switzerland) was used to perform the measurements. Master mix preparation and setting conditions were described in detail previously (Shrestha et al., 2022). The obtained Ct values were averaged from duplicates, with outliers removed. Ct values ≥ 30 were considered below the reliable detection limit of the assay.

Relative 16S rRNA gene quantities were calculated using the comparative Ct method. For each extraction method and material type, the lowest inoculation level (10^4 cells mL^{-1}) was used as the

reference sample and assigned a relative quantity of 1. Relative quantification (RQ) was calculated as (Bustin et al., 2009):

$RQ = 2^{-(Ct_sample - Ct_reference)}$,

where Ct sample represents the measured Ct value of a given sample and Ct_reference represents the mean Ct value of the corresponding 10^4 cells mL^{-1} sample within the same extraction method and material type.

Amplicon sequencing targeting the bacterial 16S rRNA gene (V3 - V4 regions) was amplified using primers 515F (5'– TGCCAGCMGCNGCGG – 3') and 802R (5'–TACNVGGGTATCTAATCC–3') (Claesson et al., 2010; Dowd et al., 2008). PCR reactions were carried out in a 50 µL final volume. Amplicons were purified using the Agencourt AMPure XP system (Beckman Coulter, Brea, CA, USA), and barcoded sequencing adapters were ligated using the Ion Xpress™ Plus gDNA Fragment Library Kit and Ion Xpress™ Barcode Adapters (Thermo Fisher Scientific, Waltham, MA, USA). Sequencing was performed using the Genexus™ Integrated Sequencer (Thermo Fisher Scientific, Waltham, MA, USA) with Ion Torrent semiconductor sequencing technology, using the Ion PGM™ Hi-Q™ Sequencing Kit and Ion 314™ Chip v2, following the manufacturer's protocols.

Bioinformatics and statistical analysis

All data processing and visualization were performed in Rstudio (version 2024.09.1 + 394). DNA concentrations and relative 16S rRNA gene quantities are presented as mean values of duplicate extractions with standard deviations. To evaluate extraction performance across the dilution series, linear regression was performed between Ct values and the logarithm (base 10) of the expected cell concentration. The coefficient of determination (R^2) was used as a measure of linearity. Apparent amplification efficiency (E) was calculated from the regression slope according to:

$E = (10^{(-1/slope)} - 1) \times 100$ (Bustin et al., 2009)

Regression analyses were performed for the full dilution series (10^4–10^9 cells mL^{-1}) as well as reduced dilution ranges (10^5–10^9 and $10^{6-1}0^9$ cells mL^{-1}) to assess the influence of low-biomass samples on extraction performance.

Raw sequencing data were processed using QIIME 2 version 2024.10 (Bolyen et al., 2019). Demultiplexing and quality assessment were performed using the q2-demux plugin, followed by denoising, quality filtering and amplicon sequence variant (ASV) inference with DADA2 implemented through q2-dada2 (Callahan et al., 2016). Taxonomic assignment was performed using the q2-feature-classifier plugin with the classify-sklearn naïve Bayes classifier (Bokulich et al., 2018) trained against the SILVA 138.2 reference database (Quast et al., 2013) Sequences assigned to chloroplasts and mitochondria were removed prior to downstream analyses. Taxonomic classification accuracy was verified using the artificial mock community included in the sequencing run. Subsequent analyses were

conducted in R using the qiime2R and phyloseq packages (McMurdie & Holmes, 2013). Samples with fewer than 100 sequencing reads were excluded, and taxa with zero counts remaining after filtering were removed. For taxonomic composition analyses, ASVs were agglomerated at the genus level, and genera with a mean relative abundance below 1% across all samples were excluded from visualization.

Alpha diversity indices, including Observed richness, Shannon, Simpson and inverse Simpson diversity, were calculated using rarefied ASV tables. Differences in alpha diversity between extraction methods were assessed using pairwise Wilcoxon rank-sum tests implemented in the ggpubr package. Beta diversity analyses were performed using Bray–Curtis dissimilarities calculated from relative abundance data without rarefaction. Community differences were visualized by Principal Coordinates Analysis (PCoA) and hierarchical clustering heatmaps. Statistical differences in community composition between extraction methods were evaluated using permutational multivariate analysis of variance (PERMANOVA; adonis2, 999 permutations) implemented in the vegan package (Oksanen et al., 2022), while homogeneity of multivariate dispersion was assessed using PERMDISP **(**betadisper**)**

Differential abundance analysis was performed at the genus level using DESeq2 (Love et al., 2014). Differentially abundant taxa between chloroform and kit extraction methods were identified using a negative binomial generalized linear model, and statistical significance was determined using Wald tests with Benjamini–Hochberg correction for multiple comparisons. Taxa with an adjusted p-value < 0.05 were considered significantly differentially abundant. Final visualizations, including relative abundance bar plots and differential abundance (lollipop) plots, were generated using ggplot2 (Oksanen, 2011).

Results

Microbial colonization and biofilm formation

SEM analysis of CEM revealed pronounced time-dependent changes in surface morphology under both biotic and abiotic conditions (Fig. 1 and Fig. 2). At T0 (1 h), CEM samples inoculated with 10^7 cells/ml exhibited a porous and heterogeneous microstructure, with rod-shaped bacteria and filamentous on the surface (white arrow in Fig. 1A), whereas abiotic controls showed a comparatively rough, compact and granular mineral surface (Fig. 2A After 20 days (T1), biotic samples showed occasional presence of needle-like crystals and microorganisms (rods) on relatively smooth surface (white arrows in Fig. 1B), while abiotic samples showed dense mineral precipitates forming an uniform crystalline coverage (Fig. 2B). By 40 days (T2), both biotic and abiotic CEM surfaces were characterized by extensive mineral crust formation, and no distinct biofilm structures or recognizable bacterial morphologies could be visually identified in biotic samples. Apparently, the surface morphology differed significantly between biotic and abiotic samples showing continuous crystallization in abiotic controls contrasting mineralization of biofilm filaments (Fig. 1C and Fig. 2C).

At T0, AFM surfaces of samples inoculated with 10^7 cells/ml exhibited heterogeneous microtopography with irregular particulate aggregates and filamentous structures and rod-shaped bacteria bridging

surface features, indicating early microbial attachment and matrix interaction (white arrows in Fig. 1D). In contrast, abiotic AFM controls displayed a more compact and no sign of microorganisms (Fig. 2D). At T1, filamentous or biofilm-like features were still locally detectable but appeared embedded within the surface layer in biotic samples (Fig. 1E). By 40 days (T2), distinct microbial structures were no longer visible in biotic samples; surfaces were dominated by dense, fine-grained mineral layers with reduced textural contrast relative to abiotic controls (Fig. 1F and 2F).

In NNM samples, SEM observations revealed rapid surface modification and progressive mineral coverage in both biotic and abiotic samples (Figs. 1 and 2). At T0, biotic samples showed rod-shaped microorganisms embedded within heterogeneous matrices on surface (Fig. 1G). At T1, biotic NNM surfaces samples became increasingly compact, with microbial morphologies no longer clearly distinguishable and the surface dominated by fine-grained aggregates or extracellular polymeric substances (EPS) with mineral deposits (Fig. 1H). By 40 days (T2), both biotic and abiotic samples were largely covered by dense granular mineral aggregates that obscured the original surface morphology (Fig. 1I and 2I). Nevertheless, a few larger embedded elongated structures remained visible within the mineral layer (Fig. 1I). They may represent mineral-encrusted microbial aggregates, biofilm remnants, or pre-existing matrix fragments incorporated into the newly precipitated mineral layer. The surrounding dense granular texture indicates that progressive mineral precipitation and grain growth dominated surface evolution during incubation. However, this does not necessarily indicate the absence of biofilm, as ongoing mineral precipitation may progressively encrust or overlay pre-existing microbial structures. The occurrence of similar mineral overgrowth in both biotic and abiotic samples further suggests that mineral-driven surface reorganization is a dominant process in NNM over time.

In GP samples, persistent microbial morphologies across all incubation times were detected. At T0, biotic samples showed clear rod-shaped cells consistent with *Bacillus*- and *Pseudomonas*-like morphology (Fig. 1J), which were absent in abiotic controls (Fig. 2J). After 20 days, *Bacillus*-like cells were frequently observed, displaying thickened, rough surface envelopes with fractured, stone-like textures, suggesting calcium carbonation encrustation (white arrow in Fig. 1K). Notably, at 40 days, rod-shaped microorganisms remained detectable on the GP surface, indicating sustained morphological preservation (white arrow in Fig. 1L). In contrast to CEM and AFM materials, GP samples showed limited evidence of extensive carbonate crust formation, and microbial structures were not fully masked by secondary mineral layers over time, unlike abiotic control (Fig. 2K and L).

Overall, colonization of the four studied materials followed distinct scenarios reflecting differences in the physicochemical properties of the cementitious and geopolymer matrices. Cement-based matrices with pH values around 13 appeared to hinder microbial colonization or promote the development of protective EPS. Encapsulation of bacteria within biofilms may have protected them from highly alkaline leachates released from the matrix but may also have limited microbial detectability after 40 days of incubation.

Extraction efficiency: comparison of materials

CEM (pH 12.4) was used to evaluate the impact of a calcium-rich environment on DNA recovery and downstream molecular analyses. GP represents a chemically distinct matrix with lower pH (11.5) and was therefore analyzed in parallel to assess material dependent effects on extraction performance.

DNA recovery differed substantially between extraction methods and material types (Fig. 3A). In CEM samples, DNA concentrations were generally low regardless of extraction method, although the kit extraction method yielded slightly higher DNA concentrations than the chloroform extraction at higher inoculation levels (10^8–10^9 cells·mL⁻¹). In contrast, chloroform extraction produced substantially higher DNA concentrations than kit extraction in GP samples, particularly at 10^9 cells·mL⁻¹, where DNA concentrations were approximately three-fold higher than those obtained using kit extraction. Overall, DNA yields were consistently lower in CEM than in GP samples.

Spectrophotometric purity assessment provided critical context for these extraction outcomes (Table S1). Samples processed via chloroform extraction exhibited high variability and severe purity anomalies; $A_{260/280}$ ratios ranged inconsistently, producing notable negative artifacts (e.g., down to -20.65) as well as artificially elevated values (up to 4.51). Furthermore, $A_{260/230}$ ratios in chloroform extracts remained persistently predominantly below 0.5, pointing to heavy carryover of organic solvents, residual reagents, or background matrix interference. Conversely, kit-extracted samples generally exhibited less variable $A_{260/280}$ ratios, although occasional anomalous values were also observed (0.18–14.30). $A_{260/230}$ ratios remained consistently low (0.06–0.74), indicating that spectrophotometric purity was limited across both extraction methods.

Quantification of bacterial 16S rRNA genes revealed a different pattern from that observed for total DNA concentration (Fig. 3B). Relative 16S rRNA gene quantities increased with increasing cell abundance for both extraction methods and matrices. However, despite producing lower total DNA concentrations, kit extracts generally yielded equal or higher relative 16S rRNA gene quantities than chloroform extracts at comparable inoculation levels. This difference was particularly evident in GP samples at 10^6–10^9 cells·mL⁻¹, where kit extraction samples consistently recovered higher quantities of amplifiable bacterial DNA despite lower total DNA concentrations. This observation may be associated with differences in DNA purity and the co-extraction of inhibitory compounds between the two extraction methods. In addition, kit-extracted samples exhibited lower variability among replicate samples, whereas chloroform extracted samples showed greater variation, especially at lower inoculation levels, consistent with the observed spectrophotometric instability.

The relationship between 16S rRNA gene abundance and the expected dilution series was further assessed using linear regression analysis. In GP samples, kit extracts showed the strongest correspondence with the expected dilution trend, achieving apparent amplification efficiencies of 95.1% and an R^2 value of 0.949 across the 10^5 – 10^9 cells·mL⁻¹ range (Table. S2). In contrast, chloroform extracts only approached optimal amplification performance after exclusion of the two lowest biomass dilutions, yielding an apparent efficiency of 101% and an R^2 value of 0.900 for the 10^6 – 10^9 cells·mL⁻¹

range (Table. S2). In CEM samples, both extraction methods exhibited poor linearity and inflated apparent amplification efficiencies, indicating substantial material-associated interference with DNA recovery and amplification.

Bray Curtis dissimilarity analysis revealed differences in bacterial community composition associated with DNA extraction method, although extraction protocol was not the primary source of variation within the dataset (Fig. 4A). Samples extracted using kit extraction generally exhibited lower pairwise dissimilarities and formed relatively coherent clusters across materials (CEM and GP) and mock concentrations. In particular, high-biomass samples (≥ 10^7 cells·mL^{-1}) extracted using either method showed substantial similarity, indicating consistent recovery of the dominant mock community members. However, several low-biomass samples extracted using chloroform, particularly CEM samples inoculated with 10^4–10^5 cells·mL^{-1}, displayed markedly higher dissimilarities and clustered separately. Similar patterns were observed for some extraction controls, which showed increased dissimilarity relative to the majority of samples.

PCoA supported the clustering patterns observed in the Bray–Curtis heatmap (Fig. 4B). High-biomass samples extracted using both protocols clustered closely together, whereas low-biomass chloroform samples were more widely dispersed along PC1. PERMANOVA analysis confirmed a significant effect of extraction method on bacterial community composition ($F = 3.16$, $R^2 = 0.057$, $p = 0.014$), indicating that extraction protocol explained approximately 5.7% of the total community variation (Table. S3). These suggest that both methods recovered broadly similar community structures at higher biomass levels, while differences between extraction methods became more pronounced as biomass decreased.

16S rRNA gene sequencing revealed distinct taxonomic recovery patterns depending on extraction method and inoculum level (Fig. 5). Across both matrices, the dominant members of the mock consortium, including *Pseudomonas*, *Staphylococcus*, and *Bacillus*-related taxa (represented by *Bacillaceae* and *Bacillales*), were consistently recovered by both extraction methods. At higher inoculation levels (≥ 10^7 cells·mL^{-1}), taxonomic profiles were largely dominated by these expected taxa, with relatively similar community compositions observed between chloroform and kit extracts. At lower inoculation levels, differences between extraction methods became more apparent (Fig. 5). kit extracted samples generally retained community profiles dominated by the expected mock taxa across the dilution series, whereas chloroform extracted samples frequently showed increased relative abundances of additional taxa, including *Enterobacteriaceae*-, *Escherichia-Shigella*-, *Corynebacterium*-, and *Comamonadaceae*-related groups. These taxa were also detected in several chloroform extraction controls and low-biomass samples, contributing to increased variability in community composition. Differential abundance analysis further demonstrated enrichment of *Enterobacteriaceae*, *Brevundimonas*, *Sphingomonas*, and *Microbacteriaceae*-related taxa in chloroform extracts, whereas *Arhodomonas*, *Sphingobacterium*, *Tepidimonas*, and *Caldimonas* were significantly enriched in kit extracts (Fig. 5).

Discussion

Scanning electron microscope observations revealed distinct colonization patterns among the investigated materials. GP samples exhibited limited carbonate crust formation, allowing microbial cells and biofilm structures to remain visible throughout the incubation period. In contrast, calcium-rich cementitious matrices showed progressive mineral precipitation and EPS development, with microbial cells frequently embedded within surface-associated biofilms. These suggest that mineral encrustation and EPS formation may reduce DNA accessibility by physically shielding microbial cells and limiting cell lysis during extraction. In addition, extracellular DNA released from damaged cells may become strongly adsorbed onto calcium-rich mineral surfaces, reducing DNA recovery as previously reported for cementitious materials (Kiledal et al., 2021, 2023).

Based on these structural observations, we propose that microbial DNA within cementitious materials is unlikely to be uniformly distributed throughout the matrix but instead exists as heterogeneous fractions differing in physical accessibility and preservation. Rather than representing a single colonization event, microbial communities within cementitious materials are therefore likely to reflect multiple episodes of colonization associated with different stages of material ageing and environmental exposure. As cementitious materials progressively undergo carbonation, mineral alteration, and environmental weathering, newly formed ecological niches may support successive microbial colonization while simultaneously preserving DNA derived from earlier communities. Consequently, a single cementitious sample may contain intracellular DNA from recently active microorganisms together with extracellular and mineral-associated DNA originating from previous colonization events.

This conceptual model provides a plausible explanation for the discrepancy observed between total DNA recovery and molecular analyses. While metabolically active microbial cells are expected to provide readily amplifiable DNA, extracellular or mineral-associated DNA may contribute substantially to total DNA yield without accurately representing the contemporary microbial community. Consequently, extraction methods recovering larger quantities of total DNA may not necessarily provide a more representative characterization of the living microbiome. Instead, DNA extraction protocol should be evaluated according to its ability to recover reproducible and biologically informative DNA, emphasizing the importance of integrating DNA quantification with qPCR and sequencing-based analyses when investigating low-biomass cementitious materials.

The molecular analyses performed in the present study are consistent with this conceptual framework. If DNA recovered from cementitious materials represents multiple preservation states rather than a single biologically active fraction, extraction performance cannot be evaluated using total DNA yield alone. Instead, the analytical value of an extraction method depends on its ability to recover DNA that remains amplifiable and accurately reflects the microbial community of interest.

Consistent with this conceptual model, the present study demonstrates that DNA extraction performance in cementitious materials should be evaluated using multiple analytical endpoints rather than DNA yield alone. Although chloroform extraction frequently yielded higher DNA concentrations, particularly in geopolymer samples, these increases were not accompanied by proportional

improvements in bacterial 16S rRNA gene recovery or microbial community reconstruction. The negative $A_{260/280}$ values observed in the chloroform method, particularly at higher cell inputs, typically point to instrument baseline zeroing errors, low total yield relative to reagent noise, or severe phase-separation solvent contamination absorbing heavily at reference wavelengths. Likewise, spectrophotometric analysis did not consistently correlate with downstream molecular performance, indicating that conventional DNA quality metrics alone are insufficient to assess DNA extract quality and suitability for downstream molecular analyses in cementitious matrices.

This interpretation was further supported by linear regression analysis of the qPCR dilution series. In GP samples, kit extracts showed a closer correspondence between the expected dilution series and recovered bacterial signal across the entire 10^5–10^9 cells·mL^{-1} range, whereas comparable regression performance for chloroform extracts was only achieved after excluding the two lowest biomass dilutions. Together with the qPCR and sequencing results, these findings indicate that the additional DNA recovered by chloroform extraction contributed little to downstream molecular analyses despite increasing total DNA concentrations.

One possible explanation is that the two extraction methods recover different fractions of DNA from cementitious matrices. As proposed above, these materials may simultaneously contain intracellular DNA from recently active microorganisms together with extracellular and mineral-associated DNA preserved within the matrix. The extensive chemical lysis and organic solvent extraction employed by the chloroform protocol may therefore recover a greater proportion of extracellular or mineral-associated DNA, increasing the measured DNA concentration without proportionally increasing amplifiable bacterial DNA. In contrast, the commercial extraction kit appeared to recover a smaller but more analytically informative DNA fraction, resulting in stronger agreement among DNA quantification, qPCR, and amplicon sequencing. This interpretation is consistent with previous reports describing the co-extraction of inhibitory compounds and non-amplifiable DNA during solvent-based extraction from mineral-rich matrices (Kiledal et al., 2021; Maresca et al., 2016; Schrader et al., 2012).

The sequencing results were consistent with trends observed in DNA quantification and qPCR analyses. Although chloroform extraction produced substantially higher total DNA concentrations in GP samples, particularly at 10^9 cells·mL^{-1}, these higher DNA yields did not correspond to improved recovery of bacterial 16S rRNA genes. Relative quantification analysis demonstrated that kit extracts generally contained equal or higher quantities of amplifiable bacterial DNA than chloroform extracts at comparable inoculation levels. This difference was most evident in GP samples between 10^6 and 10^9 cells·mL^{-1}, where PM consistently recovered higher relative 16S rRNA gene quantities despite lower total DNA concentrations. In addition, kit extracts exhibited lower variability among replicates, whereas chloroform extracts showed greater variation, particularly at lower biomass levels.

Alpha diversity metrics showed considerable variation among samples (Fig. S2). Observed richness ranged from approximately 10 to 190 ASVs, while Shannon diversity values ranged from approximately 1.0 to 3.5. The highest diversity values were generally associated with environmental samples and

several extraction controls, whereas mock communities dominated by the expected taxa displayed lower richness and diversity. Notably, low-biomass chloroform extracted samples frequently exhibited elevated diversity relative to comparable kit extracted samples, consistent with the increased representation of additional taxa observed in the sequencing profiles.

In summary, microbial biomass exerted a greater influence on analytical reliability than extraction chemistry alone. Although extraction method significantly influenced community composition, it explained only a modest proportion of the total variation (R^2 = 5.7%). Instead, reproducible qPCR signals, predictable dilution-response relationships, and stable community profiles were consistently obtained only at inoculation levels of approximately ≥ 10^6–10^7 cells·mL^{-1}. Below this threshold, community composition became increasingly variable, particularly following chloroform extraction, consistent with previous observations that decreasing microbial biomass increases stochastic amplification and the relative contribution of background DNA and laboratory contamination (Salter et al., 2014).

Conclusion

This study demonstrates that DNA extraction strategy is a critical determinant of reliable microbiome characterization in cementitious materials and that extraction performance cannot be evaluated using DNA yield alone. By integrating DNA quantification, DNA purity assessment, qPCR, mock-community validation, amplicon sequencing, and SEM observations, we show that recovery of amplifiable and biologically informative DNA provides a more meaningful measure of extraction performance than spectrophotometric DNA concentration or purity alone. Our results further suggest that microbial DNA within cementitious materials exists in heterogeneous preservation states associated with mineral phases and biofilms, meaning that higher DNA recovery does not necessarily improve reconstruction of the contemporary microbial community. These findings provide a methodological framework for selecting and evaluating DNA extraction strategies and support more reliable molecular characterization of low-biomass cementitious materials, including emerging geopolymer concretes.

Declarations

Funding

Part of the work was funded by the Technology Agency of the Czech Republic (TAČR) through the ALMARA project (No. FW01010115). The authors acknowledge the assistance provided by the Research Infrastructure NanoEnviCz, supported by the Ministry of Education, Youth and Sports of the Czech Republic under Project No. LM2023066

Acknowledgements

Authors would like to acknowledge Petr Fabián and Lukáš Grič from CHEMCOMEX a.s. company for preparation of geopolymer matrix. We are grateful to Pavel Kejzlar for his assistance using a scanning electron microscope.

Data Accessibility and Benefit Sharing

The 16S rRNA gene amplicon sequencing data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1515430.

Declaration of conflicts

Authors declare no conflicts of interest.

Author Contribution

T.D.L. - Conceptualization, Formal analysis, Investigation, Methodology, Data curation, Visualization, Writing – original draft.K.L – Methodology, InvestigationP.V. – Methodology, Investigation, Funding Acquisition.A.S. – Writing – review & editing, Funding Acquisition.J.R. – Data curation, Formal analysis, Visualization.V.H. -Conceptualization, Formal analysis, Investigation, Methodology, Writing – review & editing.

Acknowledgement

Authors would like to acknowledge Petr Fabián and Lukáš Grič from CHEMCOMEX a.s. company for preparation of geopolymer matrix. We are grateful to Pavel Kejzlar for his assistance using a scanning electron microscope.

Data Availability

The 16S rRNA gene amplicon sequencing data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1515430.

References

1. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Caporaso G, J (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome 6(1):90. https://doi.org/10.1186/s40168-018-0470-z
2. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Caporaso JG (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37(8) Article 8. https://doi.org/10.1038/s41587-019-0209-9
3. Bone JR, Stafford R, Hall AE, Herbert RJH (2022) Biodeterioration and bioprotection of concrete assets in the coastal environment. Int Biodeterior Biodegrad 175:105507.

https://doi.org/10.1016/j.ibiod.2022.105507

4. de Bruijn FJ (2011) Handbook of Molecular Microbial Ecology I: Metagenomics and Complementary Approaches. Wiley
5. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55(4):611–622. https://doi.org/10.1373/clinchem.2008.112797
6. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High resolution sample inference from Illumina amplicon data. Nat Methods 13(7):581–583. https://doi.org/10.1038/nmeth.3869
7. Claesson MJ, Wang Q, O'sullivan O, Greene-Diniz R, Cole JR, Ross RP, O'toole PW (2010) Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. Nucleic Acids Res 38(22):e200–e200
8. Clifford RJ, Milillo M, Prestwood J, Quintero R, Zurawski DV, Kwak YI, Waterman PE, Lesho EP, Mc Gann P (2012) Detection of bacterial 16S rRNA and identification of four clinically important bacteria by real-time PCR. PLoS ONE 7(11). https://doi.org/10.1371/journal.pone.0048558
9. Cwalina B (2008) Biodeterioration of concrete. Archit Civil Eng Environ 1(4):133–140
10. Danilyan E, Ekaputri JJ, Zulaika E, Risdanareni P, Snoeck D (2025) Utilization of soil microbial community to enhance mechanical properties of geopolymer paste for sustainable construction materials. Case Stud Constr Mater 22:e04727. https://doi.org/10.1016/j.cscm.2025.e04727
11. Direito SOL, Marees A, Röling WFM (2012) Sensitive life detection strategies for low-biomass environments: Optimizing extraction of nucleic acids adsorbing to terrestrial and Mars analogue minerals. FEMS Microbiol Ecol 81(1):111–123. https://doi.org/10.1111/j.1574-6941.2012.01325.x
12. Dowd SE, Callaway TR, Wolcott RD, Sun Y, McKeehan T, Hagevoort RG, Edrington TS (2008) Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). BMC Microbiol 8:125. https://doi.org/10.1186/1471-2180-8-125
13. Drugă B, Ukrainczyk N, Weise K, Koenders E, Lackner S (2018) Interaction between wastewater microorganisms and geopolymer or cementitious materials: Biofilm characterization and deterioration characteristics of mortars. Int Biodeterior Biodegrad 134:58–67. https://doi.org/10.1016/j.ibiod.2018.08.005
14. Duxson P, Fernández-Jiménez A, Provis JL, Lukey GC, Palomo A, van Deventer JSJ (2007) Geopolymer technology: The current state of the art. J Mater Sci 42(9):2917–2933. https://doi.org/10.1007/s10853-006-0637-z
15. Freeman CL, Dieudonné L, Agbaje OBA, Žure M, Sanz JQ, Collins M, Sand KK (2023) Survival of environmental DNA in sediments: Mineralogic control on DNA taphonomy. Environ DNA 5(6):1691–1705. https://doi.org/10.1002/edn3.482

16. Gaylarde CC, Ortega-Morales BO (2023) Biodeterioration and Chemical Corrosion of Concrete in the Marine Environment: Too Complex for Prediction. Microorganisms 11(10):2438. https://doi.org/10.3390/microorganisms11102438
17. Gómez-Acata ES, Centeno CM, Falcón LI (2019) Methods for extracting ’omes from microbialites. J Microbiol Methods 160:1–10. https://doi.org/10.1016/j.mimet.2019.02.014
18. Grengg C, Ukrainczyk N, Koraimann G, Mueller B, Dietzel M, Mittermayr F (2020) Long-term in situ performance of geopolymer, calcium aluminate and Portland cement-based materials exposed to microbially induced acid corrosion. Cem Concr Res 131:106034. https://doi.org/10.1016/j.cemconres.2020.106034
19. Kiledal EA, Keffer JL, Maresca JA (2021) Bacterial Communities in Concrete Reflect Its Composite Nature and Change with Weathering. mSystems 6(3):e01153–e01120. https://doi.org/10.1128/mSystems.01153-20
20. Kiledal EA, Shaw M, Polson SW, Maresca JA (2023) Metagenomic Analysis of a Concrete Bridge Reveals a Microbial Community Dominated by Halophilic Bacteria and Archaea. Microbiol Spectr 0(0):e05112–e05122. https://doi.org/10.1128/spectrum.05112-22
21. Le Duc T, Vasicek R, Cerna K, Riha J, Nemecek J, Vecernik P, Cernousek T, Halodova P, Nemecek J, Sevcu A, Cernik M, Hlavackova V (2026) Chemo-mechanical changes of low-pH concrete under deep geological repository-like conditions: Coupled biotic and abiotic effects. Npj Mater Degrad 10(1):60. https://doi.org/10.1038/s41529-026-00773-0
22. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550
23. Maresca JA, Moser P, Schumacher T (2016) Analysis of bacterial communities in and on concrete. Mater Struct 50(1):25. https://doi.org/10.1617/s11527-016-0929-y
24. McMurdie PJ, Holmes S (2013) Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE 8(4):e61217. https://doi.org/10.1371/journal.pone.0061217
25. Mijnendonckx K, Bleyen N, Van Gompel A, Coninx I, Leys N (2022) pH and microbial community determine the denitrifying activity in the presence of nitrate-containing radioactive waste. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.968220
26. Oksanen J (2011) Vegan: Community Ecology Package. R package version 2.0–2. *Http://CRAN. R-Project. Org/Package= Vegan*
27. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, O’Hara RB, Solymos P, Stevens MHH, Szoecs E, Wagner H (2022) *vegan: Community Ecology Package*(Version 2.6-4) [Computer software]. https://cran.r-project.org/package=vegan
28. Pouhet R, Cyr M (2014) Alkali–silica reaction in metakaolin-based geopolymer mortar. Mater Structures/Materiaux et Constructions 48(3):571–583. https://doi.org/10.1617/s11527-014-0445-x
29. Provis JL, Palomo A, Shi C (2015) Advances in understanding alkali-activated materials. Cem Concr Res 78:110–125. https://doi.org/10.1016/j.cemconres.2015.04.013

30. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res 41(Database issue):D590–D596. https://doi.org/10.1093/nar/gks1219
31. Romanowski G, Lorenz MG, Wackernagel W (1991) Adsorption of plasmid DNA to mineral surfaces and protection against DNase I. Appl Environ Microbiol 57(4):1057–1061. https://doi.org/10.1128/aem.57.4.1057-1061.1991
32. Ruiz-Fresneda MA, Martinez-Moreno MF, Povedano-Priego C, Morales-Hidalgo M, Jroundi F, Merroun ML (2023) Impact of microbial processes on the safety of deep geological repositories for radioactive waste. *Frontiers in Microbiology*, *14*. https://doi.org/10.3389/fmicb.2023.1134078
33. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW (2014) Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 12(1):87. https://doi.org/10.1186/s12915-014-0087-z
34. Schrader C, Schielke A, Ellerbroek L, Johne R (2012) PCR inhibitors – occurrence, properties and removal. J Appl Microbiol 113(5):1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x
35. Shrestha R, Cerna K, Spanek R, Bartak D, Cernousek T, Sevcu A (2022) The effect of low-pH concrete on microbial community development in bentonite suspensions as a model for microbial activity prediction in future nuclear waste repository. Sci Total Environ 808:151861. https://doi.org/10.1016/j.scitotenv.2021.151861
36. Singh B, Ishwarya G, Gupta M, Bhattacharyya SK (2015) Geopolymer concrete: A review of some recent developments. Constr Build Mater 85:78–90. https://doi.org/10.1016/j.conbuildmat.2015.03.036
37. Turick CE, Berry CJ (2016) Review of concrete biodeterioration in relation to nuclear waste. J Environ Radioact 151:12–21. https://doi.org/10.1016/j.jenvrad.2015.09.005
38. Yeates C, Gillings M, Davison A, Altavilla N, Veal D (1998) Methods for microbial DNA extraction from soil for PCR amplification. Biol Procedures Online 1:40–47. https://doi.org/10.1251/bpo6
39. Statements & Declarations

Figures

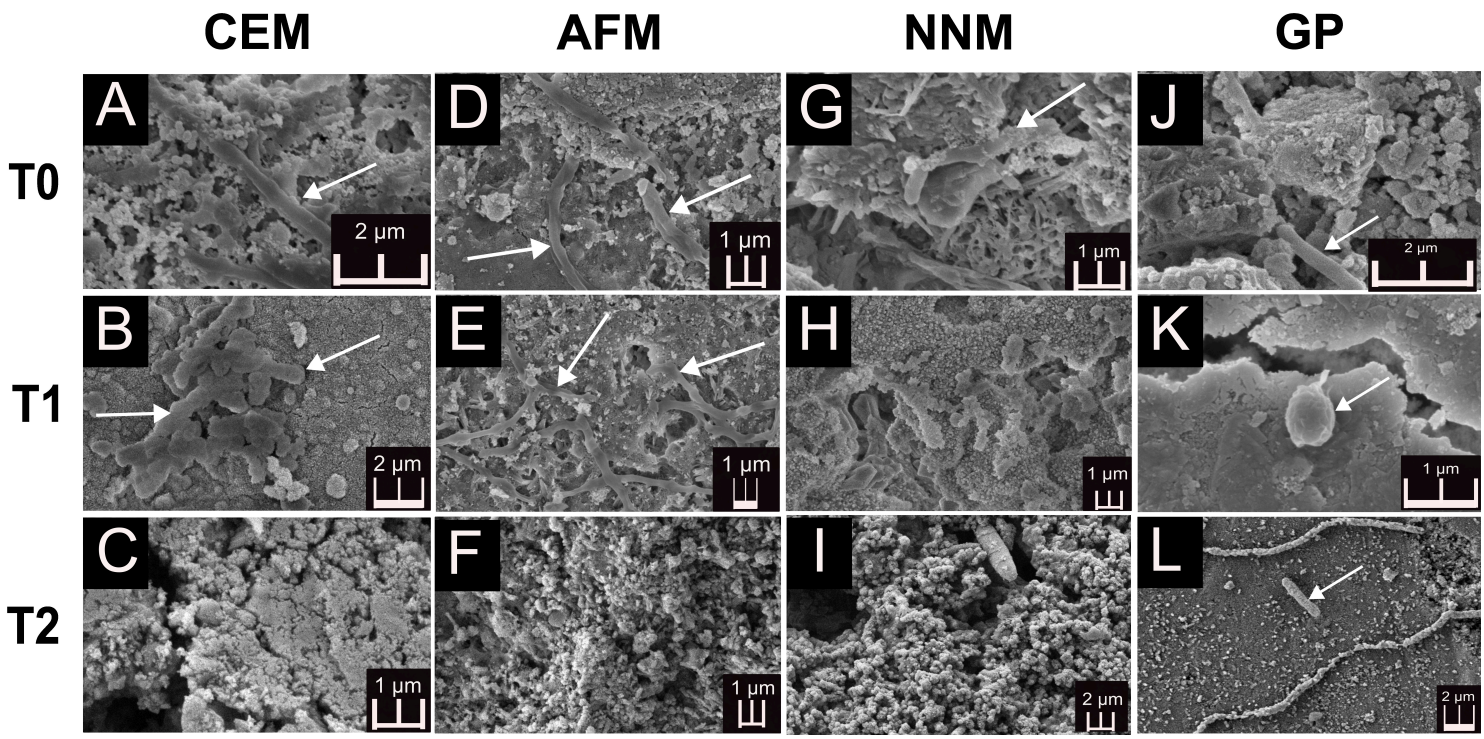

Figure 1

Legend not included with this version

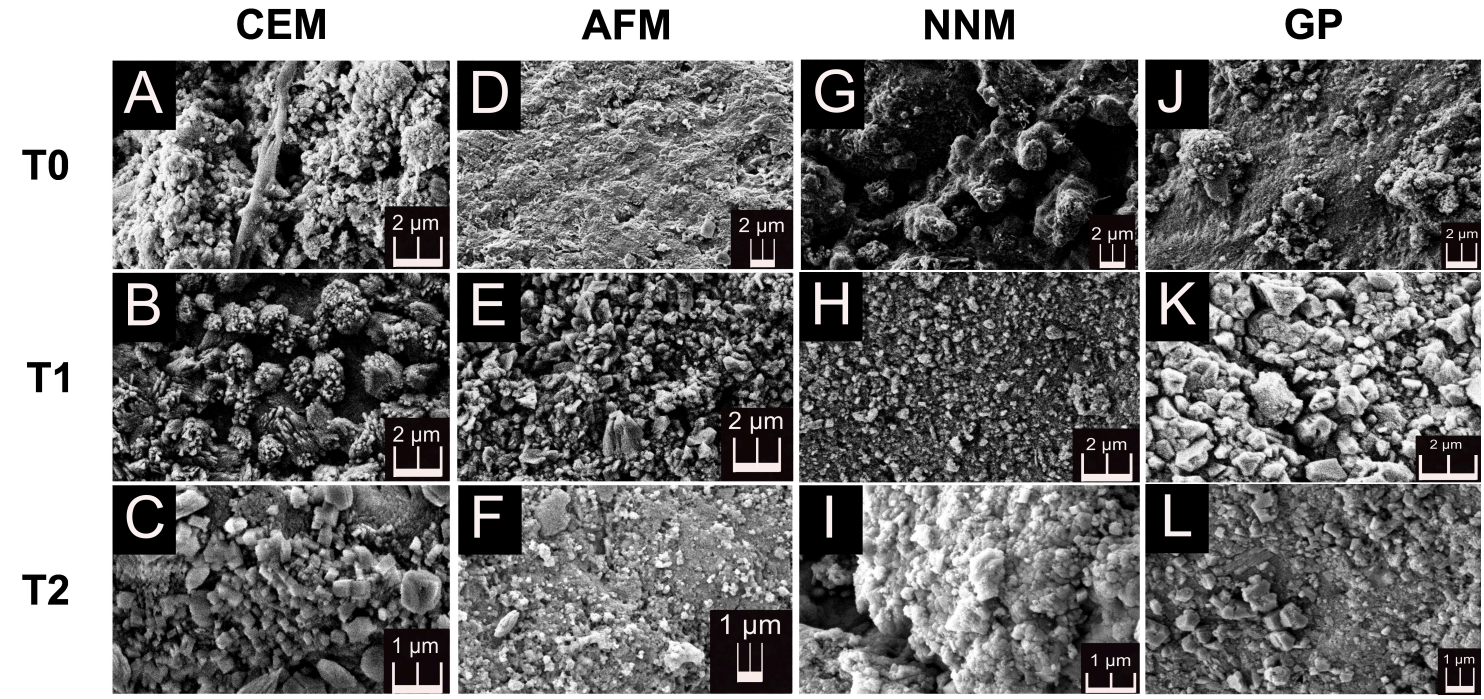

Figure 2

Legend not included with this version

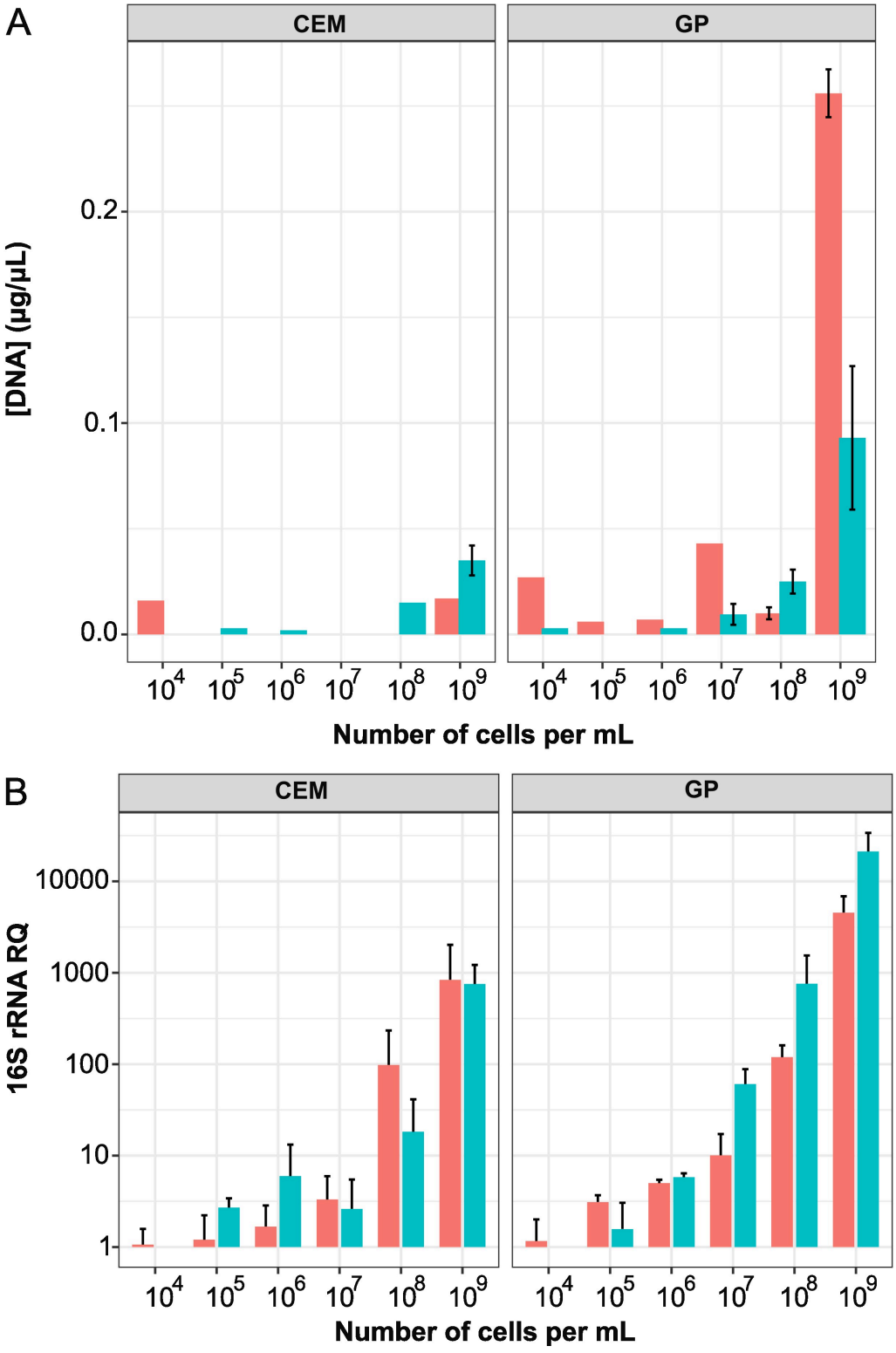

Figure 3

Legend not included with this version

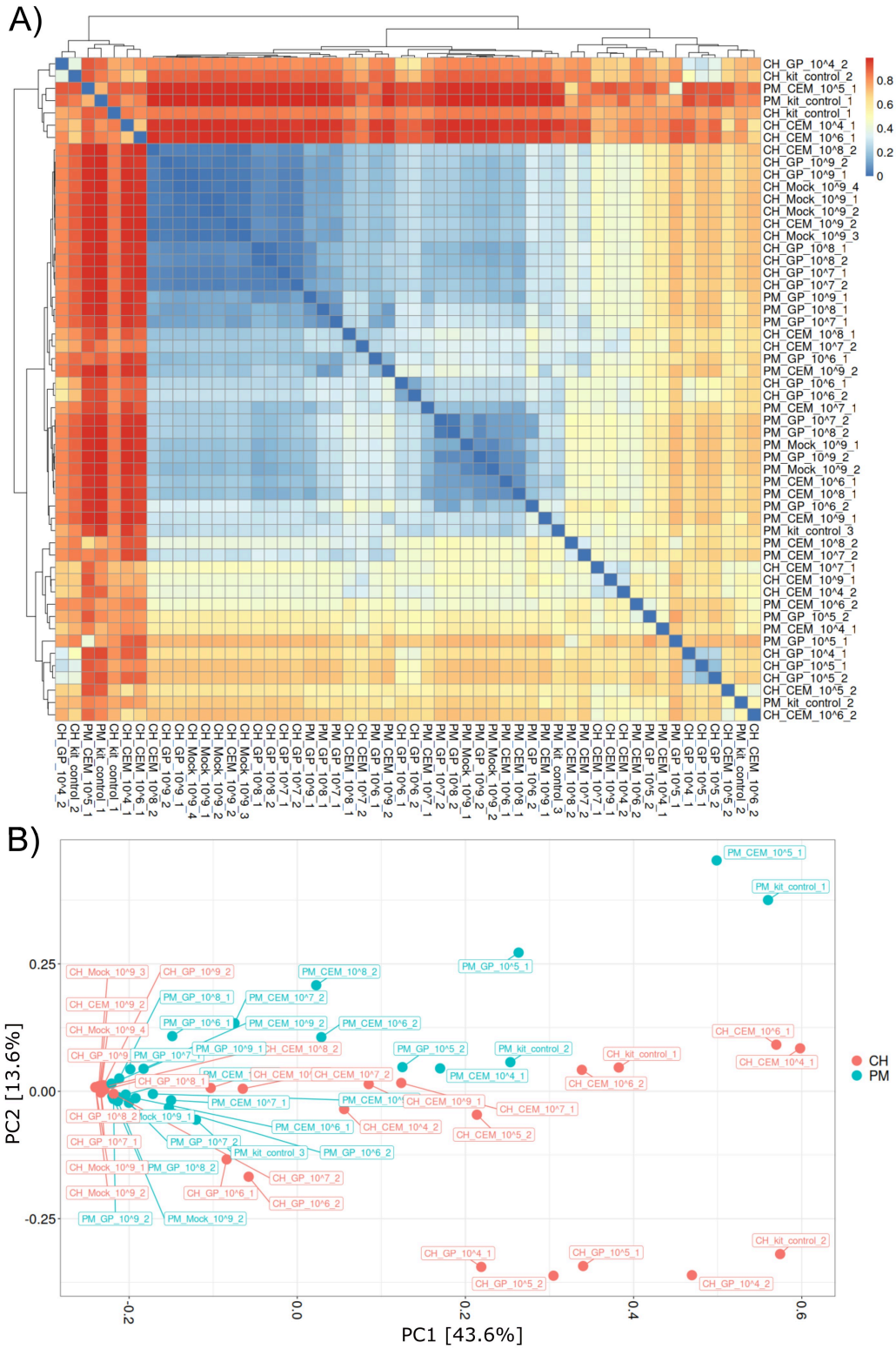

Figure 4

Legend not included with this version

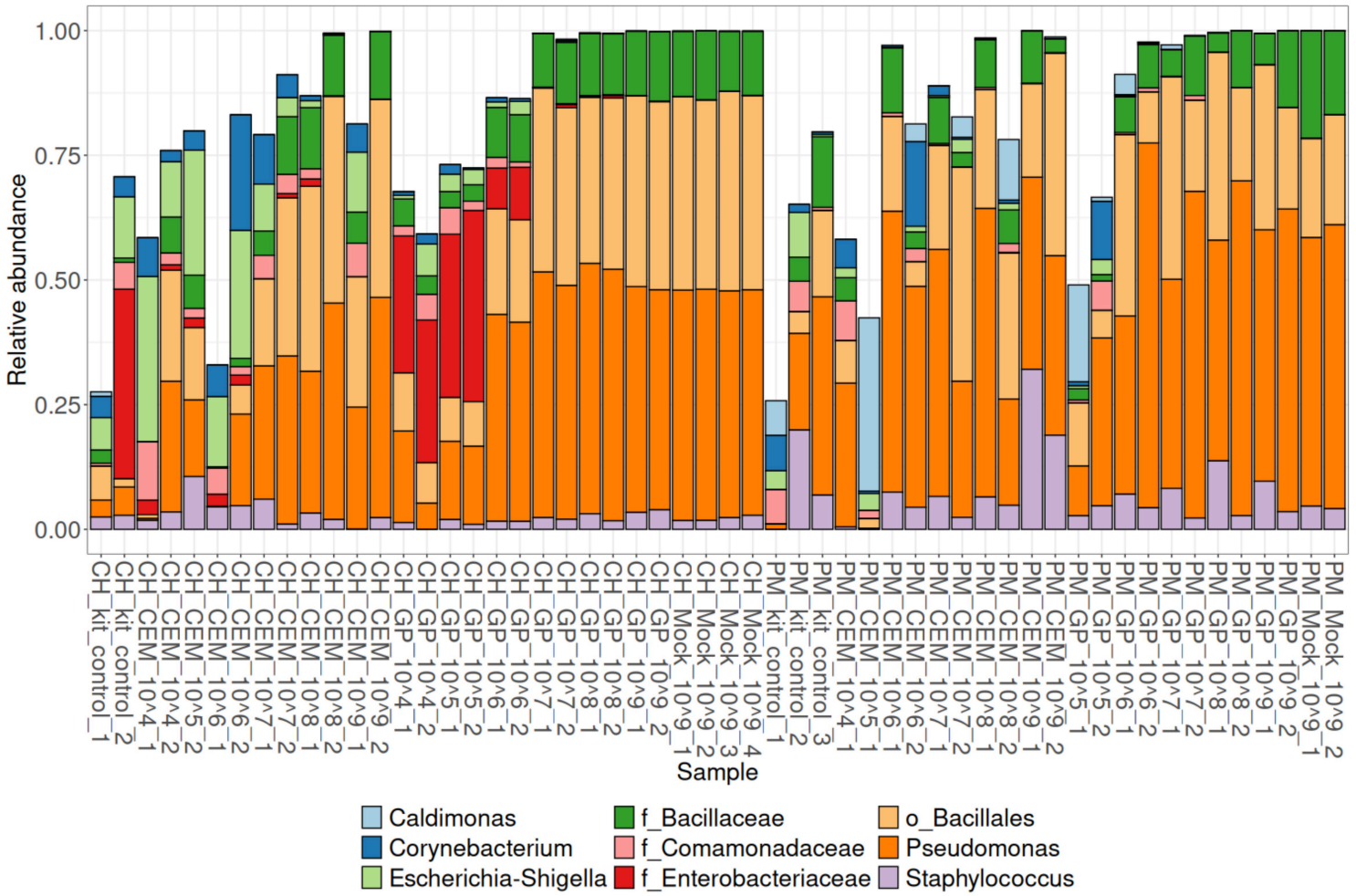

Figure 5

Legend not included with this version

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementinformation.docx