Supplementary Materials

Release and migration of Pb from Pb(II) and Pb(IV) compounds in the presence of microbiological activity

*Yaohuan Gao**[†‡], *Benjamin F. Trueman*[‡], *Bofu Li*[‡], *Martin R. Earle*[‡], *and Graham A. Gagnon*[‡]

[†]School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, 28 Xianning West Road, Xi'an, Shaanxi, 710049, P. R. China

[‡]Department of Civil and Resource Engineering, Dalhousie University, 1360 Barrington Street, Halifax, Nova Scotia Canada B3H 4R2

Description of the material:

- 17 pages
- 6 figures
- 5 tables
- Text on the reactor operation
- Text on ATP extraction and quantification
- Text on water quality of the feed tap water
- Predicting equilibrium lead solubility using *pbcusol*
- Text on DNA sequencing and analysis

Operation of the reactors

Four 200 ml glass flow-through reactors were custom-built and kept half-full throughout the experiment (Figure 1). The water level was controlled by an outlet line that passively drew water out of the reactors. At the beginning of the experiment (before inoculation), the reactors were operated in a flow-through mode for three days with 60 mg/L sodium bicarbonate in ultrapure water to stabilize the reactor and limit erratic Pb release. Then all the reactors were inoculated with detached biofilm from a filter medium (80 ml) collected from a full-scale biofilter at the J.D. Kline water treatment plant (Halifax, NS, CAN). The biofilm was detached by vortex-mixing the medium in 60 mg/L sodium bicarbonate solution at 3,000 rpm for two minutes, pelleted by centrifugation at 3,500 rpm for eight minutes, resuspended in 500 mL sodium bicarbonate solution (60 mg/L) and then pumped through the four reactors. The reactors were subsequently set in batch mode for 24 hours to allow cell attachment. Afterward, the tap water feed was initiated without filtration. The hydraulic retention time (HRT) was set to 60 minutes during the first few days and then switched to 30 minutes to minimize the loss of residual free chlorine.

The flow rate of the two control reactors was maintained to be the same as the experimental groups. The tap water fed to the control reactors was pumped from a 20 L autoclavable polypropylene carboy to the flow-through reactors. The tap water was filtered or autoclaved to decrease or eliminate biological activity. To minimize the changes in water chemistry, only the level of free chlorine was controlled using sodium hypochlorite (5.65-6%, #SS290-4, Fisher Scientific). More details are organized in the table below.

	Control reactors	
	Pb[II]	Pb[IV]
Feed	Sterile-filtered tap water (131±7 days) and autoclaved tap water (160±7 days)	
pH	Influent: 7.13±0.18; Effluent: did not measure	
Effluent chlorine (mg/L)	0.14-0.20 (median=0.16)	0.15-0.22 (median=0.18)
Influent orthophosphate (mg/L)	0.82-1.15 (median=0.94)	
Influent chloride (mg/L)	8.61-8.89 (median=8.70)	
Influent sulfate (mg/L)	8.24-8.82 (median=8.56)	

ATP extraction and quantification

ATP in effluent samples and ATP of Pb compound slurry samples were quantified using a Quench-Gone Aqueous (QGA) test kit and a Deposit & Surface Analysis (DSA) test kit (LuminUltra Tech. Ltd.), respectively, according to the instructions from the supplier. To collect the Pb compound slurry at the end of the experiment, the water inside each flow-through cell was slowly removed with a 60 mL syringe and a needle without disturbing the liquid or tilting the glass bottle. Once the liquid was removed, a sterilized spatula was used to sample the Pb compound slurry from three randomly selected spots in each bottle. When checked with prepared ATP solution (ca. 100 pg/mL) and $Pb(NO_3)_2$ solution with Pb(II) of 1000 µg/L—much higher than the detected values here—no adverse effect of Pb was

observed. The ATP from the Pb compound slurry samples was reported as pg ATP/g dry mass.

Quench-Gone Aqueous (QGA) test kit and Deposit & Surface Analysis (DSA) test kit were used for ATP quantifications. The detailed instructions of operation can be found on the LuminUltra website (https://www.luminultra.com/qga/ & https://www.luminultra.com/products/dsa/). For liquid samples, briefly, samples (60 mL or 100 mL) were filtered using a 60 mL syringe and a QGA filter. Then, the collected biomass on the filter was lysed by adding 1 mL of the UltraLyse 7 reagent and collected into the 9 mL UltraLute dilution tube. After extraction, 100 µL of diluted sample was mixed with 100 µL of Luminase enzyme (Lu-3mL-FD, LuminUltra) in a disposable plastic cuvette and measured in a Kikkoman Lumitester C-100 (Hach). The enzyme activity was calibrated using the LumiCheck reagent in the QGA test kit. Sample relative light units (RLU) were recorded along with a standard assay, and the value was recorded as ATP1. For solid samples, one gram of sample was mixed with 5ml UltraLyse 7 and the mixture was well mixed and incubated for five minutes. Then, one mL of the UltraLyse 7 reagent was collected and added to the 9 mL UltraLute dilution tube, caped and mixed. The rest of the procedure is similar to these for the QGA test kit. The six powdered samples after ATP measurement, three for each Pb oxidation state, were flushed into pre-labeled and weighed aluminum dishes (dried for 24 hours at 105 ºC and then weighed). Dry mass was quantified two days after 105 ºC treatment. The ATP readings from these six samples were reported as pg ATP/g dry mass. The biofilm mass was assumed to be negligible compared with the mass of Pb powders.

The cellular ATP concentration (QGA) and total ATP concentration (DSA) were calculated according to this equation:

$$ATP\ \left(pg\frac{ATP}{mL}\right) = \frac{RLU_{ATP}}{RLU_{ATP1}} \times \frac{Dilution\ Factor\ (pg\ ATP)}{V_{Sample}(mL)}$$

where the dilution factor is 10,000 for the QGA test kit and 50,000 for the DSA kit, and in the case of ATP quantification of powdered samples, sample dry mass (grams) is in the place of sample volume, and the unit of ATP is changed to pg ATP/g accordingly.

Water quality of the feed tap water

Major anions (sulfate, chloride, fluoride, and phosphate) in the tap water were monitored on an ion chromatograph (Dionex Aquion-AS-AP, Thermo Sci.) equipped with a Dionex IonPac AS22 column (4×250 mm). Anion concentrations were within the ranges in the annual reports from Halifax Water.[1, 2] TOC/DOC was measured according to Standard Method 5310[3] on a TOC-V CPH analyzer with a Shimadzu ASI-V autosampler and a nondispersive infrared detector (NDIR) having a method detection limit of 0.08 mg/L. Values of common water quality parameters from the annual reports of Halifax Water,[1, 2] including the measured values in this study, are tabulated in Table S1.

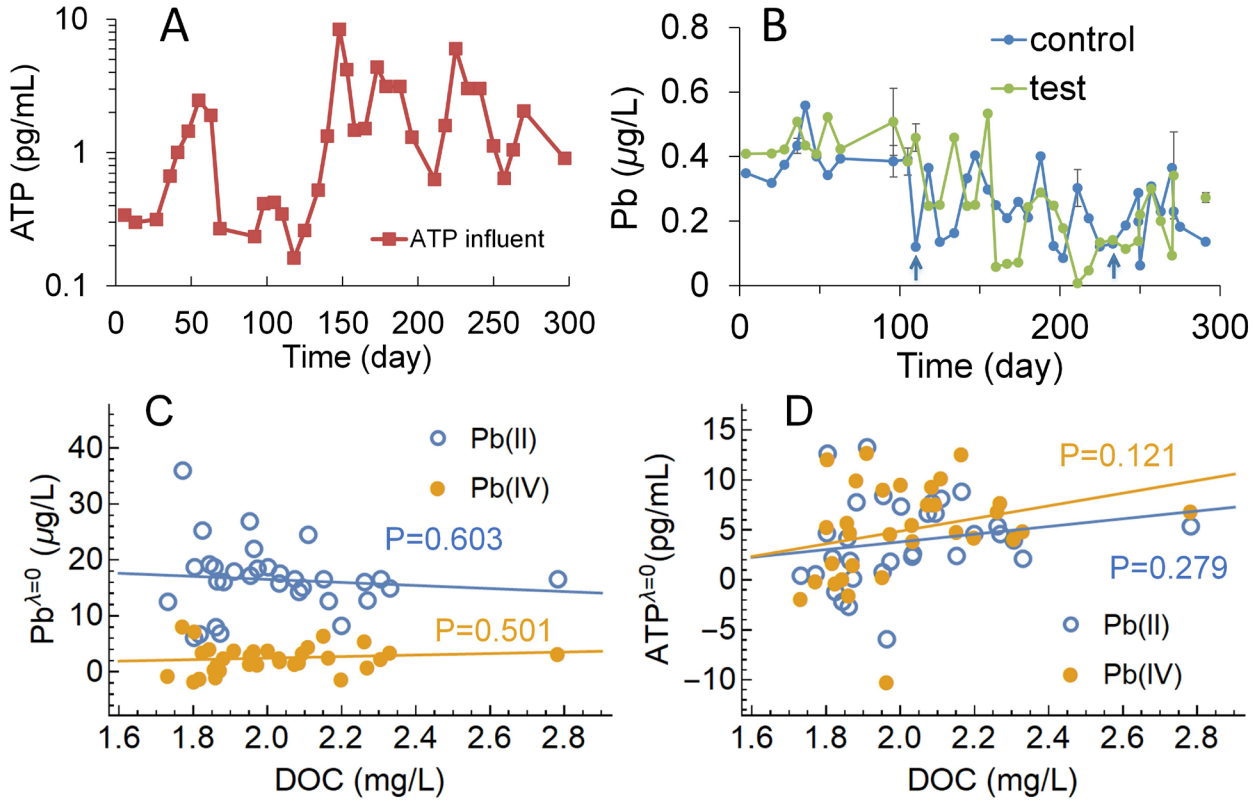


Figure S1. The ATP (A) and Pb levels (B) in the tap water during this study. Arrows in (B) mark the approximate beginning and ending of sterilization by filtration. (C) shows the correlation between DOC and the transformed average effluent Pb. (D) shows the correlation between DOC and the transformed average effluent ATP. Box-Cox power transformation was applied to treat the averaged effluent Pb values for normality and constant variance. Five power numbers (λ=-1, -0.5, 0, 0.5, and 1) were examined and they were compared for the power that yields the lowest variance. λ=0 ($Pb^{\lambda=0}$=geometric mean×ln[ave. Pb conc.]) was selected for effluent Pb.

Table S1. Water chemistry of the tested tap water including data from Halifax Water[1, 2] and the quantification conducted in this study (mg/L).

	2017-2018	2018-2019
PARAMETERS	Treated Water	Treated Water
Alkalinity (as $CaCO_3$)	21.5	21.0
Aluminum	0.108	0.105
Ammonia (N)	<0.050	<0.050
Arsenic	<0.001	<0.001
Calcium	3.8	4.2
Chloride	8.5	8.9 **(8.54-8.89)***
Chlorate	<0.1	< 0.1
Chlorite	<0.1	< 0.1
Colour (True Colour Units)	<5.0	<5.0
Conductivity (µS/cm)	84	97
Copper (Total)	< 0.002	<0.002
Fluoride	0.63	0.68 **(0.39-0.86)***
Hardness (as CaCO3)	11.3	12.0
HAA5 (avg.)	0.040	0.02
Iron (Total)	<0.050	<0.050
Langelier Index @ 4°C	-2.47	-2,43
Langelier Index @ 20°C	-2.22	-2.18
Lead (Total) (µg/L)	<0.50	<0.50
Magnesium	0.4	0.41
Manganese (Total)	0.014	0.011
Mercury (µg/L)	< 0.013	< 0.013
Nitrate & Nitrite (as N)	0.059	<0.050 **(0.03-0.08)***
Phosphate #	**(0.85-1.16)***	
pH (pH Units)	7.3	7.5
Potassium	0.31	0.310
Sodium	13.7	12.0
Solids (Total Dissolved)	46.5	63.5
Sulphate	7.5	8.0 **(8.23-8.79)***
Turbidity (NTU)	<0.08	0.11
Total Organic Carbon (TOC)	1.7	2.70
THM's (avg.)	0.06	0.034
Uranium (µg/L)	<0.10	<0.10
Zinc (Total)	0.093	0.092
PCB (µg/L)	<0.05	<0.05
Gross Alpha/ Gross Beta (Bq/L)	<0.10/<0.10	<0.10/ <0.10

***(XX-XX)** shows the measured values in our lab during the experiment, # only measured values in our lab are available.

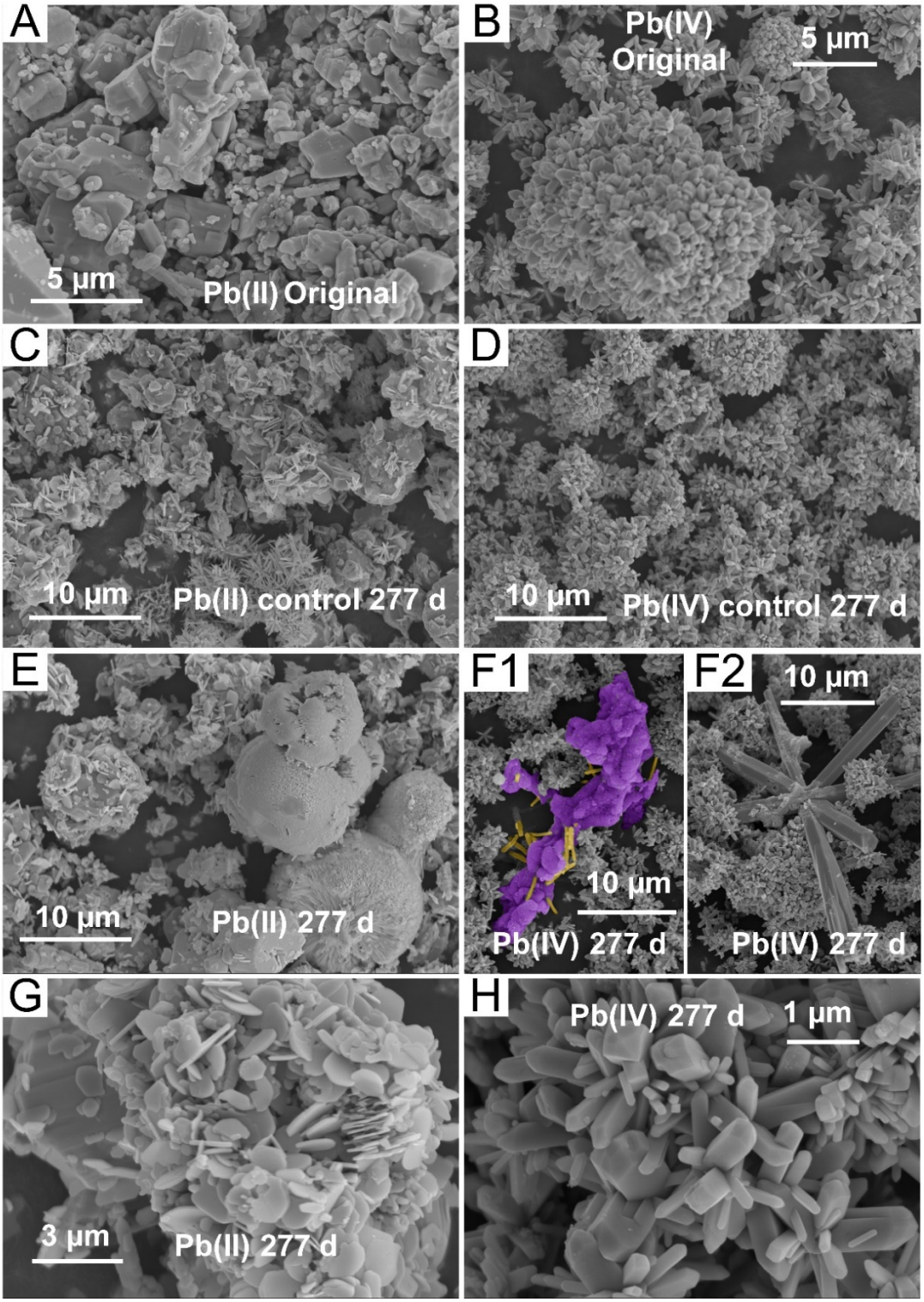


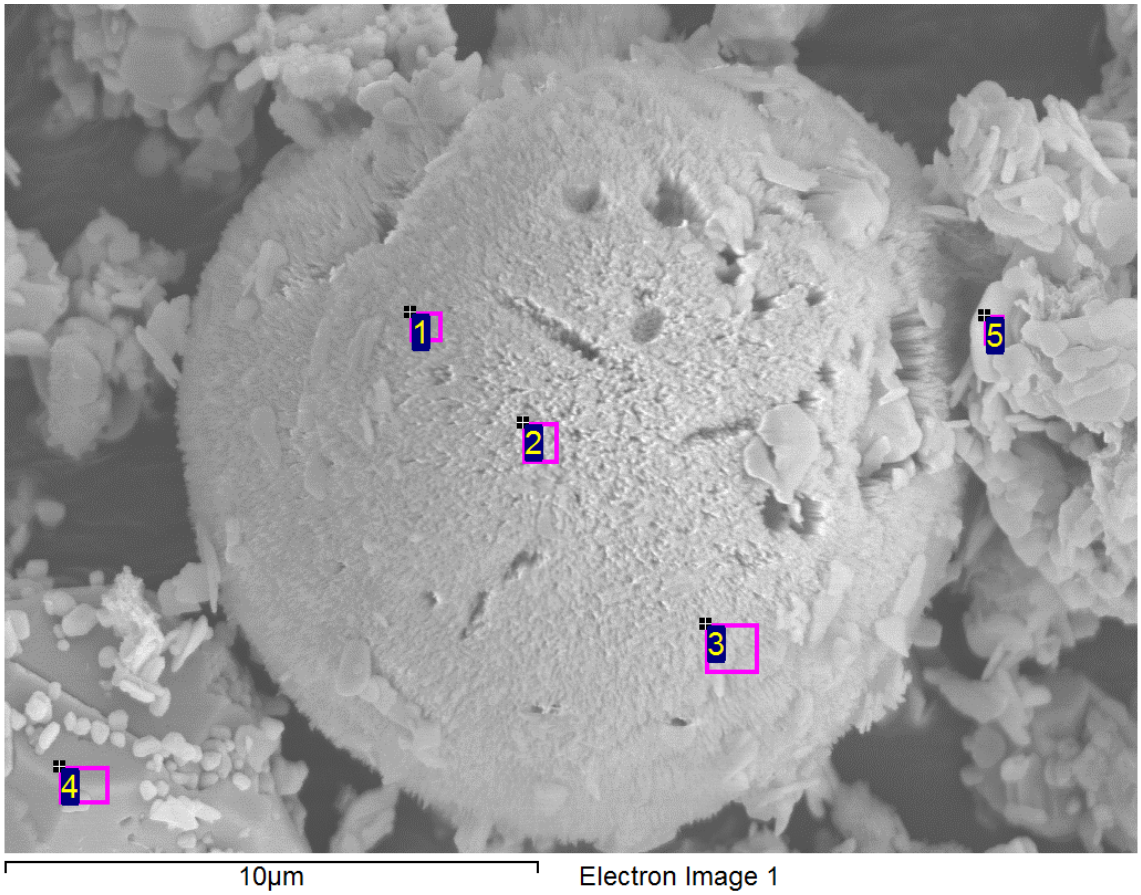
Figure S2. Scanning electron microscopy images of the cerussite and plattnerite powders in this study. Panels in the left column sequentially show the original cerussite (A), the control after 277 days (C), and

the experimental group after 277days (E). (G) shows a magnified spot of the sample in (E). Panels in the right column show the corresponding structures of plattnerite. The cell-like aggregate is highlighted in purple and yellow based on the morphologies.

The XRD pattern representing the original cerussite (Figure 6) contained signals from hydrocerussite, which reflects the dynamic equilibrium between the two compounds.[4] Hydrocerussite is also a common Pb corrosion product and is representative of distribution system conditions.[5] The observed prismatic cerussite crystals (Figure S2a) and hexagonal hydrocerussite platelets (Figure S2g) agree well with other SEM analyses of these Pb(II) compounds[6, 7] and they corroborate the XRD results. Based on the SEM results, chloropyromorphite or hydroxylpyromorphite also formed in the Pb(IV) system. When magnified, the tetragonal structure of β-PbO_2[8] (Figure S2h) and hexagonal rods of pyromorphite[9] are obvious (Figure S2f-2).

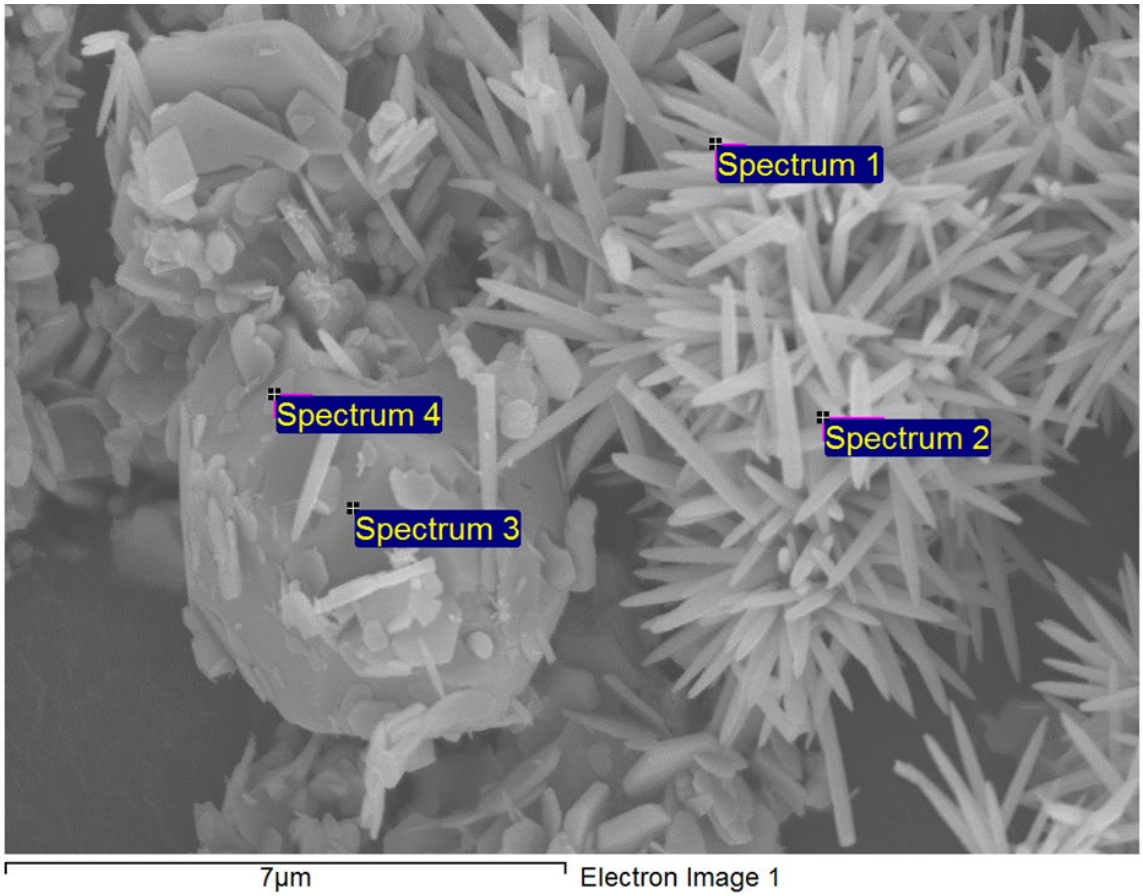
Table S2. The linear regression parameters of the variation of total Pb concentrations with stagnation time in both the experimental and the control groups (Figure 2b). The last two columns show the associated z-test parameters[10] on the slopes from inoculated and the control systems.

	slope	S.E.	intercept	S.E.	Adjusted R^2	z statistic	P value
Pb(II), experimental	0.509	0.0505	7.41	4.30	0.97	2.478	0.007
Pb(II), control	0.381	0.00993	2.768	0.69	0.99		
Pb(IV), experimental	0.126	0.0123	1.908	1.04	0.97	9.537	<0.00001
Pb(IV), control	0.00798	0.00175	0.508	0.122	0.77		



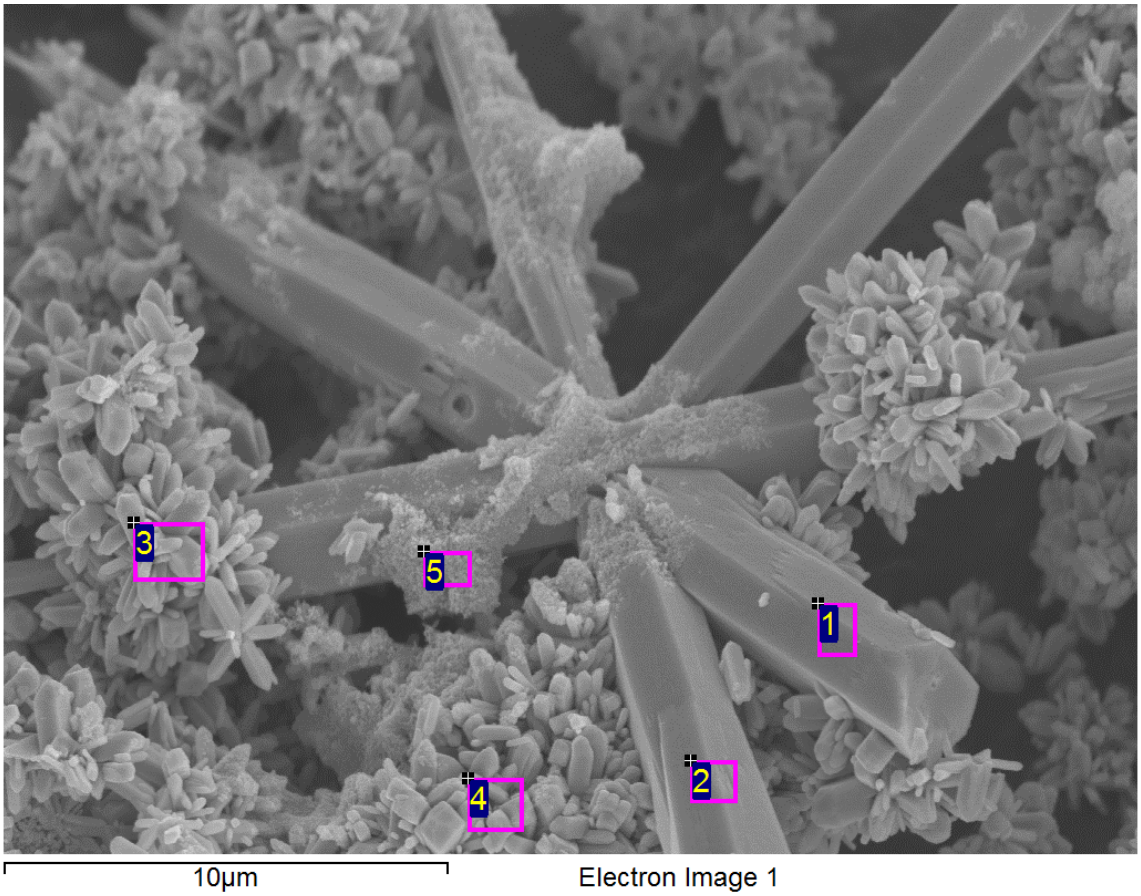
Spectrum	In stats.	C	O	Al	P	Cl	Pb	Total
1	Yes	19.15	22.46		3.79	2.27	52.33	100.00
2	Yes	15.93	14.25		5.13	2.35	62.34	100.00
3	Yes	16.03	10.32		3.97	1.72	67.97	100.00
4	Yes	19.61	26.00				54.39	100.00
5	Yes	17.04	21.57	0.73			60.66	100.00

Figure S3. The elemental composition of sphere-shaped or semisphere-shaped crystals observed from the biotic Pb(II) system and the corresponding w% of each major element. It is noteworthy that EDS analysis must have picked up signals from the carbon conductive tape beneath the powdered samples because the level of C detected exceeded both the maximum possible C content in cerussite/hydrocerussite (~4.5 w%) and the maximum level of C level for pyromorphite that incorporates carbonates (4.2 w%).[11]



Spectrum	In stats.	C	O	P	Cl	Pb	Total
Spectrum 1	Yes	20.49	26.62	3.79	2.08	47.01	100.00
Spectrum 2	Yes	14.41	19.57	4.42	2.58	59.03	100.00
Spectrum 3	Yes	17.30	16.48			66.22	100.00
Spectrum 4	Yes	20.88	25.49			53.63	100.00

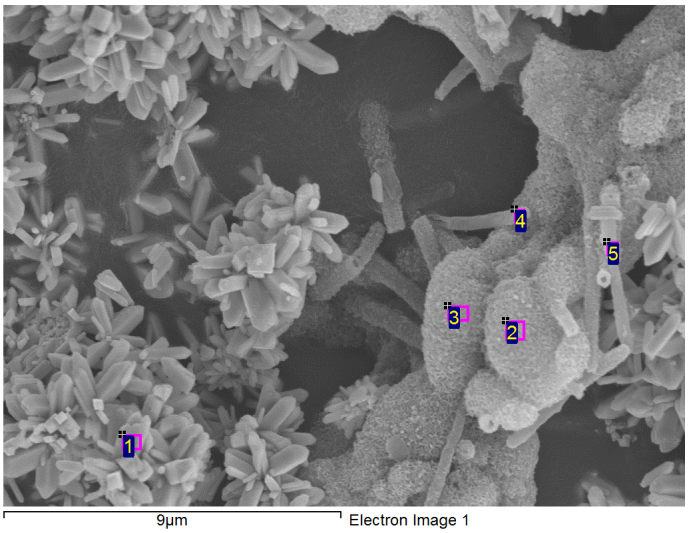
Figure S4. The elemental composition of needle-shaped crystals from the Pb(II) control system and the corresponding w% of each major element.



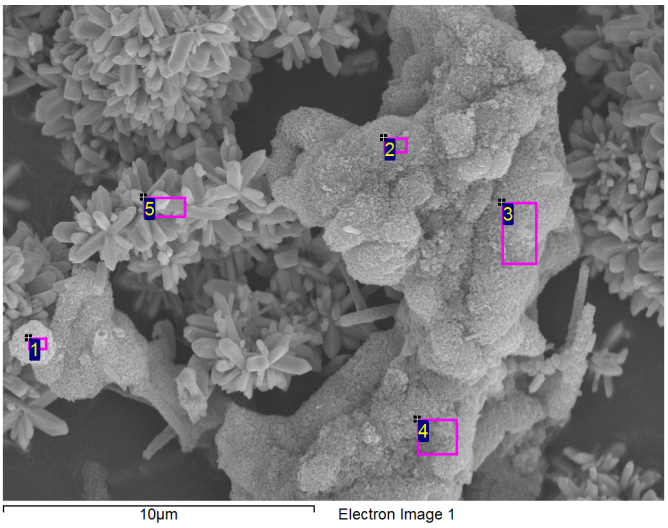
Spectrum	In stats.	C	O	Al	Si	P	Cl	Pb	Total
1	Yes	16.15	24.60			4.47	1.98	52.79	100.00
2	Yes	17.81	27.50			3.98	2.26	48.46	100.00
3	Yes	7.81	17.59					74.60	100.00
4	Yes	9.92	20.51					69.57	100.00
5	Yes	19.81	39.49	5.29	1.30	3.85		30.26	100.00

Figure S5. The elemental composition of hexagonal rods of pyromorphite in a sample collected from the inoculated Pb(IV) system and the corresponding w% of each major element.

A



B



Elemental composition of selected locations in panel A

Spectrum	In stats.	C	O	F	Al	P	Mn	Cu	Pb	Total
1	Yes	8.09	14.24						77.67	100.00
2	Yes	20.47	41.21	3.66	3.63		13.53	0.09	17.41	100.00
3	Yes	21.15	46.39		4.46		11.56	0.09	16.35	100.00
4	Yes	34.53	39.55	3.30	3.34		7.27	0.06	11.95	100.00
5	Yes	14.33	32.81		3.31	1.89	12.94		34.71	100.00

Elemental composition of selected locations in panel B

Spectrum	In stats.	C	O	F	Al	P	Mn	Fe	Cu	Pb	Total
1	Yes	19.79	36.72		3.03		10.24		0.21	30.00	100.00
2	Yes	18.67	49.47	3.8	5.17	2.43	7.94		0.05	12.48	100.00
3	Yes	16.30	26.78		4.71		20.4		0.10	31.71	100.00
4	Yes	17.73	19.93		4.23	3.93	16.11	14.47		23.60	100.00
5	Yes	9.89	18.31							71.80	100.00

Figure S6. The elemental composition of the particular spot with a cluster of microbial cells observed in a sampled originally collected from the inoculated Pb(IV) system. The tables show the corresponding w% of each major element. Under the set voltage, signals deep through the sample have been picked up as mentioned in the figure caption for Figure S6 and this explains why some spaces occupied by biological materials failed to show any phosphorus.

Predicting equilibrium lead solubility using *pbcusol*

pbcusol uses a modified version of the MINTEQ database available with PHREEQC that is consistent the thermodynamic data provided by Schock *et al.*[12], Table 4-14 (LEADSOL values). The log K value for chloropyromorphite is not listed in that source and was instead taken from Xie *et al.* [13]; as the authors point out, it reflects recent observations that chloropyromorphite is more soluble than previously thought. The relevant reactions involving lead are shown below. Activity coefficients were calculated as described in Parkhurst and Appelo [14] page 11, using the calculated ionic strength.

Sample code:

```
library("pbcusol") # remotes::install_github("bentrueman/pbcusol")
library("chemr") # remotes::install_github("paleolimbot/chemr")

pb_sol_fixed(
  ph = 7.3,
  dic = calculate_dic(7.3, 21.5),
  phosphate = .326,
  Cl = 8.5 / mass("Cl"),
  S = 7.5 / mass("SO4"),
  Ca = 3.8 / mass("Ca"),
  Mg = .4 / mass("Mg"),
  Na = 14 / mass("Na"),
  K = .3 / mass("K"),
  phase = "Hxypyromorphite"
)

pb_sol_fixed(
  ph = 7.3,
  dic = calculate_dic(7.3, 21.5),
  phosphate = .326,
  Cl = 8.5 / mass("Cl"),
  S = 7.5 / mass("SO4"),
  Ca = 3.8 / mass("Ca"),
  Mg = .4 / mass("Mg"),
  Na = 14 / mass("Na"),
  K = .3 / mass("K"),
  phase = "Chloropyromorphite",
```

```
  new_phase = list(
    "Chloropyromorphite",
    "Pb5(PO4)3Cl = 5Pb+2 + 3PO4-3 + Cl-",
    log_k = -80.4,
    delta_h = "0 kcal"
  )
)
```

Table S3. Summary of thermodynamic data used in equilibrium solubility modeling

Phase	Equation	log K
-	$Pb^{+2} + 2Cl^- = PbCl_2$	1.8
-	$Pb^{+2} + 3Cl^- = PbCl_3^-$	1.71
-	$Pb^{+2} + 4Cl^- = PbCl_4^{-2}$	1.43
-	$Pb^{+2} + Cl^- = PbCl^+$	1.59
-	$Pb^{+2} + CO_3^{-2} + H^+ = PbHCO_3^+$	12.59
-	$Pb^{+2} + 2CO_3^{-2} = Pb(CO_3)_2^{-2}$	10.33
-	$Pb^{+2} + CO_3^{-2} = PbCO_3$	7.10
-	$Pb^{+2} + 2H_2O = Pb(OH)_2 + 2H^+$	-16.91
-	$Pb^{+2} + 3H_2O = Pb(OH)_3^- + 3H^+$	-28.08
-	$Pb^{+2} + 4H_2O = Pb(OH)_4^{-2} + 4H^+$	-39.72
-	$3Pb^{+2} + 4H_2O = Pb_3(OH)_4^{+2} + 4H^+$	-23.86
-	$4Pb^{+2} + 4H_2O = Pb_4(OH)_4^{+4} + 4H^+$	-20.88
-	$6Pb^{+2} + 8H_2O = Pb_6(OH)_8^{+4} + 8H^+$	-43.62
-	$2Pb^{+2} + H_2O = Pb_2OH^{+3} + H^+$	-6.36
-	$Pb^{+2} + H_2O = PbOH^+ + H^+$	-7.22
-	$Pb^{+2} + PO_4^{-3} + H^+ = PbHPO_4$	15.41
-	$Pb^{+2} + PO_4^{-3} + 2H^+ = PbH_2PO_4^+$	21.05
-	$Pb^{+2} + 2SO_4^{-2} = Pb(SO_4)_2^{-2}$	3.50
-	$Pb^{+2} + SO_4^{-2} = PbSO_4$	2.73
-	$Pb^{+2} = Pb^{+2}$	0
Cerussite	$PbCO_3 = Pb^{+2} + CO_3^{-2}$	-13.11
Chloropyromorphite	$Pb_5(PO_4)_3Cl = 5Pb^{+2} + 3PO_4^{-3} + Cl^-$	-80.4
Hydroxylpyromorphite	$Pb_5(PO_4)_3OH + H^+ = 5Pb^{+2} + 3PO_4^{-3} + H_2O$	-62.83
Hydrocerussite	$Pb(OH)_2{:}2PbCO_3 + 2H^+ = 3Pb^{+2} + 2CO_3^{-2} + 2H_2O$	-18.00

We generated solubility predictions for Halifax tap water using the pH and major ion concentrations listed in Table S4.

Table S4. Equilibrium solubility model inputs (mg/L)

Parameter	Typical value
Ca	3.8
Cl	8.5
Mg	0.4

PO_4	1.0
SO_4	7.5
Na	14.0
K	0.3
pH	7.3
Alkalinity (mg $CaCO_3$/L)	21.5

The predictions are presented in Table S5 for both hydroxylpyromorphite and chloropyromorphite:

Table S5. Equilibrium solubility model outputs.

Solubility controlling phase	[Pb] mg/L
Hydroxylpyromorphite	0.0229
Chloropyromorphite	0.0011

DNA sequencing and analysis

Briefly, the amplification step consisted of a dual-indexing and one-step polymerase chain reaction (PCR) approach using B969F+BA1406R primers targeting the bacterial 16S rRNA V6-V8 region. PCR quality was checked using a high-throughput Invitrogen 96-well E-gel, followed by amplicon clean-up and normalization using a high-throughput Invitrogen SequalPrep 96-well plate kit. Samples were sequenced using an Illumina MiSeq instrument with paired-end 2×300 bp reads, resulting in a length of approximately 400-500 bp. The details of PCR procedures, primers and sequencing are described elsewhere.[15] An average sequence length of 250 bp was obtained from both Pb reactor samples and the biofilter sample (6-12” depth). Pb reactor samples retained significantly more sequences than the biofilter sample after processing, with >18,000 sequences in both Pb reactor samples and only 2,700 sequences in the biofilter sample.

Analysis of sequencing data was performed using QIIME2 version 2020.2 and Microbiome Helper following the workflow at https://github.com/LangilleLab/microbiome_helper/wiki/Amplicon-SOP-v2-(qiime2-2020.2).[15, 16] Read quality was evaluated using FastQC and primers were trimmed using Cutadapt.[17, 18] Sequences were denoised using DADA2 with forward and reverse reads truncated where the quality score dropped below 30.[19] Taxonomy was assigned by referencing the Silva database v. 132.[20] Base QIIME2 scripts were used to remove rare amplicon sequence variants (ASVs) and to calculate alpha rarefaction and alpha diversity estimates. The *phyloseq* package was used to analyze and visualize sequences in the R programming environment.[21, 22] Sequences below 2% relative abundance were removed from the plot (Figures 3 and S8) for clarity. Sequences are also clustered into Operational Taxonomic Units (OTUs) based on a similarity threshold of 97%. Venn diagram was plotted to show the similarities and differences among the microbial communities on Pb compound surfaces and those from the filter media. Heatmap analysis was used to further illustrate the similarities and differences among the bacterial communities following an average linkage clustering method. The top 30 most abundant OTUs were

shown.

Reference

1. J. Campbell, *Halifax Water 2017/18 Annual Report*, Halifax Canada, 2018.
2. Halifax Water, *Halifax Water 2018/19 Annual Report*, Halifax, Nova Scotia, Canada, 2019.
3. APHA, in *Standard Methods For the Examination of Water and Wastewater*, 2017, DOI: 10.2105/smww.2882.104.
4. P. Taylor and V. J. Lopata, Stability and solubility relationships between some solids in the system PbO-CO2-H2O, *Canadian Journal of Chemistry*, 1984, **62**, 395-402.
5. J. D. Noel, Y. Wang and D. E. Giammar, Effect of water chemistry on the dissolution rate of the lead corrosion product hydrocerussite, *Water Research*, 2014, **54**, 237-246.
6. Y. Li, Y. Zhu, S. Zhao and X. Liu, The weathering and transformation process of lead in Chinas shooting ranges, *Environmental Science: Processes & Impacts*, 2015, **17**, 1620-1633.
7. H. Liu, G. V. Korshin and J. F. Ferguson, Investigation of the Kinetics and Mechanisms of the Oxidation of Cerussite and Hydrocerussite by Chlorine, *Environmental Science & Technology*, 2008, **42**, 3241-3247.
8. H. Harada, Y. Sasa and M. Uda, Crystal data for [beta]-PbO2, *Journal of Applied Crystallography*, 1981, **14**, 141-142.
9. D.-Q. Ng, T. J. Strathmann and Y.-P. Lin, Role of Orthophosphate As a Corrosion Inhibitor in Chloraminated Solutions Containing Tetravalent Lead Corrosion Product PbO2, *Environmental Science & Technology*, 2012, **46**, 11062-11069.
10. R. Paternoster, R. Brame, P. Mazerolle and A. Piquero, Using The Correct Statistical Test For The Equality Of Regression Coefficients *Criminology*, 1998, **36**, 859-866.
11. M. P. Sternlieb, J. D. Pasteris, B. R. Williams, K. A. Krol and C. H. Yoder, The structure and solubility of carbonated hydroxyl and chloro lead apatites, *Polyhedron*, 2010, **29**, 2364-2372.
12. M. R. Schock, I. Wagner and R. Oliphant, The corrosion and solubility of lead in drinking water, *Internal corrosion of water distribution systems*, 1996, **4**, 131-230.
13. L. Xie and D. E. Giammar, Equilibrium Solubility and Dissolution Rate of the Lead Phosphate Chloropyromorphite, *Environmental Science & Technology*, 2007, **41**, 8050-8055.
14. D. L. Parkhurst and C. A. J. Appelo, *Users guide to PHREEQC (Version 2): A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations*, Report 99-4259, 1999.
15. A. M. Comeau, G. M. Douglas and M. G. I. Langille, Microbiome Helper: a Custom and Streamlined Workflow for Microbiome Research, *mSystems*, 2017, **2**, e00127-00116.
16. E. Bolyen, J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. K. Cope, R. Da Silva, C. Diener, P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J. M. Gauglitz, S. M. Gibbons, D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. A. Huttley, S. Janssen, A. K. Jarmusch, L. Jiang, B. D. Kaehler, K. B. Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T. Kosciolek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y.-X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. T. Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian, T. Pearson, S. L. Peoples, D. Petras, M. L. Preuss, E. Pruesse, L. B.

Rasmussen, A. Rivers, M. S. Robeson, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R. Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight and J. G. Caporaso, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2, *Nature Biotechnology*, 2019, **37**, 852-857.

17. B. Institute, Babraham Biofinformatics-FastQC, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).
18. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, *2011*, 2011, **17**, 3.
19. B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson and S. P. Holmes, DADA2: High-resolution sample inference from Illumina amplicon data, *Nature methods*, 2016, **13**, 581-583.
20. C. Quast, E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies and F. O. Glöckner, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools, *Nucleic Acids Research*, 2012, **41**, D590-D596.
21. P. J. McMurdie and S. Holmes, phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data, *PLOS ONE*, 2013, **8**, e61217.
22. H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu and H. Yutani, Welcome to the Tidyverse, *Journal of Open Source Software*, 2019, **4**, 1686.