Supplementary Information for

Nitrifying microorganisms linked to biotransformation of perfluoroalkyl sulfonamido precursors from legacy aqueous film forming foams

Bridger J. Ruyle[1,Ω,*], Lara Schultes[1,Ω], Denise M. Akob[2], Cassandra R. Harris[2], Michelle M. Lorah[3], Simon Vojta[4], Jitka Becanova[4], Shelley McCann[5], Heidi M. Pickard[1], Ann Pearson[1,5], Rainer Lohmann[4], Chad D. Vecitis[1], Elsie M. Sunderland[1,5,6]

[1]Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Boston, Massachusetts, 02134, USA
[2]U.S. Geological Survey, Geology, Energy and Minerals Science Center, Reston, Virginia, 20192, USA
[3]U.S. Geological Survey, Maryland-Delaware-DC Water Science Center, Baltimore, Maryland, 21228, USA
[4]Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island, 02882, USA
[5]Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, 02138, USA
[6]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, 02115, USA

[Ω]Equal contribution
[*]Corresponding author: bruyle@g.harvard.edu

This PDF file includes:

Tables S2, S4, S5 and Supplementary Methods Tables 1 and 2 are provided in full in accompanying Excel spreadsheet

Supplementary Methods

Biotransformation experiments

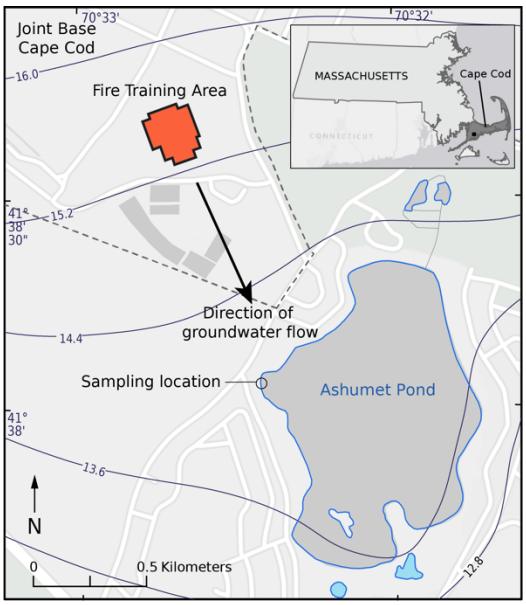

Figure S1. Ashumet Pond field map showing origin of water/sediment slurry. The elevation of the groundwater table in meters is denoted by the contour lines.

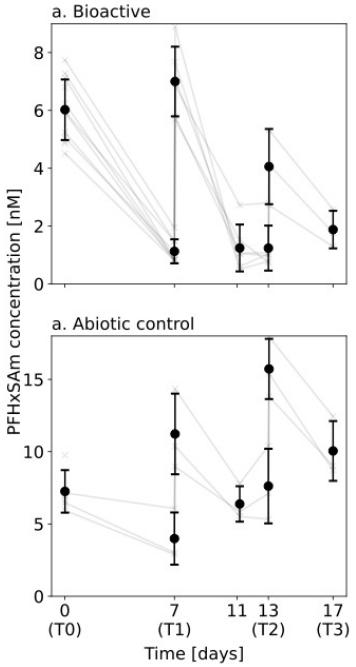

Figure S2. PFHxSAm concentrations in microbial community analysis experiments. Light gray lines represent data from individual bottles. Black dots represent the mean and error bars represent standard deviation of replicate bottles (n=12 at T0 and decreased by increments of 3 at each spiking timepoint as bottles were sacrificed for microbial community analysis).

Targeted PFAS analysis

Aqueous phase samples were analyzed for targeted per- and polyfluoroalkyl substances (PFAS) using an Agilent (Santa Clara, California, United States) 6460 triple quadrupole liquid chromatograph-tandem mass spectrometer (LC-MS/MS) equipped with an Agilent 1290 Infinity Flex Cube online SPE, following a previously published method with slight modifications.[1] Prior to instrumental injection, 750 μL aqueous sample were combined with 710 μL LCMS-grade methanol (Honeywell, Charlotte, North Carolina, United States) and 40 μL of internal standard (IS) mixture in methanol (800 ng L^{-1}; Wellington Laboratories, Guelph, Ontario, Canada; Table S1). For each sample, 300 μL was loaded onto an Agilent Zorbax SB-Aq (4.6 x 12.5 mm; 5 μm) online SPE cartridge with 0.85 mL of 0.1% aqueous formic acid at a flow rate of 1 mL min^{-1}. Analytes were eluted from the SPE cartridge and loaded onto an Agilent Poroshell 120 EC-C18 (3.9 x 50 mm; 2.7 μm) reversed-phase high pressure liquid chromatography (HPLC) column using ammonium acetate (2 mM) in methanol and ammonium acetate (2 mM) in Milli-Q water at a flow rate of 0.5 mL min^{-1} and a column temperature of 50°C. Analytes were ionized with an electrospray ionization (ESI) source switching between negative and positive ion mode and introduced to the tandem mass spectrometer at a temperature of 300°C, gas flow rate of 13 L min^{-1}, and nebulizer pressure of 45 psi.

Targeted PFAS were quantified using isotopic dilution with 9 to 11-point calibration curves between 1-10,000 ng L^{-1}. Calibration curves were prepared with 50:50 Milli-Q water:methanol to match the sample and had $R^2 > 0.99$. Calibration quality controls ranging between 100 and 10,000 ng L^{-1} were analyzed every 12 sample injections and were within ±30% of the expected value. Branched and linear isomers of perfluorohexane sulfonate (PFHxS) were quantified with individual native isomer calibration curves and the isotopically labeled PFHxS

standard. Perfluorohexane sulfonamide (FHxSA), perfluorohexane sulfonamido propyl tertiary amine (PFHxSAm), and perfluorohexane sulfonamido propyl quaternary amine (PFHxSAmS) do not have commercially available isotopically labeled standards and were therefore quantified using either the FOSA internal standard (FHxSA) or the PFHxS internal standard (PFHxSAm and PFHxSAmS) (Table S1).

Limits of detection (LODs) were calculated based on the average concentration at which the sample signal-to-noise ratio was three. Method detection limits (MDLs) were 2x the LOD accounting for the 1:1 dilution of the sample in methanol prior to instrumental injection (Table S2). MDLs were 20 ng L^{-1} for PFHxS (0.05 nM), 24 ng L^{-1} for PFHxSAm (0.05 nM), 28 ng L^{-1} for FHxSA (0.07 nM), and 60 ng L^{-1} for PFHxSAmS (0.12 nM). Two Milli-Q water procedural blanks were run every six sample injections and were always below MDL. Only data above the MDL were used for subsequent modeling and data analyses.

Table S1. LC-MS/MS method.

Analyte	Type	Internal Standard	Precursor Ion	Quantifier Ion	Quantifier Collision Energy (V)	Qualifier Ion	Qualifier Collision Energy (V)	Fragmentor Voltage (V)
PFHxS	Target	[$^{13}C_3$] PFHxS	398.9	80.0	58	98.9	34	135
FHxSA	Target	[$^{13}C_8$] FOSA	398.0	78.0	40			180
PFHxSAm	Target	[$^{13}C_3$] PFHxS	485.3	85.0	30	58.0	35	175
PFHxSAmS	Target	[$^{13}C_3$] PFHxS	499.3	73.0	35	60.0	35	170
[$^{13}C_3$] PFHxS	ISTD*		401.9	98.9	38			180
[$^{13}C_8$] FOSA	ISTD*		505.9	78.0	38			95

ISTD = internal standard

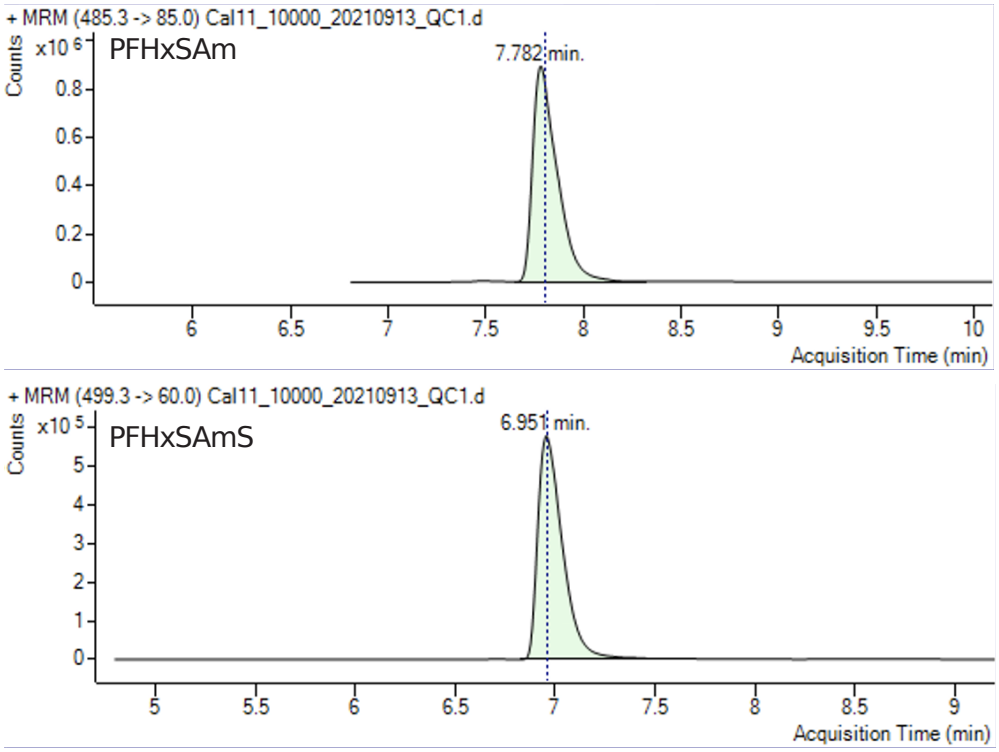

Figure S3. Typical chromatograms for PFHxSAm and PFHxSAmS

Table S2. Measured PFAS concentrations.

See accompanying Excel spreadsheet

High resolution mass spectrometry PFAS analysis

The suspect screening analysis was performed using a SCIEX ExionLC AC UHPLC system coupled to a SCIEX X500R quadrupole time-of-flight tandem mass spectrometer (QTOF MS/MS). A Phenomenex Gemini 3um C18 110Å 50x2mm LC analytical column preceded with a Phenomenex SecurityGuard cartridge was used for the analyte separation. Another Phenomenex Gemini 5um C18 110 Å 50x4.6mm LC analytical column was used to delay the PFAS instrumental contribution. The aqueous mobile phase (MPA) was 10mM ammonium

acetate in water and the organic mobile phase (MPB) was 10mM ammonium acetate in methanol. LC parameters were set as follows: flow 0.3mL/min, injection 20uL, column oven 45°C. Solvent gradient of MPB gradually increased from 40% to 80% (1 to 5.5 min), 80% to 100% (5.5 to 7 min), held for one minute and finally dropped to 40% (8 to 8.5 min) and held for another 6.5 minutes.

The MS data for the suspect screening analysis were collected using both IDA and SWATH acquisitions. ESI operated in negative and positive mode with the following parameters was used: curtain Gas at 30 psi, ion source gas 1 at 40 psi, ion source gas 2 at 60 psi, temperature 450°C. The compounds were tentatively identified using the suspect screening approach, based on precursor mass, isotope pattern, retention time, exact mass accuracy (<5ppm), and MS/MS fragmentation matching (SCIEX Fluorochemical HR-MS/MS Spectral Library 2.0) (Table S3, Figures S3 and S4). Following the conventions from Schymanski et. al.,[2] both compounds were assigned the confidence level of identification 2a (MS/MS spectral library match).

Table S3. LC-MS/MS and qTOF diagnostics for sulfonamido precursors and their metabolites

Analyte	Exact mass [Da]	Identification confidence*	Parent chemical m/z (ESI mode)	Diagnostic fragments m/z
PFHxSAmS	499.0719	1	499 (positive)	73 60
PFHxSAm	484.0490	1	485 (positive)	85 58
Secondary amine	470.0334	Not detected		
Primary amine	456.0177	Not detected		
Aldehyde	454.9861	Not detected		
Perfluorohexane sulfonamido propanoic acid (FHxSA-PrA)	470.9810	2a	469.9736 (negative)	397.9543 168.9864 77.9656
Acetic acid (FHxSAA)	456.9653	Not detected		
FHxSA	398.9599	1	398 (negative)	
PFHxSi	383.9490	2a	382.9413 (negative)	318.9814 169.9899 118.9941 82.9612
PFHxS	399.9439	1	399 (negative)	99 80

*Identification confidence according to the Schymanski scale.[2] Level 1 = confirmed structure by reference standard. Level 2a = Probable structure by library spectrum match. The diagnostic fragments for FHxSA-PrA nad PFHxSi are shown in Figures S4 and S5.

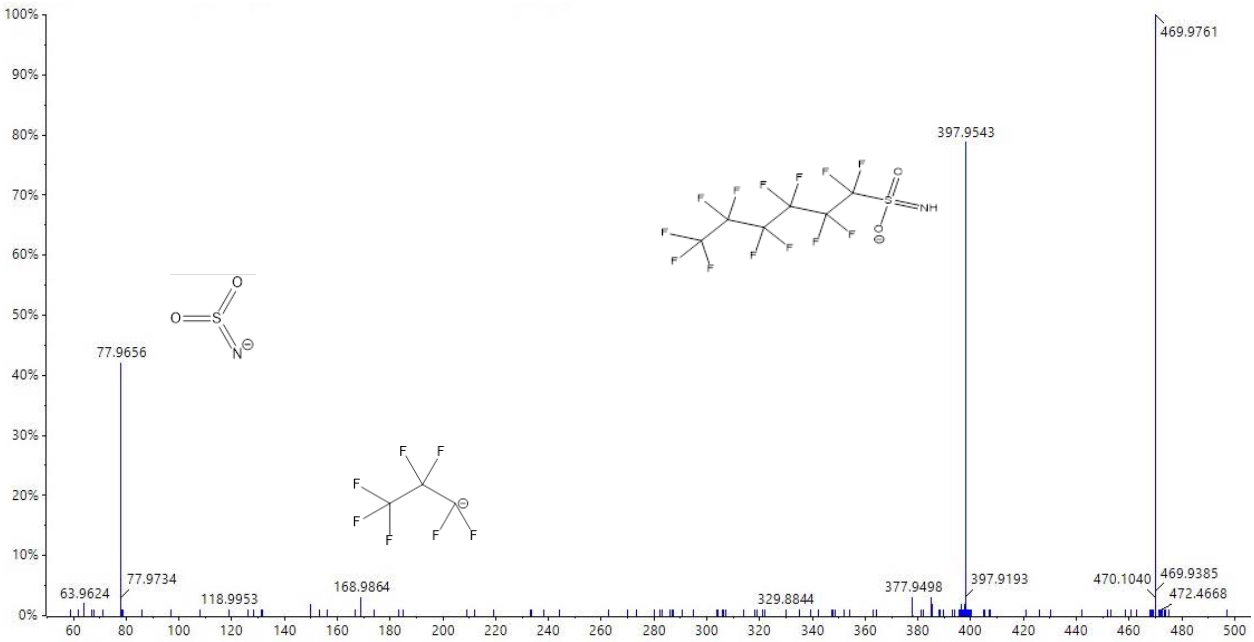

Figure S4. MS/MS spectra for perfluorohexane sulfonamido propanoic acid (FHxSA-PrA)

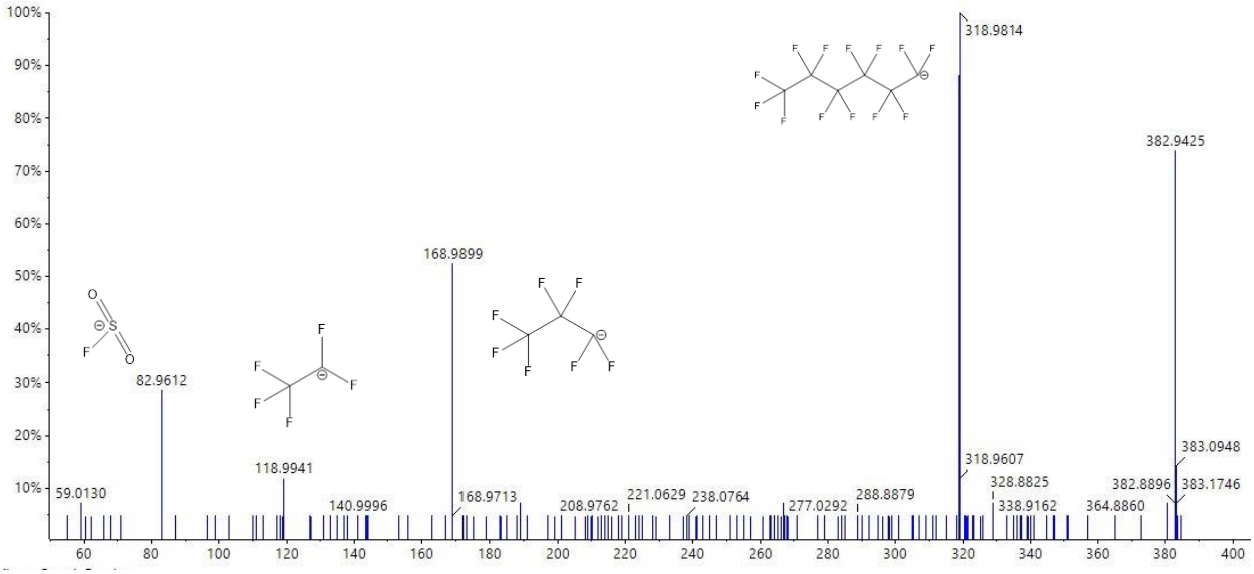

Figure S5. MS/MS spectra for perfluorohexane sulfinate (PFHxSi)

Nitrate analysis

Nitrate was measured on a Metrohm (Herisau, Switzerland) 930 Compact IC Flex ion chromatography system fitted with a Metrosep A Supp 5 (150/4.0) column. The mobile phase was a 3.2 mM sodium bicarbonate and 1 mM sodium carbonate mixture in Milli-Q water at a flowrate of 0.7 mL min^{-1}. Samples were quantified using a 7-point calibration curve ($R^2 > 0.99$) between 0.05-10 mg L^{-1} in Milli-Q water. Calibration quality controls were analyzed every 20 samples and were within ±15% of the expected value. Nitrate was never detected in Milli-Q water procedural blanks that were run every 8 sample injections. We report all values above the lowest calibration concentration = 0.05 mg L^{-1} (method reporting level) in Table S2.

Microbial community composition

DNA Extraction, Quantification, and Amplification. Frozen sediment pellets from the microbial analysis microcosms were shipped on dry ice to the USGS Reston Microbiology Lab (RML). At RML, genomic DNA (gDNA) was extracted from soil samples using the Qiagen DNeasy PowerSoil Kit (Qiagen, Germantown, Maryland, United States) according to the manufacturer's protocol. The gDNA was quantified with a Quant-iT™ double-stranded DNA (dsDNA) High Sensitivity assay kit (Life Technologies, Carlsbad, California, United States). Bacterial 16S rRNA gene abundances were determined using quantitative PCR (qPCR) for the live microcosms. Each 20 µL reaction contained 10 µL of KiCqStart® SYBR® Green qPCR ReadyMix™ with Rox (Sigma), 0.5 µM of each forward and reverse primer, and 2 µL of DNA extract diluted to 1.0 ng μL^{-1}. The primers 515f (5'-GTGCCAGCMGCCGCGGTAA-3') and 806r (5'-GGACTACHVGGGTWTCTAAT-3'), targeting the V4 region of the bacterial 16S rRNA gene of Bacteria and Archaea,[3–6] were used for amplification. Triplicate qPCR reactions

for each sample were run on a StepOne™ Real-Time PCR System (Applied Biosystems, Waltham, Massachusetts, United States) according to the following thermocycling conditions: heat inactivation at 95°C for 10 s, followed by 40 cycles of 95°C for 15 s (denaturation), 55°C for 15 s (annealing), and 72°C for 10 s (extension and acquisition), followed by a melt curve step comprising of 95°C for 15 seconds, 60°C for 1 minute, and 95°C for 15 seconds. A standard curve was constructed by serially diluting plasmid DNA containing the cloned 16S rRNA gene from *E. coli*. A 10-fold dilution series was generated, consisting of standards that were 10^0 to 10^4 gene copy numbers per µL of solution. A second, spiked standard curve with microcosm and plasmid DNA was used to account for sample-specific inhibition according to Hargreaves et al. 2013.[7] Standard curves had r^2 values > 0.99 and efficiency values of 95 and 100% for the normal and spiked standards.

Illumina 16S rRNA gene sequencing. gDNA extracts were sent to Michigan State University's Research Technology Support Facility (RTSF, East Lansing, MI, United States) for Illumina 16S iTag sequencing (Illumina, Inc., San Diego, CA, United States). Amplicon libraries of the V4 hypervariable region of the 16S rRNA gene of Bacteria and Archaea were prepared using dual indexed, Illumina compatible primers 515f and 806r[3–6] following the protocol developed by the Patrick Schloss lab.[8] Following PCR, all products were batch normalized using Invitrogen SequalPrep DNA Normalization (Invitrogen, Carlsbad, CA, United States) plates, and products were recovered from the plate pooled. The pool was quality controlled and quantified using a combination of Qubit dsDNA HS, Agilent 4200 TapeStation HS DNA1000 and Invitrogen Collibri Library Quantification qPCR assays. This pool was loaded onto an Illumina MiSeq Standard v2 flow cell, and sequencing was performed in a 2 × 250 bp paired-end format using a 500-cycle v2 reagent cartridge. Custom sequencing primers were added to appropriate

wells of the reagent cartridge as described in Kozich et al.[8] Base calling was done by Illumina Real-Time Analysis (RTA) v1.18.54, and output of RTA was demultiplexed and converted to FastQ format with Illumina Bcl2fastq v2.20.0.

Sequence Processing and Analysis. Initial quality control, alignment, and taxonomic assignment of microbial sequence data were performed using mothur v.1.43.0[9] according to the mothur MiSeq standard operating procedure[8] and using the USGS Advanced Research Computing (ARC) Yeti high-performance computing facility. Operational taxonomic units (OTUs) were assigned based on a 97% similarity cutoff, with taxonomy assigned based on similarity the Silva nr99 v132 database.[10,11] Sequencing yielded an average of 237,656 ± 44,121 reads and 6,288 ± 745 OTUs over the 24 live microcosm samples (Table S4). The killed controls had an average of 14,151 ± 6,375 reads and 345 ± 90 OTUs indicating that autoclaving and addition of azide successfully inhibited microbial activity (Supplementary Methods Table 1).

Supplementary Methods Table 1. Phylogenetic affiliation of the 16S rRNA gene sequences from the killed control microcosms that was collected at day 0 and 17.

See accompanying Excel spreadsheet

Statistical and diversity analyses were performed in R using the vegan version 2.5.6,[12] and phyloseq v. 1.26.1[13] packages. Alpha diversity statistics were calculated without rarefaction or singleton removal and prior to normalization. Coverage ranged from 0.98 to 1.00 indicating that the microbial community was well sampled (Supplementary Methods Table 2). The phylogenetic affiliation of OTUs were converted to percent (%) relative abundance. The Quantitative Sequencing (QSeq) approach [14–16] was used to normalize relative abundance data with bacterial 16S rRNA gene copy numbers to incorporate differences in count data across samples. In R and using code adapted from (https://benjjneb.github.io/dada2/tutorial.html), relative abundance data were multiplied by total 16S rRNA gene copy numbers for each sample

to calculate the estimated total abundance of microbial taxa. This generated a new phyloseq object that was used to create a Bray Curtis distance matrix which was created using the 'ordinate' function within phyloseq. Non-metric multidimensional scaling (NMDS) constrained to 2 axes was performed on this distance matrix, with a minimum of 1,000 replicates performed using the 'plot_ordination' function in the phyloseq package.

Supplementary Methods Table 2. Characteristics of the microbial community in unamended and PFHxSAm-amended microcosms including a summary of sequencing, diversity statistics, and 16S rRNA gene copy numbers.

See accompanying Excel spreadsheet

Following QSeq normalization, microbial taxonomic data were plotted using Prism version 9 (GraphPad Software, San Diego, CA). A literature review was conducted to identify taxa with known dechlorinating bacteria, PFAS-tolerance or biotransforming ability, and ammonia oxidizers. Those taxa were then identified in the QSeq normalized biom file with all the OTUs. The total abundance of these taxa was calculated for each microcosm, averaged per treatment then imported into Prism version 9 (GraphPad Software, San Diego, California, United States). Values were transformed via the standard function Y = log(Y) to allow for low abundance taxa to be visible in graphical outputs. The transformed data were graphed into a heatmap using Prism version 9.

mothur code

```
#Mothur version 1.43.0 run on 10/12/2021 by Cassandra Harris, USGS using
the USGS Advanced Research Computing (ARC) Yeti high-performance computing
facility.

set.dir(tempdefault=/home/crharris/Lara_Mothur)
make.contigs(file=LaraPFAS.files.txt, processors=40)
summary.seqs()
screen.seqs(fasta=current, group=current, maxambig=1, optimize=start-end-
minlength-maxlength, criteria=90, maxhomop=8)
```

```
summary.seqs()
unique.seqs(fasta=current)
count.seqs(name=current, group=current)
summary.seqs(count=current)
align.seqs(fasta=current,
reference=/home/crharris/Silva_Files_Cass/silva.nr_v138.align, flip=T)
summary.seqs(count=current)
screen.seqs(fasta=current, count=current, optimize=start-end-minlength-
maxlength, criteria=95)
summary.seqs(count=current)
filter.seqs(fasta=current, vertical=T, trump=.)
summary.seqs(count=current)
unique.seqs(fasta=current, count=current)
pre.cluster(fasta=current, count=current, diffs=3)
summary.seqs(count=current)
chimera.vsearch(fasta=current, count=current, dereplicate=t)
remove.seqs(fasta=current, accnos=current)
summary.seqs(count=current)
classify.seqs(fasta=current, count=current,
reference=/home/crharris/Silva_Files_Cass/silva.nr_v138.align,
taxonomy=/home/crharris/Silva_Files_Cass/silva.nr_v138.tax, cutoff=80,
probs=F)
remove.lineage(fasta=current, count=current,
taxonomy=current,taxon=Chloroplast-Mitochondria-unknown-Eukaryota)
summary.seqs(count=current)
cluster.split(fasta=current, count=current, taxonomy=current,
splitmethod=classify, taxlevel=5, cutoff=0.03)
make.shared(list=current, count=current, label=0.03)
classify.otu(list=current, count=current, taxonomy=current, label=0.03)
make.biom(shared=current, constaxonomy=current)
count.groups(shared=current)
collect.single(shared=current, calc=chao-invsimpson, freq=100)
summary.single(shared=current, calc=nseqs-coverage-sobs-invsimpson)
tree.shared(shared=current,calc=jclass-braycurtis)
```

R analysis code

```
### QSeq normalization and construction of NMDS plot
library(phyloseq)
library(ggplot2)
library(Biostrings)
library(vegan)
theme_set(theme_bw())
setwd("~/Documents/1-Projects/2-Lara_PFAS")
### Upload your metadata file
### For the cells without a qPCR value, leave them blank
Meta.data <- read.csv("Lara_metadata_qPCR.csv")
Meta.df <- data.frame(Meta.data)
row.names(Meta.df) #<-Sample_ID
View(Meta.df)
Metadata.df <- sample_data(data.frame(Treatment = Meta.df["Treatment"],
                                      Treatment = Meta.df["Treatment"],
```

```
                                    Sample_Type = Meta.df["Sample_Type"],
                                    qPCR = Meta.df["qPCR"],
                                    Time_Point = Meta.df["Time_Point"],
                                    PFAS_Additions = Meta.df["PFAS_Additions"],
                                    Library_Name = Meta.df["Library_Name"],
                                    Subset1 = Meta.df["Subset1"],
                                    Bottle = Meta.df["Bottle"],
                                    Subset2 = Meta.df["Subset2"]))
View(Metadata.df)
### Setting the row names since my original rownames were changed to sa## and not my sample names
rownames(Metadata.df) <-
c("225","226","227","228","229","230","231","232","233","234","235","236","237","238","239","240","241","242","243","244","245","246","247","248","249","250","251","252","253","254","Neg1","Neg2","RTSF_NTC")
View(Metadata.df)
otu.table.L <- as.data.frame(as.matrix(read.csv("LaraPFAS_BiomTable_Fresh.csv", header = TRUE, row.names = 1)))
View(otu.table.L)
### Setting my column names because my original column names were changed to have an X in front of the sample name and not my column names
colnames(otu.table.L) <-
c("225","226","227","228","229","230","231","232","233","234","235","236","237","238","239","240","241","242","243","244","245","246","247","248","249","250","251","252","253","254","Neg1","Neg2","RTSF_NTC")
View(otu.table.L)
### Changing the orientation of the OTU table so that it matches the metadata with samples names as the rows
otu.table.L.flipped <- t(otu.table.L)
View(otu.table.L.flipped)
### Adding Taxa
Taxa2 <- import_mothur(mothur_constaxonomy_file
="LaraPFAS.files.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.opti_mcc.0.03.cons.taxonomy")
### Created a phyloseq object
ps <- phyloseq(otu_table(otu.table.L.flipped, taxa_are_rows = FALSE),
sample_data(Metadata.df), tax_table(Taxa2))
### We used this command to turn our OTU table into a relative abundance table
rel.abund <- transform_sample_counts(ps, function(OTU)((OTU/sum(OTU))))
### Now we're going to normalize using the qPCR genes per g of sediment
orf.ra <- as.data.frame(as.matrix(otu_table(rel.abund)))
View(orf.ra)
qpcr.data <- as.matrix(Metadata.df$qPCR)
View(qpcr.data)
orf.normalized <- orf.ra*qpcr.data ### New OTU table
View(orf.normalized)
```

```
### Flipping because Excel does NOT like this orientation
orf.normalized.flipped <- t(orf.normalized)
View(orf.normalized.flipped)
write.csv(orf.normalized.flipped, "LaraPFAS_qPCR_OTU_Table_New.csv")
### Created a new phyloseq object
ps.norm <- phyloseq(otu_table(orf.normalized, taxa_are_rows = FALSE),
sample_data(Metadata.df), tax_table(Taxa2))
### From here on out, I should be able to use ps.norm as my normal phyloseq object
### Removing the negatives and the killed control from the NMDS because they skew the
figures
subset_qLaraPFAS <- subset_samples(ps.norm, Subset1=="no")
sample_data(subset_qLaraPFAS)
subset_qLaraPFAS.ord <- ordinate(subset_qLaraPFAS, "NMDS", "bray")
p51<- plot_ordination(subset_qLaraPFAS,
                      subset_qLaraPFAS.ord,
                      type = "samples", color = "Treatment", shape = "Sample_Type",
                      title = "taxa")
p51 + geom_point(size=5) +
  ggtitle("Lara PFAS Sample Treatments") +
  scale_color_manual(values = c("#332288", "#D16103", "#44AA99", "#332288", "#D16103",
"#44AA99", "#CC6677")) + stat_ellipse()
### Creating a diversity table
estimate_richness(subset_qLaraPFAS, split = TRUE, measures = NULL)
```

Model fitting and data analysis

Solving rate equations. The kinetic rates (k) in Equations 1 and 2 were fit to the measured data (X) using Markov chain Monte Carlo (MCMC) analysis (Equation S1):

$$P(k|X) \propto P(X|k) \times P(k) \qquad \textbf{(Equation S1)}$$

Where P(k|X) is the posterior probability distribution of the data-optimized kinetic rates, P(X|k) is the likelihood equation, and P(k) is the prior. For each rate (k), we implemented normal priors parameterized by the mean and standard deviation of a non-linear least squares fitting of the experimental data implemented in *scipy* version 1.7.3.[17] The likelihood was the squared error between the measurements and the modeled concentration shown in Equation S2:

$$P(X|K) = \big(X - f(k)\big)^2 \qquad \textbf{(Equation S2)}$$

Where f(k) is either Equation 1 (1st order loss of precursor spike) or Equation 2 (0th order production of PFHxS), depending on the data being modeled.

MCMC analysis was performed using *emcee* version 3.1.1 with four independently seeded walkers that update their sampling location according to the differential evolution Snooker algorithm. The analyses were run until the MCMC error was $2500^{-0.5}$. We present results for the expected mean and 90th confidence interval throughout the manuscript. The code is available for use at https://github.com/SunderlandLab/pfas-precursor-biotransformation.

Box model. Zero and first order reactions compose the four box model (Equations S3-S6) and is shown graphically in Figure S6.

$$\frac{dM_{aq}}{dt} = k_2 M_{sorb} + k_{bio} M_{bio} - k_1 c_{aq} - k_{bio} M_{aq} \quad \textbf{(Equation S3)}$$

$$\frac{dM_{sorb}}{dt} = k_1 M_{aq} - k_2 M_{sorb} \quad \textbf{(Equation S4)}$$

$$\frac{dM_{bio}}{dt} = k_{bio} M_{aq} - k_{bio} M_{bio} - k_{PFHxS} \quad \textbf{(Equation S5)}$$

$$\frac{dM_{PFHxS}}{dt} = k_{PFHxS} \quad \textbf{(Equation S6)}$$

Where M_x is the mass in nmoles of PFAS is reservoir x (where x is either aqueous, bio, or sorbed FHxSA or PFHxS), k_1 and k_2 (day^{-1}) are the forward and backwards rate constants matching $k_{control}$ (Table 1; see https://github.com/SunderlandLab/pfas-precursor-biotransformation for more details), k_{bio} (day^{-1}) is the rate of precursor microbial association (Table 1), and k_{PFHxS} (nmol day^{-1}) is rate of PFHxS production (Table 1). The model is initiated by setting $M_{aq}(t=0)$ equal to the initial mass of the precursor spike at the start of incubation.

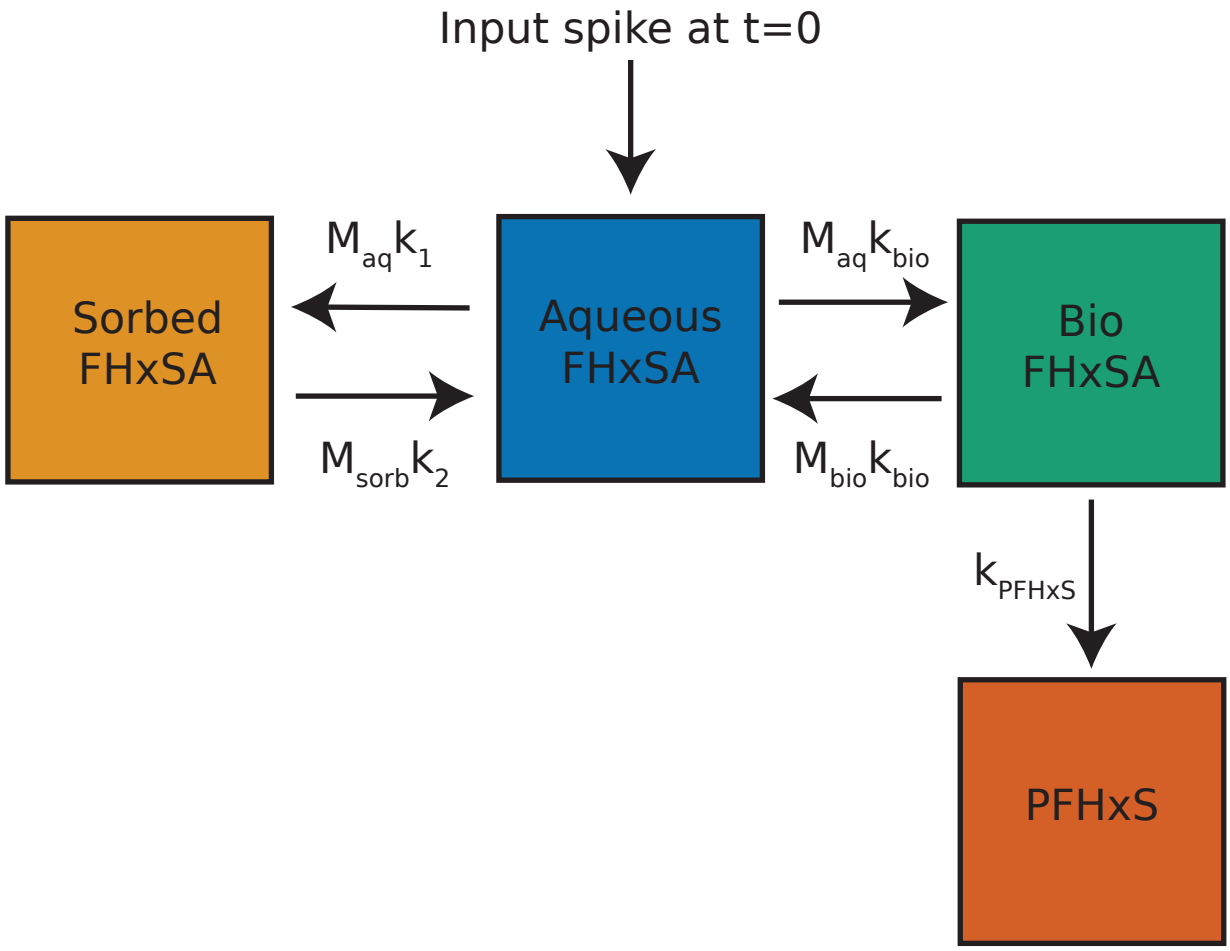

Figure S6. Schematic of four box model. The box model is used to generate results shown in Figure 5.

The sorbed (sediment/glassware) and biological reservoirs of PFHxS were neglected because they are not substantial reservoirs in our experimental setup. The PFHxS mass balance is shown in Supplementary Methods Equation 1:

$$M_{tot} = M_{aq} + M_{sorb} + M_{bio} \qquad \textbf{(Supplementary Equation 1)}$$

Where M is the mass of PFHxS is various reservoirs. Substituting partitioning coefficient (K_d for sediment/water and K_{oc} for microbial/water) into equation Supplementary Methods Equation 1 results in Supplementary Methods Equation 2:

$$M_{tot} = c_{aq}(V + k_d M_{sed} + K_{oc} M_{bio}) \qquad \textbf{(Supplementary Methods Equation 2)}$$

Where c_{aq} is the aqueous phase concentration of PFHxS, V is the water volume (35 mL), M_{sed} is the sediment mass (10 g), and M_{bio} is the microbial mass. The term for the sorbed reservoir can be neglected because $\log K_d$ of PFHxS < -1.5 to Cape Cod sediment.[18] Using the range of $\log K_{oc}$ reported in the literature (0.66 – 2.0),[18] M_{bio} would have to be 0.35–7.6 g for the biological reservoir to be greater than the aqueous reservoir. The biomass to sediment ratio in the microcosms would therefore have to be approximately 1:10 – 1:1. Although the mass of biomass was not directly measured, we assume that such large ratios are not likely because the fraction of organic carbon on Cape Cod sediment is <0.1% by weight.[18,19]

Supplementary Results

Rapid biological sorption of sulfonamido precursors

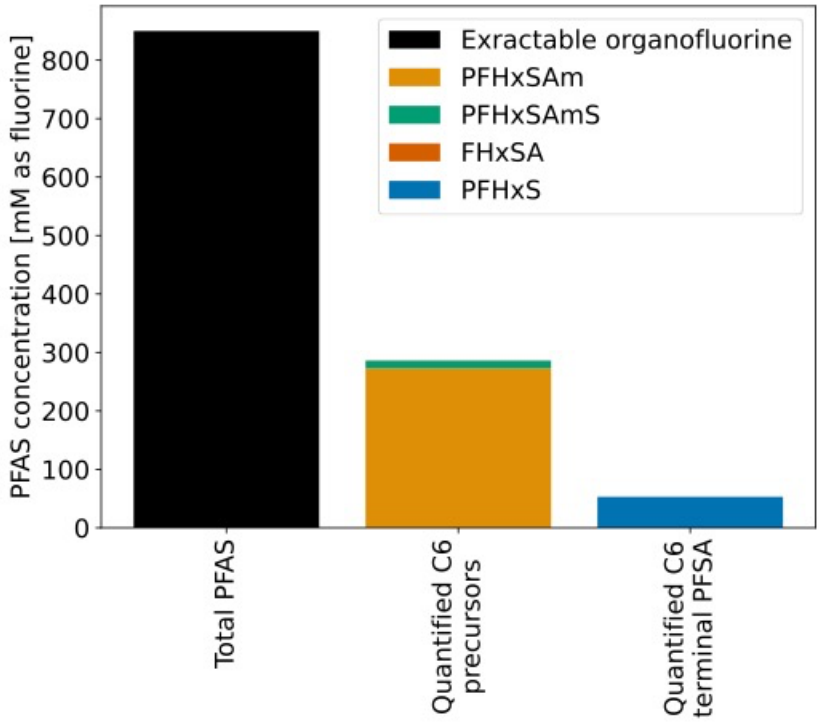

Figure S7. Comparison of C6 PFAS concentrations in a 2001 3M LightWater AFFF. Concentrations are presented in mM as fluorine. The concentration of extractable organofluorine (black bar; total PFAS proxy) and PFHxS (blue) are presented in Ruyle et al. 2021a.[1]

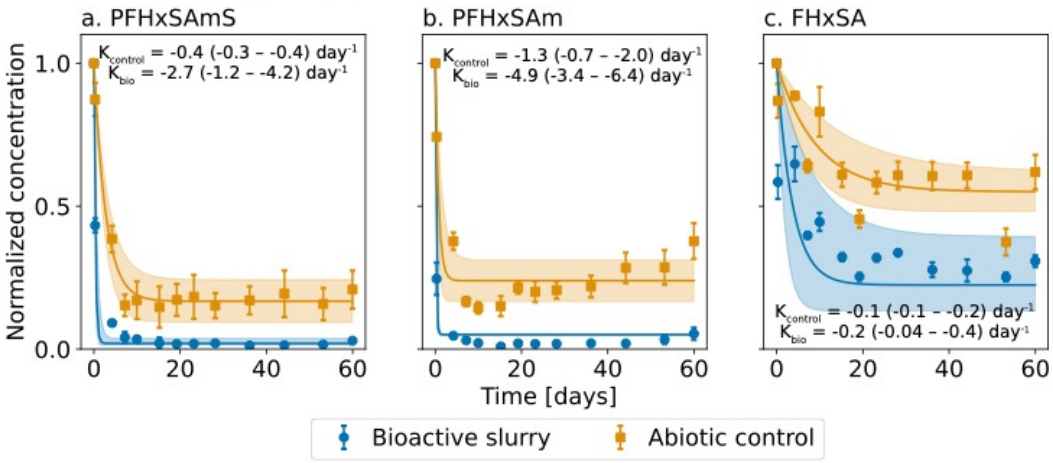

Figure S8. Time-dependent loss of precursors in water-sediment slurry at 19°C. The mean (dots) and standard deviation (error bars) of measurements (n = 3 per timepoint) were normalized to the input concentration. Data were modeled according to Equation 1 (shown above the plots) using MCMC. Lines represent the expected mean and shaded regions represent the 90th confidence interval. The expected mean and 90th confidence interval for the rates (in units [days^{-1}]) are shown for each precursor (negative sign indicates a loss process).

Links between microbial activity and PFHxSAm biotransformation

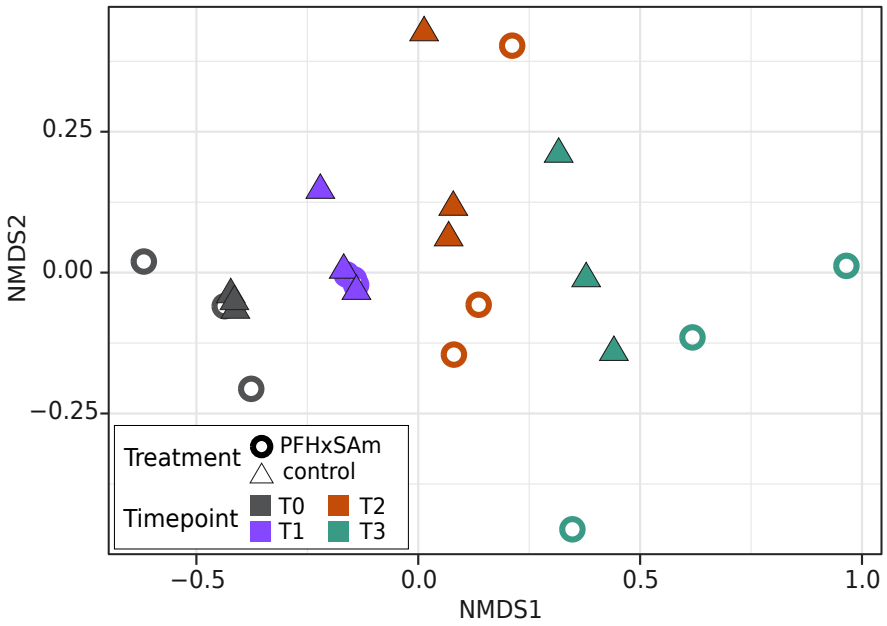

Figure S9. Nonmetric multidimensional scaling (NMDS) plot of Bray-Curtis distances among microbial communities in bioactive PFHxSAm-amended and unamended (control) microcosms over time. Data were QSeq normalized prior to calculating sample dissimilarity. Samples from PFHxSAm-amended microcosms are indicated by open circles and unamended microcosms by triangles with sampling timepoint indicated by color.

Operational taxonomic units (OTUs) were dominated by several phyla in both the PFHxSAm-amended and control microcosms including Gammaproteobacteria (44-49%), Bacteroidota (28-30%), and Acidobacteria (domain Bacteria,13-26%). Members of the domain Archaea were detected in all PFHxSAm-amended and control microcosms and ranged from 2.50 x 10^5 to 2.58 x 10^6 gene copies per g sediment. Shifts in the microbial community composition over time were assessed using nonmetric multidimensional scaling (NMDS) of the Bray-Curtis distances among OTUs (Figure S9). Dimension one reflected temporal changes (shifting to the right in time) and dimension two reflected variability among replicate microcosms. Shifts along dimension one corresponded to statistically significant decreases (test, $p<0.05$, score) in the number of OTUs and Chao1 richness (Supplementary Methods Table 2, Figure S9). Overlap

along dimension one in microcosms that did/did not receive PFHxSAm additions at each timepoint was observed (Figure S9). Inverse Simpson's and the Shannon Diversity indices were also similar between the unamended and PFHxSAm-amended microcosms at each timepoint (Supplementary Methods Table 2, Figure S9). Together these data indicate that changes in the microbial community due to the bottle effect outweighed those due to the presence of PFHxSAm. Although total bacteria 16S rRNA genes decreased over time in the PFHxSAm-amended microcosms at T3, total copy numbers were higher compared to the unamended microcosms. This indicates that PFHxSAm might be supporting microbial growth over background levels of substrates in the source sediment.

At all timepoints, members of the *Polyangia* genus (family Polyangiaceae) were dominant taxa making up a similar proportion of the community over time (Figure 4a, Figure S10a). Members of the Polyangiaceae are known to degrade complex bio-macromolecules, but little is known about the mechanisms of degradation. The potential for these organisms to contribute to the degradation of PFAS compounds is unknown.

Nitrifiers found throughout the incubations for both PFHxSAm-amended and control treatments included OTUs affiliated with autotrophic ammonia oxidizing (amines à nitrite) archaea (AOA) and bacteria (AOB) in the classes Nitrososphaeria (*Candidatus* Nitrosopumilus,[20] *Candidatus* Nitrososphaera,[21] and *Candidatus* Nitrosotalea[22]) and Gammaproteobacteria (*Nitrosomonas* and *Nitrosospira*[23]), respectively. The nitrite-oxidizing (nitrite à nitrate) bacteria detected were members of the Alphaproteobacteria (*Nitrobacter*[23]), Nitrospinia (*Nitrospina*[24]), and Nitrospiria (*Nitrospira*[24,25]) classes. Three additional nitrifying taxa increased in total abundance from 0 to 17 days in both the PFHxSAm-amended and control microcosms: *Nitrospira*, *Nitrosospira*, and *Nitrosomonas*. There were no significant differences in their total

abundance at the end of the incubation between treatments (paired t-test, *p*-value > 0.05). The largest increase in abundance over the experiment was seen for the nitrite oxidizer *Nitrospira* (phylum Nitrospirota) which increased from an average of 3.82×10^5 ($\pm 1.03 \times 10^5$) at day 0 to 3.94×10^6 ($\pm 2.71 \times 10^6$) genes per g sediment at day 17 in PFHxSAm and control microcosms (Table S5). The increase over time of *Nitrospira*, *Nitrosospira*, and *Nitrosomonas* in both the control and amended treatments indicates that their presence was not linked to degradation of PFHxSAm but was likely supported by ammonia in groundwater at the field site.[26]

Nitrobacter were also observed in the precursor-amended and control microcosms, but members of this genus were found to decrease from days 0 to 17 in the amended microcosms and detected transiently in the control bottles (Figure 4c and S10c). The other AOA detected, *Candidatus* Nitrosotalea, also was observed in both treatments and decreased in total abundance from 0 to 17 days.

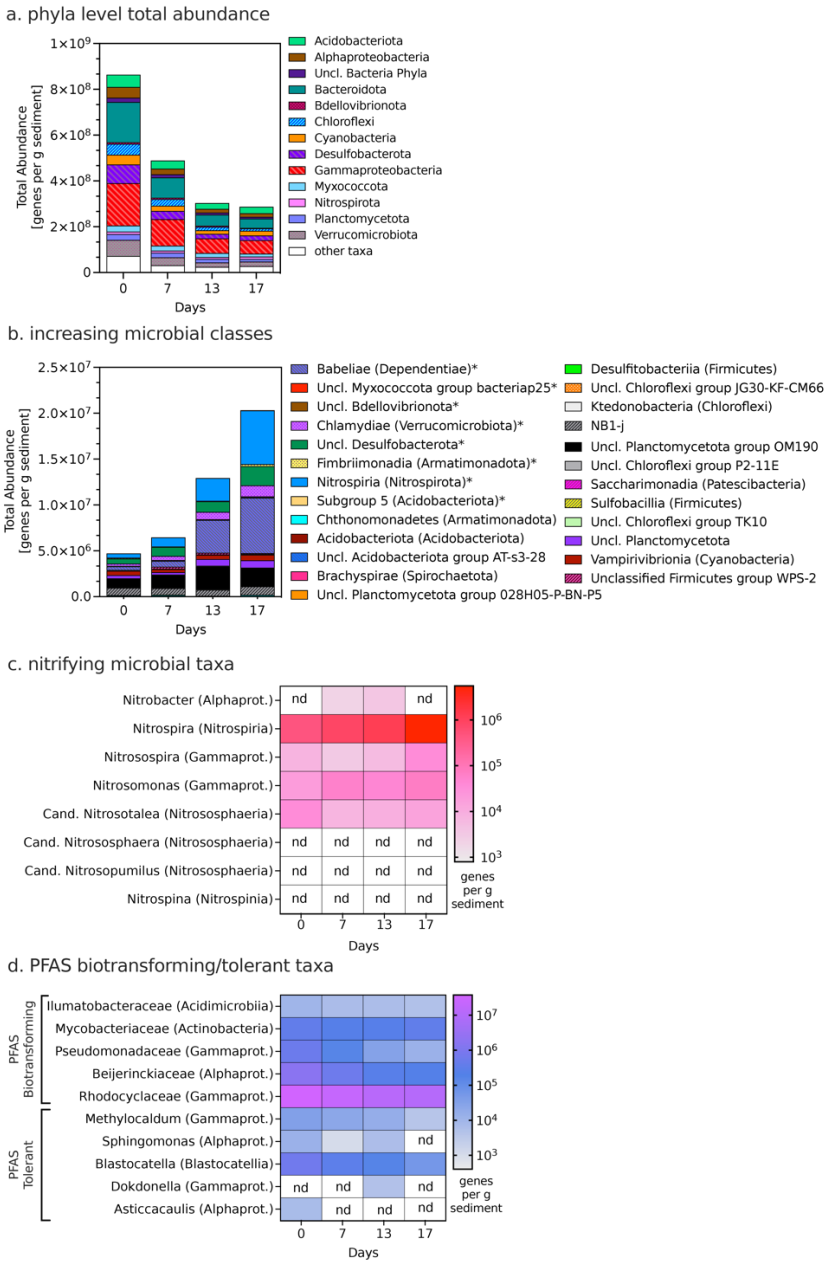

Figure S10. Composition of microbial communities in the unamended control microcosms using QSeq. (A) Total abundance of OTUs over time at the phyla level or class level for Proteobacteria. (B) Taxa at the class level that increased over time in the control microcosms (asterisks indicate classes that also increased over time in the PFHxSAm-amended microcosms). (C, D) Heat maps showing the total abundance of taxa known for their ability to (c) catalyze nitrification reactions and (D) biotransform or tolerate PFAS-compounds. Values in panels C and D were log transformed to allow visualization of low abundance taxa and 'nd' indicates not detected. In panel D, PFAS transforming taxa include both bioaccumulating and biodegrading organisms. Data are presented as average genes per g sediment for triplicate microcosms; values for each replicate are presented in Table S4 and references used to assign putative function for each taxon are presented in Table S5. "Uncl." refers to unclassified taxa.

Table S4. Phylogenetic affiliation of the 16S rRNA gene sequences from the PFHxSAm and unamended microcosms at day 0, 6, 12, and 17 sediment.

See accompanying Excel spreadsheet

Table S5. Total abundance of known nitrifying, PFAS-related, and dechlorinating taxa.

See accompanying Excel spreadsheet

Supplementary Discussion

Experimental and modeling results agree with field observations

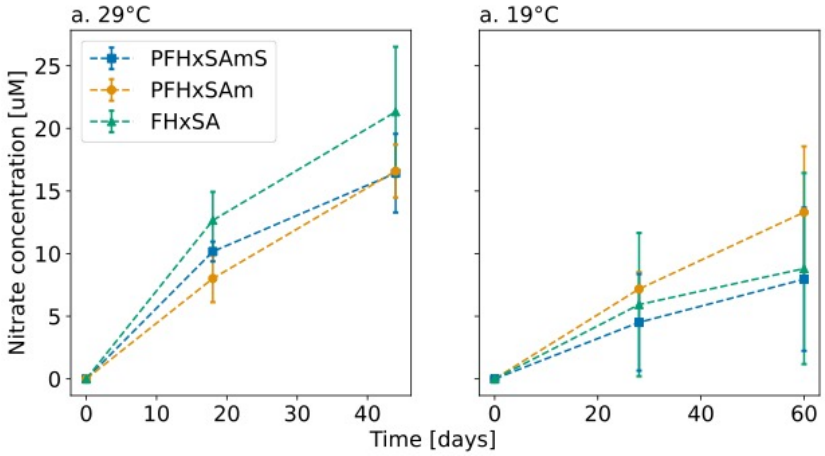

Figure S11. Concentration of nitrate in bioactive experiments. Dots represent the mean and error bars represent the standard deviation from three replicate bottles. Dashed lines connect the temporal data. Panel a shows data from the experiment conducted at 29°C corresponding to PFAS data presented in Figure 1. Panel b shows data from the experiment conducted at 19°C corresponding to PFAS data presented in Figure S8.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

(1) Ruyle, B. J.; Thackray, C. P.; McCord, J. P.; Strynar, M. J.; Mauge-Lewis, K. A.; Fenton, S. E.; Sunderland, E. M. Reconstructing the Composition of Per- and Polyfluoroalkyl Substances in Contemporary Aqueous Film-Forming Foams. *Environ. Sci. Technol. Lett.* **2021**, *8* (1), 59–65. https://doi.org/10.1021/acs.estlett.0c00798.

(2) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48* (4), 2097–2098. https://doi.org/10.1021/es5002105.

(3) Caporaso, J. G.; Lauber, C. L.; Walters, W. A.; Berg-Lyons, D.; Huntley, J.; Fierer, N.; Owens, S. M.; Betley, J.; Fraser, L.; Bauer, M.; Gormley, N.; Gilbert, J. A.; Smith, G.; Knight, R. Ultra-High-Throughput Microbial Community Analysis on the Illumina HiSeq and MiSeq Platforms. *ISME J* **2012**, *6* (8), 1621–1624. https://doi.org/10.1038/ismej.2012.8.

(4) Apprill, A.; McNally, S.; Parsons, R.; Weber, L. Minor Revision to V4 Region SSU RRNA 806R Gene Primer Greatly Increases Detection of SAR11 Bacterioplankton. *Aquatic Microbial Ecology* **2015**, *75* (2), 129–137. https://doi.org/10.3354/ame01753.

(5) Parada, A. E.; Needham, D. M.; Fuhrman, J. A. Every Base Matters: Assessing Small Subunit RRNA Primers for Marine Microbiomes with Mock Communities, Time Series and Global Field Samples. *Environmental Microbiology* **2016**, *18* (5), 1403–1414. https://doi.org/10.1111/1462-2920.13023.

(6) Walters, W.; Hyde, E. R.; Berg-Lyons, D.; Ackermann, G.; Humphrey, G.; Parada, A.; Gilbert, J. A.; Jansson, J. K.; Caporaso, J. G.; Fuhrman, J. A.; Apprill, A.; Knight, R. Improved Bacterial 16S RRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* **2015**, *1* (1), e00009-15. https://doi.org/10.1128/mSystems.00009-15.

(7) Hargreaves, S. K.; Roberto, A. A.; Hofmockel, K. S. Reaction- and Sample-Specific Inhibition Affect Standardization of QPCR Assays of Soil Bacterial Communities. *Soil Biology and Biochemistry* **2013**, *59*, 89–97. https://doi.org/10.1016/j.soilbio.2013.01.007.

(8) Kozich, J. J.; Westcott, S. L.; Baxter, N. T.; Highlander, S. K.; Schloss, P. D. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Applied and Environmental Microbiology* **2013**, *79* (17), 5112–5120. https://doi.org/10.1128/AEM.01043-13.

(9) Schloss, P. D.; Westcott, S. L.; Ryabin, T.; Hall, J. R.; Hartmann, M.; Hollister, E. B.; Lesniewski, R. A.; Oakley, B. B.; Parks, D. H.; Robinson, C. J.; Sahl, J. W.; Stres, B.; Thallinger, G. G.; Van Horn, D. J.; Weber, C. F. Introducing Mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology* **2009**, *75* (23), 7537–7541. https://doi.org/10.1128/AEM.01541-09.

(10) Pruesse, E.; Quast, C.; Knittel, K.; Fuchs, B. M.; Ludwig, W.; Peplies, J.; Glöckner, F. O. SILVA: A Comprehensive Online Resource for Quality Checked and Aligned Ribosomal RNA Sequence Data Compatible with ARB. *Nucleic Acids Res* **2007**, *35* (21), 7188–7196. https://doi.org/10.1093/nar/gkm864.
(11) Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F. O. The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools. *Nucleic Acids Res* **2013**, *41* (Database issue), D590–D596. https://doi.org/10.1093/nar/gks1219.
(12) Oksanen, J.; Simpson, G. L.; Blanchet, F. G.; Kindt, R.; Legendre, P.; Minchin, P. R.; O'Hara, R. B.; Solymos, P.; Stevens, M. H. H.; Szoecs, E.; Wagner, H.; Barbour, M.; Bedward, M.; Bolker, B.; Borcard, D.; Carvalho, G.; Chirico, M.; Caceres, M. D.; Durand, S.; Evangelista, H. B. A.; FitzJohn, R.; Friendly, M.; Furneaux, B.; Hannigan, G.; Hill, M. O.; Lahti, L.; McGlinn, D.; Ouellette, M.-H.; Cunha, E. R.; Smith, T.; Stier, A.; Braak, C. J. F. T.; Weedon, J. Vegan: Community Ecology Package, 2022. https://CRAN.R-project.org/package=vegan (accessed 2022-09-12).
(13) McMurdie, P. J.; Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **2013**, *8* (4), e61217. https://doi.org/10.1371/journal.pone.0061217.
(14) Jian, C.; Luukkonen, P.; Yki-Järvinen, H.; Salonen, A.; Korpela, K. Quantitative PCR Provides a Simple and Accessible Method for Quantitative Microbiota Profiling. *PLOS ONE* **2020**, *15* (1), e0227285. https://doi.org/10.1371/journal.pone.0227285.
(15) Tettamanti Boshier, F. A.; Srinivasan, S.; Lopez, A.; Hoffman, N. G.; Proll, S.; Fredricks, D. N.; Schiffer, J. T. Complementing 16S RRNA Gene Amplicon Sequencing with Total Bacterial Load To Infer Absolute Species Concentrations in the Vaginal Microbiome. *mSystems* **2020**, *5* (2), e00777-19. https://doi.org/10.1128/mSystems.00777-19.
(16) Epp Schmidt, D.; Dlott, G.; Cavigelli, M.; Yarwood, S.; Maul, J. E. Soil Microbiomes in Three Farming Systems More Affected by Depth than Farming System. *Applied Soil Ecology* **2022**, *173*, 104396. https://doi.org/10.1016/j.apsoil.2022.104396.
(17) SciPy 1.0 Contributors; Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **2020**, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2.
(18) Weber, A. K.; Barber, L. B.; LeBlanc, D. R.; Sunderland, E. M.; Vecitis, C. D. Geochemical and Hydrologic Factors Controlling Subsurface Transport of Poly- and Perfluoroalkyl Substances, Cape Cod, Massachusetts. *Environmental Science & Technology* **2017**, *51* (8), 4269–4279. https://doi.org/10.1021/acs.est.6b05573.
(19) Barber, L. B.; Thurman, E. M.; Runnells, D. D. Geochemical Heterogeneity in a Sand and Gravel Aquifer: Effect of Sediment Mineralogy and Particle Size on the Sorption of Chlorobenzenes. *Journal of Contaminant Hydrology* **1992**, *9* (1–2), 35–54. https://doi.org/10.1016/0169-7722(92)90049-K.

(20) Qin, W.; Martens-Habbena, W.; Kobelt, J. N.; Stahl, D. A. Candidatus Nitrosopumilales. In *Bergey's Manual of Systematics of Archaea and Bacteria*; John Wiley & Sons, Ltd, 2016; pp 1–2. https://doi.org/10.1002/9781118960608.obm00122.
(21) Stieglmeier, M.; Klingl, A.; Alves, R. J. E.; Rittmann, S. K.-M. R.; Melcher, M.; Leisch, N.; Schleper, C. 2014. Nitrososphaera Viennensis Gen. Nov., Sp. Nov., an Aerobic and Mesophilic, Ammonia-Oxidizing Archaeon from Soil and a Member of the Archaeal Phylum Thaumarchaeota. *International Journal of Systematic and Evolutionary Microbiology* **2014**, *64* (Pt_8), 2738–2752. https://doi.org/10.1099/ijs.0.063172-0.
(22) Prosser, J. I.; Nicol, G. W. Candidatus Nitrosotalea. In *Bergey's Manual of Systematics of Archaea and Bacteria*; John Wiley & Sons, Ltd, 2016; pp 1–7. https://doi.org/10.1002/9781118960608.gbm01292.
(23) Prosser, J. I.; Head, I. M.; Stein, L. Y. The Family Nitrosomonadaceae. In *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*; Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer: Berlin, Heidelberg, 2014; pp 901–918. https://doi.org/10.1007/978-3-642-30197-1_372.
(24) Daims, H. The Family Nitrospiraceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*; Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer: Berlin, Heidelberg, 2014; pp 733–749. https://doi.org/10.1007/978-3-642-38954-2_126.
(25) Daims, H.; Lebedeva, E. V.; Pjevac, P.; Han, P.; Herbold, C.; Albertsen, M.; Jehmlich, N.; Palatinszky, M.; Vierheilig, J.; Bulaev, A.; Kirkegaard, R. H.; von Bergen, M.; Rattei, T.; Bendinger, B.; Nielsen, P. H.; Wagner, M. Complete Nitrification by Nitrospira Bacteria. *Nature* **2015**, *528* (7583), 504–509. https://doi.org/10.1038/nature16461.
(26) Stoliker, D. L.; Repert, D. A.; Smith, R. L.; Song, B.; LeBlanc, D. R.; McCobb, T. D.; Conaway, C. H.; Hyun, S. P.; Koh, D.-C.; Moon, H. S.; Kent, D. B. Hydrologic Controls on Nitrogen Cycling Processes and Functional Gene Abundance in Sediments of a Groundwater Flow-Through Lake. *Environ. Sci. Technol.* **2016**, *50* (7), 3649–3657. https://doi.org/10.1021/acs.est.5b06155.