Supplementary Information: Read-me instructions

Supplementary Methods S1 In-detail description of inocula preparation, inoculation, and start of the experiment. PDF format.

Supplementary Methods S2 In-depth description of electrospray ionisation mass spectrometry (ESI-MS). PDF format.

Supplementary Table S1 Overview of the physico-chemical characteristics for the four environmental sampling sites. PDF format.

Supplementary Table S2 Overview of the 454-pyrosequencing effort.

Supplementary Figure S1 Schematic figure visualising the experimental set-up. PDF format.

Supplementary Figure S2 Rarefaction curves for the twelve experimental and four environmental bacterial communities. PDF format.

Supplementary Figure S3 Non-metric multidimensional scaling (NMDS) representation of bacterial communities from the four natural environments and experimental treatments. PDF format.

Supplementary Figure S4 Taxonomic composition of bacterial experimental (a) and environmental (b) communities as relative abundances of each phylum. PDF format.

Supplementary Methods S1 *In-detail description of inocula preparation, inoculation, and start of the experiment.*

The experiment was performed as a common garden experiment, applying a batch culture approach in which a uniform, soil-derived yet artificial medium was inoculated with bacterial cells from the four environments (henceforth called experimental treatments). The experiment was run with four experimental treatments (Ba, GW, HL, VA) and a control (note: no bacterial inoculum was added to the control): n = 5.

Preparation of inocula, inoculation, and start of the experiment was done as follows. Samples for bacterial abundance and community composition taken in the field (i.e., from the four aquatic environments EnvBa, EnvGW, EnvHL, and EnvVA) **(1)** were immediately pre-sieved (225 µm; nylon net filter) and filtered with a GF/F filter (0.7 µm, pre-combusted at 400°C for six hours; Whatman, Maidstone, UK) **(2)** to avoid capturing larger particles and remove grazers, respectively. The part of each of the four respective filtrates to be used as inoculum was then stored at 4°C and in the dark for six days (, while the other part was used at once to proceed with analyses of bacterial abundance and community composition on the four original, environmental samples EnvBa, EnvGW, EnvHL, and EnvVA; **(3)**). Immediately prior to inoculation and, thus, the start of the experiment, the four inocula were concentrated to ~1 x 10^6 cells mL^{-1} via tangential flow filtration (Merck Millipore, Billerica, USA) **(4)**. At time point zero (i.e., the beginning of the experiment), five individual and independent batches (i.e., one for each experimental treatment [Ba, GW, HL, and VA] and one for the control) were prepared, consisting of a uniform, soil-derived yet artificial medium (99% of final volume) and the inoculum (1% of final volume; no inoculum was added to the control): 5 x 1, n = 5 **(5)**. Samples for bulk dissolved organic carbon (DOC) concentration, UV-visible absorbance and fluorescence, electrospray ionisation mass spectrometry (ESI-MS), and bacterial abundance were taken immediately following inoculation. Finally, each of the five batch cultures was divided up into three experimental replicate batch cultures each. Hence, each experimental treatment plus the control was run in three independent, experimental replicates after time point zero: 5 x 3, n = 15 **(6)**. See Supplementary Figure S1 for a schematic visualisation of how the experiment was set up and carried out.

Numbers in bold and in brackets refer to numbers found in Supplementary Figure S1.

Supplementary Methods S2 *In-depth description of electrospray ionisation mass spectrometry (ESI-MS).*

ESI-MS measurements were conducted at the Department of Plant and Environmental Sciences, University of Copenhagen (Denmark). Samples were run either by size exclusion chromatography (SEC) or direct injection (DI); both, SEC and DI, were connected in series with diode array detector (DAD), fluorescence detector (FLD), and positive and negative ion mode electrospray quadrupole time-of-flight mass spectrometric detection (ESI-QTOF-MS). Respective experimental conditions are given below. With regard to the present study, only the negative ion mode ESI-QTOF-MS data were analysed.

Samples
Samples were stored at -18°C prior to analysis. First, aliquots of 400 µL were evaporated to complete dryness and reconstituted in an equivalent volume of mobile phase. Two samples served as reference samples and were analysed four times throughout the measurements.

Instrument Calibration
The instrument was calibrated using a solution of 10 mL 0.1 mM sodium hydroxide, 10 mL 1% formic acid, and 980 mL methanol. The calibration was performed on smoothed and centred data, with resolution and Np multiplier set to 10000 and 0.7, respectively; enabling dead-time correction of data of varying high and low intensity peaks. Sodium formate adducts of mass to charge ratios (*m/z*) 181-1133 and a 5^{th} order polynomial fit were used.

Chemical Analysis
The samples were analysed using an Acquity UPLC (Waters, Milford, MA, USA) equipped with a binary solvent delivery system and either a size-exclusion column thermostated at 25°C (Shodex OHpak802.5 HQ; Showa Denko, Munich, Germany) (SEC) or not (DI), and connected in series to a diode array (DAD), fluorescence (FLD), and electrospray ionisation (ESI) quadrupole/orthogonal acceleration time-of-flight mass spectrometer (ESI-QTOF-MS) (Micromass Q-TOF Ultima Global; Waters).

The DAD measured absorbance from 220-499 nm at 1 nm intervals, while the FLD emission spectra ranged from 320-520 nm and was recorded in 10 nm intervals. Ionisation was performed in both positive and negative ion mode. MS operated in QTOF scan mode (100-1999 *m/z*) and the quadrupole served as an ion-focusing device (RF only). The ESI source operating conditions were as follows: ion-source temperature, 120°C; de-solvation gas temperature, 250°C; cone gas flow, 25 L h^{-1} (ESI^+) and 50 L h^{-1} (ESI^-); de-solvation gas flow, 700 L h^{-1}; capillary voltage, 2.5 kV (ESI^+) and 2.0 kV (ESI^-); cone voltage, 20 V; scan time, 1 s; and inter-scan delay, 0.1 s. The mobile phase consisted of 10 mmol NH_4-HCO_3 dissolved in 80:20 (v/v) glass-distilled water:methanol, which was adjusted to pH 8.2 with droplets of NH_4-OH. The injection volume was 1 µL and the flow operated at 0.032 mL min^{-1} based on a column maximum pressure tolerance of 725 psi (SEC) or 0.100 mL min^{-1} (DI).

Transfer of Data
Data were retrieved using the Masslynx software (v4.2; Waters) and exported as NetCDF-files with the DataBridge application. The NetCDF-files were imported into MATLAB and Statistics Toolbox Release (R2009b; The MathWorks, Inc., Natick, USA), applying in-house programming routines (courtesy of G. Tomasi and J. Christensen), thereby binning the QTOF *m/z*-axis to nominal *m/z*-values. The FLD instrument software erroneously interpolated between the actually measured emissions every 10^{th} nm and in doing so introduced nine fake data points per every real data point. The artefact data points were removed from the data.

Data

If, for instance, *m/z* peaks representing carbon isotopes are 1 *m/z* unit apart from each other, ions are single-charged so that the molecular weight is approximately equivalent to the *m/z*-value (i.e., -1 for proton adducts, -18 for ammonium adducts, -23 for sodium adducts in positive ion mode or +1 for proton abstracted ions, -35 for chloride adducts in negative ion mode, etc.). Should *m/z* peaks be ½ *m/z* unit apart from each other, ions are double-charged so that their respective molecular weight is approximately twice that of the *m/z*-value (×2-2 for proton adducts, ×2-36 for ammonium adducts, ×2-46 for sodium adducts in positive ion mode or ×2+2 for proton abstracted ions, ×2-70 for chloride adducts in negative ion mode, etc.). And so forth. Because these types of spectra are generally dominated by single-charged ions and ESI-QTOF-MS was carried out in negative ionisation scan mode, an *m/z*-value of 200 would correspond to a molecular weight of 201 (*m/z*-value +1), assuming that the ion observed at *m/z* 200 is a proton adduct.

Supplementary Table S1. Overview of the physico-chemical characteristics for the four environmental sampling sites.

Name	Acro-nym	Location		Altitude	Tempe-rature	Salinity	pH	DOC	DIC	$Abs_{250/365}$	Abs_{436}	$SUVA_{254}$
		North	East	[m.a.s.l]	[°C]	[PSU]		[mg L^{-1}]	[mg L^{-1}]			
Baltic	EnvBa	63° 33′ 51″	19° 50′ 01″	0	7.8	3.4	7.64	2.68	11.71	5.78	0.009	0.009
Groundwater	EnvGW	64° 15′ 13″	19° 46′ 34″	275	3.9	*na*	5.38	28.95	6.59	5.03	0.069	0.479
Headwater Lake	EnvHL	64° 15′ 38″	19° 45′ 38″	289	11.2	*na*	4.97	17.98	0.83	4.04	0.095	0.035
Vindelälven	EnvVA	63° 57′ 08″	19° 52′ 38″	74	11.4	*na*	7.02	7.72	1.79	4.79	0.019	0.038

Abbreviations: DOC, Dissolved Organic Carbon; DIC, Dissolved Inorganic Carbon; $Abs_{250/365}$, Absorbance measured at wavelength 250 nm divided by that of 365; Abs_{436}, Absorbance measured at wavelength 436 nm; $SUVA_{254}$, Specific Ultraviolet Absorption measured at wavelength 254 nm normalised to the concentration of DOC; *na*, not applicable.

Supplementary Table S2. Overview of 454-pyrosequencing effort.

Sample	Total Number of Reads[A]	Number of Reads after AmpliconNoise[B]	Number of Reads after Perseus[C]	Number of Chimeric Reads	Number of "Good" Reads[D]	Number of unique Reads[E]	Final Number of "Good" Reads[F,H]	Final Number of unique OTUs[G,H]
Ba1	47784	41085	40863	222	40862	57	40862	43
Ba2	34589	29720	29527	193	29526	55	29526	44
Ba3	43899	37734	37665	69	37665	62	37665	46
EnvBa[I]	34724	29877	29868	9	29868	477	29847	396
GW1	45626	39357	39103	254	39097	117	39097	74
GW2	38517	32984	32684	300	32681	115	32681	82
GW3	44391	38141	37863	278	37862	121	37862	89
EnvGW[I]	32662	27671	27512	159	27500	2870	27361	1816
HL1	35054	30240	30042	198	30026	84	30026	60
HL2	42424	36349	36159	190	36153	91	36153	59
HL3	43553	37649	37383	266	37381	107	37381	78
EnvHL[I]	28738	23375	23304	71	23201	1562	23163	1140
VA1	45697	39863	39685	178	39671	78	39671	58
VA2	42895	36500	35585	915	35585	91	35585	66
VA3	33538	29375	29213	162	29206	85	29206	63
EnvVA[I]	26113	21092	20660	432	19373	2142	19340	1684

[A] Number of reads matching multiplex identifiers (MIDs) and primer sequences.

[B] Number of reads after flowgram pre-filtering (AmpliconNoise; Quince et al. 2011).

[C] Number of reads after flowgram pre-filtering (AmpliconNoise) and removal of chimeric reads via Perseus (Quince et al. 2011).

[D] Number of reads after flowgram pre-filtering (AmpliconNoise) as well as removal of chimeric and short (< 200 bp) reads.

[E] Number of unique reads after flowgram pre-filtering (AmpliconNoise) as well as removal of chimeric and short (< 200 bp) reads.

[F] Number of reads after flowgram pre-filtering (AmpliconNoise), removal of chimeric and short (< 200 bp) reads, as well as reads of Archaea, Eukaryotes, and "no hits".

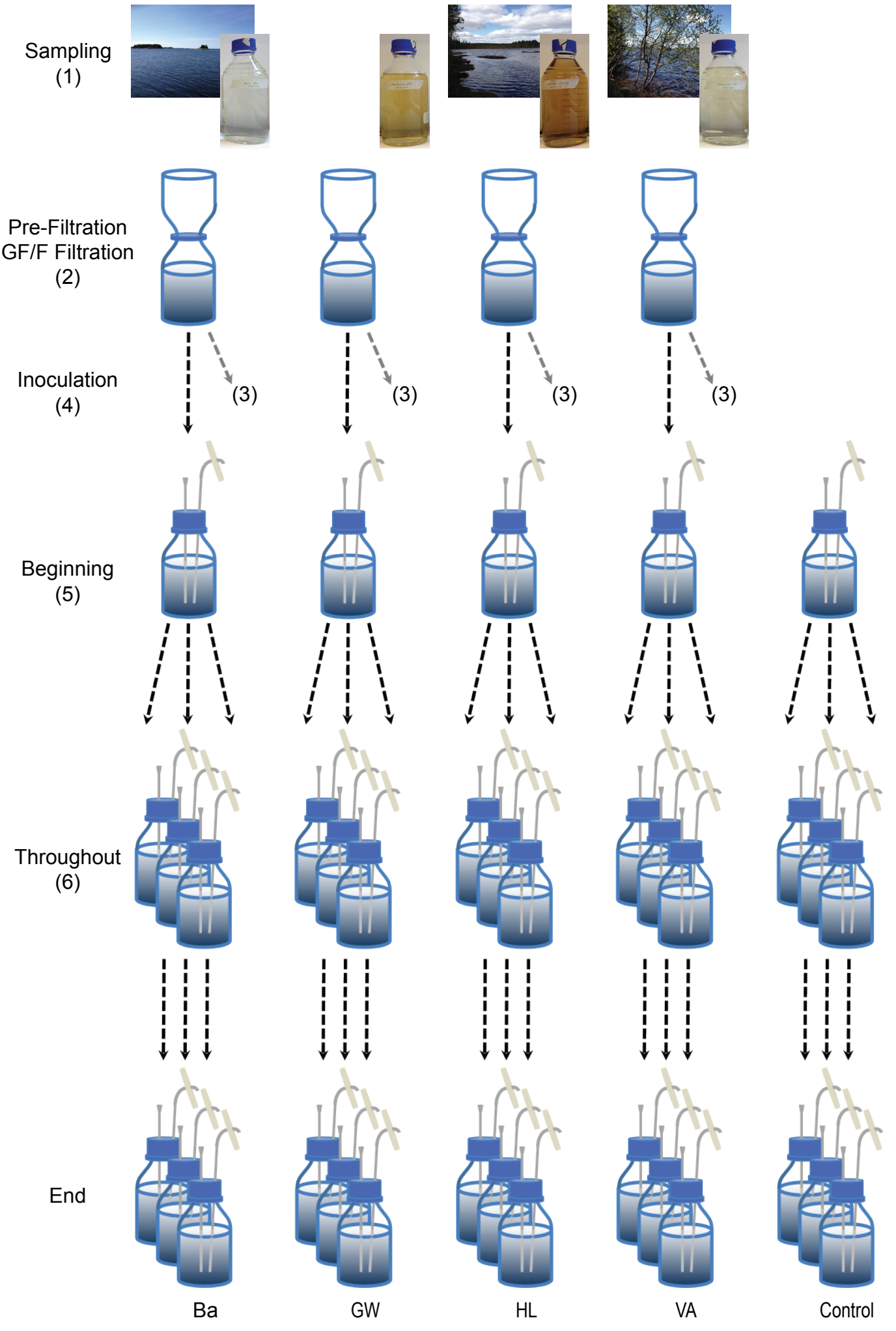
[G] Number of unique OTUs after flowgram pre-filtering (AmpliconNoise), removal of chimeric and short (< 200 bp) reads, as well as reads of Archaea, Eukaryotes, and "no hits".

[H] Reads were clustered into operational taxonomic units (OTUs) at a level of 97% sequence identity and classified based on the RDP naïve Bayesian rRNA Classifier (v2.6; Wang et al. 2007).

[I] Envrionmental samples taken at the four environmental sampling sites.

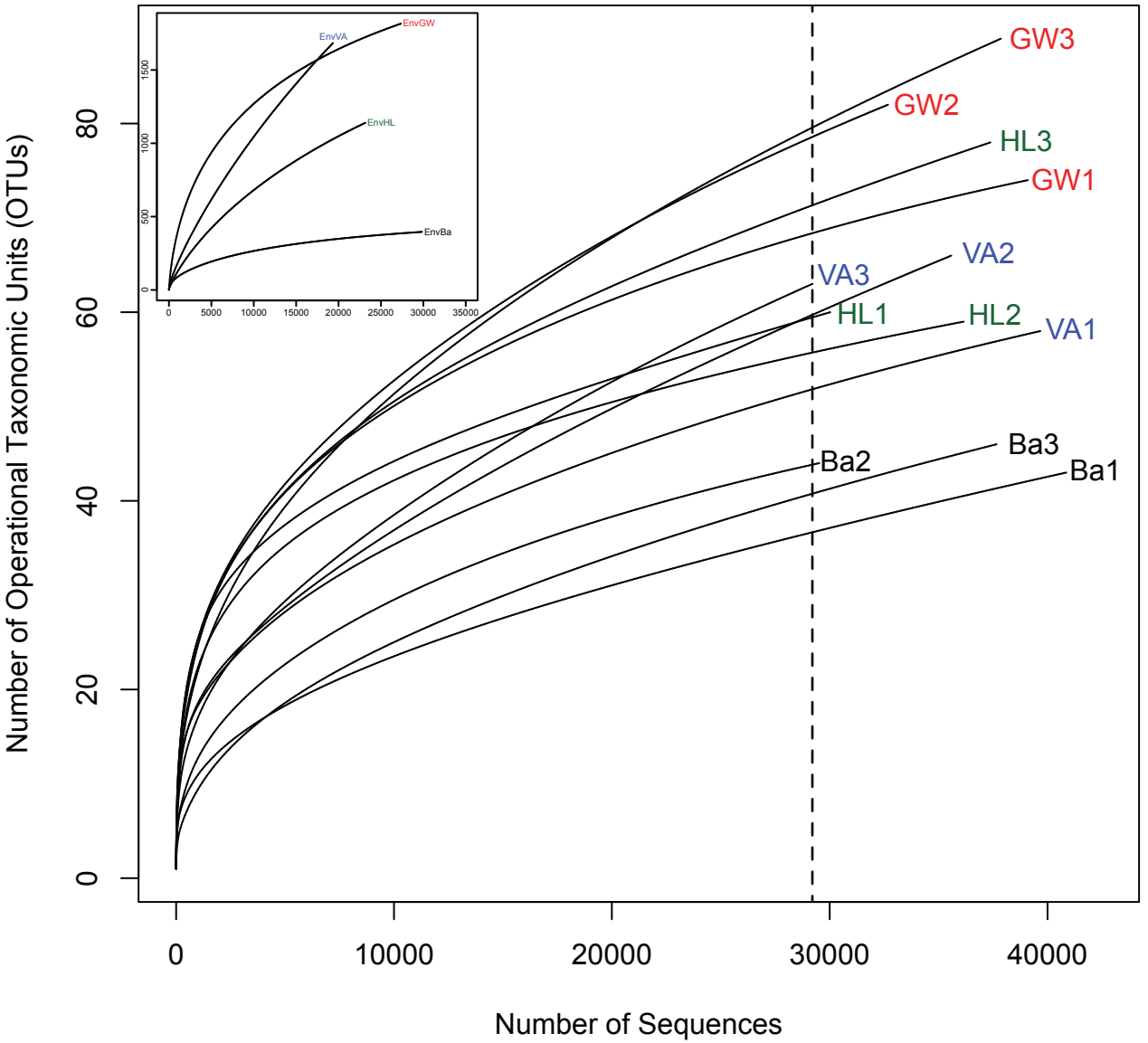
Abbreviations: Ba, Baltic; EnvBa, Environmental Baltic; GW, Groundwater; EnvGW, Environmental Groundwater; HL, Headwater Lake; EnvHL, Environmental Headwater Lake; VA, Vindelälven; EnvVA, Environmental Vindelälven.

Supplementary Figure S1



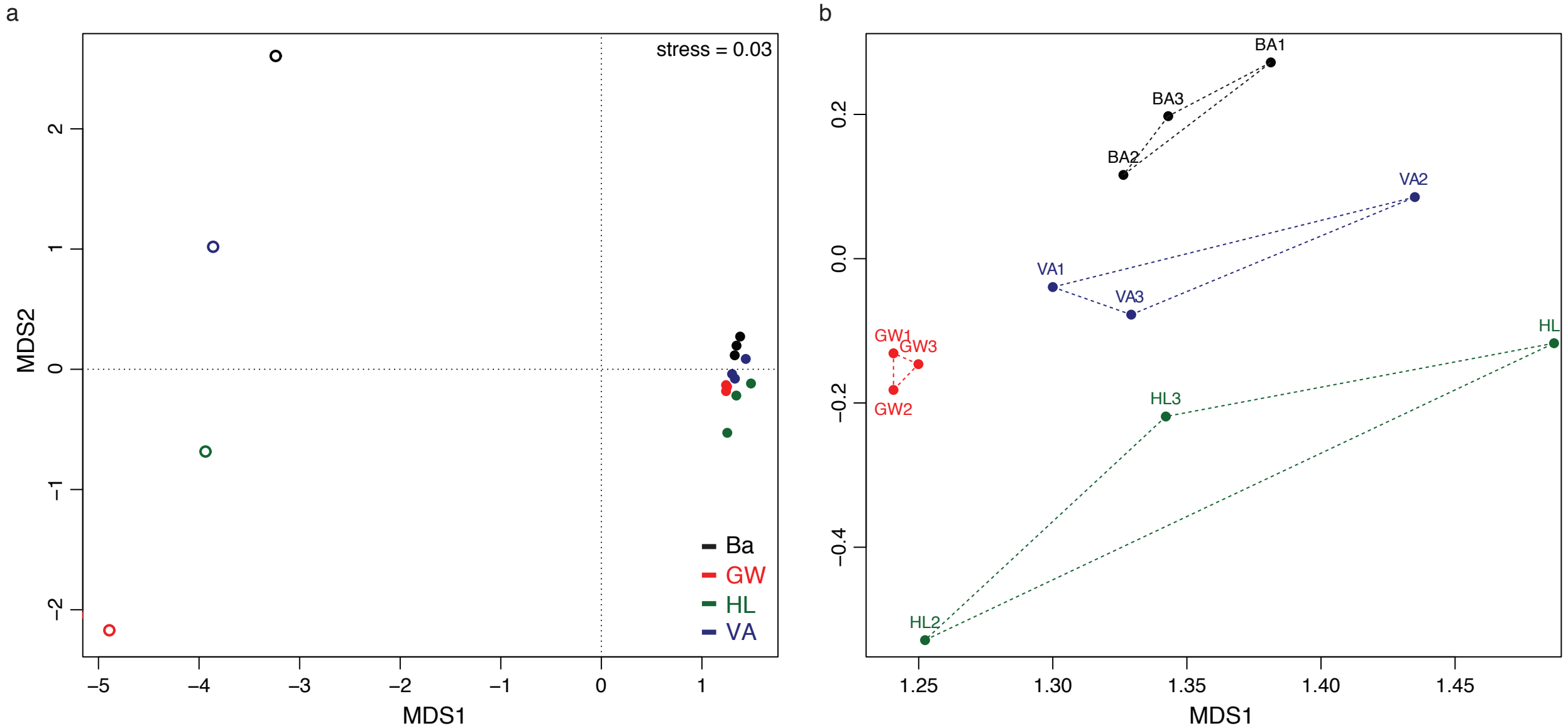
Schematic figure visualising inocula preparation, inoculation, and running the experiment. An in-detail description can be found in Supplementary Methods S1. Abbreviations are as follows: Ba, Baltic; GW, Groundwater; HL, Headwater Lake; VA, Vindelälven.

Supplementary Figure S2



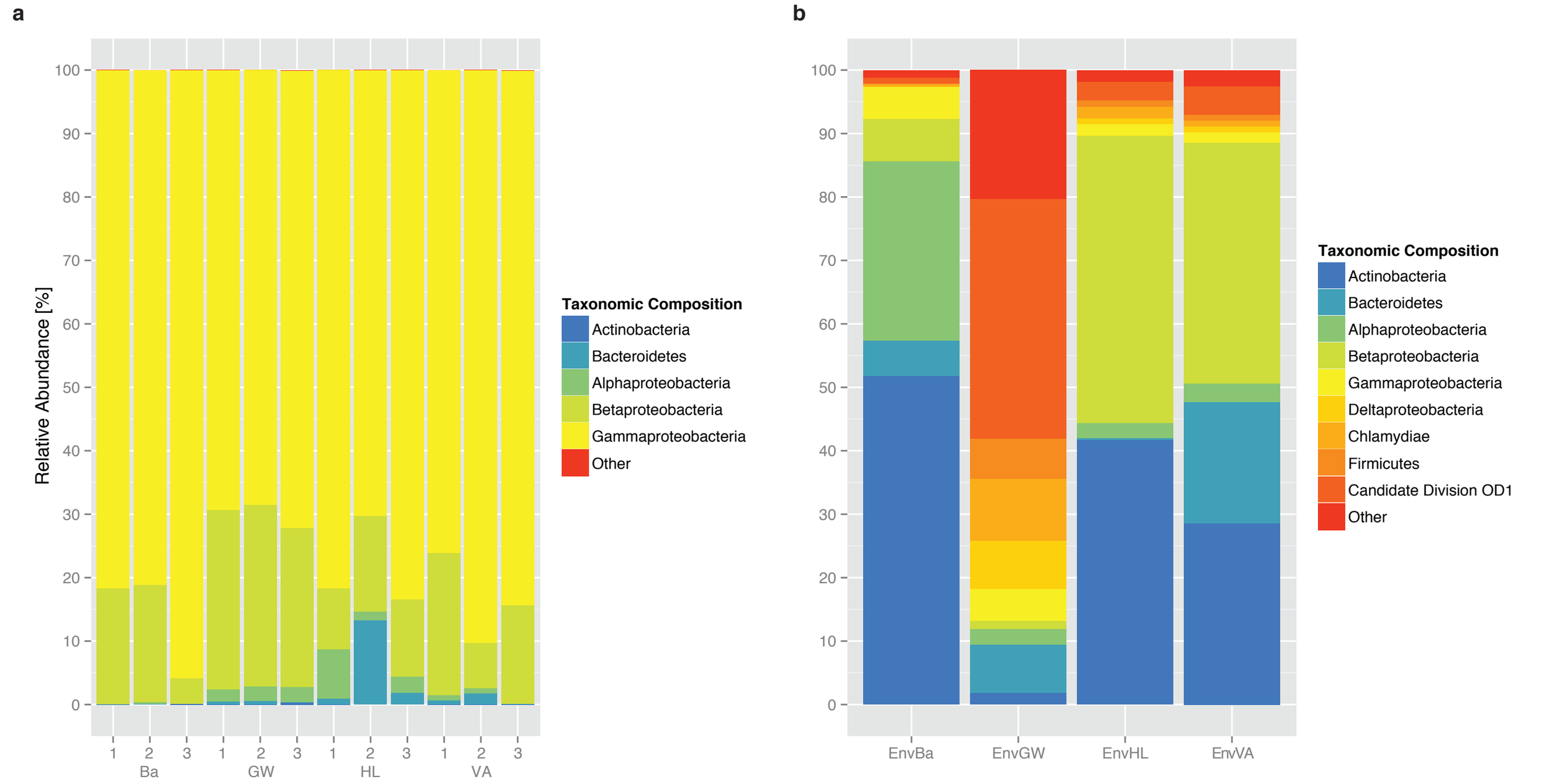
Rarefaction curves for the 12 experimental and four environmental (inset) bacterial communities, showing the cumulative number of taxa recorded as a function of sampling effort (i.e., number of pyrosequences). The dashed line visualises the lowest sampling effort (29206) for which all experimental samples were normalised during data analysis. Abbreviations are as follows: Ba, Baltic; GW, Groundwater; HL, Headwater Lake; VA, Vindelälven; EnvBa, Environmental Ba; EnvGW, Environmental GW; EnvHL, Environmental HL; EnvVA, Environmental VA.

Supplementary Figure S3



Non-metric multidimensional scaling (NMDS) representation of bacterial communities from the four natural environments (open circles) and experimental treatments (filled circles) (a). (b) depicts a close-up of only the four experimental treatments and their three replicates each (numbers represent replicates one, two, and three, and hulls were drawn to group replicates within an experimental treatment). Note that experimental replicates are represented differently in the ordination space in comparison to Figure 1 in the article. Here, NMDS ordination is based on 16 samples and sub-sampled data of 19340 reads randomly drawn from each environmental and experimental sample (compared to the 12 samples and a sub-sampling level of 29206 reads per sample on which Figure 1 in the article is based). NMDS representation was derived from pairwise Bray-Curtis distances. Abbreviations are as follows: Ba, Baltic; GW, Groundwater; HL, Headwater Lake; VA, Vindelälven.

Supplementary Figure S4



Taxonomic composition of bacterial experimental (a) and environmental (b) communities as relative abundances of each phylum. Proteobacteria are identified to class level. Abbreviations are as follows: Ba, Baltic; GW, Groundwater; HL, Headwater Lake; VA, Vindelälven; EnvBa, Environmental Ba; EnvGW, Environmental GW; EnvHL, Environmental HL; EnvVA, Environmental VA.