Electronic Supplementary Material (ESI) for Environmental Science: Water Research & Technology.
This journal is © The Royal Society of Chemistry 2016

Table S1. Summary of datasets used in this study

No	DOI number for paper	Publication year	Disinfectant residual type	Country of sampling	Sampling points	Sequencing instrument	Hypervariable region amplified	Reference number	Number of treatment plants	Average % of matches in SILVA database
1	10.1016/j.watres.2015.12.010	2015	Chlorinated	UK	POU	Illumina Miseq	V4	15	1	89.02% ± 3.79%
2	doi: 10.1016/j.watres.2014.01.049	2014	Chlorinated	UK	DWDS	Roche 454 FLX	V1-V3	28	1	71.15% ± 6.22%
3	10.1016/j.watres.2013.11.027	2014	Chlorinated/Chloraminated	USA	POU	Roche 454 FLX Titanium	V1-V2	32	15	79.69% ± 12.42%
4	10.1016/j.ecoenv.2014.07.029	2014	Chlorinated	China	DWTP outlet, POU	Roche 454 FLX Titanium	V3-V4	37	1*	61.61% ± 12.48%
5	10.1128/AEM.01892-12	2012	Chlorinated/Chloraminated	USA	DWDS	Roche 454 FLX Titanium	V4-V5	33	1	89.56% ± 3.88%
6	10.1371/journal.pone.0141087	2015	Chlorinated/Chloraminated	USA	DWTP outlet, POU	Illimina Miseq	V4	34	5	67.2% ± 8.08%
7	10.1021/acs.est.5b03521	2015	Chlorinated	China	DWTP outlet, POU	Roche 454 FLX Titanium	V1-V3	38	1*	85.28% ± 8.29%
8	10.1016/j.watres.2013.10.071	2013	Chlorinated	China	DWTP outlet	Roche 454 FLX	V1-V3	10	1	96.92%
9	10.1021/es5009467	2014	No disinfectant residual	Netherlands	POU	Roche 454 FLX Titanium	V4-V6	26	1**	63.61% ± 4.24%
10	10.1128/mBio.01135-14	2014	Chloraminated	USA	DWTP outlet, POU	Roche 454 GS-FLX	V4-V5	7	1	69.98% ± 21.13%
11	10.1111/1462-2920.12739	2015	No disinfectant residual	Netherlands	DWDS	Illumina MiSeq/Roche 454 GS-FLX	V4 and V5-V6	6	32**	45.22% ± 6.65%
12	10.1021/es502646d	2014	Chlorinated/Chloraminated	USA	DWDS	Illumina MiSeq	V4	35	1	68.93% ± 2.65%
13	10.1007/s11274-013-1321-5	2013	Chloraminated	China	DWTP outlet, POU	Roche 454 GS-FLX	V3-V5	39	1	29.91% ± 0.64%
14	10.1128/AEM.01297-15	2015	Chloraminated	Australia	DWTP outlet, DWDS	Ion Torrent	V3	43	2	90.24% ± 6.82%

*Both studies sampled at the same treatment plant

**Plant from study No.9 may have been sampled in study No. 11

Figure S1: The number of samples retained (primary Y-axes, blue squares) with increasing subsampling depths and their corresponding Good's coverage (secondary Y-axes, red squares) is shown. A subsampling depth of 500 sequences per sample was chosen as this allowed for an average Good's coverage 0.8 while retaining maximum number of samples (142 out of 145) in the analyses.

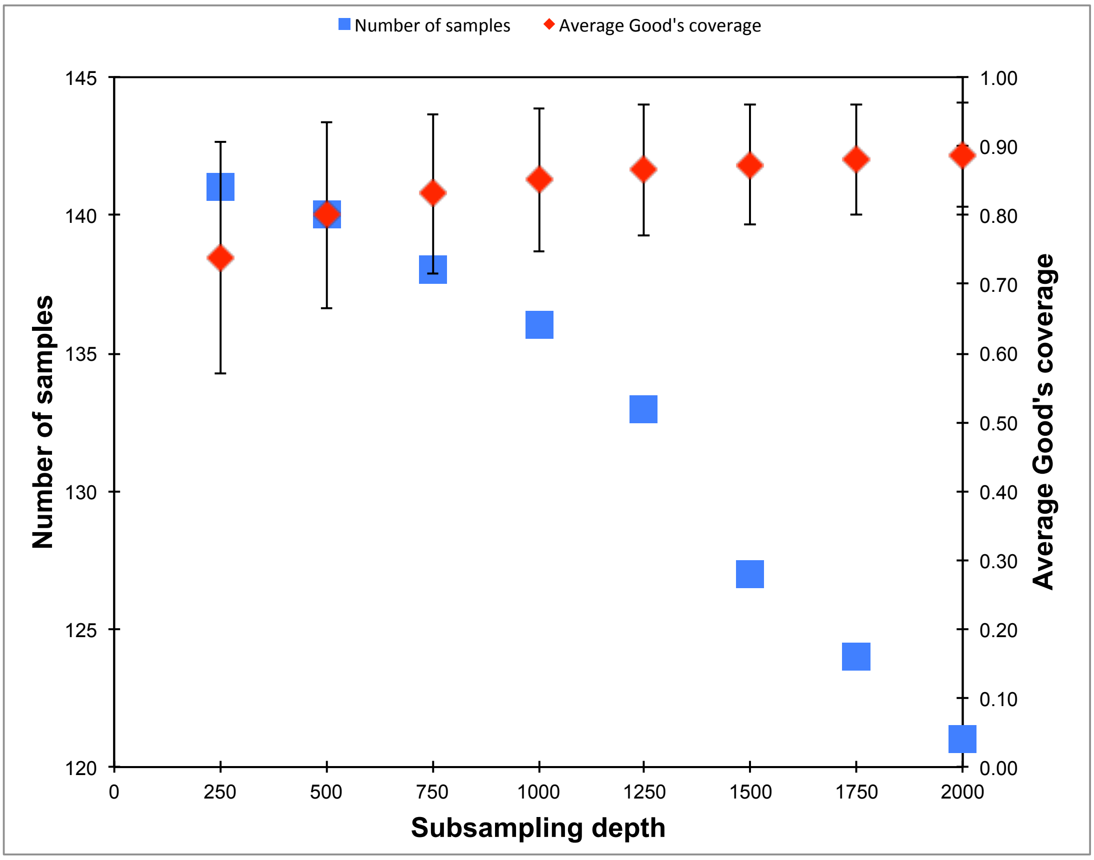

Figure S2: Workflow illustrating data collection, data processing and data analysis steps.

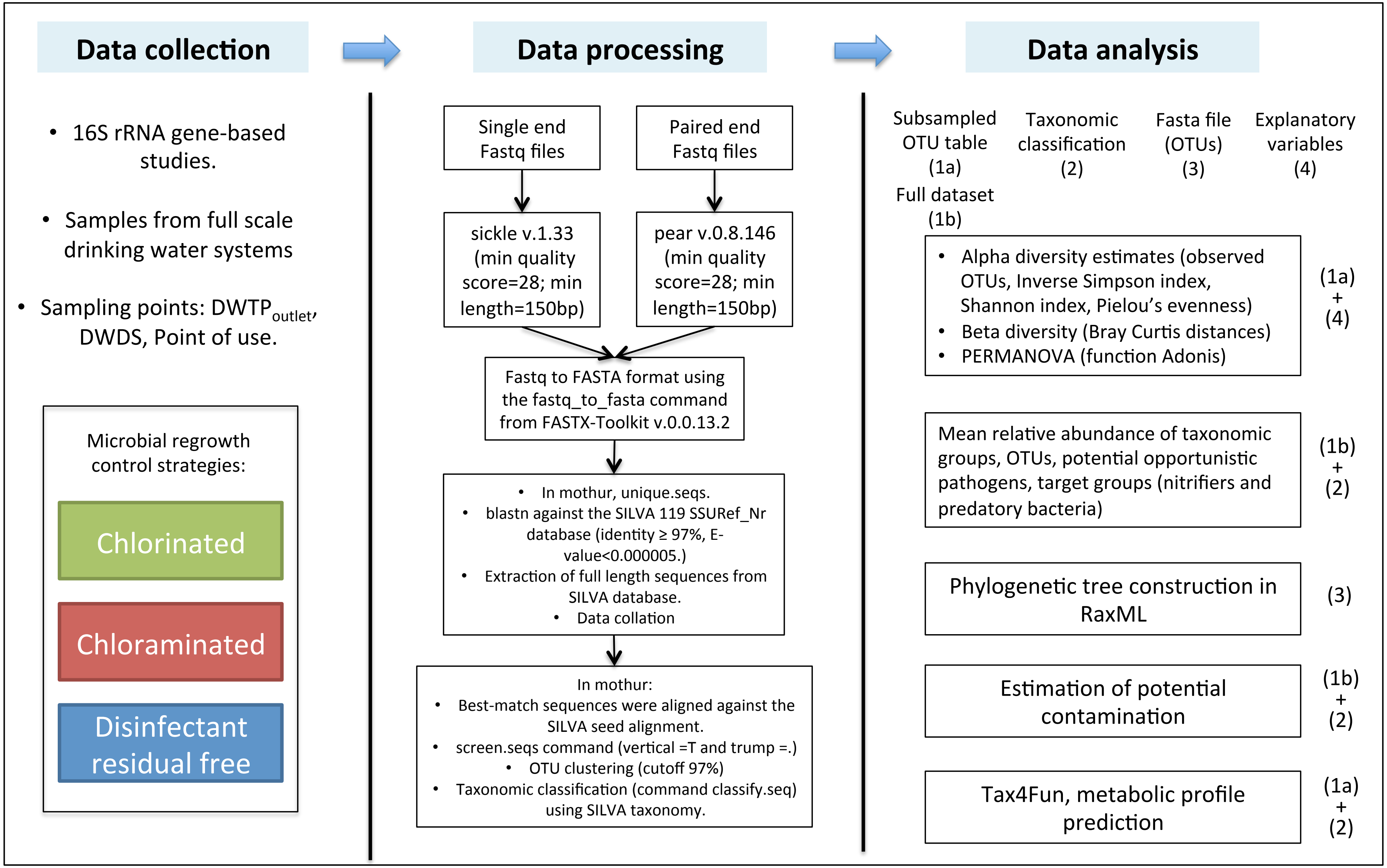

Figure S3: Heatmap of OTU abundance (OTUs with relative abundance >0.01% in the subsampled OTU table). Abundances were scaled with a Z-score transformation to improve visualization. The sampling location dendrogram was generated with Bray Curtis distances and UPGMA clustering method. Grouping information is indicated by the color legend (Chl: chlorinated, Chm: chloraminated, Drf: disinfectant residual-free).

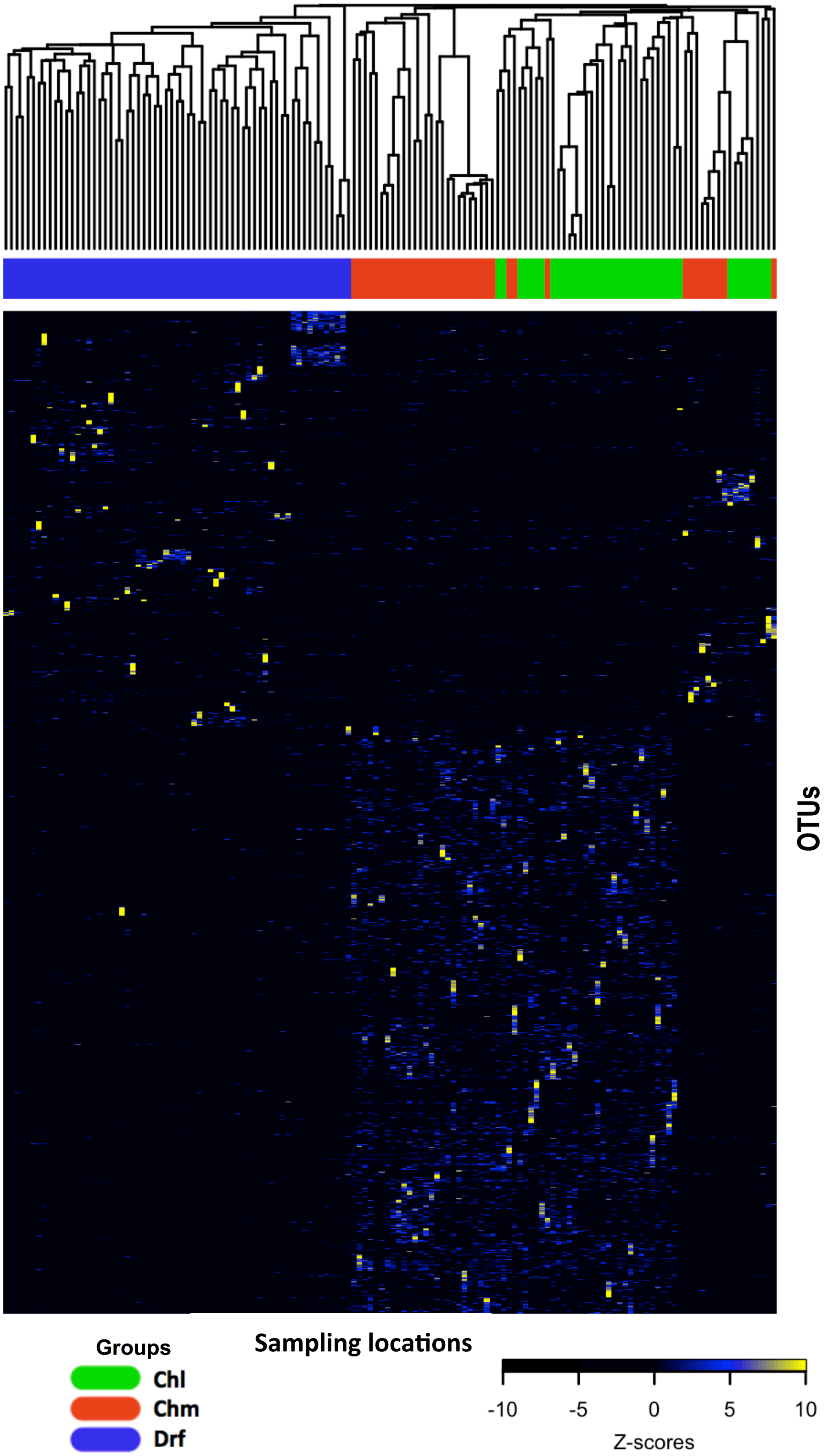

Figure S4: Dendrogram of sampling locations generated with Bray Curtis distances and UPGMA clustering method. Color legends indicate type of source water (SW: surface water; GW: groundwater; Mix: mixed source water including surface water, groundwater and desalinated seawater), and disinfection group (Chl: chlorinated, Chm: chloraminated, Drf: disinfectant residual-free). Reference numbers of each dataset are at the bottom of the plot.