Supporting information for

Bacterial community structure in the drinking water microbiome is governed by filtration processes.

Ameet J. Pinto [1], Chuanwu Xi [2], Lutgarde Raskin [1]*

[1] Department of Civil and Environmental Engineering, University of Michigan.

[2] Department of Environmental Health Sciences, University of Michigan.

* Phone number: +1-(734)-647-6920. Email address: raskin@umich.edu

Summary: 10 pages with five figures and three tables are included below.
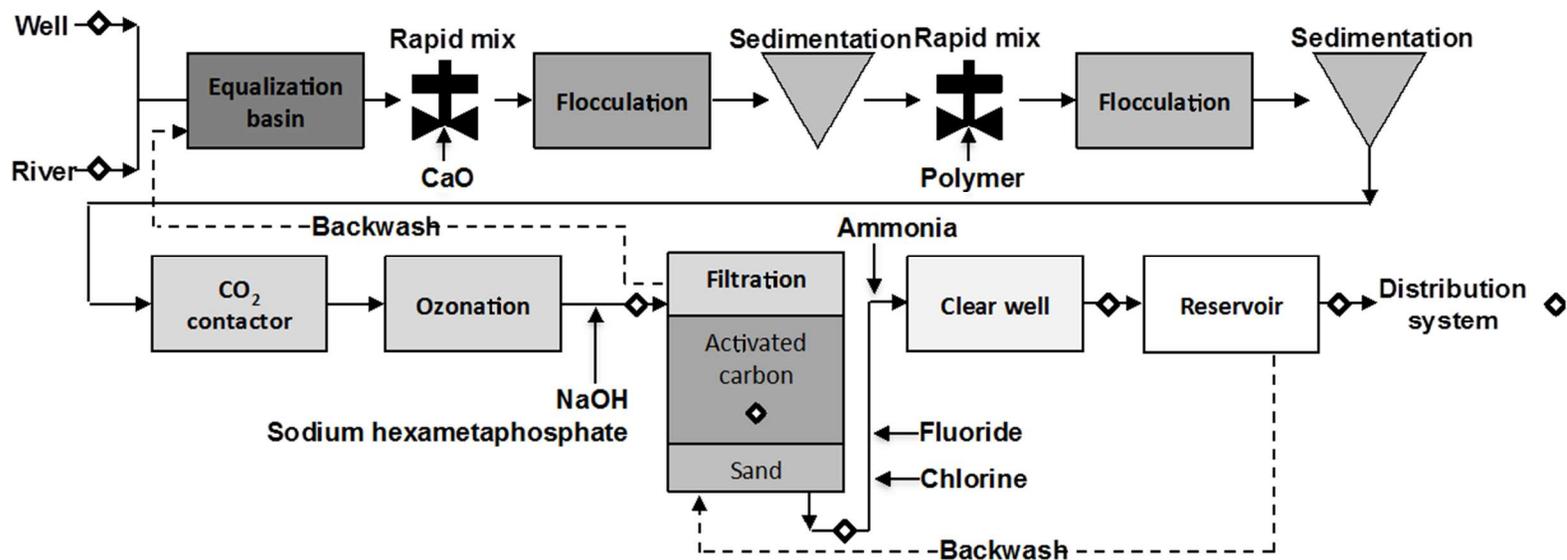
Figure S1. Schematic of the Ann Arbor Drinking Water Treatment Plant. Open diamonds mark the sampling locations. The dual media filters are backwashed with finished water and the backwash water is introduced into the equalization basin at the head of the plant. The distribution system was sampled at 13 different monitoring points.
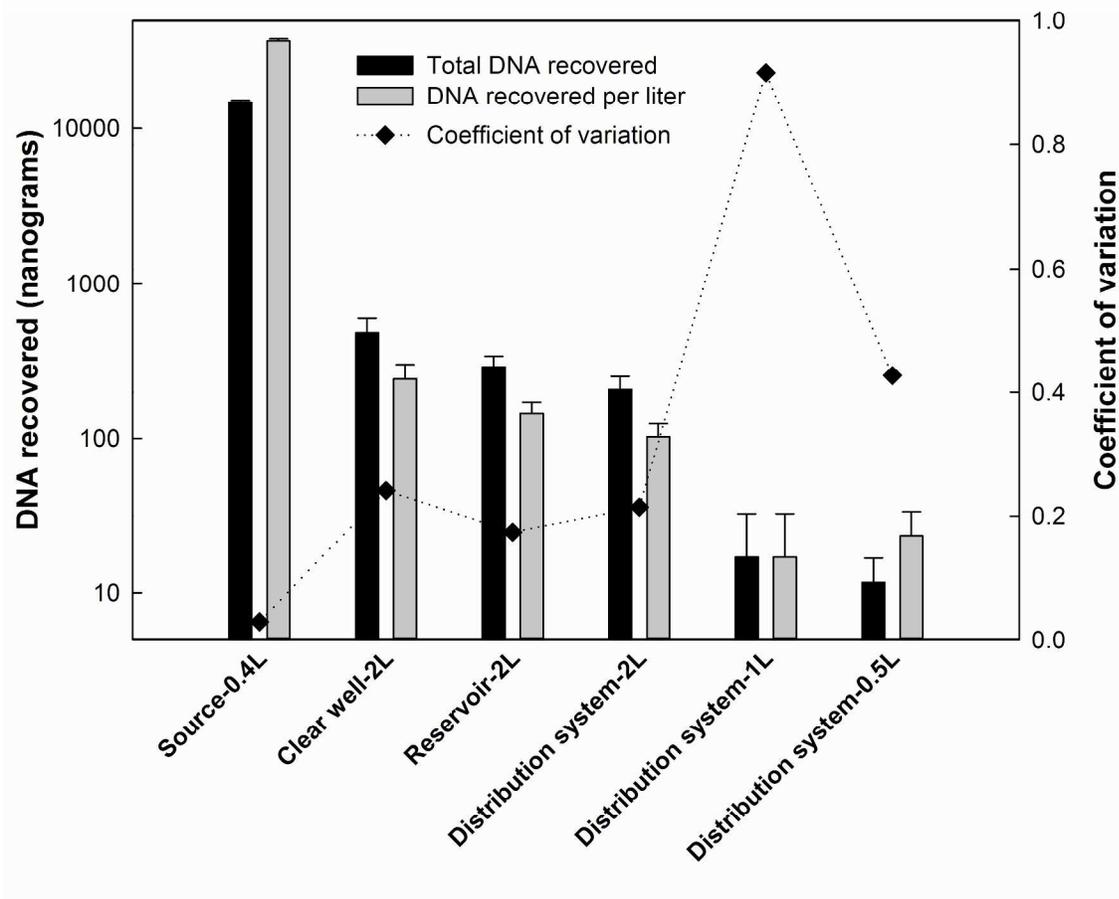
Figure S2. Amount of DNA extracted on a mass basis (black bars) and mass per volume of original sample (gray bars). Samples and their respective filtered volumes are shown on the x-axis. Error bars indicate variability between extractions from triplicate samples. Black diamonds show the coefficients of variation for the DNA recovered from triplicate membrane filters of the same sample.
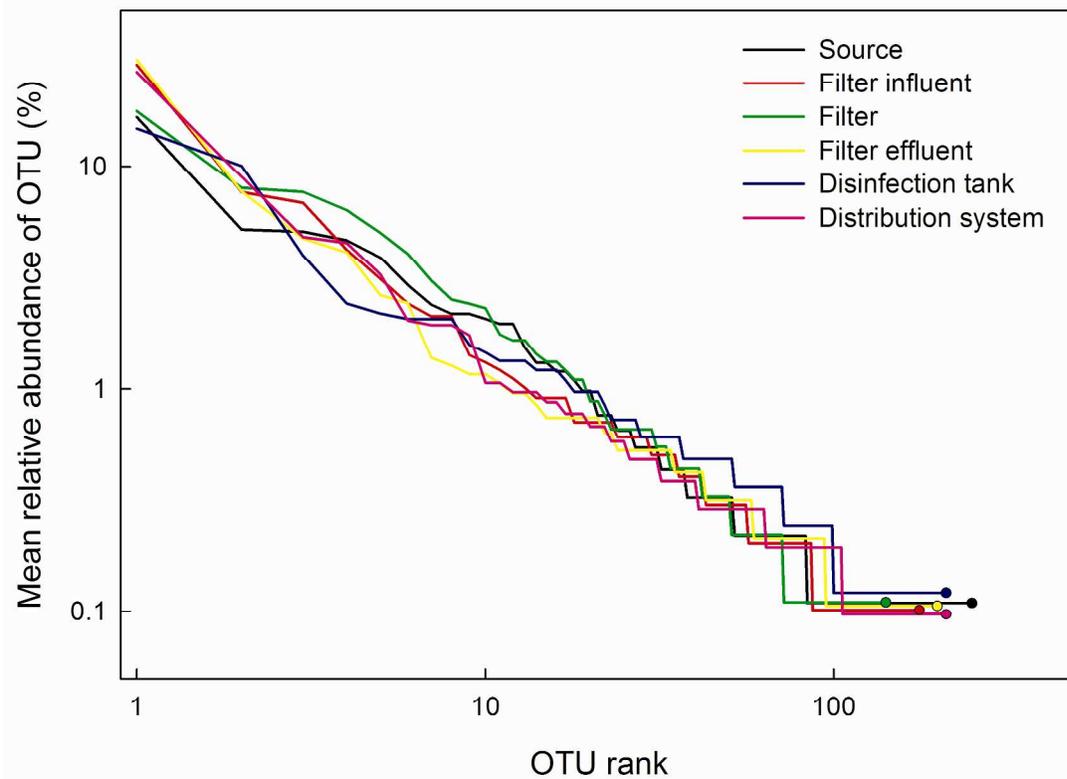
Figure S3. Rank abundance distributions (averaged over three seasons) for the OTUs detected in each sampling location. Circles show the mean relative abundance of the least abundant taxa detected in this study. The mean relative abundance of the rarest OTUs was between 0.1-0.12% for all sampling locations. Hence, deeper sequencing would only result in greater detection of rare taxa with mean relative abundance equal to or less than 0.1-0.12%.
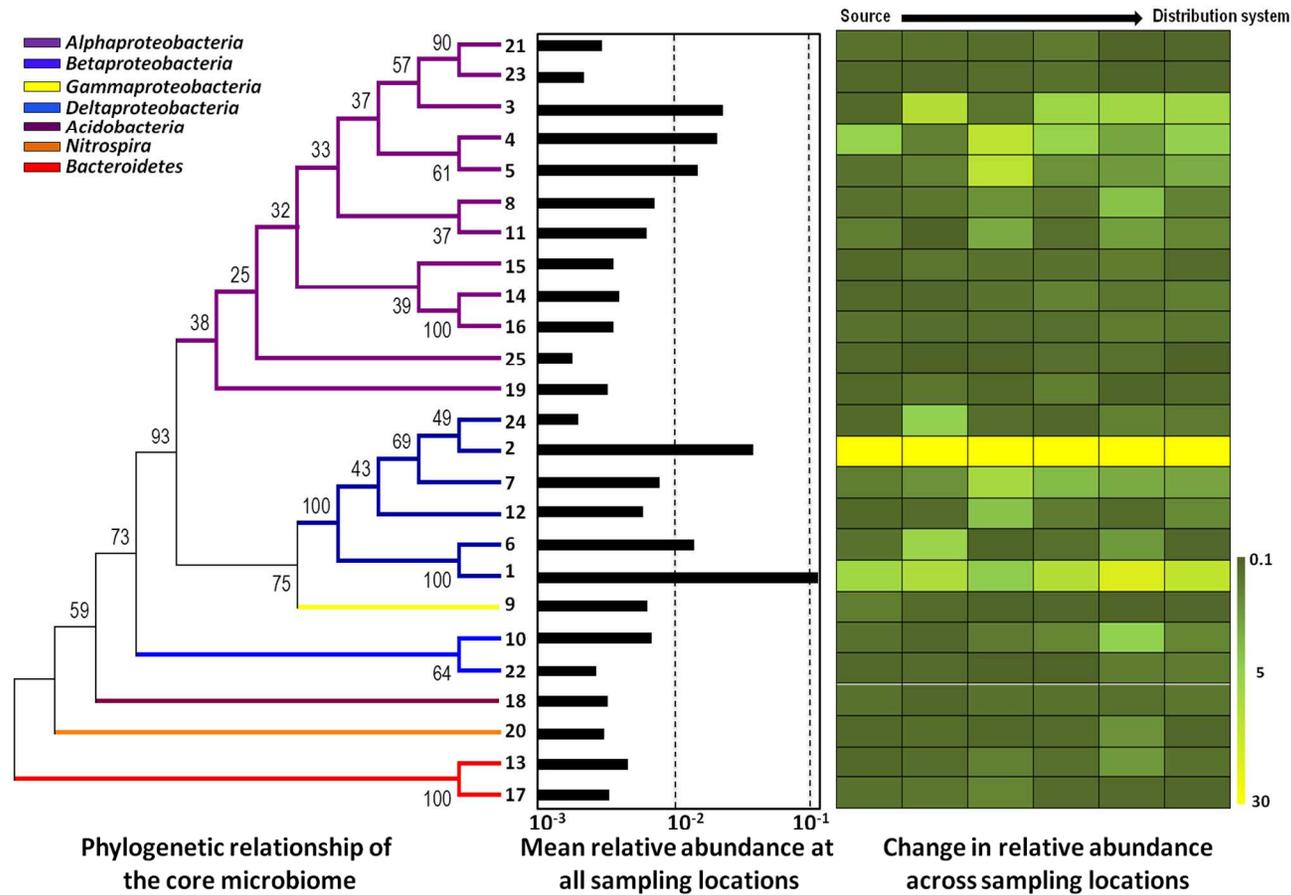
Figure S4. The left panel shows a neighbor-joining (1,000 iterations) tree highlighting the phylogenetic relationship among the 25 core OTUs. The terminal leafs of the phylogenetic tree represent these OTUs labeled in descending order of their mean relative abundance as observed across all sampling locations. The middle panel shows the mean abundance of each of the core OTUs across the six sampling locations, and the right panel represents the heat-map of changes in relative abundance at the six sampling locations.
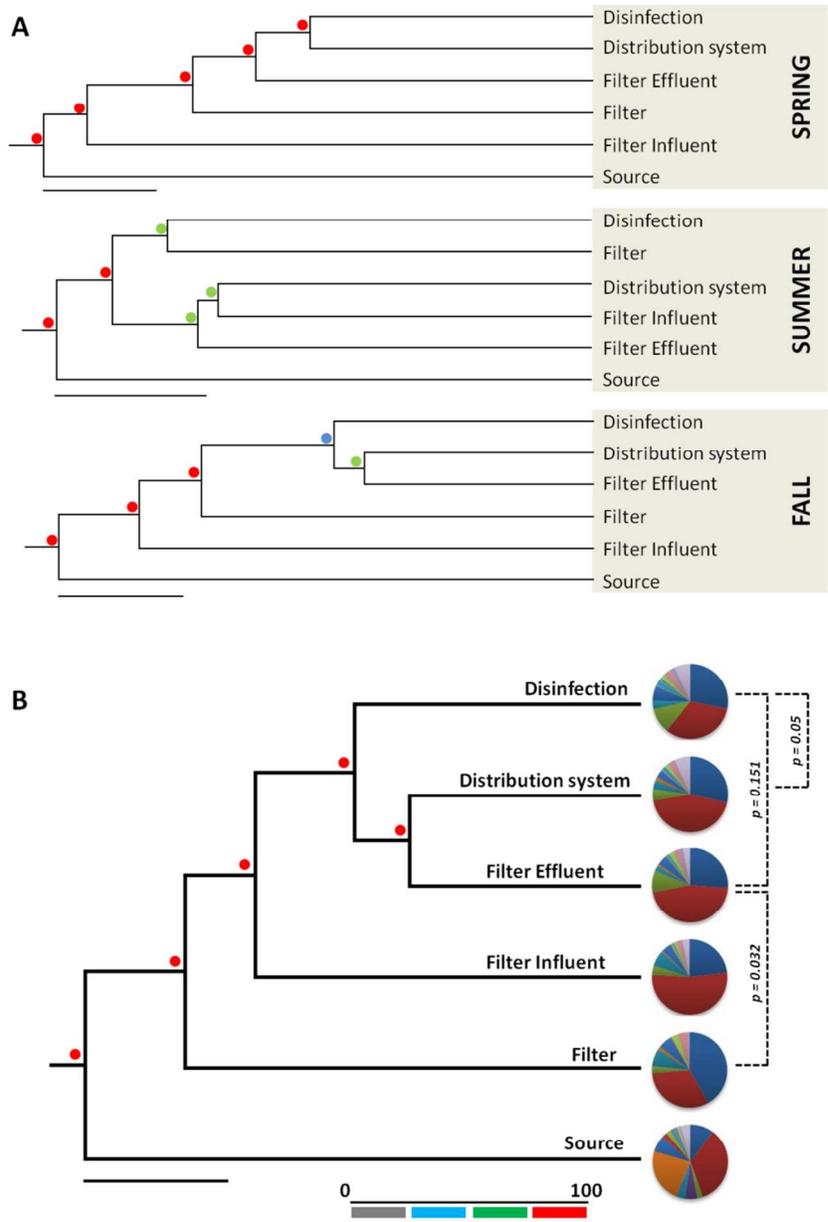
Figure S5. Jackknife clustering was performed with 1,000 iterations for sampling locations within each season (A) and for all seasons combined (B) using the WUnF metric. The node colors indicate the Jackknife fractions, i.e., the percentage of iterations for which each node was conserved. Parsimony p values are indicated in Figure S6B for pairwise comparisons for filter-filter effluent, filter effluent-disinfection tank, and disinfection tank-DWDS and show that these sampling locations had bacterial community structures that were not significantly different from each other (*p>0.001*).

Table S1. A summary of samples collected (n=120) and pooling strategy resulting in a total of 18 samples for sequencing. The left column shows the sample collected, and the right column indicates the final, pooled sampless sequenced for each season.

Collected samples							Sequenced samples
Month	Apr	Jun	Jul	Aug	Sep	Oct	
	Spring		Summer		Fall		**Season**
Location							**Location**
River	Source-Spring		Source-Summer		Source-Fall		Source
Well							
Filter Influent	Filter influent-Spring		Filter influent-Summer		Filter influent-Fall		Filter influent
Filter	Filter-Spring		Filter-Summer		Filter-Fall		Filter
Filter effluent	Filter effluent-Spring		Filter effluent-Summer		Filter effluent-Fall		Filter effluent
Clear well	Disinfection-Spring		Disinfection-Summer		Disinfection-Fall		Disinfection
Reservoir							
Distribution system	Distribution system-Spring		Distribution system-Summer		Distribution system-Fall		Distribution system

Table S2. Summary of quality-filtered and chimera free reads for each seasonal sample.

Sample	Season	Quality-score filtered reads	Chimera free reads
Source	Spring	243	238
	Summer	358	355
	Fall	329	326
Filter influent	Spring	207	204
	Summer	381	376
	Fall	417	410
Filter	Spring	333	327
	Summer	428	418
	Fall	171	165
Filter effluent	Spring	322	319
	Summer	334	331
	Fall	302	295
Disinfection	Spring	316	312
	Summer	263	257
	Fall	263	256
Distribution system	Spring	239	236
	Summer	351	344
	Fall	460	454

Table S3: A summary of water quality parameters for the five bulk water sampling locations across the three seasons. Data are shown as averages of replicate analyses with standard deviations in parentheses. ND: Non-detectable. NQ: Detectable but not quantifiable. NA: Not applicable.

	Spring	Summer	Fall
Total organic carbon (mg/l)			
Source	10.4 (4.8)	5.1 (3.4)	6.1 (3.5)
Filter influent	6.7 (0.5)	4.6 (1.6)	4.2 (1.8)
Filter effluent	5.1 (0.2)	4.2 (0.8)	3.4 (1.4)
Disinfection tank	5.5 (0.7)	3.7 (0.2)	3.3 (1.4)
Distribution system	6.4 (2.6)	3.5 (0.6)	3.7 (0.6)
pH			
Source	7.6 (0.4)	7.8 (0.4)	7.9 (0.4)
Filter influent	9.2 (0.2)	9.5 (0)	9.4 (0.2)
Filter effluent	9.0 (0.2)	9.3 (0.1)	9.4 (0.2)
Disinfection tank	9.1 (0.1)	9.3 (0.1)	9.3 (0.2)
Distribution system	9.1 (0.1)	9.3 (0.1)	9.3 (0.1)
Ammonia (µg-N/l)			
Source	59 (85)	63 (39)	8 (16)
Filter influent	123 (14)	134 (17)	51 (2)
Filter effluent	17 (10)	11 (3)	ND
Disinfection tank	387 (91)	304 (58)	305 (79)
Distribution system	474 (294)	334 (68)	253 (68)
Nitrite (µg-N/l) *NQ < 10 µg-N/l			
Source	249 (490)	NQ*	ND
Filter influent	NQ*	NQ*	ND
Filter effluent	50 (50)	ND	NQ*
Disinfection tank	NQ*	NQ*	ND
Distribution system	10 (9)	45 (9)	NQ*
Nitrate (µg-N/l) *NQ < 20 µg-N/l			
Source	NQ*	230 (160)	191 (286)
Filter influent	NQ*	258 (258)	494 (70)
Filter effluent	NQ*	209 (56)	709 (312)
Disinfection tank	NQ*	197 (163)	590 (481)
Distribution system	NQ*	250 (224)	264 (224)

	Phosphate (μg-N/l) *NQ < 50 μg-P/l		
Source	ND	ND	ND
Filter influent	77 (42)	39 (30)	59 (59)
Filter effluent	57 (4)	67 (12)	92 (71)
Disinfection tank	53 (45)	96 (40)	936 (36)
Distribution system	153 (41)	148 (60)	99 (60)
	Temperature (°C)		
Source	16 (3)	21 (4)	17 (3)
Filter influent	16 (3)	23 (1)	15 (3)
Filter effluent	17 (2)	23 (0)	15 (4)
Disinfection tank	18 (3)	24 (1)	17 (3)
Distribution system	18 (2)	24 (2)	21 (2)
	Sulfate (mg/l)		
Source	21 (19)	21 (17)	51 (42)
Filter influent	18 (4)	10 (3)	28 (6)
Filter effluent	17 (3)	9 (3)	25 (6)
Disinfection tank	17 (3)	10 (3)	24 (8)
Distribution system	18 (4)	10 (2)	23 (2)
	Chloride (mg/l)		
Source	93 (53)	95 (48)	115 (36)
Filter influent	118 (13)	43 (20)	118 (24)
Filter effluent	116 (7)	51 (18)	96 (28)
Disinfection tank	116 (7)	51 (18)	96 (28)
Distribution system	117 (9)	51 (8)	117 (8)
	Chlorine (mg Cl_2/l)		
Source	NA	NA	NA
Filter influent	NA	NA	NA
Filter effluent	NA	NA	NA
Disinfection tank	3.2 (0.2)	3.2 (0.2)	3.2 (0.2)
Distribution system	2.6 (0.3)	2.5 (0.3)	2.6 (0.3)