SUPPLEMENTARY MATERIAL

Characterization of biofilms developing on hospital shower hoses and implications for nosocomial infections

Maria J. Soto-Giron, Luis M. Rodriguez-R, Chengwei Luo, Michael Elk, Hodon Ryu, Jill Hoelle, Jorge W. Santo Domingo, Konstantinos T. Konstantinidis

Table S1. General statistics of five biofilm metagenomes from the shower hoses

	Sample				
	M1	**M3**	**M4**	**M5**	**M6**
Assembly					
No. of reads (millions)	18.01	14.96	29.48	37.03	23.43
No. of contigs*	15,323	2,878	9,394	18,311	21,910
N50 of contigs (b)	3,242	19,478	6,469	7,483	6,226
Average size (Mb)	34.86	15.95	25.23	58.99	51.11
No. of predicted genes	45,366	17,676	31,656	70,196	64,737
Taxonomy classification (%)**					
Mycobacterium	45.3	58.5	52.5	26.3	30
Erythrobacter	13.8	6.71	9.94	9.6	5.81
Sphingomonas	3.87	6.34	10.7	6.41	6.24
Novosphingobium	3.83	5.99	5.4	3.65	2.06
Bradyrhizobium	2.66	0.22	0.18	15.45	8.24
Citromicrobium	2.09	2.02	1.73	1.47	1.02
Sphingobium	1.41	6	8.98	2.28	2.21

*Contigs longer than 5000 bases were counted.

** Relative abundance of taxa at genus level based on annotated proteins recovered in each sample and classified by MyTaxa.

Table S2. Statistics of the 94 isolate genomes recovered from the shower hose biofilms.

Isolate ID	OTU designation	Identity (%)	GG content (%)	Assembly size (Mb)	No of contigs	No of protein -coding genes
CCH3-A3	*Blastomonas*	100	64.24	4.34	162	4,238
CCH9-F3	*Blastomonas*	100	64.26	4.18	154	4,078
CCH6-A6	*Blastomonas*	100	64.16	4.39	169	4,301
CCH8-E1	*Blastomonas*	100	64.33	4.27	145	4,167
CCH10-E1	*Blastomonas*	100	64.18	4.35	230	4,292
CCH9-A1	*Blastomonas*	100	64.19	4.20	95	4,083
CCH13-E1	*Blastomonas*	97.4	64.2	4.36	159	4,265
CCH2-E1	*Blastomonas*	100	64.24	4.40	164	4,344
CCH4-A2	*Acidovorax delafieldii*	100	64.29	4.28	192	4,155
CCH5-A5	*Sphingomonas*	98.2	66.88	3.55	81	3,410
CCH3-E3	*Blastomonas*	100	64.27	4.35	153	4,256
CCH8-A3	*Blastomonas*	100	64.11	4.57	217	4,567
CCH18-B1	*Sphingomonas*	99.7	64.23	4.32	173	4,198
CCH7-E1	*Blastomonas*	100	64.18	4.19	111	4,067
CCH15-G10	*Blastomonas*	100	64.22	4.35	192	4,300
CCH6-E2	*Blastomonas*	100	64.07	4.49	1088	5,127
CCH13-A3	*Blastomonas*	100	64.23	4.33	148	4,197
CCH21-G11	*Sphingomonas*	99.8	66.73	3.65	59	3,483
CCH5-E3	*Blastomonas*	100	64.31	4.3	189	4,254
CCH4-D12	*Blastomonas*	100	64.29	4.28	173	4,148
CCH12-A3	*Novosphingobium*	100	63.27	5.48	328	5,404
CCH3-A5	*Rhizobiales*	99.7	66.5	5.10	208	5,042
CCH5-F6	*Bradyrhizobium*	99.7	64.12	8.15	206	7,766
CCH12-B7	*Dermacoccus*	99.8	67.66	4.75	279	4,559
CCH16-	*Sphingomonas*		66.85	3.56	74	

B10		99.8				3,386
CCH9-G4	*Pseudoxanthomonas mexicana*	99.8	66.55	3.98	105	3,718
CCH10-E5	*Rhizobiales*	99.6	67.66	5.32	252	5,333
CCH3-G3	*Acidovorax delafieldii*	100	64.78	5.79	634	5,698
CCH6-A11	*Sphingomonas*	100	68.67	3.87	261	3,839
CCH6-A12	*Neisseria perflava*	100	41.73	2.18	105	2,121
CCH10-B3	*Sphingomonas*	99.8	65.15	3.83	98	3,779
CCH9-A3	*Rhizobiales*	99.7	67.34	5.69	281	5,658
CCH4-E10	*Chryseobacterium*	100	36.66	4.41	195	4,120
CCH4-A6	*Bradyrhizobiaceae*	99.8	60.89	5.28	159	5,103
CCH10-C7	*Bradyrhizobiaceae*	99.8	60.72	5.60	201	5,427
CCH5-A3	*Blastomonas*	100	64.26	4.43	132	4,334
CCH19-E1	*Porphyrobacter donghaensis*	100	64.32	3.96	237	3,914
CCH12-A10	*Comamonadaceae*	100	63.34	4.87	470	4,791
CCH7-A10	*Porphyrobacter donghaensis*	100	51.11	2.52	564	2,521
CCH1-A6	*Blastomonas*	100	64.37	4.80	168	4,740
CCH10-H12	*Neisseria perflava*	100	51.12	2.52	561	2,495
CCH4-C5	*Comamonadaceae*	100	63.19	4.61	456	4,509
CCH4-E1	*Caulobacter*	99.8	66.39	4.12	216	4,128
CCH5-D3	*Streptococcus*	99.8	39.94	2.23	102	2,099
CCH5-D2	*Methylobacterium*	99.8	71.09	6.09	420	6,144
CCH1-B1	*Bradyrhizobiacea*	99.7	67.36	5.81	302	5,764
CCH8-H5	*Streptococcus*	100	39.8	2.20	105	2,105
CCH9-E1	*Caulobacter*	99.8	67.66	4.73	260	4,521
CCH3-A4	*Porphyrobacter donghaensis*	100	66.58	4.16	236	4,162
CCH20-B6	*Sphingomonas*	99.8	66.62	4.08	100	3,963
CCH5-E12	*Caulobacter*	99.5	66.74	4.95	189	4,741
CCH8-G7	*Streptococcus*	99.3	40.43	2.24	81	2,215
CCH9-H8	*Sphingomonas*	99.2	66.06	4.57	220	4,529

CCH7-A1	*Porphyrobacter*	100	64.23	4.41	182	4,351
CCH5-B3	*Xylophilus ampelinus*	100	68.27	6.21	320	6,056
CCH9-H3	*Phenylobacterium*	99.2	69.48	5.62	149	5,528
CCH1-A1	*Porphyrobacter donghaensis*	99.8	66.77	4.21	215	4,178
CCH6-E1	*Porphyrobacter donghaensis*	100	66.45	4.28	275	5,127
CCH5-A9	*Bosea*	99.2	67.39	5.65	278	5,628
CCH9-E2	*Sphingomonas*	99.2	66.09	4.31	203	4,260
CCH8-C6	*Streptococcus*	100	43.37	2.21	107	2,205
CCH9-F2	*Sphingomonas*	100	68.51	4.08	261	4,082
CCH6-A4	*Rhizobiales*	99.5	66.82	6.25	380	6,339
CCH8-A2	*Porphyrobacter donghaensis*	100	66.47	4.33	274	4,449
CCH8-D1	*Rhizobiales*	99.7	66.78	6.16	295	6,160
CCH17-B8	*Porphyrobacter donghaensis*	100	66.58	4.17	223	4,136
CCH2-D9	*Dermacoccus*	99.7	69.19	3.01	58	2,733
CCH12-G6	*Porphyrobacter donghaensis*	100	66.92	6.13	454	6,291
CCH11-D2	*Rhizobiales*	99.7	66.91	6.12	340	6,181
CCH5-D1	*Microbacterium*	100	68.25	4.01	110	3,918
CCH2-A4	*Rhizobiales*	99.7	66.86	6.17	350	6,246
CCH11-B1	*Sphingobium*	99.5	64	4.77	212	4,713
CCH6-A1	*Porphyrobacter donghaensis*	100	66.64	4.39	246	4,384
CCH10-A2	*Mycobacterium mucogenicum*	100	66.8	6.72	531	6,883
CCH11-A4	*Blastomonas*	100	64.18	4.23	152	4,122
CCH3-E2	*Micrococcus luteus*	100	73.18	2.51	237	2,369
CCH19-C6	*Sphingomonas*	99.8	66.7	3.93	109	3,819
CCH12-B4	*Phenylobacterium*	99.7	69.56	5.55	199	5,487
CCH12-C2	*Erythrobacteraceae*	99.5	63.86	4.18	124	4,053
CCH15-F11	*Sphingomonas*	99.8	66.77	4.35	109	4,236
CCH5-D11	*Sphingomonas*	99.1	65.55	4.44	116	4,235

CCH18-H6	*Sphingomonas*	99.8	66.66	4.22	107	4,108
CCH5-H10	*Rhodospirillaceae*	99.7	66.24	6.05	240	5,941
CCH2-A2	*Blastomonas*	100	64.26	4.41	180	4,357
CCH12-A2	*Mycobacterium mucogenicum*	100	66.84	6.65	370	6,653
CCH7-E3	*Porphyrobacter donghaensis*	100	66.85	4.10	212	4,064
CCH7-A2	*Blastomonas*	100	67.36	6.02	279	5,959
CCH15-A1	*Sphingomonas*	99.2	67.87	5.11	3796	7,882
CCH13-B11	*Sphingomonas*	99.7	66.76	3.89	64	3,766
CCH5-A1	*Porphyrobacter donghaensis*	99.7	66.4	4.33	241	4,309
CCH6-D9	*Corynebacterium durum*	99.3	49.23	4.69	3475	6,946
CCH7-B2	*Sphingomonas*	99.8	66.77	9.82	5483	10,000
CCH15-E2	*Porphyrobacter donghaensis*	100	57.23	6.21	251	6,054
CCH12-A4	*Blastomonas*	100	63.77	9.12	592	5,487

Table S3. General statistics of the binned populations recovered from the shower hose biofilm metagenomes.

	Binned population genomes	
	***Blastomonas* sp.**	***Mycobacterium* sp.**
Average sequencing depth	41.40	115.9
Size of the genome (Mb)	7.36	6.4
Completeness (%)	94.9	100
Contamination (%)	39.3	2.0
GC content (%)	7.36	6.4
Average nucleotide identity (%)	84.2	85.9
Virulence-associated genes*	0	15
Antibiotic resistance profile	Aminoglycoside, polymyxin, kanamycin, macrolide, bacitracin	Fluoroquinolone, penicillin, cephalosporin, gentamicin, Netilmicin

* BLASTp searches against the Virulence Factors of Pathogenic Bacteria (1) and PATRIC databases (2).

Table S4. Functional categories and relative abundance of protein sequences recovered from the biofilm metagenomes (assembled contigs) with similarities to the annotated proteins at UniProt database. Relative abundance was calculated as the number of metagenomic proteins annotated to a specific function divided by the total number of annotated proteins in the respective sample.

UniProt id	Name	Gene Ontology	Sample				
			M1	M3	M4	M5	M6
Q50615	PE-PGRS family protein PE_PGRS33	Possible role in virulence and antigenic variation	0.23	0.01	0.48	0.21	0.18
P50360	Protein y4hP	Transposase activity	0.16	0.42	0.1	0.12	0.13
Q79FV4	pyridoxal phosphate-dependent protein	Catalytic activity	0.12	0.17	0.15	0.09	0.09
P0A690	PE-PGRS family protein PE_PGRS46	Virulence and antigenic variation	0.12	0.01	0.22	0.15	0.07
Q10637	PE-PGRS family protein PE_PGRS24	Virulence and antigenic variation	0.12	0.27	0.24	0.12	0.11
P35883	Transposase for insertion element IS6120	Transposase activity	0.12	0.21	0.15	0.01	0.09
O86034	D-beta-hydroxybutyrate dehydrogenase	Metabolic process	0.07	0.14	0.05	0.09	0.11
P55501	Uncharacterized protein y4jA/y4nE/y4sE	Transposase activity	0.06	0.14	0.03	0.07	0.06
P08080	Transposase for insertion element	Transposase activity	0.04	0.14	0.01	0.12	0.14
A5TY80	Insertion element IS6110 protein	Transposase activity	0.06	0.14	0.06	0.02	0.04
Q2G6U3	Protein translocase subunit SecA	Protein transport	0.03	0.14	0.05	0.09	0.07
P17985	Insertion element ISR1 protein A3	Transposase activity	0.04	0.07	0.01	0.13	0.06
P38054	Cation efflux system protein CusA	Resistance to copper and silver	0.06	0.13	0.12	0.09	0.08
A1KQG0	Phthioceranic/hydroxyphthioceranic acid synthase	Drug target	0.2	0.01	0.12	0.1	0.07
P72003	Serine/threonine-protein kinase PknF	Stress response	0.16	0.01	0.2	0.01	0.09
O65934	ABC transporter ATP-binding/permease protein Rv1747	Required for virulence	0.16	0.01	0.14	0.01	0.21
P9WPS6.	Probable cation-transporting ATPase G	Required for virulence	0.14	0.01	0.01	0.01	0.01

1							
P96218	Glutamate synthase [NADPH]	Biosynthesis	0.14	0.1	0.06	0.01	0.06
P60230	Transposase for insertion element IS1081	Transposition	0.13	0.13	0.13	0.11	0.11
Q02251	Mycocerosic acid synthase	Lipid metabolism	0.13	0.01	0.05	0.01	0.05
P9WPS2.1	Probable copper-exporting P-type ATPase V	Required for virulence	0.13	0.01	0.01	0.01	0.01
O53735	Putative membrane protein mmpL4	Growth	0.13	0.07	0.12	0.01	0.01
O53303	Putative alcohol dehydrogenase D	Stress response	0.12	0.2	0.13	0.01	0.01
Q57307	Cholesterol oxidase	Required for virulence	0.12	0.07	0.09	0.01	0.01
Q2G480	Phosphoenolpyruvate carboxykinase	Biofilm formation	0.043	0.1	0.12	0.01	0.01
P55390	Probable cold shock protein y4cH	Regulation of transcription	0.01	0.01	0.01	0.08	0.11
Q5NRH4	Glutamine-fructose-6-phosphate aminotransferase	Biosynthetic process	0.01	0.001	0.09	0.08	0.07
	Disinfectant mechanisms						
Q9KU26	Extracellular polymeric substance	Signal transducer activity	0.086	0.03	0.03	0.14	0.084
P37578	60 kDa chaperonin GroEL1	Protein refolding	0.302	0.17	0.33	0.32	0.253
L8F435	GDP-mannose 6-dehydrogenase AlgD	Polysaccharide biosynthetic process	0.014	0.42	0.01	0.01	0.01
R4R145	Alginate biosynthesis protein AlgA	Polysaccharide biosynthetic process	0.000	0.01	0.02	0.01	0.011
P04425	Glutathione synthetase gshB	Response to oxidative stress	0.043	0.07	0.1	0.1	0.095
Q73VT8	Putative glutathione reductase gorA	reductase activity	0.014	0.03	0.02	0.01	0.011
A0R692	DNA protection during starvation protein	Response to stress	0.331	0.001	0.38	0.42	0.242
P52214	Thioredoxin reductase trxB	Oxidoreductase activity	0.475	0.56	0.58	0.5	0.442
Q51506	Redox-sensitive transcriptional activator SoxR	Response to oxidative stress	0.043	0.07	0.03	0.04	0.032

P9WGE7	Superoxide dismutase sodA	Response to oxidative stress	0.029	0.01	0.01	0.07	0.053
Q1I657	RNA polymerase sigma factor RpoS	Response to stress	0.000	0.07	0.17	0.12	0.179
P9WGE9	Superoxide dismutase SodC	Response to oxidative stress	0.058	0.07	0.09	0.06	0.084
P96273	Exodeoxyribonuclease III protein XthA	DNA repair	0.000	0.14	0.12	0.18	0.168
A0PSD2	Hydrogen peroxide-inducible genes activator, OxyR	Binding transcription factor activity	0.173	0.1	0.15	0.21	0.116
A0QYP1	Catalase-peroxidase katG	Response to oxidative stress	0.274	0.21	0.19	0.28	0.2
Q9I6Z2	Alkyl hydroperoxide reductase ahpF	Response to reactive oxygen species	0.086	0.07	0.09	0.05	0.042
P31224	Multidrug efflux pump subunit AcrB	Response to drug	0.070	0.014	0.034	0.002	0.011
P52599	Probable multidrug resistance protein emrK	Response to antibiotic	0.070	0.029	0.034	0.001	0.011
Q98D15	Probable multidrug resistance protein NorM	Drug transmembrane transporter activity	0.070	0.014	0.017	0.001	0.021
P52002	Multidrug resistance protein MexB	Response to antibiotic	0.035	0.014	0.017	0.001	0.021
P78966	Mating factor M secretion protein mam1	Transmembrane transport	0.035	0.000	0.00	0.003	0.00
Q6D2B1	Multidrug resistance protein MdtB	Transporter activity	0.011	0.01	0.00	0.003	0.01
Q73V87	Multidrug resistance protein mmr	Response to antibiotic	0.035	0.00	0.017	0.001	0.01
P34713	Multidrug resistance protein PGP3	Response to antibiotic	0.035	0.014	0.000	0.00	0.00
Q1GR76	UPP (Bacitracin resistance protein) (Undecaprenyl pyrophosphate phosphatase) UPPP	Response to antibiotic	0.070	0.029	0.051	0.003	0.032

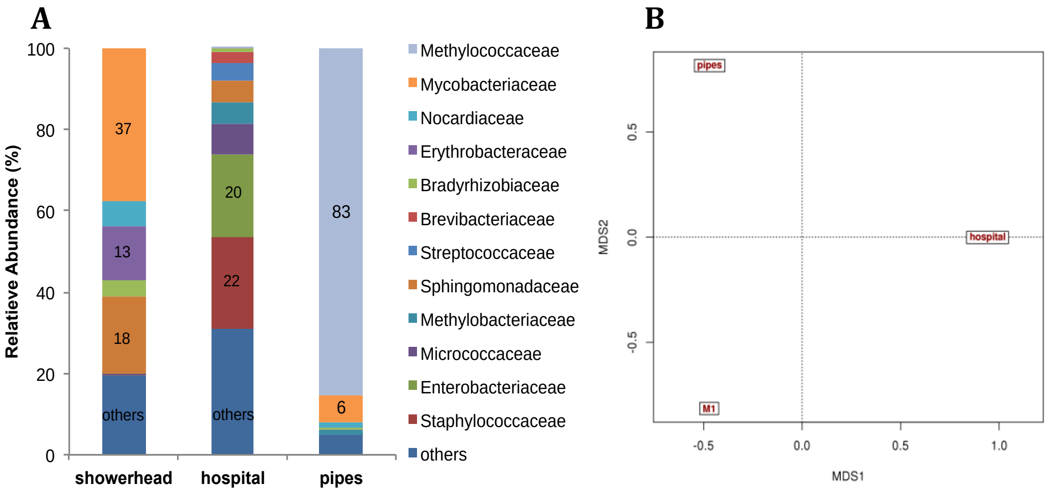

Figure S1. Comparison of the taxonomic profile of the shower hose metagenome with other microbial communities at the family level. (A) The bacterial community structures were assessed using 16S rRNA gene-encoding metagenomic reads. The dataset called “Hospital” was collected from an ICU ward surface of the University Hospital A Coruña, Spain, (SRA ID: SRX099356). The dataset called “Pipes” was collected from drinking water pipes in Florida, USA (SRA ID: SRX472092). (B) Multidimensional scaling (MDS) plot based on the relative abundance of taxa on each dataset using the Bray–Curtis distance.

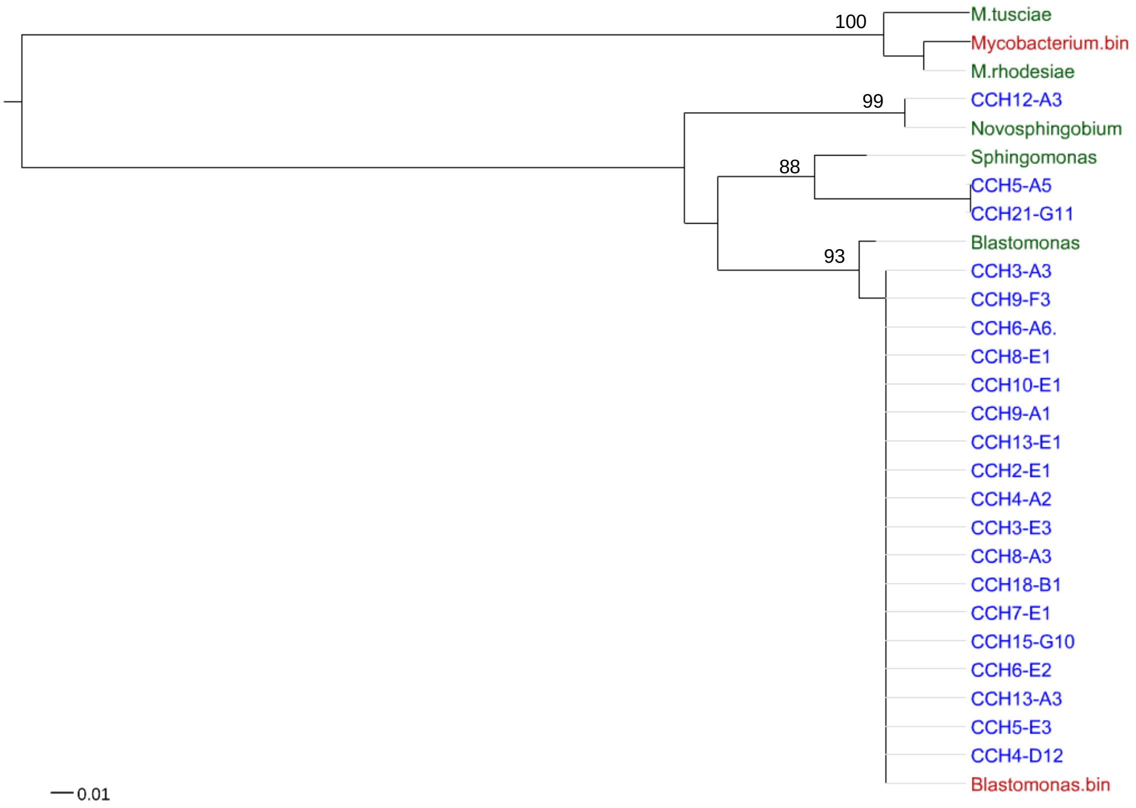

Figure S2. Phylogenetic relationships among population bins and isolate genomes based on 30S ribosomal protein S20 sequences. The isolates from the shower hoses are colored in blue; the recovered bins in red, and the reference genomes in green. The phylogeny was generated using the Neighbor-joining algorithm with 1000 bootstrap replicates using MEGA V.5. The number at nodes indicates the bootstrap support. Scale bar represents substitutions per site.

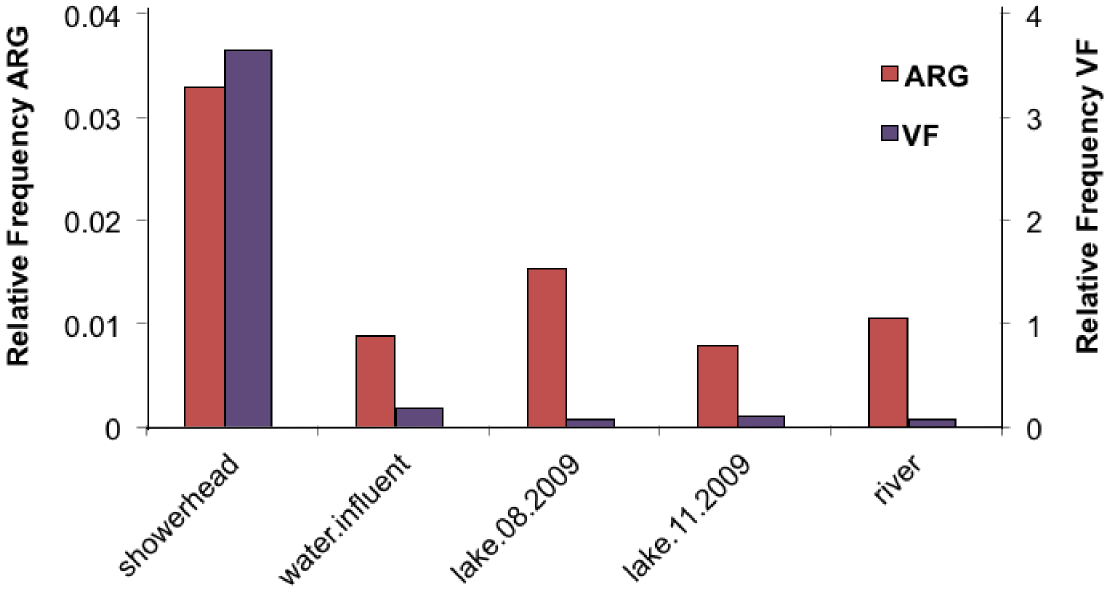

Figure S3. Metagenomic comparison of the genes involved in antibiotic resistance (ARG) and virulence (VF) mechanisms between the shower hose and other water-associated microbial communities. The “river” dataset corresponds to samples from river water (SRA ID: SRR1022353), the “drinking.water” dataset from a drinking water treatment plant in China (SRA ID: SRR835363), the “lake.08.2009” dataset collected during summertime (SRA ID: SRR096386), and the “lake.11.2009” dataset collected during falltime (SRA ID: SRR096389) from Lake Lanier in Georgia, USA (3). Pearson's Chi-squared test values between the shower hose metagenome and each water metagenome were significant (p-value < 0.05).

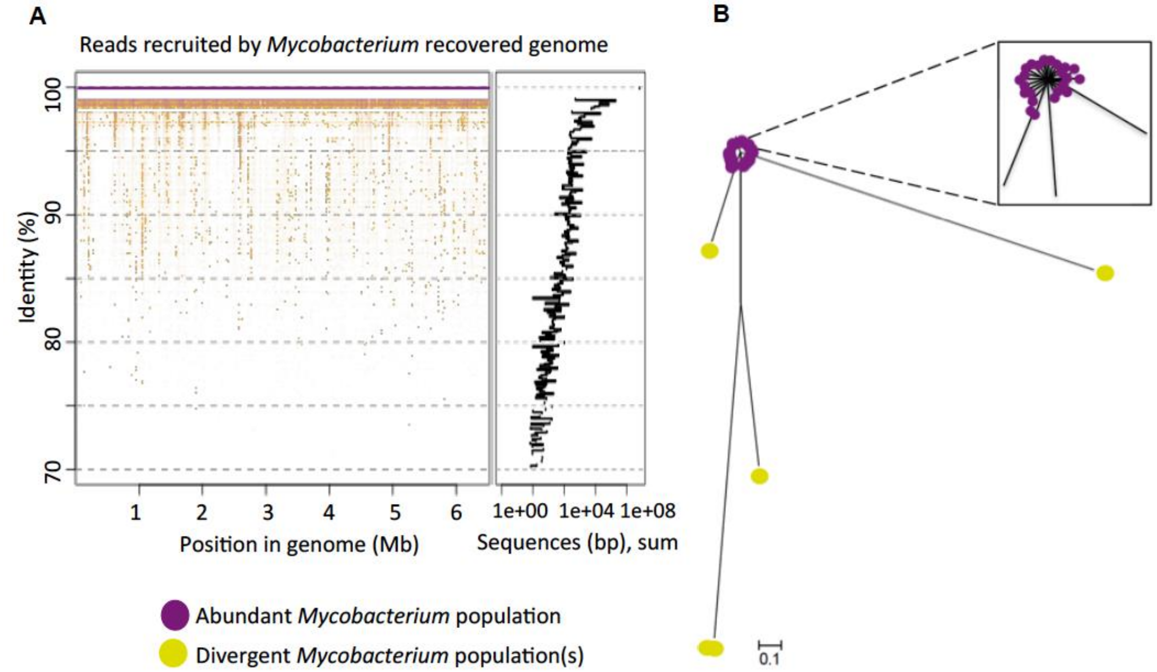

Fig S4. Intra-population diversity of the abundant *Mycobacterium* population recovered from the shower hose metagenomes. (A) Fragment recruitment plot of the recovered *Mycobacterium* population versus the M1 metagenome. Metagenomic reads were searched against the recovered genome sequence using a cut-off of at least 70% nucleotide identity and complete alignment to the genome reference. The y-axis corresponds to the identity of each read and the x-axis to the position of the read mapped on the genome. The histogram on the right represents the sum of the total base pairs of the reads recruited per unit of nucleotide identity. (B) Neighbor-joining phylogenetic tree of metagenomic reads that mapped on the single-copy 30S ribosomal protein S9 encoded on the recovered genome. Inset represents a zoomed in view of the tree where the reads representing the abundant *Mycobacterium* sp. population clustered together (in purple color). Note the star-like phylogeny formed by the latter reads. Scale bar represents substitutions per site.

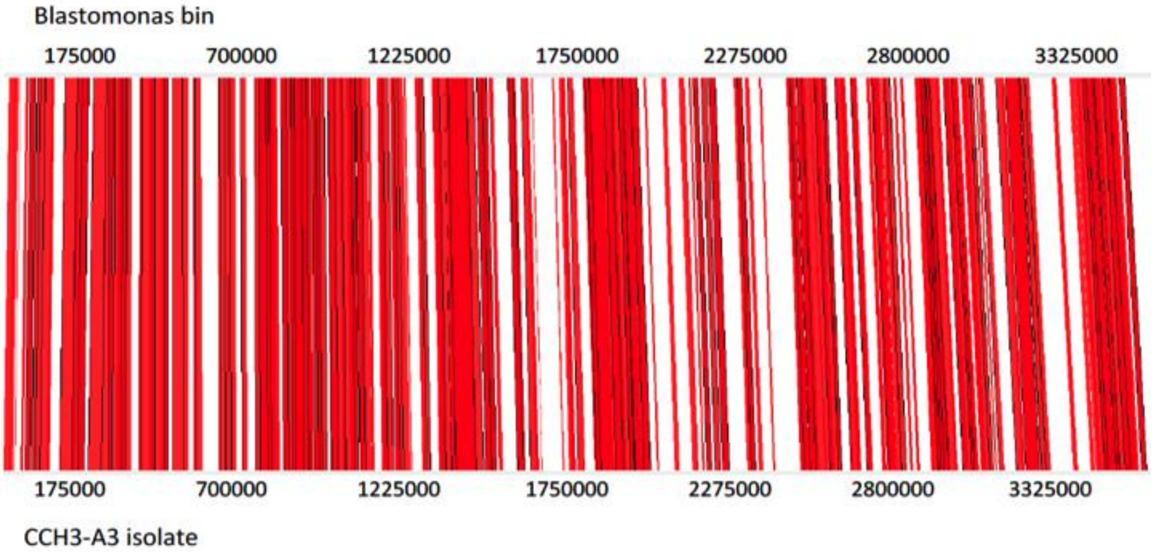

Figure S5. Genome alignment of the (binned) *Blastomonas* population genome against the *Blastomonas* isolate genome. The Artemis Comparison Tool (ACT) (4) was used to visualize the alignment of the two genomes. Contigs of the *Blastomonas* bin (top) were ordered based on homology searches and (assumed) synteny with a *Blastomonas* isolate genome available in GenBank (Accession number: GCA_000331245.1). Red bars indicate regions of similarity with the same orientation; empty/white bars indicate regions of gene content differences. Note, however, that most of the gene-content differences probably represent sequencing gaps (e.g., neither genome was complete) as opposed to real gene-content differences.

REFERENCES CITED

1. **Chen L, Xiong Z, Sun L, Yang J, Jin Q.** 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic acids research **40:**D641-645.
2. **Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJ, Yoo HS, Zhang C, Zhang Y, Sobral BW.** 2014. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic acids research **42:**D581-591.
3. **Oh S, Caro-Quintero A, Tsementzi D, DeLeon-Rodriguez N, Luo C, Poretsky R, Konstantinidis KT.** 2011. Metagenomic insights into the evolution, function, and complexity of the planktonic microbial community of Lake Lanier, a temperate freshwater ecosystem. Applied and environmental microbiology **77:**6000-6011.
4. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944-945.