Science Advances

www.advances.sciencemag.org/cgi/content/full/1/5/e1400127/DC1

Supplementary Materials for

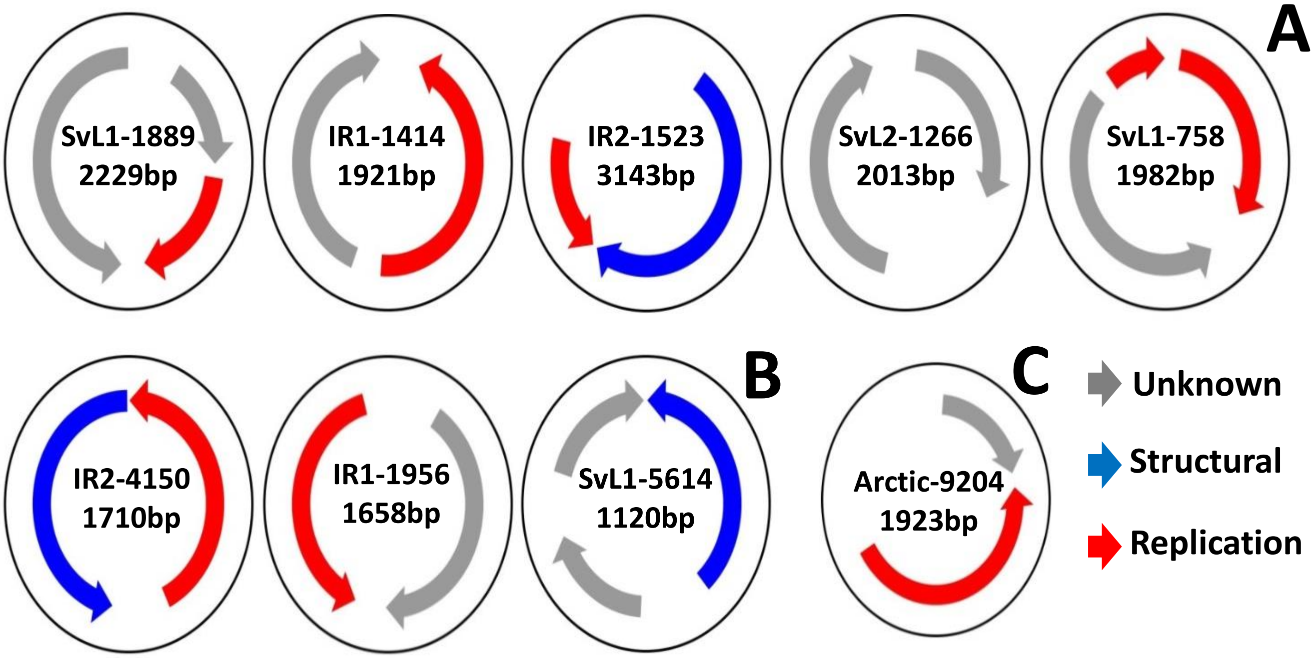
Biodiversity and distribution of polar freshwater DNA viruses

Daniel Aguirre de Cárcer, Alberto López-Bueno, David A. Pearce, Antonio Alcamí

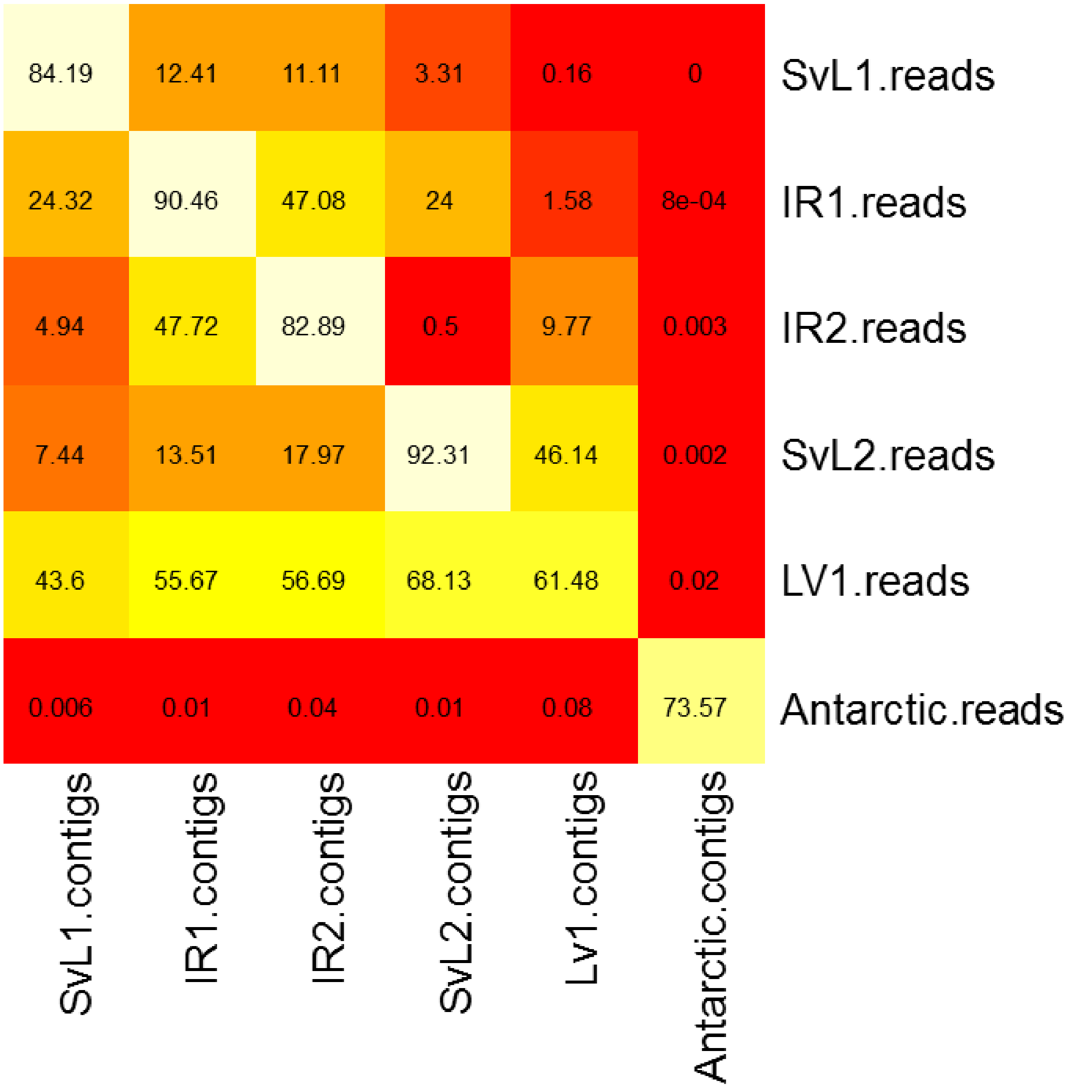
Published 19 June 2015, *Sci. Adv.* **1**, e1400127 (2015)
DOI: 10.1126/sciadv.1400127

This PDF file includes:

Supplementary Materials

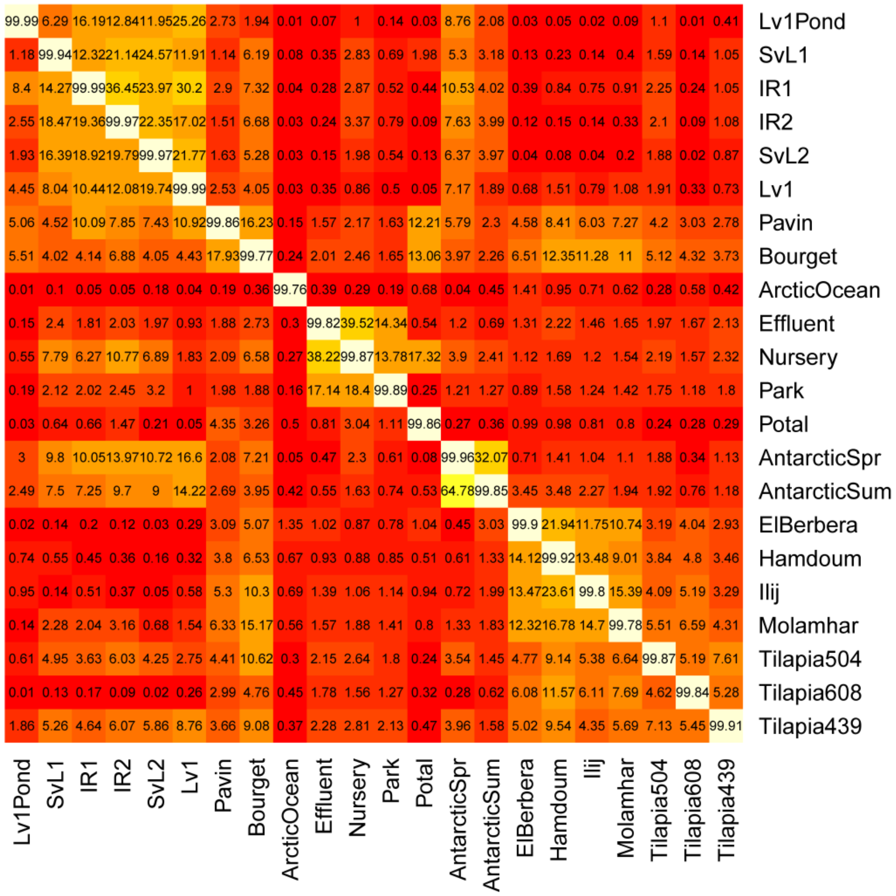


Supplementary figure 1. Significant circular ssDNA-related contigs. Contig name and length are shown. ORFs with no major similarities are shown in gray, while those with significant similarities to known viral structural or replication-related proteins are shown in blue and red, respectively. **(A)** Most abundant circular contigs >1000 bp found in each Arctic virome. **(B)** Representatives of the three highly similar contig groups found in at least four of the five Arctic viromes. **(C)** Representative of a shallow lineage found in both Antarctic and Arctic environments.



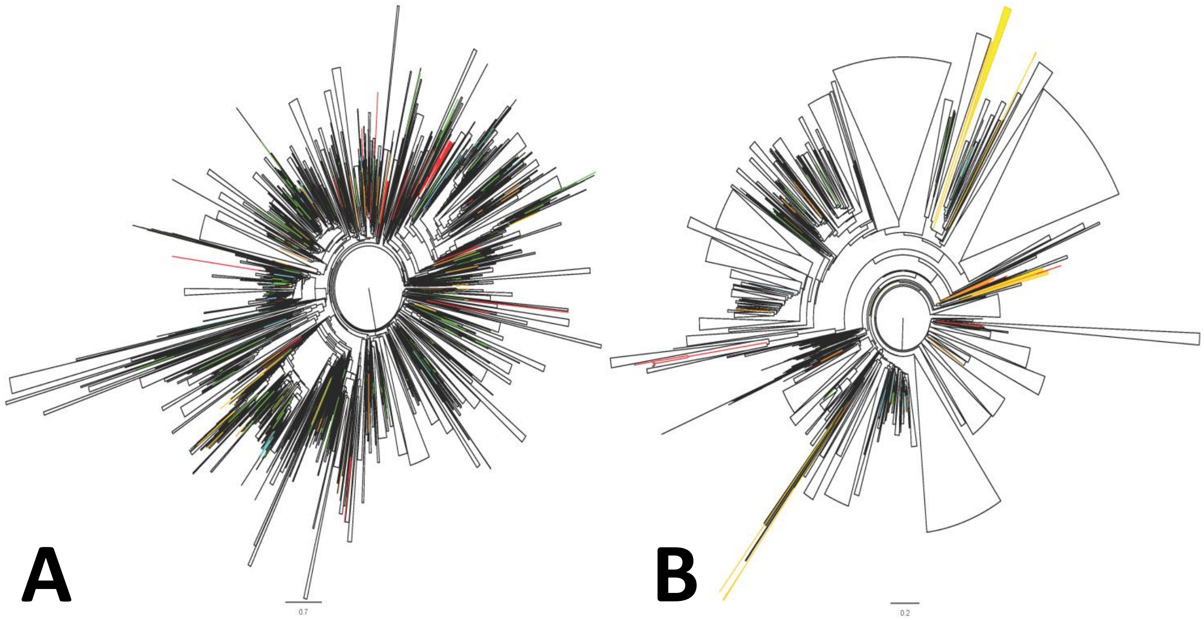
	SvL1.contigs	IR1.contigs	IR2.contigs	SvL2.contigs	Lv1.contigs	Antarctic.contigs
SvL1.reads	84.19	12.41	11.11	3.31	0.16	0
IR1.reads	24.32	90.46	47.08	24	1.58	8e-04
IR2.reads	4.94	47.72	82.89	0.5	9.77	0.003
SvL2.reads	7.44	13.51	17.97	92.31	46.14	0.002
LV1.reads	43.6	55.67	56.69	68.13	61.48	0.02
Antarctic.reads	0.006	0.01	0.04	0.01	0.08	73.57

Supplementary figure 2. Fine-grain genetic overlap between polar freshwater viromes. The heatplot represents the results of reads-to-contigs mappings with *bowtie2*. Values (% of total reads with hits with --*score-min* L,0,-0.2) are registered within the boxes and color-coded.



	Lv1Pond	SvL1	IR1	IR2	SvL2	Lv1	Pavin	Bourget	ArcticOcean	Effluent	Nursery	Park	Potal	AntarcticSpr	AntarcticSum	ElBerbera	Hamdoum	Ilij	Molamhar	Tilapia504	Tilapia608	Tilapia439
Lv1Pond	99.99	6.29	16.19	12.84	11.95	25.26	2.73	1.94	0.01	0.07	1	0.14	0.03	8.76	2.08	0.03	0.05	0.02	0.09	1.1	0.01	0.41
SvL1	1.18	99.94	12.32	21.14	24.57	11.91	1.14	6.19	0.08	0.35	2.83	0.69	1.98	5.3	3.18	0.13	0.23	0.14	0.4	1.59	0.14	1.05
IR1	8.4	14.27	99.99	36.45	23.97	30.2	2.9	7.32	0.04	0.28	2.87	0.52	0.44	10.53	4.02	0.39	0.84	0.75	0.91	2.25	0.24	1.05
IR2	2.55	18.47	19.36	99.97	22.35	17.02	1.51	6.68	0.03	0.24	3.37	0.79	0.09	7.63	3.99	0.12	0.15	0.14	0.33	2.1	0.09	1.08
SvL2	1.93	16.39	18.92	19.79	99.97	21.77	1.63	5.28	0.03	0.15	1.98	0.54	0.13	6.37	3.97	0.04	0.08	0.04	0.2	1.88	0.02	0.87
Lv1	4.45	8.04	10.44	12.08	19.74	99.99	2.53	4.05	0.03	0.35	0.86	0.5	0.05	7.17	1.89	0.68	1.51	0.79	1.08	1.91	0.33	0.73
Pavin	5.06	4.52	10.09	7.85	7.43	10.92	99.86	16.23	0.15	1.57	2.17	1.63	12.21	5.79	2.3	4.58	8.41	6.03	7.27	4.2	3.03	2.78
Bourget	5.51	4.02	4.14	6.88	4.05	4.43	17.93	99.77	0.24	2.01	2.46	1.65	13.06	3.97	2.26	6.51	12.35	11.28	11	5.12	4.32	3.73
ArcticOcean	0.01	0.1	0.05	0.05	0.18	0.04	0.19	0.36	99.76	0.39	0.29	0.19	0.68	0.04	0.45	1.41	0.95	0.71	0.62	0.28	0.58	0.42
Effluent	0.15	2.4	1.81	2.03	1.97	0.93	1.88	2.73	0.3	99.82	39.52	14.34	0.54	1.2	0.69	1.31	2.22	1.46	1.65	1.97	1.67	2.13
Nursery	0.55	7.79	6.27	10.77	6.89	1.83	2.09	6.58	0.27	38.22	99.87	13.78	17.32	3.9	2.41	1.12	1.69	1.2	1.54	2.19	1.57	2.32
Park	0.19	2.12	2.02	2.45	3.2	1	1.98	1.88	0.16	17.14	18.4	99.89	0.25	1.21	1.27	0.89	1.58	1.24	1.42	1.75	1.18	1.8
Potal	0.03	0.64	0.66	1.47	0.21	0.05	4.35	3.26	0.5	0.81	3.04	1.11	99.86	0.27	0.36	0.99	0.98	0.81	0.8	0.24	0.28	0.29
AntarcticSpr	3	9.8	10.05	13.97	10.72	16.6	2.08	7.21	0.05	0.47	2.3	0.61	0.08	99.96	32.07	0.71	1.41	1.04	1.1	1.88	0.34	1.13
AntarcticSum	2.49	7.5	7.25	9.7	9	14.22	2.69	3.95	0.42	0.55	1.63	0.74	0.53	64.78	99.85	3.45	3.48	2.27	1.94	1.92	0.76	1.18
ElBerbera	0.02	0.14	0.2	0.12	0.03	0.29	3.09	5.07	1.35	1.02	0.87	0.78	1.04	0.45	3.03	99.9	21.94	11.75	10.74	3.19	4.04	2.93
Hamdoum	0.74	0.55	0.45	0.36	0.16	0.32	3.8	6.53	0.67	0.93	0.88	0.85	0.51	0.61	1.33	14.12	99.92	13.48	9.01	3.84	4.8	3.46
Ilij	0.95	0.14	0.51	0.37	0.05	0.58	5.3	10.3	0.69	1.39	1.06	1.14	0.94	0.72	1.99	13.47	23.61	99.8	15.39	4.09	5.19	3.29
Molamhar	0.14	2.28	2.04	3.16	0.68	1.54	6.33	15.17	0.56	1.57	1.88	1.41	0.8	1.33	1.83	12.32	16.78	14.7	99.78	5.51	6.59	4.31
Tilapia504	0.61	4.95	3.63	6.03	4.25	2.75	4.41	10.62	0.3	2.15	2.64	1.8	0.24	3.54	1.45	4.77	9.14	5.38	6.64	99.87	5.19	7.61
Tilapia608	0.01	0.13	0.17	0.09	0.02	0.26	2.99	4.76	0.45	1.78	1.56	1.27	0.32	0.28	0.62	6.08	11.57	6.11	7.69	4.62	99.84	5.28
Tilapia439	1.86	5.26	4.64	6.07	5.86	8.76	3.66	9.08	0.37	2.28	2.81	2.13	0.47	3.96	1.58	5.02	9.54	4.35	5.69	7.13	5.45	99.91

Supplementary figure 3. Coarse-grain genetic overlap between viromes. The heatplot represents the results of cross-tBLASTx results. Values (% of total reads with hits with E-value < 0.001) are registered within the boxes and color-coded. Horizontal labels represent viromes acting as query and vertical labels represent viromes acting as reference.



Supplementary figure 4. NJ trees depicting the relationships between phylogenetic marker genes. *G*ene sequences were obtained from Arctic viromes (black and solid white branches), Antarctic viromes (green), Temperate lakes viromes (blue), recently published global sequences (orange), and reference sequences (red). **(A) *Circo-Nano-Geminiviridae* (*rep* gene).** (B) ***Microviridae (vp1*** **gene).** Arctic ssDNA viruses are broadly distributed across the known gene diversity spectra. Moreover, it seems that these contigs may expand the known diversity of *vp1* genes as the trees presents several deep–branching clusters with no previously described representatives (B).

Supplementary table 1. Cross-contig analysis of Arctic viromes. Number of contigs from each Arctic virome sharing at least 500bp at 98% similarity.

	SvL1	*IR1*	*IR2*	*SvL2*
IR1	169			
IR2	252	2408		
SvL2	25	37	26	
Lv1	21	100	135	49

Supplementary table 2. Taxonomic distribution (%) of assigned metagenomic reads from subsampled freshwater viromes*. Results represent BLASTx hits against a viral genomes database summarized by MEGAN. Only categories reaching 1% of assigned reads in at least one virome are shown.

	Lv1p	***Bou***	***SvL1***	***IR1***	***IR2***	***SvL2***	***Lv1***	***Effl***	***A.Oce***	***Nur***	***Par***
Ascoviridae	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6	0.0	0.0
Myoviridae	0.0	17.7	0.7	0.9	0.4	0.0	0.8	16.4	17.3	9.5	7.7
Podoviridae	0.0	3.7	0.3	0.1	0.2	0.0	0.6	13.7	4.1	9.7	5.5
Siphoviridae	0.0	21.4	3.8	1.0	0.3	0.0	1.3	37.6	25.5	29.0	20.7
Unclass. *Caudovirales*	0.0	0.7	0.1	0.0	0.0	0.0	0.0	0.9	1.2	0.3	0.6
Iridoviridae	0.0	0.8	0.1	0.1	0.0	0.0	0.4	0.4	0.8	0.4	0.1
Mimiviridae	0.0	0.2	0.0	0.1	0.0	0.0	0.0	0.2	4.5	0.0	0.2
Phycodnaviridae	0.0	4.2	0.2	0.5	0.1	0.0	0.0	0.9	14.8	0.9	1.1
Unclass. dsDNA phages	0.0	11.2	0.5	0.1	0.1	0.0	0.3	5.1	2.1	3.5	5.1
Unclass. dsDNA viruses	0.0	0.4	0.1	0.2	0.0	0.0	0.4	0.8	2.9	0.3	0.2
Satellites	0.5	0.0	1.7	1.2	1.7	0.5	0.6	0.0	0.0	0.0	0.0
Circoviridae	13.5	0.2	37.0	41.3	22.0	56.9	42.9	0.5	0.0	2.5	0.7
Geminiviridae	4.3	0.0	3.3	5.6	6.0	0.9	0.6	0.1	0.0	0.3	0.1
Inoviridae	0.0	0.0	0.4	0.1	0.0	0.0	0.0	0.1	4.1	0.1	0.0
Microviridae	1.4	13.2	8.3	11.2	14.1	1.0	1.3	0.1	0.0	3.8	0.0
Nanoviridae	9.2	0.2	7.2	2.0	6.4	3.6	11.9	0.3	0.0	1.0	0.0
Unclass. ssDNA viruses	70.0	0.2	28.4	28.4	39.4	35.7	34.2	2.3	0.0	22.7	48.5
ssRNA viruses	0.5	0.3	1.2	0.7	1.2	0.3	0.4	0.0	0.4	0.0	0.0
Unclass. phages	0.0	7.0	0.2	0.2	0.0	0.0	0.9	9.1	6.6	5.9	3.6
%Reads assigned	0.5	4.6	3.9	3.2	5.2	6.8	2.0	2.9	0.6	3.7	4.2

	Pav	***Pot***	***A.Spr***	***A.Sum***	***ElBer***	***Ham***	***Ilij***	***Mol***	***T504***	***T608***	***T439***
Ascoviridae	0.0	0.0	0.0	0.6	0.1	0.1	0.1	0.0	0.0	0.0	0.0
Myoviridae	20.4	13.4	2.5	5.4	31.0	24.9	28.6	45.4	9.5	14.1	14.0
Podoviridae	4.8	9.4	2.2	0.8	10.7	11.4	6.9	5.0	23.4	27.4	25.2
Siphoviridae	27.8	27.2	6.2	5.4	20.6	25.7	28.4	26.9	18.6	24.1	23.1
Unclass. *Caudovirales*	0.5	0.9	0.0	0.1	2.2	2.4	2.9	2.1	1.4	1.9	1.9
Iridoviridae	0.5	0.4	0.2	0.8	1.0	1.5	0.4	1.5	0.9	1.2	0.6
Mimiviridae	0.4	1.8	0.0	2.2	2.1	0.7	0.9	0.3	0.4	0.2	0.3
Phycodnaviridae	2.4	3.1	1.3	41.8	8.9	9.3	5.6	4.6	4.6	2.8	2.1
Unclass. dsDNA phages	4.4	4.5	1.0	0.4	4.1	3.4	6.1	4.9	6.9	10.9	7.0
Unclass. dsDNA viruses	0.5	0.0	0.2	15.4	3.4	2.2	1.6	0.7	0.5	1.2	0.6
Satellites	0.3	0.0	0.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Circoviridae	3.0	0.0	31.9	2.8	0.0	0.0	0.0	0.0	0.3	0.0	1.4
Geminiviridae	0.0	0.0	1.2	0.5	0.0	0.0	0.0	0.0	0.1	0.0	2.7
Inoviridae	0.0	2.7	0.2	0.1	0.2	0.1	0.0	0.0	0.4	0.0	0.3

Microviridae	0.3	3.6	4.0	0.2	0.0	0.0	0.0	0.2	0.1	0.0	0.3
Nanoviridae	0.7	0.0	1.8	0.5	0.0	0.0	0.0	0.0	0.1	0.0	0.5
Unclass. ssDNA viruses	10.1	0.0	37.2	16.2	0.0	0.1	0.1	0.1	18.5	0.0	1.7
ssRNA viruses	0.8	0.4	1.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.1
Unclass. phages	5.7	26.8	1.0	0.9	3.3	4.2	4.0	2.2	3.3	4.1	6.0
%Reads assigned	3.0	0.6	1.5	2.7	2.5	3.8	3.6	5.8	3.6	2.7	3.9

*****Virome names have been unambiguously shortened to fit.

Supplementary table 3. Cross-contig analysis of pooled freshwater environments; Number of contigs from each pooled environment sharing at least 500bp at 98%, 95% and 90% similarity. Results are influenced by total information available for contig reconstruction in each case*.

98%	Antarctic	Acuacult.	Arctic	ArctOcean	Sahara	Temperate	Reclaimed
Antarctic	-	-	-	-	-	-	-
Aquacult.	0	-	-	-	-	-	-
Arctic	0	0	-	-	-	-	-
ArctOcean	0	0	0	-	-	-	-
Sahara	26	0	0	0	-	-	-
Temperate	0	0	5	0	0	-	-
Reclaimed	0	0	0	0	0	0	-

95%	Antarctic	Acuacult.	Arctic	ArctOcean	Sahara	Temperate	Reclaimed
Antarctic	-	-	-	-	-	-	-
Aquacult.	0	-	-	-	-	-	-
Arctic	2	0	-	-	-	-	-
ArctOcean	0	0	0	-	-	-	-
Sahara	28	0	0	0	-	-	-
Temperate	0	0	59	0	1	-	
Reclaimed	0	0	0	0	0	1	-

90%	Antarctic	Acuacult.	Arctic	ArctOcean	Sahara	Temperate	Reclaimed
Antarctic	-	-	-	-	-	-	-
Aquacult.	0	-	-	-	-	-	-
Arctic	7	0	-	-	-	-	-
ArctOcean	0	0	0	-	-	-	-
Sahara	32	0	3	0	-	-	-
Temperate	13	1	286	0	23	-	-
Reclaimed	1	0	0	0	1	5	-

*Antarctic (ca. 240Mbp), Aquaculture (ca. 42Mbp), Arctic (ca. 3,200Mbp), Arctic Ocean (ca. 74Mbp), Sahara (ca. 65Mbp), Temperate lakes (ca. 250Mbp), Reclaimed water (ca. 238Mbp).