

Supplementary material

Supplementary Table 1: Variance partitioning results of the distance-based redundancy analysis (dbRDA) for each taxonomic group. The table reports the proportion of community variation explained by the environmental variables (R^2 , adjusted R^2), and permutation test P-values for the global constrained model and individual environmental terms.

	R^2	R^2_{adj}	P_global	P_T	P_EC	P_NO3
Archaea	0.405	-0.189	0.742	0.924	0.800	0.275
Bacteria	0.476	-0.048	0.649	0.527	0.668	0.514
Fungi	0.403	-0.193	0.882	0.385	0.902	0.981
Eukaryota	0.486	-0.028	0.587	0.599	0.833	0.315
Viruses	0.443	-0.113	0.849	0.717	0.609	0.908

Supplementary Table 2: Number of sequencing reads obtained by shotgun high throughput sequencing from cave water samples (raw reads and reads after trimming) and relative amount of the reads classified on the level of domains using Diamond/Megan.

	CW 1	CW 2	CW 3	CW 4	CW 5	CW 6	CW 7
Raw reads	2380710	3043646	2183998	2368210	2503142	2394770	2300082
Reads after trimming	2367224	3031850	2176246	2358548	2496658	2386872	2292508
Bacterial reads (%)	26.27	33.54	18.22	23.03	32.98	17.03	12.67
Archaeal reads (%)	0.12	0.10	0.26	0.25	0.16	0.30	0.26
Eukaryotic reads (%)	12.77	22.58	1.12	2.56	6.18	6.26	4.55
Viral reads (%)	2.02	0.41	3.72	2.29	3.19	3.36	2.62
Unclassified reads (%)	58.81	43.37	76.68	71.88	57.50	73.06	79.91

Supplementary Table 3: Number of sequencing reads classified on the level of viral realms obtained from cave water samples using Diamond/Megan. Viral reads classified as “environmental samples <viruses>” and “unclassified viruses” represent sequences not classified at realm level with used classification pipeline.

	CW 1	CW 2	CW 3	CW 4	CW 5	CW 6	CW 7
<i>Duplodnaviria</i>	6760	2792	17364	17273	6176	10068	14083
<i>Monodnaviria</i>	7007	2696	29009	18990	48615	51537	19791
<i>Varidnaviria</i>	1517	821	926	1020	559	1228	4134
<i>Riboviria</i>	13045	1548	24897	4366	7711	2468	8663
environmental samples <viruses>	964	303	2630	2788	1896	2488	2246
unclassified viruses	18609	1635	13139	9684	10555	11682	12987

Supplementary Table 4: (SupplementaryTable3_families_host.xlsx): Number of reads classified on the level of viral families detected in Diamond/Megan analysis together with host information based on usual host associations of viral families as identified by International Committee on Taxonomy of Viruses (ICTV).

Supplementary Table 5: The number of generated ASVs per sample after bioinformatic analyses of partial sequences 16S rDNA (*Archaea*, *Bacteria*), ITS2 region of rDNA (fungi) and 18S rDNA (*Eukarya*), obtained by amplicon sequencing.

Sample	<i>Archaea</i>	<i>Bacteria</i>	Fungi	<i>Eukarya</i>
CW1 Pivka River – Črna jama	15605	32080	44625	28800
CW2 Vir	29419	30615	18411	36808
CW3 Krka River – Krška jama	26986	29213	37936	47
CW4 Pivka River – Planinska jama	12409	29213	31084	923
CW5 Jelševniščica	8137	23224	30660	35462
CW6 Kompoljska jama	7341	31488	46483	2177
CW7 Reka River – Škocjanske jame	17159	20938	34837	31179

Supplementary Table 6: Abundance and diversity of fungi in Slovenian cave waters reported as the number of assigned ASVs.

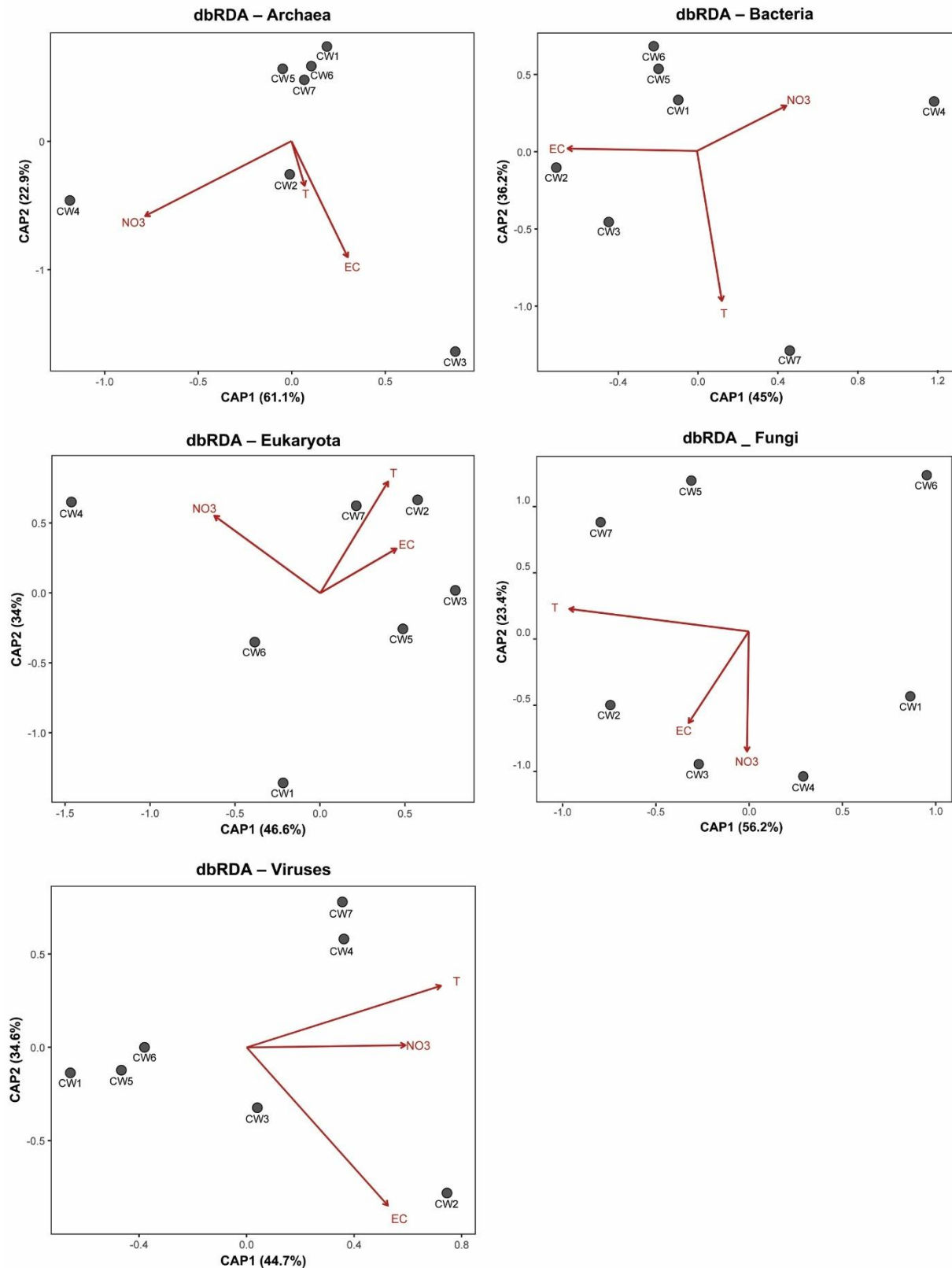
		Cave water samples						
		CW1	CW2	CW3	CW4	CW5	CW6	CW7
Fungal abundance (number of ASVs)	Total	44625	18411	37963	31084	30660	46483	34837
	Basidiomycota	43639	17799	33405	30802	23683	29918	31544
	Ascomycota	783	437	304	259	2061	7935	1905
	Mortierellomycota	105	37	4	12	631	8350	0
	Mucoromycota	1	0	4243	0	136	1	0
	Chytridiomycota	0	0	0	0	1297	210	0
	Other	97	138	7	11	2852	69	1388
Fungal diversity (number of different genera)	Total	31	23	29	26	48	77	36
	Basidiomycota	13	9	12	15	16	28	17
	Ascomycota	15	12	14	9	26	44	18
	Mortierellomycota	1	1	1	1	1	1	0
	Mucoromycota	1	0	1	0	1	1	0
	Chytridiomycota	0	0	0	0	3	2	0
	Other	No data	No data	No data	No data	No data	No data	No data

Supplementary Table 7: (SupplementaryTable6_protozoan_taxa): Abundance and diversity of eukaryotes in cave waters on phylum level shown as the number of assigned ASVs.

Supplementary Table 8: Richness (Chao1) and diversity (Shannon) index values of different communities in each sample. For alpha diversity analyses, samples were normalized to an equal number of sequences by rarefaction. All samples with a total number of reads, less than the chosen rarefaction value, were removed from these analyses. 7,000 reads were retained for archaeal, 20,000 reads for

bacterial, 2,000 reads for protozoan and 18,000 reads for fungal samples, to maximize the number of samples rather than the number of reads. The highest values are in bold, and the lowest values are underlined, / - no data.

Sample	<i>Archaea</i>		<i>Bacteria</i>		<i>Fungi</i>		<i>Protozoa</i>	
	Chao1	Shanno n	Chao1	Shanno n	Chao1	Shanno n	Chao1	Shanno n
Pivka River – Črna jama (CW1)	205.38	6.73	428.96	7.31	102.50	3.86	75	<u>3.19</u>
Vir (CW2)	319.71	7.51	353.35	6.66	<u>72</u>	<u>3.37</u>	99.20	5.09
Krka River – Krška jama (CW3)	30	<u>3.28</u>	257	<u>6.12</u>	100.33	3.44	/	/
Pivka River – Planinska jama (CW4)	<u>28</u>	3.54	304.36	6.69	104.18	4.41	/	/
Jelševniščica (CW5)	106.33	5.96	360.87	7.25	148.50	6.06	84.08	3.32
Kompoljska jama (CW6)	117.25	6.25	294	6.37	215	7.04	<u>36</u>	3.70
Reka River – Škocjanske jame (CW7)	170	7.12	<u>252.77</u>	6.77	143	5.10	132.83	6.17



56

57 **Supplementary Figure 1:** Influence of environmental variables on microbial community composition
58 across cave water samples, shown by distance-based redundancy analysis (dbRDA) of Bray–Curtis
59 dissimilarities. Separate ordinations are presented for archaea, bacteria, fungi, other eukaryotes, and
60 viruses, based on Hellinger-transformed sequence abundance data. The vectors represent environmental
61 variables (T – temperature; EC – electrical conductivity; NO₃ – nitrate) and indicate the direction and
62 strength of their correlations with the canonical axes (CAP1–CAP2).

63



64

65 **Supplementary Figure 2:** Overview of the percentage of viral reads according to usual host
66 associations based on known associations of corresponding viral families. Viral reads classified as
67 “Other – not assigned to family” represent sequences not classified at family level with used
68 classification pipeline.

69