

Evaluation of the Viral Diversity of Artemia Cysts from Saline Lakes in Kazakhstan Using Viral Metagenomics Analysis

Marat Kumar ^{1,2,*}, Kobey Karamendin ¹, Zhanara Mazhibayeva ³, Yermukhammet Kassymbekov ¹, Temirlan Sabyrzhan ^{1,2}, Kuanysh Isbekov ³, Saule Assylbekova ³ and Aidyn Kydyrmanov ^{1,*}

(Supplementary Table S1)

Table S1. Information about the sampling location in the four regions (sampling site, coordination, temperature, salinity, sampling dates, and Aquatic organisms).

Sampling site (Region, Lakes)		Coordination	Temp °C	Salinity g/L	Sampling Date	Aquatic organisms
North Kazakhstan Region	Pasynki Lake	54°47'18.0"N, 67°5'7.0"E	20,2 - 22,4	136,4	08/2022	<i>Cletocamptus retrogressuss</i>
	Smirnovka Lake	54°47'7.0"N, 68°22'26.0"E	18,0 - 22,0	107,0	08/2022	<i>Cletocamptus retrogressus</i> , <i>Cyclopidae gen.sp.</i>
	Zhamantuz Lake	54°1'58.0"N, 69°14'28.0"E	17,5-19,8	89,7	08/2022	<i>Cletocamptus retrogressuss</i>
	Kalibek Lake	53°52'23.0"N, 70°37'40.0"E	18,2 - 23,1	235,4	08/2022	
Pavlodar Region	Kyzyltuz Lake	53°47'20.0"N, 76°46'28.0"E	22,0–24,2	288,0	08/2021	
	Ashchitakyr Lake	52°33'27.0"N, 78°19'35.0"E	21,5–23,0	79,9	08/2021	
	TuzKala Lake	51°52'38.0"N, 77°29'22.0"E	22,0–23,7	236,0	08/2021	
	Kazy Lake	51°42'09.0"N, 78°03'34.9"E	21,3–24,1	243,0	08/2021	
Almaty Region	Ray Lake	46°02'08.0"N, 78°08'45.0"E	21,5-24,2	140,8	08/2022	<i>Brachionus plicatilis</i> , <i>Hexarthra fennica</i> , <i>Moina brachiate</i> , <i>Diaptomus sp.</i> , <i>Cletocamptus retrogressus</i>
	Tuzkol Lake	43°00'21.0"N, 79°59'27.0"E	22,0-24,7	75,7	08/2022	<i>Cletocamptus retrogressuss</i> , <i>Cyclopidae gen.sp.</i> , <i>Diaptomidae gen.sp.</i>
Kyzylorda Region	Tushchybas Bay	46°15'00.0"N, 59°40'00.0"E	23,6-27,4	72,3	08/2021	
	Chernyshev Bay	45°58'21.7"N, 59°16'05.6"E	24,5-27,8	206,2	08/2021	

(Supplementary Table S2)

Table S2: Overview of Sequencing Reads Classification at the Domain Level (DIAMOND Blastx and MEGAN-LCA).

Domain of life	North_Kz (Assigned contigs No.)	Pavlodar (Assigned contigs No.)	Almaty (Assigned contigs No.)	Kyzylorda (Assigned contigs No.)
Bacteria	1251	2155	188	2448
Archaea	78	216	0	23
Eukaryota	2700	492	308	856
Viruses	56	39	20	53
Not Assigned	2310	1059	615	1452
Total	6395	3961	1131	4832

(Supplementary Figure S1)

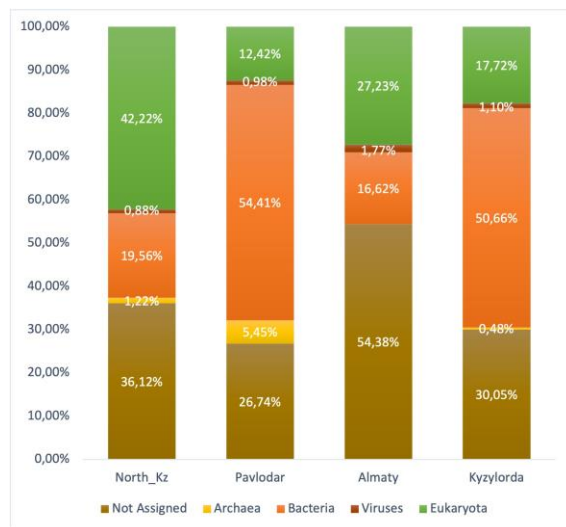


Figure S1. Proportion of the classification level of sequencing reads for different regions;based on DIAMOND blastx similarity search, followed by MEGAN-LCA binning to the classification level "domain".

Table S3. Summary of the total count of viral readings obtained from 15 distinct viral families and unclassified viruses.					Sampling regions			
Taxonomy Name	Viral Genome	Classification	Naturalhosts	Host	North_KZ	Pavlodar	Almaty	Kyzylorda
Autographiviridae	Group I: dsDNA viruses	Bacteriophage	Bacteria	Bacteria	2	0	0	0
Myoviridae	Group I: dsDNA viruses	Bacteriophage	Bacteria	Bacteria	1	0	0	3
Podoviridae	Group I: dsDNA viruses	Bacteriophage	Bacteria	Bacteria	0	0	0	1
Straboviridae	Group I: dsDNA viruses	Bacteriophage	Bacteria	Bacteria	2	0	0	0
Schitoviridae	Group I: dsDNA viruses	Bacteriophage	Bacteria	Bacteria	1	0	0	0
Zobellviridae	Group I: dsDNA viruses	Bacteriophage	Bacteria	Bacteria	2	0	0	0
unclassified Caudovirales	Group I: dsDNA viruses	Bacteriophage	Bacteria	Bacteria	8	2	2	14
Tombusviridae	Group IV: ssRNA positive-strand viruses	Plant viruses	Plant	Plant	0	0	4	0
Solemoviridae	Group IV: ssRNA positive-strand viruses	Plant viruses	Plant	Plant	0	0	1	0
Picornaviridae	Group IV: ssRNA positive-strand viruses	Vertebrates viruses	Vertebrates	Vertebrates	0	2	0	0
Astroviridae	Group IV: ssRNA positive-strand viruses	Vertebrates viruses	Vertebrates	Vertebrates	0	0	2	0
Reoviridae	Group III: dsRNA viruses	Unclear	Vertebrates- Invertebrates	Invertebrates	26	14	0	0
Nodaviridae	Group IV: ssRNA positive-strand viruses	Invertebrates viruses	Vertebrates- Invertebrates	Invertebrates	2	0	0	0
Dicistroviridae	Group IV: ssRNA positive-strand viruses	Insect viruses	Invertebrates	Invertebrates	0	3	4	0
unclassified viruses	Unclear	Unclassified	Unclear	Unclear	0	3	5	1

Table S4: Overview of contigs of Bacteriophage species identified using BLASTx in Artemia cysts collected from four regions.

No.	region	coverage	length	bitscore	ident	length.1	stitle	qlen	staxid	species	genus	family	order	class	phylum	superkingdom
1	North_Kz	2	569	124	61.9	97	YP_009800545.1 terminase large subunit [Pseudomonas phage Njord]	569	2163985	Uliginivirus njord	Uliginivirus	Autographiviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
2	North_Kz	2	390	161	65.6	125	UJH95814.1 major capsid protein [Pantoea phage Nifs112]	390	2894289	Pantoea phage Nifs112	Eracentumvirus	Autographiviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
3	North_Kz	2	358	137	64.9	111	WAI96127.1 portal protein [Vibrio phage vB_VhaP_PG11]	358	2996149	Vibrio phage vB_VhaP_PG11	unknown	Schitoviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
4	North_Kz	2	356	196	74.4	117	YP_010090488.1 NrdA-like aerobic NDP reductase large subunit [Pseudomonas phage PspYZU05]	356	1983556	Jiangsuivirus pspyzu05	Jiangsuivirus	Straboviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
5	North_Kz	2	327	132	63.8	94	QPI18018.1 ribonucleotide reductase large subunit [Pectobacterium phage POP12]	327	2741065	Pectobacterium phage POP12	unknown	Straboviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
6	North_Kz	1	419	278	92.1	139	YP_007010563.1 DNA packaging enzyme (terminase) [Salinivibrio phage CW02]	419	1161935	Salinivirus utanense	Salinivirus	Zobellviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
7	North_Kz	2	378	215	81.5	124	YP_007010534.1 DNA polymerase [Salinivibrio phage CW02]	378	1161935	Salinivirus utanense	Salinivirus	Zobellviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
8	North_Kz	2	348	202	86.1	115	ASV44548.1 thymidine kinase-like protein [Salicola phage SCTP-2]	348	2015814	Salicola phage SCTP-2	unknown	Myoviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
9	North_Kz	2	363	208	90.1	111	AUV61410.1 small terminase [Pontimonas phage phiPsal1]	363	2079347	Pontimonas phage phiPsal1	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
10	North_Kz	2	410	199	73.5	136	APC44505.1 hypothetical protein [Pseudoalteromonas phage PH357]	410	1913046	Pseudoalteromonas phage PH357	Qingdaovirus	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
11	North_Kz	2	366	141	70.5	88	YP_009595559.1 major head protein [Pseudomonas phage pf16]	366	1815630	Chakrabartyvirus pf16	Chakrabartyvirus	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
12	North_Kz	2	522	229	64.7	173	YP_009595646.1 DNA topoisomerase II [Pseudomonas phage pf16]	522	1815630	Chakrabartyvirus pf16	Chakrabartyvirus	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
13	North_Kz	3	338	144	64.2	109	YP_009595528.1 DNA polymerase [Pseudomonas phage pf16]	338	1815630	Chakrabartyvirus pf16	Chakrabartyvirus	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
14	North_Kz	2	535	211	60.2	176	CAB4157494.1 GP4d_helicase domain containing protein [uncultured Caudovirales phage]	535	2100421	uncultured Caudovirales phage	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
15	North_Kz	2	349	122	58	112	WAX23091.1 tail sheath protein [Vibrio phage 13VT501A]	349	3003693	Vibrio phage 13VT501A	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
16	North_Kz	2	365	129	51.5	130	UGV20061.1 DNA polymerase [Salmonella phage ph2-2]	365	2900271	Salmonella phage ph2-2	Felixovavirus	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
17	Pavlodar	4.6083	1397	124	73.3	86	DAF34309.1 MAG TPA: hypothetical protein [Caudoviricetes sp.]	1397	2832643	Caudoviricetes sp.	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
18	Pavlodar	2.2209	829	137	81.9	83	DAF34309.1 MAG TPA: hypothetical protein [Caudoviricetes sp.]	829	2832643	Caudoviricetes sp.	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
19	Almaty	1	367	163	83.5	97	DAK91763.1 MAG TPA: hypothetical protein, partial [Caudoviricetes sp.]	367	2832643	Caudoviricetes sp.	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
20	Almaty	1	349	140	66.1	115	CAB4154309.1 tape_meas_TP901, phage tail tape measure protein, TP901 family, core region [uncultured Caudovirales phage]	349	2100421	uncultured Caudovirales phage	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
21	Kyzylorda	1	412	190	78	123	ATV46261.1 major capsid protein [Lake Baikal phage Baikal-20-5m-C28]	412	2047876	Lake Baikal phage Baikal-20-5m-C28	unknown	Myoviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
22	Kyzylorda	2	331	159	74.8	103	AUR89446.1 P-loop containing nucleoside triphosphate hydrolase [Vibrio phage 1.123.O_10N.286.48.F3]	331	2070727	Vibrio phage 1.123.O_10N.286.48.F3	unknown	Myoviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
23	Kyzylorda	2	334	124	69.5	95	ATV46261.1 major capsid protein [Lake Baikal phage Baikal-20-5m-C28]	334	2047876	Lake Baikal phage Baikal-20-5m-C28	unknown	Myoviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage
24	Kyzylorda	3	523	166	66.4	119	YP_009009736.1 anti-repressor Ant [Vibrio phage PVA1]	523	1461743	Vibrio phage PVA1	unknown	Podoviridae	unknown	Caudoviricetes	Uroviricota	Bacteriophage

25	Kyzylord a	2	421	173	84.2	101	QJG62451.1 terminase large subunit [Pseudoalteromonas phage AL]	421	2712933	Pseudoalteromonas phage AL	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
26	Kyzylord a	1	409	232	83.8	136	YP_009810889.1 ribonucleoside-diphosphate reductase large subunit [Synechococcus phage S-T4]	409	2268578	Tamkungvirus ST4	Tamkungvirus	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
27	Kyzylord a	2	313	165	81.7	104	QEG08956.1 P22 coat-protein S family protein [Aeromonas phage 4_4512]	313	2588516	Aeromonas phage 4_4512	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
28	Kyzylord a	2	466	260	80.6	155	UUW39915.1 putative DNA polymerase II [Vibrio phage VPMCC14]	466	2968382	Vibrio phage VPMCC14	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
29	Kyzylord a	2	456	241	79.6	152	CAB4125971.1 clpA ATP-binding subunits of Clp protease and DnaK/DnaJ chaperones [uncultured Caudovirales phage]	456	2100421	uncultured Caudovirales phage	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
30	Kyzylord a	1	432	140	75	92	AUR95823.1 ERF superfamily protein [Vibrio phage 1.213.O_10N.222.54.F10]	432	1881410	Vibrio phage 1.213.O_10N.222.54.F10	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
31	Kyzylord a	2	341	157	74.6	114	AUR92199.1 coil containing protein [Vibrio phage 1.170.O_10N.261.52.C3]	341	1881390	Vibrio phage 1.170.O_10N.261.52.C3	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
32	Kyzylord a	3	479	121	73.3	86	UAW58936.1 putative tail fiber protein [Roseobacter phage CRP-207]	479	2873404	Roseobacter phage CRP-207	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
33	Kyzylord a	3	693	225	71.8	142	DAJ54684.1 MAG TPA: DNA repair protein-like protein [Caudoviricetes sp.]	693	2832643	Caudoviricetes sp.	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
34	Kyzylord a	2	427	129	71.1	90	CAB4143225.1 hypothetical protein UFOVP449_138 [uncultured Caudovirales phage]	427	2100421	uncultured Caudovirales phage	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
35	Kyzylord a	2	443	213	70.7	150	WAX23081.1 major capsid protein [Vibrio phage 13VT501A]	443	3003693	Vibrio phage 13VT501A	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
36	Kyzylord a	2	454	122	70.6	85	WAX23101.1 hypothetical protein pA_gene0060 [Vibrio phage 13VT501A]	454	3003693	Vibrio phage 13VT501A	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
37	Kyzylord a	1	405	122	70.3	91	CAB5221546.1 hypothetical protein UFOVP240_190 [uncultured Caudovirales phage]	405	2100421	uncultured Caudovirales phage	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage
38	Kyzylord a	3	405	147	68.4	98	CAB4133809.1 COG0720 6-pyruvoyl-tetrahydropterin synthase [uncultured Caudovirales phage]	405	2100421	uncultured Caudovirales phage	unknown	unknown	unknown	Caudoviricetes	Uroviricota	Bacteriophage