

Supplemental Material to:

**Ana Senčilo, Deborah Jacobs-Sera, Daniel A. Russell,
Ching-Chung Ko, Charles A. Bowman, Nina S. Atanasova,
Eija Österlund, Hanna M. Oksanen, Dennis H. Bamford,
Grahm F. Hatfull, Elina Roine and Roger W. Hendrix**

Snapshot of haloarchaeal tailed virus genomes

2013; 10(5)

<http://dx.doi.org/10.4161/rna.24045>

<http://www.landesbioscience.com/journals/rna/article/24045>

Supplementary Data

For

The Genomes of Haloarchaeal Tailed Viruses

Sencilo A.¹, Jacobs-Sera D.², Russell DA.², Ko C.-C.², Atanasova NS.¹, Österlund E.¹,
Oksanen HM.¹, Bamford DH.¹, Hatfull GF.², Roine E.¹, Hendrix RW.²

1. Department of Biosciences and Institute of Biotechnology, University of Helsinki, Helsinki, Finland.
2. Department of Biological Sciences, University of Pittsburgh, Pennsylvania, United States of America.

SUPPLEMENTARY RESULTS

Provirus regions in euryarchaeal genomes

Related archaeal proviruses were identified by first searching for homologues of terminase large subunit of our viruses in archaeal non-redundant protein sequence database. Then genome regions around the putative large terminase subunit genes were inspected for other putative viral genes. In this way four new proviral elements residing in the genomes of euryarchaea belonging to three different classes were identified (Table S1, Fig. S1).

In order to determine the borders of the proviral sequences, regions close to the integrase gene on one side and other putative proviral genes on the other side were analyzed for the direct repeats produced during virus integration into the host chromosome.^{1, 2} Such direct repeats ranging in size from 16 to 55 bp were identified in three out of the four proviruses (Table S1).

Putative genes coding for virion structural and assembly proteins, proteins involved in DNA packaging and DNA and RNA metabolism were annotated in the genomes of proviruses (Fig. S1). Additional putative homologues of bacteriophage proteins have been identified. For instance, putative Mu gpG-like and Lambda head decoration D-like protein homologues have been found in haloarchaeal proviral regions from our studied set. P2 baseplate GpU-like and GpW-like protein homologues were predicted to be encoded by Aprof. Aprof also codes for late control GPD, P2 V-like protein and baseplate assembly J-like protein putative homologues, which have been identified in HGTV-1 too. Thus, Aprof seems to encode a full set of P2 baseplate protein homologues (Fig. S2).

Arrangements of haloarchaeal proviral regions are similar to earlier described proviruses (Krupovic et al., 2010). *M. okinawensis* proviral region is organized in a similar way as *Methanocaldococcus vulcanius* M7 provirus Mvul-Pro1.² Its genes coding for structural and DNA packaging proteins are encoded by two separate gene clusters on opposite strands. One of the clusters codes for terminase large subunit, portal-Mu gpF fusion protein and the other one codes for putative minor and major capsid proteins.

Putative prohead proteases in studied haloarchaeal tailed viruses

Putative prohead proteases have been identified viruses HCTV-2, HHTV-2, HGTV-1, HRTV-5, HRTV-7 and HRTV-8. Predictions were based either on similarity to earlier identified prohead proteases² or based on the position of the encoding gene, conservation of

secondary structure and a putative catalytic triad.^{3,4} All of our analyzed putative prohead proteases appear to have a conserved catalytic triad (aspartate, histidine, serine), except for that of HRTV-8 virus. Instead of aspartate HRTV-8 prohead protease has glutamate at a corresponding position, despite over 90% of shared amino acids with putative prohead protease of HF2 virus.²

Proteins with Zeta toxin-like domains in HCTV-1 and HCTV-5 viruses.

HCTV-1 and HCTV-5 viruses were predicted to encode protein with Zeta-toxin like domain in a genome region coding for capsid structural proteins. In bacteria Zeta toxin was shown to impair peptidoglycan synthesis by phosphorylating peptidoglycan precursor UDP-N-acetylglucosamine (UNAG).^{5,6} Bacterial Zeta toxins have two highly conserved motifs – one for ATP binding (Walker A motif) and the other one for UNAG binding.^{5,6} Zeta toxin-like domains predicted in HCTV-1 and HCTV-5 proteins have essentially the same secondary structure as bacterial ones with recognizable Walker A motif, but lacking key residues essential for contact with UNAG. In both HCTV-1 and HCTV-5 putative Zeta toxin-like domain is an N-terminal part of an approximately 1200 amino acid protein. In addition to Zeta toxin-like domain, HCTV-5 protein is predicted to have a ParB-like nuclease domain in the central part. HCTV-1 protein lacks this region.

Cluster of QueCDEF genes encoded by HCTV-1 and HCTV-5 viruses.

Both bacteria and archaea share the first step of the queuosine biosynthesis, during which GTP is converted into 7-cyano-7-deazaguanine (Pre-Q₀) with the help of three enzymes – QueC, QueD and QueE. All of these enzymes are predicted to be encoded in HCTV-1 and HCTV-5 genomes. Next step in archaeosine synthesis involves Pre-Q₀ insertion into tRNA and conversion to archaeosine by addition of ammonia by yet unknown enzyme.⁷ In bacteria, Pre-Q₀ is reduced to 7-aminomethyl-7-deazaguanine (pre-Q₁) prior to its addition to tRNA.⁷ This step is catalyzed by the enzyme QueF, which is also predicted to be encoded in HCTV-1 and HCTV-5 genomes. Thus, HCTV-1 and HCTV-5 seem to encode bacterial queuosine synthesis enzymes.

Ribozymes in HRTV-5 and HRTV-8 genomes.

HRTV-5 ribozyme at 65998-66064 nt has identical catalytic site residues as HF1 and HF2 ribozymes, whereas T3 residue is missing in HRTV-8 variant at 63409-63474 nt (Fig. S3). Differently from the above mentioned hammerhead ribozymes, HRTV-8 ribozyme at 65506-

65572 nt has canonical active site residues C3, U4, G5, A6 without 2, 2a, 2b insertion (Fig. S3). It has been shown that 2, 2a, 2b insertion in HF2 ribozyme destabilizes catalytic site thereby decreasing ribozyme activity.⁸ The stability and consequent activity of restored upon increasing Mg²⁺ concentrations.⁸ It was suggested that 2, 2a, 2b insertion acts to regulate ribozyme activity in a salt-dependent manner.⁸ Since HRTV-8 has both variants of hammerhead ribozyme, it may have both Mg²⁺-dependent and independent activities.

HNH homing endonucleases in HCTV-5 genome.

Based on similarity at the amino acid level HCTV-5 HNH homing endonucleases can be divided into three groups. The first group includes ORF 89, 104, 106, 108 and 111 products. The second group is comprised of ORFs 42, 49 and 56, whereas the third one includes ORFs 22 and 77. In order to distinguish the most conserved amino acids, we have done the alignments of HCTV-5 HNH HEs with additional related HNH sequences from non-redundant protein database (Fig. S4). Sequence logos of the alignments (Fig. S5) showed that in addition to the conserved HNH motif, group I members have two conserved cysteine dyads (Cys dyads). In the second Cys dyad of ORF89 histidine substitutes for one of the cysteines residues. The same substitution has been observed in other HNHs.⁹ It was suggested that it may not be functionally detrimental and histidine may still act in zinc ion binding⁹. The II group members have one Cys dyad with less usual 4 nucleotide insertion between the two cysteines. The third group lacks cysteine dyads. Similar combinations of conserved HNH motif and cysteine dyads have been described before for other HNH homing endonucleases.⁹

REFERENCES

1. Casjens S. Prophages and bacterial genomics: what have we learned so far? *Mol Microbiol* 2003; 49:277-300.
2. Krupović M, Forterre P, Bamford DH. Comparative analysis of the mosaic genomes of tailed archaeal viruses and proviruses suggests common themes for virion architecture and assembly with tailed viruses of bacteria. *J Mol Biol* 2010; 397:144-60.
3. Cheng H, Shen N, Pei J, Grishin NV. Double-stranded DNA bacteriophage prohead protease is homologous to herpesvirus protease. *Protein science : a publication of the Protein Society* 2004; 13:2260-9.
4. Chang JR, Spilman MS, Rodenburg CM, Dokland T. Functional domains of the bacteriophage P2 scaffolding protein: identification of residues involved in assembly and protease activity. *Virology* 2009; 384:144-50.
5. Mutschler H, Gebhardt M, Shoeman RL, Meinhart A. A novel mechanism of programmed cell death in bacteria by toxin-antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biol* 2011; 9:e1001033.

6. Mutschler H, Meinhart A. epsilon/zeta systems: their role in resistance, virulence, and their potential for antibiotic development. *J Mol Med (Berl)* 2011; 89:1183-94.
7. Phillips G, El Yacoubi B, Lyons B, Alvarez S, Iwata-Reuyl D, de Crécy-Lagard V. Biosynthesis of 7-deazaguanosine-modified tRNA nucleosides: a new role for GTP cyclohydrolase I. *J Bacteriol* 2008; 190:7876-84.
8. Perreault J, Weinberg Z, Roth A, Popescu O, Chartrand P, Ferbeyre G, et al. Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS Comput Biol* 2011; 7:e1002031.
9. Mehta P, Katta K, Krishnaswamy S. HNH family subclassification leads to identification of commonality in the His-Me endonuclease superfamily. *Protein science : a publication of the Protein Society* 2004; 13:295-300.

SUPPLEMENTARY TABLES

Table S1. Properties of proviral regions

Proviral element	Host (GenBank accession number)	attR (bp) or approximate right hand region coordinates	attL (bp) or approximate left hand region coordinates	Proviral region size (bp)	Integration site
Ngreg	<i>Natronobacterium gregoryi</i> SP2 ,ctg168 contig (NZ_AGIM01000003)	43418-43433	872-887	42,621	tRNA-Cys
Aprof	<i>Archaeoglobus profundus</i> DSM 5631 (NC_013741)	1262636-1262687	1296263-1296314	33,678	Intragenic
Hjeo	<i>Halalkalicoccus jeotgali</i> B3 (NC_014297)	1055217	1001828	ND	ND
Mokin	<i>Methanothermococcus okinawensis</i> IH1 (NC_015636)	312061-312115	340975-341029	28,968	Intergenic

Table S2. Putative frameshifts in predicted tail assembly factor genes

Virus	Frameshift type	Frameshift position (nt)	Slippery sequence
HCTV-1	-1	83350-83356	GGGAAAC
HCTV-2	-1	24643-24649	AAAAAAC
HCTV-5	-1	83612-83618	GGGAAAG

Table S3. HCTV-1 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	F	120	464	345	
2	R	891	1283	393	
3	R	1267	1512	246	
4	R	1502	1765	264	
5	R	1808	2116	309	
6	R	2182	2445	264	
7	R	2445	2690	246	
8	R	2690	3052	363	
9	R	3045	3161	117	
10	R	3154	3480	327	
11	R	3532	3717	186	
12	R	3714	3968	255	
13	R	3961	4158	198	
14	R	4158	4643	486	
15	R	4732	5013	282	
16	R	5080	5217	138	
17	R	5210	5398	189	
18	R	5466	5717	252	
19	R	5710	5952	243	
20	R	5945	6142	198	
21	R	6139	6333	195	
22	F	6409	6738	330	
23	F	6722	6934	213	
24	F	7047	7421	375	
25	F	7418	7810	393	
26	F	7803	10751	2949	Ribonucleotide reductase
27	F	10799	11338	540	
28	F	11335	11721	387	
29	F	11824	13005	1182	
30	F	13114	14388	1275	
31	F	14526	14936	411	
32	F	14936	15073	138	
33	F	15070	15312	243	
34	F	15397	16626	1230	DNA polymerase II, small subunit
35	F	16632	17000	369	
36	R	17041	17637	597	
37	R	17634	17843	210	
38	R	17854	18603	750	Thymidylate synthase
39	F	18734	19180	447	
40	F	19192	20409	1218	
41	F	20411	21961	1551	

42	F	21963	22373	411	
43	R	22500	23240	741	
44	R	23242	23436	195	
45	R	23429	24010	582	
46	R	24013	24324	312	NTP-PPase MazG-like domain
47	R	24321	25772	1452	RtcB
48	R	25769	26458	690	Hef nuclease
49	R	26455	26610	156	
50	R	26603	26869	267	
51	R	26862	27119	258	
52	R	27116	27577	462	
53	R	27570	27956	387	
54	R	27958	28383	426	
55	R	28432	28506	75	tRNA-Gln-ttg
56	R	28638	30827	2190	Von Willebrand factor type A domain
57	R	30920	32257	1338	AAA domain
58	R	32333	32782	450	
59	R	32760	33179	420	
60	R	33181	33642	462	
61	R	33698	34357	660	SprT
62	F	34617	35315	699	
63	R	35319	35534	216	
64	R	35642	36805	1164	
65	R	36807	37433	627	GTP cyclohydrolase I
66	R	37446	38141	696	Integral membrane protein
67	R	38208	38372	165	
68	R	38362	38484	123	
69	R	38486	39274	789	QueE
70	R	39271	39927	657	HNH endonuclease domain
71	R	39924	40367	444	
72	R	40360	40515	156	
73	R	40508	40966	459	QueD
74	R	40959	41726	768	QueC
75	R	41785	42189	405	
76	R	42186	42647	462	QueF
77	R	42735	43859	1125	
78	R	43923	44129	207	
79	R	44218	44463	246	
80	R	44456	44935	480	
81	R	44938	45138	201	
82	R	45107	46369	1263	Replication factor C, large subunit
83	R	46366	46527	162	XapX
84	R	46530	47027	498	N-Acyltransferase domain
85	R	47021	49747	2727	DNA polymerase, elongation subunit (family B)

86	R	49769	50065	297	HNH endonuclease, C terminus
87	R	50068	50865	798	PD-(D/E)XK nuclease
88	R	50862	51173	312	
89	R	51177	51380	204	
90	R	51380	52408	1029	Replication factor C, small subunit
91	R	52436	52621	186	
92	R	52618	52842	225	
93	R	52835	53425	591	TMP kinase
94	R	53428	53904	477	
95	R	53939	55090	1152	
96	R	55112	55969	858	
97	R	55962	56474	513	RadA
98	R	56467	57951	1485	AAA domain
99	R	57954	59045	1092	
100	F	59230	60141	912	RadA
101	R	60145	60297	153	
102	R	60297	60557	261	
103	F	60816	61517	702	
104	F	61514	62962	1449	Terminase, large subunit
105	F	62969	64837	1869	Portal
106	F	64830	67970	3141	
107	F	67967	69994	2028	
108	F	69991	70881	891	Mu gpF-like protein
109	F	70928	74560	3633	Zeta toxin domain
110	F	74790	74951	162	
111	F	74951	76948	1998	
112	F	76960	78342	1383	Major Capsid
113	F	78361	78738	378	
114	F	78756	79058	303	
115	F	79065	79508	444	
116	F	79510	79839	330	
117	F	79839	80510	672	
118	F	80510	81058	549	
119	F	81055	81543	489	
120	F	81540	81785	246	
121	F	81834	82814	981	
122	F	82817	83386	570	Tail Assembly Chaperone
123	F	82817	83544	729	Tail Assembly Chaperone
124	F	83537	83677	141	
125	F	83677	84474	798	
126	F	84476	86338	1863	Tape Measure
127	F	86338	88431	2094	Endopeptidase tail
128	F	88438	89757	1320	
129	F	89757	92894	3138	

130	F	92891	93703	813	
131	R	93678	94103	426	
132	R	94324	94581	258	
133	F	94650	94811	162	
134	F	94808	95239	432	RNAse HI
135	F	95481	96335	855	
136	F	96335	96580	246	
137	F	96577	96801	225	
138	F	96858	96965	108	
139	F	97077	97451	375	
140	F	97448	97621	174	
141	F	97618	97812	195	
142	F	97815	98291	477	
143	F	98288	98527	240	
144	F	98524	99066	543	
145	F	99063	99266	204	
146	R	99257	99904	648	Trimeric dUTPase
147	F	100005	100139	135	
148	F	100132	100254	123	
149	F	100251	100394	144	
150	F	100391	100618	228	
151	F	100615	100812	198	
152	F	100802	100954	153	
153	F	100951	101166	216	
154	F	101156	101311	156	
155	F	101301	101444	144	
156	F	101437	101691	255	
157	F	101684	101869	186	
158	F	101859	101993	135	
159	F	101980	102105	126	
160	F	102089	102250	162	
161	F	102662	102982	321	

Table S4. HCTV-2 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	F	1	426	426	
2	F	423	755	333	
3	F	752	1258	507	
4	F	1262	1468	207	
5	F	1465	2055	591	DnaJ-like protein
6	F	2052	2564	513	

7	F	2554	2751	198	
8	F	2748	3158	411	
9	F	3158	3331	174	
10	F	3367	3609	243	
11	F	3554	5134	1581	ParB-like nuclease and AdoMet-MTase domains
12	F	5141	5782	642	AdoMet-MTase
13	F	5779	6123	345	
14	F	6250	6405	156	
15	F	6398	7009	612	
16	F	6999	7100	102	
17	F	7113	7439	327	
18	F	7439	8254	816	AdoMet-MTase
19	F	8247	8603	357	
20	F	8642	9007	366	
21	F	9138	9461	324	
22	F	9466	11130	1665	Terminase, large subunit
23	F	11207	13567	2361	Portal and Mu gpF-like protein
24	F	13567	13866	300	
25	F	13866	15224	1359	
26	F	15228	15494	267	
27	F	15494	15865	372	SNase-like protein
28	F	15804	16007	204	
29	F	15997	16185	189	
30	F	16116	17564	1449	
31	F	17636	17821	186	
32	F	17932	19533	1602	Prohead protease
33	F	19535	20716	1182	Major Capsid
34	F	20720	20995	276	
35	F	21043	21789	747	
36	F	21792	22346	555	
37	F	22343	22822	480	
38	F	22833	23159	327	
39	F	23162	24142	981	Tail Tube
40	F	24176	24652	477	Tail Assembly Chaperone
41	F	24176	25143	969	Tail Assembly Chaperone
42	F	25147	26454	1308	Tape Measure
43	F	26470	26913	444	
44	F	26921	28390	1470	
45	F	28390	29193	804	
46	F	29196	30191	996	
47	R	30313	31185	873	
48	R	31192	31332	141	
49	R	31396	31521	126	
50	R	31564	32070	507	

51	R	32078	32455	378	
52	R	32455	32625	171	
53	R	32622	32939	318	
54	R	32936	33157	222	
55	R	33154	33408	255	
56	R	33405	33500	96	
57	R	33500	34135	636	
58	F	34405	35997	1593	
59	F	35997	37094	1098	Laminin G domain
60	F	37096	38349	1254	PQQ-like domain
61	F	38346	38834	489	
62	R	38835	39236	402	Holliday junction resolvase
63	R	39236	40048	813	
64	R	40038	40340	303	
65	R	40333	40827	495	
66	R	40829	41005	177	
67	R	40998	41279	282	
68	R	41289	41696	408	
69	R	41689	42003	315	
70	R	42000	42488	489	
71	R	42584	43003	420	
72	R	43000	43230	231	
73	R	43223	45076	1854	
74	R	45070	45414	345	
75	R	45418	46611	1194	
76	R	46618	46776	159	
77	R	46754	46924	171	
78	F	47003	47266	264	
79	F	47270	49624	2355	Von Willebrand factor type A domain
80	F	49663	50199	537	HNH endonuclease domain
81	F	50199	50924	726	
82	F	51055	51264	210	
83	F	51264	52817	1554	AAA domain
84	F	52835	53299	465	
85	F	53361	53540	180	
86	R	53624	54184	561	

Table S5. HCTV-5 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	F	399	620	222	
2	F	653	1138	486	
3	R	1148	1315	168	
4	R	1312	1518	207	

5	R	1515	1832	318	
6	R	1895	2467	573	
7	F	2543	2800	258	
8	R	2806	2979	174	
9	R	3022	3282	261	
10	R	3279	3431	153	
11	R	3424	3708	285	
12	R	3734	3925	192	
13	R	3922	4059	138	
14	R	4056	4223	168	
15	R	4508	4759	252	
16	R	4756	4857	102	
17	R	4854	5339	486	
18	R	5339	5596	258	
19	F	5648	5857	210	
20	F	5966	6292	327	
21	F	6294	6563	270	
22	F	6642	7115	474	HNH endonuclease domain
23	F	7112	7429	318	
24	F	7419	7811	393	
25	F	7814	10792	2979	Ribonucleotide reductase
26	F	10861	11298	438	
27	F	11332	11790	459	
28	F	11783	12055	273	
29	F	12057	12224	168	
30	F	12310	13494	1185	
31	F	13638	14957	1320	
32	F	15001	15504	504	
33	F	15504	15647	144	
34	F	15644	15898	255	
35	F	16016	17245	1230	DNA polymerase II, small subunit
36	R	17305	17688	384	
37	R	17690	18271	582	
38	R	18268	18993	726	
39	R	19032	19985	954	Thymidylate synthase
40	R	20163	20903	741	
41	R	20906	21478	573	
42	R	21480	21962	483	HNH endonuclease domain
43	R	21946	22257	312	NTP-PPase MazG-like domain
44	R	22250	22393	144	
45	R	22390	23850	1461	RtcB
46	R	23843	24073	231	
47	R	24070	24240	171	
48	R	24224	25132	909	AdoMet-MTase

49	R	25125	26303	1179	Hef and HNH endonuclease
50	R	26300	26437	138	
51	R	26427	26579	153	
52	R	26572	26832	261	
53	R	26825	27082	258	
54	R	27079	27543	465	
55	R	27536	27925	390	
56	R	27925	28401	477	HNH endonuclease domain
57	R	28404	28838	435	
58	R	28885	28959	75	tRNA Gln (ttg)
59	R	29049	29210	162	
60	R	29210	31441	2232	Von Willebrand factor type A domain
61	R	31513	31701	189	
62	R	31729	33066	1338	AAA domain
63	R	33140	33601	462	
64	R	33579	34031	453	
65	R	33998	34444	447	
66	R	34502	35164	663	SprT
67	R	35235	35438	204	
68	R	35454	36569	1116	
69	R	36571	37200	630	GTP cyclohydrolase I
70	R	37213	37392	180	
71	R	37385	37513	129	
72	R	37510	38295	786	QueE
73	R	38292	38735	444	
74	R	38728	38862	135	
75	R	38862	39320	459	QueD
76	R	39317	40081	765	QueC
77	R	40139	40621	483	HNH endonuclease domain
78	R	40634	41050	417	
79	R	41050	41265	216	
80	R	41270	41611	342	
81	R	41611	42072	462	QueF
82	R	42116	42601	486	LAGLIDADG endonuclease
83	R	42598	42732	135	
84	R	43038	43892	855	AdoMet-MTase
85	R	43895	45139	1245	Replication factor C, large subunit
86	R	45132	45329	198	
87	R	45332	45829	498	N-Acyltransferase domain
88	R	45823	48564	2742	DNA polymerase, elongation subunit (family B)
89	R	48586	49197	612	HNH endonuclease domain
90	R	49197	49994	798	PD-(D/E)XK nuclease
91	R	49987	50304	318	
92	R	50306	50494	189	

93	R	50491	51474	984	Replication factor C, small subunit
94	R	51541	52131	591	TMP kinase
95	R	52134	52586	453	
96	R	52588	53739	1152	
97	R	53754	54611	858	
98	R	54604	55101	498	RadA
99	R	55094	56614	1521	
100	R	56611	57654	1044	
101	F	57888	58799	912	RadA
102	F	58893	59141	249	
103	F	59388	60050	663	
104	F	60047	60604	558	HNH endonuclease domain
105	F	60594	62039	1446	Terminase, large subunit
106	F	62033	62620	588	HNH endonuclease domain
107	F	62624	64486	1863	Portal
108	F	64479	65225	747	HNH endonuclease domain
109	F	65222	66190	969	ADP-ribosyltransferase
110	F	66183	67196	1014	Mu gpF-like protein
111	F	67233	67976	744	HNH endonuclease domain
112	F	67992	69929	1938	Sialidase
113	R	69989	70690	702	Integrase
114	F	70811	74692	3882	Zeta toxin and ParB nuclease domains
115	R	74693	75166	474	LAGLIDADG endonuclease
116	R	75365	75532	168	
117	F	75392	75583	192	
118	F	75620	77494	1875	
119	F	77510	78883	1374	Major Capsid
120	F	78902	79279	378	
121	F	79298	79630	333	
122	F	79637	80050	414	
123	F	80053	80385	333	
124	F	80385	81065	681	
125	F	81065	81676	612	
126	F	81676	82146	471	
127	F	82164	83177	1014	Tail Tube
128	F	83232	83648	417	Tail Assembly Chaperone
129	F	83232	83986	756	Tail Assembly Chaperone
130	F	83988	85430	1443	Tape Measure
131	F	85427	85882	456	
132	F	85889	87319	1431	
133	F	87312	88094	783	
134	F	88094	89086	993	
135	F	89100	90596	1497	
136	F	90599	91708	1110	Laminin G domain

137	F	91710	92423	714	
138	F	92479	92655	177	
139	R	92727	92918	192	
140	R	92911	93267	357	
141	F	93351	93539	189	
142	F	93536	93967	432	RNase HI
143	R	93964	94341	378	
144	F	94594	95460	867	
145	F	95457	95696	240	
146	F	95728	95919	192	
147	F	95910	96092	183	
148	F	96085	96348	264	
149	F	96350	96535	186	
150	F	96532	96918	387	
151	F	96915	97352	438	
152	F	97349	97558	210	
153	R	97555	98196	642	Trimeric dUTPase
154	F	98298	98627	330	
155	F	98624	98857	234	
156	F	98835	99035	201	
157	F	99029	99202	174	
158	F	99199	99390	192	
159	F	99387	99578	192	
160	F	99571	99726	156	
161	F	99713	99874	162	
162	F	99890	100111	222	
163	F	100164	100643	480	
164	F	100646	101233	588	GIY-YIG nuclease
165	R	101287	101430	144	
166	R	101427	101741	315	
167	R	101738	101890	153	

Table S6. HGTV-1 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	F	1	114	114	
2	F	162	764	603	
3	F	961	1569	609	
4	F	1572	1940	369	
5	F	1933	3246	1314	Terminase, large subunit
6	F	3250	3603	354	
7	F	3605	5821	2217	Portal and Mu gpF-like protein

8	F	5847	6587	741	
9	F	6602	6937	336	
10	F	6949	8727	1779	
11	F	8822	11488	2667	Prohead protease
12	F	11506	11976	471	
13	F	11995	12831	837	Major Capsid
14	F	12894	13319	426	
15	F	13320	13838	519	
16	F	13842	14402	561	
17	F	14463	15818	1356	Tail Sheath
18	F	15830	16231	402	
19	F	16397	16843	447	
20	F	16852	17058	207	
21	F	17062	20013	2952	Tape Measure
22	F	20021	20650	630	
23	F	20650	21900	1251	Late control GPD
24	F	21903	22565	663	Baseplate V-like protein
25	F	22649	23026	378	
26	F	23023	24432	1410	Baseplate J-like protein
27	F	24433	25599	1167	
28	F	25651	26217	567	
29	F	26256	27323	1068	
30	F	27362	27550	189	
31	F	27560	28534	975	
32	F	28531	29358	828	
33	F	29361	30197	837	
34	F	30256	30792	537	
35	F	30785	31027	243	
36	F	31024	31194	171	
37	F	31200	31589	390	
38	R	31623	32231	609	RadA
39	F	32330	33073	744	PD-(D/E)XK nuclease
40	F	33063	33374	312	
41	F	33401	34459	1059	Carbamoyl phosphate synthase, small subunit
42	R	34444	35016	573	Holliday junction resolvase
43	R	35019	35687	669	RecB-like nuclease
44	R	35680	36099	420	
45	R	36154	36516	363	
46	R	36524	36751	228	
47	R	36753	36947	195	
48	R	36944	37105	162	
49	R	37147	40026	2880	DNA polymerase, elongation subunit (family B)
50	R	40035	40826	792	LAGLIDADG endonuclease
51	R	40869	41096	228	

52	R	41136	41585	450	
53	R	41582	41788	207	
54	R	41792	42220	429	
55	R	42266	42352	87	tRNA Ser (gga)
56	R	42353	42595	243	
57	R	42609	42683	75	tRNA Val (gac)
58	R	42688	42762	75	tRNA Val (tac)
59	R	42770	42840	71	tRNA Pro (tgg)
60	R	42885	43220	336	
61	R	43249	43431	183	
62	R	43406	43648	243	
63	R	43645	44151	507	
64	R	44129	44374	246	
65	R	44371	44673	303	
66	R	44675	44815	141	
67	R	44841	45485	645	AdoMet-MTase
68	R	45441	45526	86	tRNA Leu (taa)
69	R	45530	45613	84	tRNA Leu (caa)
70	R	45617	45700	84	tRNA Leu (gag)
71	R	45786	45869	84	tRNA Leu (tag)
72	R	45900	46352	453	
73	R	46355	46430	76	tRNA Arg (tct)
74	R	46465	46698	234	
75	R	46701	46774	74	tRNA Arg (tcg)
76	R	46894	47427	534	
77	R	47537	47609	73	tRNA Cys (gca)
78	R	47619	47909	291	
79	R	48011	48085	75	tRNA Ile (gat)
80	R	48088	48162	75	tRNA Gln (ctg)
81	R	48250	48320	71	tRNA Ile (gat)
82	R	48474	48544	71	tRNA Trp (cca)
83	R	48548	48620	73	tRNA Gln (ttg)
84	R	48804	48876	73	tRNA Lys (ttt)
85	R	48877	48952	76	tRNA Thr (ggt)
86	R	48955	49029	75	tRNA Thr (cgt)
87	R	49031	49103	73	tRNA Thr (tgt)
88	R	49226	49328	103	tRNA Ala (tgc) + Intron
89	R	49346	49759	414	
90	R	50108	50193	86	tRNA Met (cat)
91	R	50195	50271	77	tRNA Lys (ctt)
92	R	50274	50347	74	tRNA Tyr (gta)
93	R	50351	50423	73	tRNA Asn (gtt)
94	R	50474	50544	71	tRNA Gly (gcc)
95	R	50549	50619	71	tRNA Gly (tcc)

96	R	50624	50697	74	tRNA His (gtg)
97	R	50742	50951	210	
98	R	50948	51316	369	
99	R	51342	51551	210	
100	R	51556	51632	77	tRNA Glu (ctc)
101	R	51634	51710	77	tRNA Phe (gaa)
102	R	51714	52079	366	
103	R	52305	52375	71	tRNA Asp (gtc)
104	R	52408	52998	591	LAGLIDADG endonuclease
105	R	53348	53420	73	tRNA Arg (gcg)
106	R	53450	54052	603	TMP kinase
107	F	54188	54736	549	Trimeric dUTPase
108	F	54838	57198	2361	Ribonucleotide reductase
109	F	57198	57290	93	
110	F	57333	57485	153	
111	F	57528	57665	138	zinc-ribbon domain
112	R	57662	58156	495	DUF488
113	R	58197	58607	411	
114	R	58604	58843	240	
115	R	58840	59145	306	
116	R	59145	59360	216	
117	R	59357	60160	804	Thymidylate synthase
118	R	60157	60420	264	
119	R	60459	60701	243	
120	R	60727	60999	273	
121	R	60996	61343	348	
122	R	61383	61892	510	
123	R	61948	62277	330	
124	R	62277	64100	1824	
125	R	64131	64574	444	
126	R	64543	65424	882	Nucleotidyltransferase
127	R	65450	65752	303	
128	R	65749	66069	321	
129	R	66069	67598	1530	TROVE domain
130	R	67732	67878	147	
131	R	67871	68254	384	
132	R	68224	68304	81	
133	R	68447	68749	303	
134	R	68827	69825	999	
135	R	70053	70253	201	
136	R	70440	70571	132	tRNA Asp (gtc) + Intron
137	R	70671	71060	390	
138	R	71053	71307	255	
139	R	71304	71675	372	Cysteine hydrolase

140	R	71725	73179	1455	RtcB
141	R	73235	73597	363	
142	R	73597	73701	105	
143	R	73742	74392	651	
144	R	74414	74527	114	
145	R	74537	75004	468	RNAse HI
146	R	75001	75240	240	
147	R	75237	75578	342	
148	F	75620	75811	192	
149	R	75902	76192	291	
150	R	76185	76457	273	
151	R	76454	76795	342	
152	R	76804	77067	264	
153	R	77073	77579	507	
154	R	77582	77812	231	
155	R	77809	78012	204	
156	R	77981	78295	315	
157	R	78292	78669	378	
158	R	78701	78838	138	
159	R	78878	79201	324	
160	R	79230	79553	324	
161	R	79550	79906	357	
162	R	79908	80237	330	
163	R	80360	80890	531	
164	R	80890	81093	204	
165	R	81090	81323	234	
166	R	81310	81951	642	LAGLIDADG endonuclease
167	R	81995	82777	783	
168	R	82809	83153	345	
169	R	83193	83456	264	
170	R	83447	83605	159	
171	R	83602	84276	675	
172	R	84276	85865	1590	HSP90 chaperone
173	R	85894	86187	294	
174	R	86212	86478	267	Thioredoxin
175	R	86505	86810	306	
176	R	86836	87066	231	
177	R	87152	87391	240	
178	R	87384	87701	318	
179	R	87694	88098	405	Ser/Thr and Tyr protein phosphatase
180	R	88138	89007	870	
181	R	89040	89633	594	Restriction endonuclease EcoRV
182	R	89623	89928	306	
183	R	89925	90212	288	

184	R	90244	90534	291	
185	R	90565	90795	231	
186	R	90799	90933	135	
187	R	90971	91315	345	
188	R	91312	91488	177	
189	R	91513	91908	396	PD-(D/E)XKnuclase
190	R	91926	92150	225	
191	F	92225	92335	111	
192	F	92332	92568	237	
193	F	92595	92933	339	
194	F	92930	93205	276	
195	F	93237	93554	318	
196	F	93551	93928	378	
197	F	93946	95073	1128	class I lysyl tRNA synthetase
198	F	95098	95406	309	
199	F	95390	95665	276	
200	F	95658	95888	231	
201	F	95911	96576	666	
202	F	96593	96730	138	
203	F	96766	99378	2613	
204	F	99413	99553	141	
205	F	99546	100172	627	LAGLIDADG endonuclease
206	F	100220	100471	252	
207	F	100513	101529	1017	
208	F	101568	102287	720	Metallophosphatase
209	F	102284	102541	258	
210	F	102531	102803	273	
211	F	102865	103083	219	
212	F	103094	103816	723	Endonuclease III
213	F	103818	104048	231	
214	F	104102	104284	183	
215	F	104316	104990	675	
216	F	104990	105136	147	
217	F	105172	106134	963	RNA ligase
218	F	106131	106391	261	
219	F	106396	106815	420	
220	F	106819	107265	447	
221	F	107272	107577	306	
222	F	107574	107822	249	
223	F	107844	109271	1428	DNA ligase
224	F	109307	109486	180	
225	F	109617	109754	138	
226	F	109751	110017	267	
227	F	110056	110256	201	

228	F	110282	110683	402	
229	F	110692	111594	903	SPFH/Band 7 domain
230	F	111681	111845	165	
231	F	111875	112333	459	
232	F	112330	112491	162	
233	F	112501	113565	1065	HNH endonuclease domain
234	F	113585	113716	132	
235	F	113713	113985	273	
236	F	113989	114933	945	AdoMet-MTase
237	F	114926	115144	219	
238	F	115147	115398	252	
239	F	115395	115589	195	
240	F	115610	116092	483	
241	F	116127	116897	771	SprT
242	F	116909	118210	1302	tRNA adenylyltransferase
243	F	118232	118465	234	
244	F	118475	119101	627	
245	F	119145	119291	147	
246	F	119281	119862	582	zinc finger
247	F	119884	120456	573	
248	F	120461	120613	153	
249	F	120615	121022	408	DNA polymerase IV (family X), NT domain
250	F	121047	121265	219	
251	F	121266	121499	234	
252	F	121525	121788	264	
253	F	121793	122053	261	
254	F	122050	122274	225	
255	F	122309	122551	243	
256	F	122551	122820	270	
257	F	122855	123091	237	
258	F	123091	123330	240	
259	F	123361	123825	465	
260	F	123864	124262	399	
261	F	124342	124503	162	
262	F	124528	124704	177	
263	F	124704	124985	282	
264	F	125008	125205	198	
265	F	125230	125382	153	
266	F	125445	125654	210	
267	F	125864	125953	90	
268	F	126063	126158	96	
269	F	126293	126421	129	
270	F	126438	126656	219	
271	F	126695	126976	282	

272	F	127022	127396	375	
273	F	127533	127754	222	
274	F	127803	127976	174	
275	R	128086	128202	117	
276	R	128372	128620	249	
277	R	128755	129183	429	
278	R	129226	129468	243	
279	R	129604	130428	825	
280	R	130462	130713	252	
281	R	130739	131074	336	
282	R	131256	131378	123	
283	R	131421	131786	366	
284	R	131899	132057	159	
285	R	132045	132311	267	
286	R	132449	132754	306	
287	R	132889	133302	414	
288	R	133345	133470	126	
289	R	133597	133920	324	
290	R	134032	134178	147	
291	R	134175	134510	336	
292	R	134655	134870	216	
293	R	134881	134991	111	
294	R	135168	135851	684	Endonuclease III
295	R	135988	136401	414	
296	R	136567	137034	468	
297	R	137166	137558	393	
298	R	137680	137946	267	
299	R	138080	138325	246	
300	R	138450	138887	438	
301	R	139018	139137	120	
302	R	139137	139430	294	
303	R	139464	139673	210	
304	R	139678	139761	84	
305	R	139834	140121	288	
306	R	140281	140463	183	
307	R	140592	140852	261	
308	R	140988	141239	252	
309	R	141425	141697	273	
310	R	141715	141873	159	
311	R	141915	142055	141	
312	R	142080	142283	204	
313	R	142387	143091	705	
314	R	143417	143635	219	
315	R	48164	48246	83	tRNA Ser (gct)

316	R	48384	48468	85	tRNA Ser (cga)
317	R	52081	52168	88	tRNA Ala (agc)

Table S7. HHTV-1 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	F	1	594	594	
2	F	594	818	225	
3	F	815	1495	681	
4	F	1492	1821	330	
5	F	1818	2087	270	
6	F	2155	2457	303	
7	F	2447	2848	402	
8	F	2848	3039	192	
9	F	3036	4901	1866	Terminase, large subunit
10	F	4947	6284	1338	
11	F	6366	6557	192	
12	F	6605	9439	2835	
13	F	9436	12738	3303	
14	F	12735	13022	288	
15	F	13022	14008	987	Mu gpF-like protein
16	F	14114	14566	453	
17	F	14563	14904	342	
18	R	14886	15119	234	
19	F	15199	16659	1461	
20	F	16659	17066	408	
21	F	17073	18173	1101	Major Capsid
22	F	18179	18595	417	
23	F	18602	18958	357	
24	F	18958	19437	480	
25	F	19430	19957	528	
26	F	19957	20505	549	
27	F	20512	20985	474	
28	F	20986	22005	1020	
29	F	22013	22762	750	
30	R	22832	22990	159	
31	R	23156	23311	156	
32	R	23314	23928	615	
33	F	24105	24617	513	
34	F	24625	25020	396	
35	F	25013	27325	2313	Tape Measure
36	F	27332	29917	2586	Laminin G domain
37	F	29925	31043	1119	
38	F	31107	31847	741	

39	F	31844	32461	618	
40	F	32482	34218	1737	
41	F	34312	34575	264	
42	F	34585	34842	258	
43	F	35069	35248	180	
44	F	35245	36072	828	
45	F	36225	36338	114	
46	F	36335	36532	198	
47	F	36529	36783	255	
48	F	36776	37126	351	
49	F	37126	37317	192	
50	F	37320	37475	156	
51	F	37472	37669	198	
52	F	37802	39052	1251	
53	F	39049	39399	351	
54	F	39482	40285	804	PCNA
55	F	40408	40647	240	
56	F	40640	41245	606	
57	F	41238	41399	162	
58	F	41392	41808	417	GatB domain
59	F	41805	41930	126	
60	F	41927	42586	660	
61	F	42583	42978	396	
62	F	42971	43408	438	
63	F	43398	43625	228	
64	F	43625	44113	489	
65	F	44106	44531	426	Holliday junction resolvase
66	F	44518	44736	219	
67	F	44733	45263	531	DUF4326
68	F	45263	45796	534	
69	F	45793	46326	534	
70	F	46323	46637	315	
71	F	46732	46893	162	
72	F	46924	47712	789	
73	F	47712	48311	600	AdoMet-MTase
74	F	48308	48622	315	

Table S8. HHTV-2 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	F	1	333	333	
2	F	416	1483	1068	

3	F	1703	2989	1287	Von Willebrand factor type A domain
4	F	3024	3308	285	
5	F	3305	3607	303	
6	F	3604	3894	291	
7	F	3891	4286	396	
8	F	4283	5626	1344	
9	F	5648	5941	294	
10	F	5983	7365	1383	AAA domain
11	F	7504	7845	342	
12	F	7842	8186	345	
13	F	8183	8539	357	
14	F	8517	8720	204	
15	F	8734	9123	390	
16	F	9120	9608	489	
17	F	9850	10272	423	
18	F	10259	10465	207	
19	F	10465	10617	153	
20	F	10614	11072	459	
21	F	11128	11937	810	ParB-like nuclease domain
22	F	11927	12790	864	
23	F	12973	13530	558	
24	F	13523	13645	123	
25	F	13638	14093	456	
26	F	14090	14389	300	
27	F	14389	14523	135	
28	F	14520	14936	417	
29	F	15020	15235	216	
30	F	15232	15666	435	
31	F	15670	17364	1695	Terminase, large subunit
32	F	17425	19815	2391	Portal and Mu gpF-like protein
33	F	19815	20069	255	
34	F	20070	21443	1374	
35	F	21449	21730	282	
36	F	21723	22085	363	SNase-like protein
37	F	22336	23463	1128	
38	F	23711	25339	1629	Prohead protease
39	F	25341	26525	1185	Major Capsid
40	F	26537	26857	321	
41	F	26902	27402	501	
42	F	27408	28187	780	
43	F	28187	28783	597	
44	F	28780	29250	471	
45	F	29258	29569	312	
46	F	29578	30549	972	Tail Tube

47	F	30609	31067	459	
48	F	31076	31552	477	
49	F	31552	32811	1260	Tape Measure
50	F	32814	33272	459	
51	F	33279	34730	1452	
52	F	34730	35509	780	
53	F	35510	36502	993	
54	R	36551	36859	309	
55	R	36846	37040	195	
56	R	37033	37401	369	
57	R	37398	37622	225	
58	R	37726	37914	189	
59	R	37911	38162	252	
60	R	38159	38392	234	
61	R	38397	38639	243	
62	F	38916	40463	1548	
63	F	40463	41593	1131	Laminin G domain
64	F	41595	42368	774	
65	F	42373	42846	474	
66	R	42843	43247	405	Holliday junction resolvase
67	R	43240	43947	708	
68	R	43940	44155	216	
69	R	44148	44600	453	
70	R	44597	44749	153	
71	R	44730	45512	783	
72	R	45502	45777	276	
73	R	45770	46000	231	
74	R	45963	46334	372	
75	R	46304	46486	183	
76	R	46476	46829	354	
77	R	46822	46983	162	
78	R	46976	47242	267	
79	R	47332	47517	186	
80	R	47507	47794	288	
81	R	47787	48692	906	SPFH/Band 7 domain
82	R	48689	48991	303	HTH domain
83	R	48984	49418	435	
84	R	49415	49639	225	
85	R	49632	50516	885	
86	R	50503	50724	222	
87	R	50721	51344	624	
88	R	51344	52507	1164	

Table S9. HRTV-4 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	F	1	198	198	
2	F	195	380	186	
3	F	377	493	117	
4	F	614	862	249	
5	F	855	1487	633	
6	F	1484	1807	324	
7	F	1849	2388	540	
8	F	2563	2877	315	
9	F	2874	3125	252	
10	F	3118	3690	573	
11	F	3687	4262	576	
12	F	4259	4570	312	
13	F	4567	5283	717	HTH domain
14	F	5280	5504	225	
15	F	5501	5731	231	
16	F	5724	5960	237	
17	F	5953	6102	150	
18	F	6099	6497	399	
19	F	6494	7036	543	
20	F	7073	7579	507	
21	F	7579	8919	1341	Terminase, large subunit
22	F	8916	9197	282	
23	F	9225	9446	222	
24	F	9446	9580	135	
25	F	9583	9741	159	
26	F	9788	10003	216	
27	F	10000	10119	120	
28	F	10112	10210	99	
29	F	10207	10389	183	
30	F	10427	12916	2490	Portal and Mu gpF-like protein
31	F	12963	13097	135	
32	F	13090	13452	363	
33	F	13445	13618	174	
34	F	13693	14040	348	
35	F	14030	14152	123	
36	F	14190	14387	198	Endonuclease domain
37	F	14387	15970	1584	
38	F	16035	16442	408	
39	F	16442	17701	1260	Major Capsid
40	F	17783	18406	624	HNH endonuclease domain
41	F	18403	18786	384	

42	F	18876	19409	534	
43	F	19453	19698	246	
44	F	19701	19865	165	
45	F	19862	20077	216	
46	F	20074	20307	234	
47	F	20350	21825	1476	HTH domain
48	F	21943	22158	216	
49	F	22199	22624	426	
50	F	22624	23043	420	DUF646
51	F	23090	23236	147	
52	F	23233	23610	378	
53	F	23603	23866	264	
54	F	23859	24128	270	
55	F	24165	24377	213	
56	F	24356	24517	162	
57	F	24593	25102	510	
58	F	25121	25522	402	
59	F	25525	25845	321	
60	F	25845	26378	534	
61	F	26375	26587	213	
62	F	26696	27019	324	
63	F	27057	27527	471	Tail Tube
64	F	27531	27998	468	
65	F	27967	28140	174	
66	F	28182	28475	294	
67	F	28517	28729	213	
68	F	28777	29784	1008	Tape measure
69	F	29785	30636	852	
70	F	30687	31139	453	
71	F	31142	32074	933	
72	F	32078	35350	3273	Fibronectin type 3 domain
73	R	35347	35646	300	RelE

Table S10. HRTV-5 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	R	363	845	483	
2	F	1241	2017	777	
3	F	2010	2273	264	
4	F	2336	2491	156	
5	F	2494	2688	195	
6	F	2726	2932	207	

7	F	2936	3007	72	tRNA Pro (tgg)
8	F	3024	4715	1692	Terminase, large subunit
9	F	4730	6169	1440	Portal
10	F	6172	6456	285	
11	F	6453	7217	765	Prohead protease
12	F	7224	7652	429	
13	F	7675	8190	516	
14	F	8232	9188	957	Major Capsid
15	F	9234	9677	444	
16	F	9677	10015	339	
17	F	10018	13134	3117	
18	F	13136	13357	222	
19	F	13394	13903	510	Tail Tube
20	F	13912	15204	1293	Tail Sheath
21	F	15233	15778	546	
22	F	15835	16305	471	
23	F	16253	16615	363	
24	F	16618	18435	1818	Tape Measure
25	F	18436	18966	531	
26	F	18963	20405	1443	
27	F	20407	20844	438	
28	F	20841	21206	366	
29	F	21271	21597	327	
30	F	21594	22829	1236	Baseplate J-like protein
31	F	22834	23484	651	
32	F	23488	24897	1410	
33	F	24912	26117	1206	
34	F	26119	26502	384	
35	F	26524	27015	492	
36	R	27031	27873	843	
37	F	27934	28059	126	
38	F	28302	30722	2421	Ribonucleotide reductase
39	R	30729	31256	528	
40	R	31274	32041	768	
41	R	32132	32299	168	
42	R	32312	36178	3867	Restriction endonuclease and methylase domains
43	R	36276	36476	201	
44	F	36667	36738	72	tRNA Thr (ggt)
45	F	36922	38115	1194	Integrase
46	R	38116	38649	534	
47	R	38677	40038	1362	RtcB
48	R	40285	40497	213	
49	R	40494	40751	258	
50	R	40748	40933	186	

51	R	40930	41358	429	
52	R	41345	41902	558	Trimeric dUTPase
53	R	41899	42603	705	Thymidylate synthase
54	R	42600	42767	168	
55	R	42857	45448	2592	DNA polymerase, elongation subunit (family B)
56	R	45539	46012	474	
57	R	46005	46145	141	
58	R	46194	47108	915	
59	R	47122	47295	174	
60	R	47292	47411	120	
61	R	47404	47604	201	
62	R	47604	48620	1017	PD-(D/E)XK nuclease
63	R	48613	48771	159	
64	R	48768	48929	162	
65	R	48926	49207	282	
66	R	49191	49577	387	
67	R	49596	49946	351	
68	R	49948	50373	426	
69	R	50462	51643	1182	DNA polymerase II, small subunit
70	R	51640	52767	1128	
71	R	52764	52976	213	
72	R	53001	54482	1482	
73	R	54560	54811	252	
74	R	54843	55790	948	
75	R	55871	55944	74	tRNA Asn (gtt)
76	R	56107	56319	213	
77	R	56316	56447	132	
78	R	56448	56858	411	
79	R	57040	57174	135	
80	R	57176	58048	873	SPFH/Band 7 domain
81	R	58092	58168	77	tRNA Arg (tct)
82	F	58365	58916	552	
83	R	58913	59398	486	
84	F	59689	60168	480	
85	F	60165	60629	465	Endonuclease
86	R	60940	61128	189	
87	R	61121	61453	333	
88	R	61450	61746	297	
89	R	61743	62195	453	
90	R	62192	62467	276	
91	R	62464	62895	432	
92	R	62895	63323	429	HTH domain
93	R	63328	63462	135	
94	R	63473	63739	267	

95	R	63740	64138	399	
96	R	64188	65984	1797	Rad3-related DNA helicase
97	R	65988	66080	93	
98	R	66065	66238	174	
99	R	66239	66559	321	
100	R	66565	67326	762	
101	R	67357	68208	852	
102	R	68225	68413	189	
103	R	68413	68589	177	DNA ligase
104	R	68589	68741	153	
105	R	68732	68878	147	
106	R	68871	69140	270	
107	R	69137	69502	366	
108	R	69521	70189	669	
109	R	70189	70308	120	
110	R	70308	70556	249	
111	R	70560	70667	108	
112	R	70667	70942	276	
113	R	70939	71082	144	
114	R	71079	71300	222	
115	R	71313	71453	141	
116	R	71461	71829	369	
117	R	71855	72214	360	
118	R	72315	72443	129	
119	R	72534	72893	360	
120	R	72890	73219	330	
121	R	73322	74278	957	
122	F	75714	76049	336	

Table S11. HRTV-7 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	F	898	1785	888	
2	F	1778	2065	288	
3	F	2072	2194	123	
4	F	2187	2516	330	
5	F	2545	2772	228	
6	F	2769	2978	210	
7	F	2991	4763	1773	Terminase, large subunit
8	F	4770	6287	1518	Portal
9	F	6290	6574	285	
10	F	6571	7731	1161	Prohead protease

11	F	7740	8213	474	
12	F	8217	8753	537	
13	F	8800	9774	975	Major Capsid
14	F	9788	10264	477	
15	F	10251	10640	390	
16	F	10647	12119	1473	
17	F	12109	12339	231	
18	F	12340	12555	216	
19	F	12559	13071	513	Tail Tube
20	F	13079	14392	1314	Tail Sheath
21	F	14421	14909	489	
22	F	14949	15482	534	
23	F	15445	15780	336	
24	F	15784	17748	1965	Tape Measure
25	F	17748	18275	528	
26	F	18280	19800	1521	
27	F	19849	20259	411	
28	F	20259	20624	366	
29	F	20626	20955	330	
30	F	20957	22246	1290	Baseplate J-like protein
31	F	22252	22944	693	
32	F	22948	24363	1416	
33	F	24366	25355	990	
34	F	25355	25699	345	
35	F	25701	26207	507	
36	F	26222	26356	135	
37	R	26517	26654	138	
38	R	26718	27608	891	
39	R	28026	28757	732	Thymidylate synthase
40	R	28966	29082	117	
41	R	29060	29251	192	
42	R	29251	31902	2652	DNA polymerase, elongation subunit (family B)
43	R	31899	32330	432	
44	R	32297	32686	390	
45	R	32683	32868	186	
46	R	32961	33716	756	AAA domain
47	R	33716	33874	159	
48	R	33983	34432	450	Ser/Thr and Tyr protein phosphatase
49	R	34429	34551	123	
50	R	34586	35251	666	HNH endonuclease domain
51	R	35248	36309	1062	PD-(D/E)XK nuclease
52	R	36302	36541	240	
53	R	36534	36698	165	
54	R	36688	36906	219	

55	R	36906	37052	147	
56	R	37042	37455	414	
57	R	37457	37798	342	
58	R	37887	38267	381	
59	R	38269	38595	327	
60	R	38702	39964	1263	DNA polymerase II, small subunit
61	R	40029	41291	1263	
62	R	41288	41479	192	
63	R	41476	42957	1482	Winged-helix domain
64	R	43036	43335	300	
65	R	43328	43750	423	
66	R	43750	44847	1098	
67	R	45281	45787	507	
68	R	45813	45885	73	tRNA Gln (ttg)
69	F	46461	47069	609	
70	R	47066	47554	489	
71	F	47897	48415	519	
72	F	48415	48900	486	ERCC4-type nuclease
73	F	49220	50383	1164	Integrase
74	R	50380	50652	273	
75	R	50649	50834	186	
76	R	50831	51253	423	HTH domain
77	R	51250	51471	222	
78	R	51468	51647	180	
79	R	51652	52083	432	
80	R	52076	52333	258	
81	R	52380	52679	300	
82	R	52750	52914	165	
83	R	52911	53153	243	
84	R	53160	54584	1425	
85	R	54610	55254	645	LAGLIDADG endonuclease
86	R	55299	55493	195	
87	R	55490	55669	180	
88	R	55666	55875	210	
89	R	55872	56330	459	
90	R	56335	56631	297	
91	R	56771	57679	909	
92	R	57776	58465	690	
93	R	58458	58559	102	
94	R	58561	61401	2841	HSP90 chaperone
95	R	61391	61705	315	
96	R	61769	61942	174	
97	R	61942	62736	795	
98	R	62733	62969	237	

99	R	62966	63256	291	
100	R	63360	63542	183	
101	R	63543	63866	324	
102	R	63859	64137	279	
103	R	64137	64595	459	
104	R	64631	65680	1050	
105	R	66025	66168	144	
106	R	67667	68428	762	

Table S12. HRTV-8 genome features.

ORF number	Direction	Start	Stop	Length (nt)	Function
1	R	394	945	552	
2	F	1305	2081	777	
3	F	2083	2337	255	
4	F	2400	2555	156	
5	F	2558	2752	195	
6	F	2789	2995	207	
7	F	3000	3074	75	tRNA Pro (tgg)
8	F	3088	4785	1698	Terminase, large subunit
9	F	4800	6239	1440	Portal
10	F	6239	6529	291	
11	F	6526	7287	762	Prohead protease
12	F	7294	7725	432	
13	F	7749	8264	516	
14	F	8306	9262	957	Major Capsid
15	F	9308	9748	441	
16	F	9748	10092	345	
17	F	10095	13208	3114	
18	F	13210	13431	222	
19	F	13468	13977	510	Tail Tube
20	F	13987	15279	1293	Tail Sheath
21	F	15308	15853	546	
22	F	15908	16378	471	
23	F	16455	16691	237	
24	F	16694	18514	1821	Tape Measure
25	F	18515	19078	564	
26	F	19075	20508	1434	
27	F	20511	20951	441	
28	F	20948	21310	363	
29	F	21375	21701	327	
30	F	21698	22933	1236	Baseplate J-like protein

31	F	22938	23591	654	
32	F	23592	25001	1410	
33	F	25016	26221	1206	
34	F	26223	26606	384	
35	F	26628	27119	492	
36	R	27135	27971	837	
37	F	28038	28163	126	
38	F	28401	30818	2418	Ribonucleotide reductase
39	R	30916	31104	189	
40	F	31306	31376	71	tRNA Thr (ggt)
41	F	31560	32753	1194	Integrase
42	R	32755	33240	486	
43	R	33371	33661	291	HNH endonuclease
44	R	33663	35027	1365	RtcB
45	R	35275	35487	213	
46	R	35477	35671	195	
47	R	35668	35880	213	
48	R	35877	36308	432	
49	R	36305	36853	549	Trimeric dUTPase
50	R	36850	37554	705	Thymidylate synthase
51	R	37557	37745	189	
52	R	37738	37902	165	
53	R	37996	38376	381	
54	R	38456	38554	99	
55	R	38551	38727	177	
56	R	38736	41327	2592	DNA polymerase, elongation subunit (family B)
57	R	41418	41891	474	
58	R	42073	42987	915	
59	R	43001	43174	174	
60	R	43171	43299	129	
61	R	43292	43492	201	
62	R	43492	43704	213	
63	R	43697	43876	180	
64	R	43876	44004	129	
65	R	44004	45020	1017	PD-(D/E)XK nuclease
66	R	45001	45171	171	
67	R	45168	45329	162	
68	R	45326	45607	282	
69	R	45591	45977	387	
70	R	45993	46343	351	
71	R	46345	46770	426	
72	R	46859	48037	1179	archeal DNA polymerase II, small subunit
73	R	48034	49272	1239	
74	R	49398	50882	1485	

75	R	50960	51205	246	
76	R	51237	52190	954	
77	R	52419	52496	78	tRNA Asn (gtt)
78	R	52659	52868	210	
79	R	52865	52996	132	
80	R	52997	53539	543	
81	R	53632	53707	76	tRNA Gly (tcc)
82	R	53726	53860	135	
83	R	53862	54929	1068	SPFH/Band 7 domain
84	F	55034	55585	552	
85	R	55582	56067	486	
86	F	56357	56836	480	
87	F	56833	57297	465	Endonuclease
88	F	57338	57799	462	
89	R	57811	57972	162	
90	R	57965	58303	339	
91	R	58303	58599	297	
92	R	58596	59120	525	
93	R	59154	59435	282	
94	R	59432	59569	138	
95	R	59573	59998	426	
96	R	60001	60423	423	HTH domain
97	R	60423	60599	177	
98	R	60596	60733	138	
99	R	60733	60882	150	
100	R	60884	61150	267	
101	R	61151	61549	399	
102	R	61599	63395	1797	Rad3-related DNA helicase
103	R	63475	63684	210	
104	R	63875	64027	153	
105	R	64027	64356	330	
106	R	64364	65017	654	
107	R	65025	65474	450	
108	R	65715	66566	852	
109	R	66576	66761	186	
110	R	66761	66937	177	DNA ligase
111	R	66937	67089	153	
112	R	67080	67226	147	
113	R	67219	67488	270	
114	R	67485	67808	324	
115	R	67805	68176	372	
116	R	68195	68863	669	
117	R	68863	68982	120	
118	R	68982	69230	249	

119	R	69234	69341	108	
120	R	69341	69613	273	
121	R	69610	69762	153	
122	R	69759	69977	219	
123	R	69990	70130	141	
124	R	70139	70429	291	
125	R	70716	71063	348	
126	R	71069	71398	330	
127	R	71398	71529	132	
128	R	71632	72588	957	

Table S13. Haloarchaeal tailed virus ORF phamilies.

pham	number of members	HCTV-1	HCTV-2	HCTV-5	HF1	HF2	HGTV-1	HHTV-1	HHTV-2	HRTV-4	HRTV-5	HRTV-7	HRTV-8
1	4				HF1_106	HF2_119					HRTV-5_1		HRTV-8_1
2	5				HF1_105	HF2_118					HRTV-5_2	HRTV-7_1	HRTV-8_2
3	5				HF1_104	HF2_117					HRTV-5_3	HRTV-7_2	HRTV-8_3
4	5				HF1_102	HF2_116					HRTV-5_4	HRTV-7_3	HRTV-8_4
5	4				HF1_101	HF2_115					HRTV-5_5		HRTV-8_5
6	4				HF1_100	HF2_114					HRTV-5_6		HRTV-8_6
7	5				HF1_98	HF2_112					HRTV-5_8	HRTV-7_7	HRTV-8_8
8	5				HF1_97	HF2_111					HRTV-5_9	HRTV-7_8	HRTV-8_9
9	4				HF1_96	HF2_110					HRTV-5_10		HRTV-8_10
10	5				HF1_95	HF2_109					HRTV-5_11	HRTV-7_10	HRTV-8_11
11	5				HF1_94	HF2_108					HRTV-5_12	HRTV-7_11	HRTV-8_12
12	5				HF1_93	HF2_107					HRTV-5_13	HRTV-7_12	HRTV-8_13
13	6				HF1_92	HF2_105 HF2_106					HRTV-5_14	HRTV-7_13	HRTV-8_14
14	9				HF1_91	HF2_103 HF2_104					HRTV-5_15, HRTV-5_16	HRTV-7_14, HRTV-7_15	HRTV-8_15, HRTV-8_16
15	4				HF1_90	HF2_102					HRTV-5_17		HRTV-8_17
16	4				HF1_89	HF2_101					HRTV-5_18		HRTV-8_18
17	5				HF1_88	HF2_100					HRTV-5_19	HRTV-7_19	HRTV-8_19
18	5				HF1_87	HF2_99					HRTV-5_20	HRTV-7_20	HRTV-8_20
19	5				HF1_86	HF2_98					HRTV-5_21	HRTV-7_21	HRTV-8_21
20	5				HF1_85	HF2_97					HRTV-5_22	HRTV-7_22	HRTV-8_22
21	5				HF1_84	HF2_96					HRTV-5_23	HRTV-7_23	HRTV-8_23
22	4				HF1_83	HF2_95					HRTV-5_24		HRTV-8_24
23	5				HF1_82	HF2_94					HRTV-5_25	HRTV-7_25	HRTV-8_25
24	5				HF1_81	HF2_93					HRTV-5_26	HRTV-7_26	HRTV-8_26
25	5				HF1_80	HF2_92					HRTV-5_27	HRTV-7_27	HRTV-8_27

26	5				HF1 79	HF2 90						HRTV-5 28	HRTV-7 28	HRTV-8 28
27	5				HF1 78	HF2 89						HRTV-5 29	HRTV-7 29	HRTV-8 29
28	5				HF1 77	HF2 88						HRTV-5 30	HRTV-7 30	HRTV-8 30
29	5				HF1 76	HF2 87						HRTV-5 31	HRTV-7 31	HRTV-8 31
30	5				HF1 75	HF2 86						HRTV-5 32	HRTV-7 32	HRTV-8 32
31	5				HF1 74	HF2 85						HRTV-5 33	HRTV-7 33	HRTV-8 33
32	4					HF2 83						HRTV-5 34	HRTV-7 34	HRTV-8 34
33	5				HF1 73	HF2 82						HRTV-5 35	HRTV-7 35	HRTV-8 35
34	5				HF1 72	HF2 81						HRTV-5 36	HRTV-7 38	HRTV-8 36
35	3											HRTV-5 37	HRTV-7 36	HRTV-8 37
36	7	HCTV-1 26		HCTV-5 25	HF1 70	HF2 79	HGTV-1 108					HRTV-5 38		HRTV-8 38
37	1											HRTV-5 39		
38	1											HRTV-5 40		
39	1											HRTV-5 41		
40	3				HF1 69	HF2 78						HRTV-5 42		
41	4				HF1 68	HF2 77						HRTV-5 43		HRTV-8 39
42	5				HF1 67	HF2 74						HRTV-5 45	HRTV-7 73	HRTV-8 41
43	4				HF1 66	HF2 73						HRTV-5 46		HRTV-8 42
44	7	HCTV-1 47		HCTV-5 45	HF1 65	HF2 72	HGTV-1 140					HRTV-5 47		HRTV-8 44
45	3					HF2 71						HRTV-5 48		HRTV-8 45
46	5				HF1 64	HF2 70	HGTV-1 256			HHTV-2 80		HRTV-5 49		
47	4				HF1 63	HF2 69						HRTV-5 50		HRTV-8 46
48	4				HF1 61	HF2 67						HRTV-5 51		HRTV-8 48
49	7	HCTV-1 146		HCTV-5 153	HF1 60	HF2 66	HGTV-1 107					HRTV-5 52		HRTV-8 49
50	8	HCTV-1 38		HCTV-5 39	HF1 59	HF2 65	HGTV-1 117					HRTV-5 53	HRTV-7 39	HRTV-8 50
51	4				HF1 58	HF2 64						HRTV-5 54		HRTV-8 52
52	8	HCTV-1 85		HCTV-5 88	HF1 57	HF2 63	HGTV-1 49					HRTV-5 55	HRTV-7 42	HRTV-8 56
53	4				HF1 55	HF2 61						HRTV-5 56		HRTV-8 57
54	1											HRTV-5 57		
55	4				HF1 54	HF2 60						HRTV-5 58		HRTV-8 58

56	2										HRTV-5 59		HRTV-8 59
57	2										HRTV-5 60		HRTV-8 60
58	4				HF1 53	HF2 59					HRTV-5 61		HRTV-8 61
59	5				HF1 49	HF2 55					HRTV-5 62	HRTV-7 51	HRTV-8 65
60	5				HF1 47	HF2 53					HRTV-5 63	HRTV-7 52	HRTV-8 66
61	5				HF1 46	HF2 52					HRTV-5 64	HRTV-7 53	HRTV-8 67
62	4				HF1 45	HF2 51					HRTV-5 65		HRTV-8 68
63	4				HF1 44	HF2 50					HRTV-5 66		HRTV-8 69
64	5				HF1 43	HF2 49					HRTV-5 67	HRTV-7 58	HRTV-8 70
65	4				HF1 42	HF2 48					HRTV-5 68		HRTV-8 71
66	5				HF1 41	HF2 47					HRTV-5 69	HRTV-7 60	HRTV-8 72
67	5				HF1 40	HF2 46					HRTV-5 70	HRTV-7 61	HRTV-8 73
68	1										HRTV-5 71		
69	5				HF1 39	HF2 45					HRTV-5 72	HRTV-7 63	HRTV-8 74
70	4				HF1 38	HF2 44					HRTV-5 73		HRTV-8 75
71	4				HF1 37	HF2 43					HRTV-5 74		HRTV-8 76
72	4				HF1 34	HF2 40					HRTV-5 76		HRTV-8 78
73	6				HF1 33	HF2 39					HRTV-5_114, HRTV-5 77		HRTV-8_122, HRTV-8 79
74	4				HF1 32	HF2 38					HRTV-5 78		HRTV-8 80
75	3					HF2 37					HRTV-5 79		HRTV-8 82
76	6				HF1 31	HF2 36	HGTV-1 229				HHTV-2_81		HRTV-8 83
77	3										HRTV-5 82	HRTV-7 69	HRTV-8 84
78	4				HF1 29	HF2 34					HRTV-5 83		HRTV-8 85
79	5				HF1 28	HF2 33					HRTV-5 84	HRTV-7 71	HRTV-8 86
80	5				HF1 27	HF2 32					HRTV-5 85	HRTV-7 72	HRTV-8 87
81	2										HRTV-5 86		HRTV-8 89
82	4				HF1 26	HF2 31					HRTV-5 87		HRTV-8 90
83	4				HF1 25	HF2 30					HRTV-5 88		HRTV-8 91
84	5				HF1 24	HF2 29					HRTV-5 89	HRTV-7 79	HRTV-8 92

85	6		HCTV-5 148	HF1 23	HF2 28	HGTV-1 235				HRTV-5 90		HRTV-8 93
86	4			HF1 22	HF2 27					HRTV-5 91		HRTV-8 95
87	5			HF1 21	HF2 26					HRTV-5 92	HRTV-7 76	HRTV-8 96
88	1									HRTV-5 93		
89	3				HF2 23					HRTV-5 94		HRTV-8 100
90	4			HF1 20	HF2 22					HRTV-5 95		HRTV-8 101
91	4			HF1 19	HF2 21					HRTV-5 96		HRTV-8 102
92	1									HRTV-5 97		
93	1									HRTV-5 98		
94	4			HF1 18	HF2 20					HRTV-5 99		HRTV-8 105
95	4			HF1 17	HF2 19					HRTV-5 100		HRTV-8 106
96	4			HF1 15	HF2 17					HRTV-5 101		HRTV-8 108
97	2									HRTV-5 102		HRTV-8 109
98	2									HRTV-5 103		HRTV-8 110
99	3									HRTV-5 104	HRTV-7 82	HRTV-8 111
100	2									HRTV-5 105		HRTV-8 112
101	4			HF1 14	HF2 16					HRTV-5 106		HRTV-8 113
102	4			HF1 13	HF2 15					HRTV-5 107		HRTV-8 115
103	4			HF1 12	HF2 14					HRTV-5 108		HRTV-8 116
104	4			HF1 11	HF2 13					HRTV-5 109		HRTV-8 117
105	5			HF1 10	HF2 12					HRTV-5 110	HRTV-7 95	HRTV-8 118
106	4			HF1 9	HF2 11					HRTV-5 111		HRTV-8 119
107	4			HF1 8	HF2 10					HRTV-5 112		HRTV-8 120
108	4			HF1 7	HF2 9					HRTV-5 113		HRTV-8 121
109	3				HF2 8					HRTV-5 115		HRTV-8 123
110	4			HF1 6	HF2 7					HRTV-5 116		HRTV-8 124
111	3			HF1 5	HF2 6					HRTV-5 117		
112	3			HF1 4	HF2 4					HRTV-5 118		
113	4			HF1 3	HF2 3					HRTV-5 119		HRTV-8 125
114	4			HF1 2	HF2 2					HRTV-5 120		HRTV-8 126

145	1		HCTV-2_29										
146	1		HCTV-2_30										
147	1		HCTV-2_31										
148	2		HCTV-2_32						HHTV-2_38				
149	2		HCTV-2_33						HHTV-2_39				
150	1		HCTV-2_34										
151	2		HCTV-2_35						HHTV-2_42				
152	3		HCTV-2_36	HCTV-5_125					HHTV-2_43				
153	3		HCTV-2_37	HCTV-5_126					HHTV-2_44				
154	2		HCTV-2_38						HHTV-2_45				
155	3		HCTV-2_39	HCTV-5_127					HHTV-2_46				
156	3		HCTV-2_40, HCTV-2_41						HHTV-2_47				
157	3		HCTV-2_42	HCTV-5_130					HHTV-2_49				
158	3		HCTV-2_43	HCTV-5_131					HHTV-2_50				
159	3		HCTV-2_44	HCTV-5_132					HHTV-2_51				
160	3		HCTV-2_45	HCTV-5_133					HHTV-2_52				
161	3		HCTV-2_46	HCTV-5_134					HHTV-2_53				
162	1		HCTV-2_47										
163	1		HCTV-2_48										
164	1		HCTV-2_49										
165	1		HCTV-2_50										
166	1		HCTV-2_51										
167	1		HCTV-2_52										
168	1		HCTV-2_53										
169	1		HCTV-2_54										
170	1		HCTV-2_55										
171	1		HCTV-2_56										
172	1		HCTV-2_57										
173	3		HCTV-2_58	HCTV-5_135					HHTV-2_62				
174	3		HCTV-2_59	HCTV-5_136					HHTV-2_63				

411	1	HCTV-1 130												
412	1	HCTV-1 131												
413	1	HCTV-1 136												
414	1	HCTV-1 138												
415	1	HCTV-1 139												
416	1	HCTV-1 140												
417	1	HCTV-1 142												
418	1	HCTV-1 143												
419	1	HCTV-1 144												
420	1	HCTV-1 147												
421	1	HCTV-1 148												
422	1	HCTV-1 149												
423	1	HCTV-1 150												
424	1	HCTV-1 152												
425	1	HCTV-1 155												
426	1	HCTV-1 156												
427	1	HCTV-1 157												
428	1	HCTV-1 158												
429	1	HCTV-1 159												
430	1	HCTV-1 160												
431	1						HF2_91							
432	1						HF2_84							
433	2					HF1_71	HF2_80							
434	1						HF2_75							
435	2					HF1_56	HF2_62							
436	2					HF1_48	HF2_54							
437	2					HF1_16	HF2_18							
438	1						HF2_5							
439	1					HF1_103								
440	1							HGTV-1_1						

441	1					HGTV-1 2							
442	1					HGTV-1 3							
443	1					HGTV-1 4							
444	1					HGTV-1 5							
445	1					HGTV-1 6							
446	1					HGTV-1 7							
447	1					HGTV-1 8							
448	1					HGTV-1 9							
449	1					HGTV-1 10							
450	1					HGTV-1 11							
451	1					HGTV-1 12							
452	1					HGTV-1 13							
453	1					HGTV-1 14							
454	1					HGTV-1 15							
455	1					HGTV-1 16							
456	1					HGTV-1 17							
457	1					HGTV-1 18							
458	1					HGTV-1 19							
459	1					HGTV-1 20							
460	1					HGTV-1 21							
461	1					HGTV-1 22							
462	1					HGTV-1 23							
463	1					HGTV-1 24							
464	1					HGTV-1 25							
465	1					HGTV-1 26							
466	1					HGTV-1 27							
467	1					HGTV-1 28							
468	2					HGTV-1_29, HGTV-1 31							
469	1					HGTV-1 30							
470	1					HGTV-1 32							

471	1					HGTV-1 33							
472	1					HGTV-1 34							
473	1					HGTV-1 35							
474	1					HGTV-1 36							
475	1					HGTV-1 37							
476	1					HGTV-1 38							
477	1					HGTV-1 39							
478	1					HGTV-1 40							
479	1					HGTV-1 41							
480	1					HGTV-1 42							
481	1					HGTV-1 43							
482	1					HGTV-1 44							
483	1					HGTV-1 45							
484	1					HGTV-1 46							
485	1					HGTV-1 47							
486	1					HGTV-1 48							
487	2					HGTV-1_205, HGTV-1 50							
488	1					HGTV-1 51							
489	1					HGTV-1 52							
490	1					HGTV-1 53							
491	1					HGTV-1 54							
492	1					HGTV-1 56							
493	1					HGTV-1 60							
494	1					HGTV-1 61							
495	1					HGTV-1 62							
496	1					HGTV-1 63							
497	1					HGTV-1 64							
498	1					HGTV-1 65							
499	1					HGTV-1 66							

500	1					HGTV-1 67							
501	1					HGTV-1 72							
502	1					HGTV-1 74							
503	1					HGTV-1 76							
504	1					HGTV-1 78							
505	2					HGTV-1_102, HGTV-1 89							
506	1					HGTV-1 97							
507	1					HGTV-1 98							
508	1					HGTV-1 99							
509	1					HGTV-1 104							
510	1					HGTV-1 109							
511	1					HGTV-1 110							
512	1					HGTV-1 111							
513	1					HGTV-1 112							
514	1					HGTV-1 113							
515	1					HGTV-1 114							
516	1					HGTV-1 115							
517	1					HGTV-1 116							
518	1					HGTV-1 118							
519	1					HGTV-1 119							
520	1					HGTV-1 120							
521	1					HGTV-1 121							
522	1					HGTV-1 122							
523	1					HGTV-1 123							
524	1					HGTV-1 124							
525	1					HGTV-1 125							
526	1					HGTV-1 126							
527	1					HGTV-1 127							
528	1					HGTV-1 128							

529	1					HGTV-1 129							
530	1					HGTV-1 130							
531	1					HGTV-1 131							
532	1					HGTV-1 132							
533	1					HGTV-1 133							
534	1					HGTV-1 134							
535	1					HGTV-1 135							
536	1					HGTV-1 137							
537	1					HGTV-1 138							
538	1					HGTV-1 139							
539	1					HGTV-1 141							
540	1					HGTV-1 142							
541	1					HGTV-1 143							
542	1					HGTV-1 144							
543	1					HGTV-1 145							
544	1					HGTV-1 146							
545	1					HGTV-1 147							
546	1					HGTV-1 148							
547	1					HGTV-1 149							
548	1					HGTV-1 150							
549	1					HGTV-1 151							
550	1					HGTV-1 152							
551	1					HGTV-1 153							
552	1					HGTV-1 154							
553	1					HGTV-1 155							
554	1					HGTV-1 156							
555	1					HGTV-1 157							
556	1					HGTV-1 158							
557	1					HGTV-1 159							
558	1					HGTV-1 160							

559	1					HGTV-1 161							
560	1					HGTV-1 162							
561	1					HGTV-1 163							
562	1					HGTV-1 164							
563	1					HGTV-1 165							
564	1					HGTV-1 166							
565	1					HGTV-1 167							
566	1					HGTV-1 168							
567	1					HGTV-1 169							
568	1					HGTV-1 170							
569	1					HGTV-1 171							
570	1					HGTV-1 172							
571	1					HGTV-1 173							
572	1					HGTV-1 174							
573	1					HGTV-1 175							
574	1					HGTV-1 176							
575	1					HGTV-1 177							
576	1					HGTV-1 178							
577	1					HGTV-1 179							
578	1					HGTV-1 180							
579	1					HGTV-1 181							
580	1					HGTV-1 182							
581	1					HGTV-1 183							
582	1					HGTV-1 184							
583	1					HGTV-1 185							
584	1					HGTV-1 186							
585	1					HGTV-1 187							
586	1					HGTV-1 188							
587	1					HGTV-1 189							
588	1					HGTV-1 190							

589	1					HGTV-1 191							
590	1					HGTV-1 192							
591	1					HGTV-1 193							
592	1					HGTV-1 194							
593	1					HGTV-1 195							
594	1					HGTV-1 196							
595	1					HGTV-1 197							
596	1					HGTV-1 198							
597	1					HGTV-1 199							
598	1					HGTV-1 200							
599	1					HGTV-1 201							
600	1					HGTV-1 202							
601	1					HGTV-1 203							
602	1					HGTV-1 206							
603	1					HGTV-1 207							
604	1					HGTV-1 208							
605	2					HGTV-1 209, HGTV-1 210							
606	1					HGTV-1 211							
607	1					HGTV-1 212							
608	1					HGTV-1 213							
609	1					HGTV-1 214							
610	1					HGTV-1 215							
611	1					HGTV-1 216							
612	1					HGTV-1 217							
613	1					HGTV-1 218							
614	1					HGTV-1 219							
615	1					HGTV-1 220							
616	1					HGTV-1 221							
617	1					HGTV-1 222							

618	1					HGTV-1 223							
619	1					HGTV-1 224							
620	1					HGTV-1 225							
621	1					HGTV-1 226							
622	1					HGTV-1 227							
623	1					HGTV-1 228							
624	1					HGTV-1 230							
625	1					HGTV-1 231							
626	1					HGTV-1 232							
627	1					HGTV-1 233							
628	1					HGTV-1 234							
629	1					HGTV-1 236							
630	1					HGTV-1 237							
631	1					HGTV-1 238							
632	1					HGTV-1 239							
633	1					HGTV-1 240							
634	1					HGTV-1 241							
635	1					HGTV-1 242							
636	1					HGTV-1 243							
637	1					HGTV-1 244							
638	1					HGTV-1 246							
639	1					HGTV-1 247							
640	1					HGTV-1 249							
641	1					HGTV-1 250							
642	1					HGTV-1 251							
643	1					HGTV-1 252							
644	1					HGTV-1 253							
645	1					HGTV-1 254							
646	1					HGTV-1 255							
647	1					HGTV-1 257							

648	1					HGTV-1 258							
649	1					HGTV-1 259							
650	1					HGTV-1 260							
651	1					HGTV-1 261							
652	1					HGTV-1 262							
653	1					HGTV-1 263							
654	1					HGTV-1 264							
655	1					HGTV-1 265							
656	1					HGTV-1 266							
657	1					HGTV-1 267							
658	1					HGTV-1 268							
659	1					HGTV-1 269							
660	1					HGTV-1 270							
661	1					HGTV-1 271							
662	1					HGTV-1 272							
663	1					HGTV-1 273							
664	1					HGTV-1 274							
665	1					HGTV-1 275							
666	1					HGTV-1 276							
667	1					HGTV-1 277							
668	1					HGTV-1 278							
669	1					HGTV-1 279							
670	1					HGTV-1 280							
671	1					HGTV-1 281							
672	1					HGTV-1 282							
673	1					HGTV-1 283							
674	1					HGTV-1 284							
675	1					HGTV-1 285							
676	1					HGTV-1 286							
677	1					HGTV-1 287							

678	1					HGTV-1 288							
679	1					HGTV-1 289							
680	1					HGTV-1 290							
681	1					HGTV-1 291							
682	1					HGTV-1 292							
683	1					HGTV-1 293							
684	1					HGTV-1 294							
685	1					HGTV-1 295							
686	1					HGTV-1 296							
687	1					HGTV-1 297							
688	1					HGTV-1 298							
689	1					HGTV-1 299							
690	1					HGTV-1 300							
691	1					HGTV-1 301							
692	1					HGTV-1 302							
693	1					HGTV-1 303							
694	1					HGTV-1 304							
695	1					HGTV-1 305							
696	1					HGTV-1 306							
697	1					HGTV-1 307							
698	1					HGTV-1 308							
699	1					HGTV-1 309							
700	1					HGTV-1 310							
701	1					HGTV-1 311							
702	1					HGTV-1 312							
703	1					HGTV-1 313							
704	1					HGTV-1 314							
705	1									HHTV-2 1			
706	1									HHTV-2 2			
707	1									HHTV-2 4			

708	1								HHTV-2 5				
709	1								HHTV-2 6				
710	1								HHTV-2 7				
711	1								HHTV-2 8				
712	1								HHTV-2 9				
713	1								HHTV-2 11				
714	1								HHTV-2 12				
715	1								HHTV-2 13				
716	1								HHTV-2 14				
717	1								HHTV-2 15				
718	1								HHTV-2 16				
719	1								HHTV-2 17				
720	1								HHTV-2 19				
721	1								HHTV-2 20				
722	1								HHTV-2 21				
723	1								HHTV-2 22				
724	1								HHTV-2 24				
725	1								HHTV-2 25				
726	1								HHTV-2 26				
727	2								HHTV-2 28	HRTV-4 18			
728	1								HHTV-2 29				
729	1								HHTV-2 30				
730	1								HHTV-2 33				
731	1								HHTV-2 40				
732	1								HHTV-2 41				
733	1								HHTV-2 48				
734	1								HHTV-2 54				
735	1								HHTV-2 55				
736	1								HHTV-2 56				
737	1								HHTV-2 57				

738	1								HHTV-2 58				
739	1								HHTV-2 59				
740	1								HHTV-2 60				
741	1								HHTV-2 61				
742	1								HHTV-2 64				
743	1								HHTV-2 65				
744	1								HHTV-2 67				
745	1								HHTV-2 68				
746	1								HHTV-2 69				
747	1								HHTV-2 70				
748	1								HHTV-2 71				
749	1								HHTV-2 72				
750	1								HHTV-2 73				
751	1								HHTV-2 74				
752	1								HHTV-2 75				
753	1								HHTV-2 76				
754	1								HHTV-2 77				
755	1								HHTV-2 78				
756	1								HHTV-2 79				
757	1								HHTV-2 82				
758	1								HHTV-2 84				
759	1								HHTV-2 85				
760	1								HHTV-2 86				
761	1								HHTV-2 87				
762	1									HRTV-4 1			
763	1									HRTV-4 2			
764	1									HRTV-4 3			
765	1									HRTV-4 4			
766	1									HRTV-4 5			
767	1									HRTV-4 6			

768	1									HRTV-4 7			
769	1									HRTV-4 8			
770	1									HRTV-4 9			
771	1									HRTV-4 10			
772	1									HRTV-4 11			
773	1									HRTV-4 12			
774	1									HRTV-4 13			
775	1									HRTV-4 14			
776	1									HRTV-4 15			
777	1									HRTV-4 16			
778	1									HRTV-4 17			
779	1									HRTV-4 19			
780	1									HRTV-4 20			
781	1									HRTV-4 21			
782	1									HRTV-4 22			
783	1									HRTV-4 23			
784	1									HRTV-4 24			
785	1									HRTV-4 25			
786	1									HRTV-4 26			
787	1									HRTV-4 27			
788	1									HRTV-4 28			
789	1									HRTV-4 29			
790	1									HRTV-4 30			
791	1									HRTV-4 31			
792	1									HRTV-4 32			
793	1									HRTV-4 33			
794	1									HRTV-4 34			
795	1									HRTV-4 35			
796	1									HRTV-4 36			
797	1									HRTV-4 37			

798	1									HRTV-4 38			
799	1									HRTV-4 39			
800	1									HRTV-4 40			
801	1									HRTV-4 41			
802	1									HRTV-4 42			
803	1									HRTV-4 43			
804	1									HRTV-4 44			
805	1									HRTV-4 45			
806	1									HRTV-4 46			
807	1									HRTV-4 47			
808	1									HRTV-4 48			
809	1									HRTV-4 49			
810	1									HRTV-4 50			
811	1									HRTV-4 51			
812	1									HRTV-4 52			
813	1									HRTV-4 53			
814	1									HRTV-4 54			
815	1									HRTV-4 55			
816	1									HRTV-4 56			
817	1									HRTV-4 57			
818	1									HRTV-4 58			
819	1									HRTV-4 59			
820	1									HRTV-4 60			
821	1									HRTV-4 61			
822	1									HRTV-4 62			
823	1									HRTV-4 63			
824	1									HRTV-4 64			
825	1									HRTV-4 65			
826	1									HRTV-4 66			
827	1									HRTV-4 67			

948	1							HHTV-1 56					
949	1							HHTV-1 57					
950	1							HHTV-1 58					
951	1							HHTV-1 59					
952	1							HHTV-1 60					
953	1							HHTV-1 61					
954	1							HHTV-1 62					
955	1							HHTV-1 63					
956	1							HHTV-1 64					
957	1							HHTV-1 65					
958	1							HHTV-1 66					
959	1							HHTV-1 67					
960	1							HHTV-1 68					
961	1							HHTV-1 69					
962	1							HHTV-1 70					
963	1							HHTV-1 71					
964	1							HHTV-1 72					
965	1							HHTV-1 73					
966	1							HHTV-1 74					

Table S14. Interspersed repeats in haloarchaeal tailed virus genomes

Virus	Repeat positions (nt)	Repeat position (ORF)
HCTV-1	11800-11839/32323-32284	upstream ORF29 and ORF57
	85535-85571/85454-85490	ORF126 (tail tape measure protein)
HCTV-2	16138-16026/31163-31274	ORF29 and ORF48
	15867-15902/15902-15867	ORF28
	2474-2430/43739-43783	ORF6 and ORF73
HCTV-5	74327-74452/72671-72796	ORF114 (Zeta toxin-ParB)
	72830-72903/72116-72189	
	72815-72950/65471-65606	ORF114 (Zeta toxin-ParB) and ORF109 (ADP-ribosyltransferase)
	21729-21780/25413-25464	ORF42 (HNH) and ORF49 (Hef-HNH)
HHTV-1	40328-40376/37691-37739	upstream of ORF55 and ORF52
HHTV-2	121-158/158-121	ORF1
	352-418/1662-1729	Upstream of ORF2 and ORF3
HRTV-4	-	-
HRTV-5	27867-27921/58256-58310	upstream of ORF36/37 and ORF81/82
	28077-28119/58255-58296/59465-59507/59465-59507/	upstream of ORF38 (ribonucleotide reductase), ORF54, ORF81/82 and ORF83/84
HRTV-7	46376-46318/47617-47675	Upstream of ORF68/69 and ORF70/71
HRTV-8	28180-28223/56132-56175	Upstream of ORF38 (ribonucleotide reductase) and ORF85/86
	27971-28023/37944-37997/52522-52574/54923-54975	upstream of ORF36/37, ORF52, ORF77, ORF83/84
HGTV-1	128680-128730/130402-130452/131835-131885/132361-132411/132816-132866/133529-133579/133972-134022/134543-134593/135052-135102/135906-135956/136499-136549/137094-137144/137618-137668/138006-138056/138385-138435/138952-139002/139404-139454/139735-139785/140207-140257/140525-140575/141300-141350/141758-141808/142116-142166	Upstream of or close to ORF276, ORF279, ORF283, ORF285, ORF286, ORF288, ORF289, ORF291, ORF293—ORF300, ORF302, ORF304-306, ORF308, ORF309, ORF311
	123325-123364/123828-123867	upstream of ORF259 and ORF260
	125979-126006/141781-141754	upstream of ORF268 and in ORF310
	124967-125016/126403-126452	upstream of ORF264 and ORF270
	867-809/22564-22622	upstream of ORF3 and ORF25
	26254-26653/27558-27960	ORF29 and ORF31
	25087-25133/25276-25322	ORF27
	40319-40260/99900-99959	ORF50 (LAGLIDADG endonuclease) and ORF205 (LAGLIDADG endonuclease)

Table S15. Tandem repeats in haloarchaeal tailed virus genomes

Virus	Nr. of repeats	Size of repeat	Consensus sequence	Start position	Stop position	Region
HCTV-1	2.2	22	ACGACCGAAGAGACGAAAGACG	13738	13785	ORF30
	2.5	12	TCGGTGCAGGAC	85281	85310	ORF126 Tail tape measure protein
	4.5	27	CTGTGGGACACGTTTATCGTCGACCTG	85457	85578	
	4.3	27	AAGATATCATCTGGGAAGCGTTCATCA	85698	85812	
HCTV-2	3.1	9	ACGAGGACG	16714	16741	ORF30
	1.9	21	GGTGGCGCAGGACGGCGGCGT	16751	16790	ORF30
	2.1	12	CAGCATGGGCCG	25284	25308	ORF42 tail tape measure protein
	3.1	18	TCCTGCTCGTCCTCGGCC	33804	33859	ORF57
	1.9	21	CTCGTCTCGGCCACGTCGGC	33845	33884	ORF57
	2.0	26	CGGGCGTCTCGTCTGCCCGGCTGC	41537	41586	ORF68
HCTV-5	2.1	12	GGAGGACGAGGA	77026	77050	ORF118
	2.0	20	CGCCTCGTCCTCAGCACGGA	93542	93582	intergenic ORF141-142
HHTV-1	2.1	14	CACGTATGTGGCCC	43547	43577	ORF63
HHTV-2	2.3	15	GAGGACGACGACCGG	22993	23026	ORF37
HRTV-4	6.7	12	GAGGACGACGAC	15584	15660	ORF37
HRTV-5	2.8	12	GACGAGGAGGAG	7515	7548	ORF12
	3.1	14	TAGTTATACATTAG	59606	59650	intergenic ORF83-84
	3.2	10	GGAGAGACGG	75824	75855	ORF122
HRTV-7	2.2	18	CGAAAGCCGGGGTGAGC	1646	1685	ORF1
	4.3	7	GGGTCGG	1799	1828	ORF2
	2.0	20	CGCCTCAACGACATCGTGAA	15605	15644	ORF23
	2.3	16	ACTCTAAGAGATAGAA	21220	21254	ORF30 (baseplate J-like protein)
HRTV-8	7.5	6	CCGGGG	159	203	Intergenic ORF128-ORF1
	2.1	15	GAGGAAGACGAGGAG	7582	7613	ORF12
	7.5	6	CCGGGG	74332	74376	Intergenic ORF128-ORF1
HGTV-1	2.5	16	CCGAAATGTCGAAATA	142214	142253	ORF312

Table S16. Amino acid identity percentages of halovirus proteins belonging to superfamily cl00641 and other members of that superfamily from haloarchaea derived from non-redundant protein database at NCBI. Identities $\geq 19.5\%$ are marked in orange.

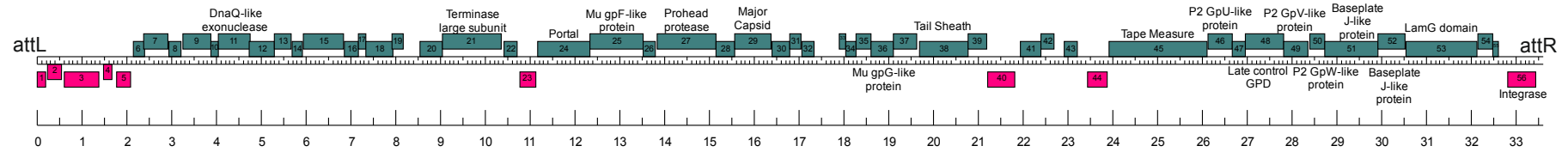
Accession number or Virus ORF number	HCTV-1_ ORF87	HCTV-5_ ORF90	HGTV-1_ ORF189	HGTV-1_ ORF39	HRTV-5_ ORF62	HRTV-7_ ORF51	HRTV-8_ ORF65
AAV44425.1	15.1	17.0	3.9	13.1	10.6	10.6	10.7
AAV45925.1	6.2	7.5	3.1	7.4	7.6	10.3	7.6
ACM57283.1	5.9	6.6	3.8	7.4	8.1	8.8	7.6
ACM57919.1	6.0	4.1	3.0	3.9	2.2	4.2	6.2
ACM58849.1	11.6	11.8	18.5	13.7	13.4	10.1	14.4
ACM59027.1	8.1	4.5	4.8	7.9	9.2	7.2	9.8
ACM59084.1	13.4	18.3	12.3	12.3	16.7	9.0	15.4
ACV10642.1	6.3	7.1	3.7	5.5	7.4	8.9	7.1
ACV11626.1	13.8	16.2	10.8	12.1	17.3	13.1	16.5
ACV11887.1	4.8	5.4	3.2	5.2	7.7	6.1	7.8
ACV12090.1	5.4	5.2	1.7	5.3	7.9	7.4	5.7
ADB58945.1	13.3	17.7	13.4	10.9	16.2	12.2	15.9
ADB61785.1	6.2	6.2	4.4	7.5	7.9	7.6	7.1
ADB62071.1	5.4	5.2	2.8	5.3	5.3	5.7	6.9
ADD04643.1	5.6	5.8	3.9	7.6	8.2	7.8	7.4
ADD05285.1	14.5	15.8	14.5	10.8	9.4	6.1	9.5
ADD05961.1	5.4	5.1	2.9	5.0	6.8	5.2	5.7
ADE02138.1	15.1	13.6	18.0	12.7	11.0	11.7	10.9
ADE04273.1	15.2	17.3	9.2	2.5	13.2	12.9	12.5
ADE04881.1	6.5	6.0	4.3	7.5	7.9	8.2	7.5
ADE04903.1	5.7	4.7	3.1	4.3	6.3	5.6	6.6
ADJ13477.1	16.1	19.5	9.7	11.7	15.4	15.5	17.3
ADJ14856.1	8.4	10.0	20.2	3.2	10.0	12.2	7.0
ADJ15210.1	6.4	6.8	3.6	8.2	8.2	9.2	8.7
ADQ66828.1	6.6	6.5	3.6	7.1	8.4	8.7	7.3
ADQ66884.1	5.4	4.5	3.4	5.0	5.7	5.6	4.6
ADQ66885.1	19.0	19.0	12.9	16.1	15.6	19.6	18.3
ADQ68575.1	4.7	6.0	3.7	4.3	6.1	6.4	6.7
ADQ68576.1	7.3	5.9	4.1	7.7	9.3	7.8	10.1
AEM58700.1	5.4	5.9	3.6	7.5	6.0	8.3	5.5
CAI48803.1	16.6	18.5	13.2	14.1	3.3	14.7	16.8
CAI49563.1	5.5	6.7	3.0	7.0	8.6	8.6	8.2
CAI50018.1	7.5	7.6	7.8	9.9	1.4	13.5	10.0
CAI50025.1	4.9	4.8	2.6	6.6	6.7	7.8	6.6
CAI50026.1	6.2	5.7	2.6	4.1	7.9	7.6	7.9
CAI50902.2	18.7	16.7	10.0	15.3	17.3	13.9	17.5
CAP14083.1	6.7	6.8	2.9	7.5	10.1	9.1	8.1
CAP15640.1	5.0	4.6	2.4	4.2	6.4	6.6	6.8
CAP15641.2	6.5	6.6	3.7	5.3	8.3	8.5	6.0
EHB56722.1	17.1	22.6	10.1	14.9	20.0	17.5	19.3
EHB56793.1	14.3	15.4	18.9	12.2	12.0	9.6	12.2

HCTV-1_ORF87	100.0	61.7	13.6	10.8	19.5	20.5	19.1
HCTV-5_ORF90	61.7	100.0	10.7	19.5	20.8	17.8	18.9
NP_861634	20.8	20.5	10.1	17.1	89.7	37.9	84.1
NP_542546	20.8	20.5	10.1	17.1	89.7	37.9	84.1
HGTV-1_ORF189	13.7	10.7	100.0	8.4	8.8	7.8	7.7
HGTV-1_ORF39	10.8	19.5	8.4	100.0	16.0	16.0	16.1
HRTV-5_ORF62	19.5	20.8	8.8	16.0	100.0	36.8	91.7
HRTV-7_ORF51	20.5	17.8	7.8	16.0	36.8	100.0	36.8
HRTV-8_ORF65	19.1	18.9	7.7	16.1	91.7	36.8	100.0
NP_280319.1	7.2	7.2	3.1	7.8	10.3	8.9	8.4
YP_001690750.1	6.8	6.8	3.9	5.6	8.6	8.8	6.2
YP_002564620.1	6.6	6.0	2.7	4.7	8.3	5.3	5.3
YP_002564621.1	7.3	6.7	3.7	7.0	9.2	9.5	8.8
YP_002564646.1	6.6	6.5	2.4	3.7	8.1	8.3	8.1
YP_002567430.1	11.6	9.5	2.8	13.4	0.9	5.9	0.6
YP_002567714.1	17.6	12.0	7.2	13.7	9.7	6.1	5.3
YP_002567746.1	11.1	9.2	2.9	15.4	15.0	1.6	16.2
YP_003130454.1	13.2	16.1	8.5	6.2	16.0	13.0	15.8
YP_003130621.1	7.2	8.1	4.1	6.1	9.9	9.1	9.3
YP_003130824.1	7.4	7.5	4.2	6.1	9.5	9.1	9.7
YP_003175762.1	5.3	4.8	2.3	4.8	5.7	7.1	5.8
YP_003176726.1	13.5	13.3	14.0	10.0	17.1	15.0	16.5
YP_003177931.1	6.5	6.1	3.1	7.1	7.7	9.6	7.6
YP_003178649.1	16.4	12.6	18.6	11.4	10.9	11.7	12.3
YP_003533884.1	12.2	11.2	8.2	3.1	7.0	3.6	8.3
YP_003535784.1	8.6	9.8	5.8	7.4	1.3	13.3	8.5
YP_003738491.1	4.2	4.5	2.6	4.7	5.7	6.3	5.3
YP_003738492.1	4.9	4.6	3.2	4.5	8.2	8.8	7.9
YP_004586129.1	19.3	19.7	11.1	15.7	13.9	8.7	14.2
YP_004595509.1	7.4	7.7	3.2	3.3	9.7	8.2	10.0
YP_004595510.1	5.0	4.8	3.1	3.6	7.2	6.3	6.1
YP_004596959.1	5.0	6.1	4.1	7.1	7.8	7.8	7.7
YP_004598292.1	5.7	5.4	2.5	4.5	6.0	6.4	5.3
YP_004785685.1	6.6	7.1	3.5	7.2	8.4	8.8	8.7
YP_004785822.1	15.1	16.4	17.8	12.1	6.0	11.5	11.4
YP_004796178.1	6.3	7.6	3.2	7.5	7.7	8.5	7.7
YP_005839106.1	19.3	21.1	15.6	11.0	16.6	9.7	14.4
YP_005839926.1	7.6	7.2	4.4	7.9	8.8	8.8	9.0
YP_005839927.1	6.2	6.2	3.1	4.5	6.3	5.5	6.3
YP_005839934.1	13.7	12.7	17.5	11.9	11.2	11.0	12.1
YP_005839955.1	14.0	5.9	3.3	10.3	6.0	6.1	16.5
YP_005840739.1	5.7	6.4	3.2	6.8	7.9	6.5	7.5
YP_006347987.1	16.2	19.5	12.1	16.5	13.9	12.6	12.8
YP_006350441.1	6.5	5.9	4.2	7.8	7.3	9.0	7.6
YP_006351428.1	2.4	15.8	4.0	11.5	11.9	8.2	11.2
YP_134235.1	17.0	5.0	4.0	15.6	2.5	12.8	2.8

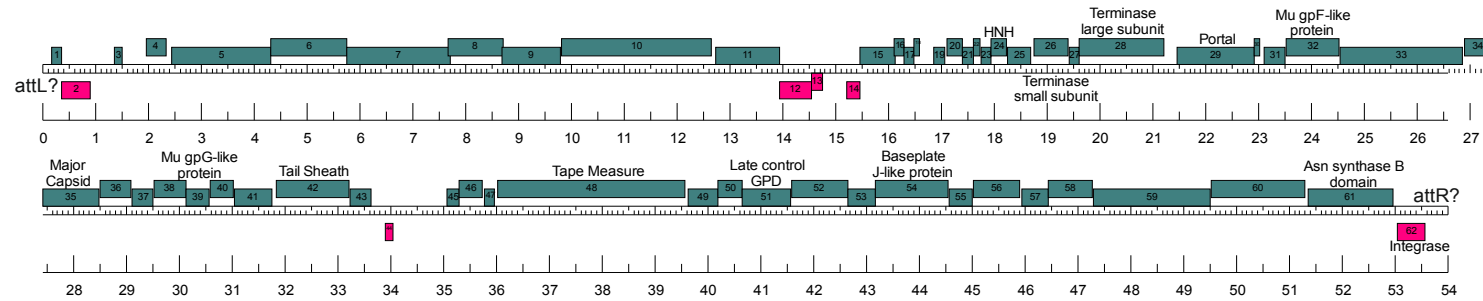
YP_137099.1	13.1	18.2	8.1	12.0	14.7	12.9	14.6
YP_327283.1	21.1	22.2	10.3	11.2	19.0	11.6	13.1
YP_657900.1	7.3	5.5	4.4	7.8	8.9	8.9	9.0
YP_657901.1	5.1	6.2	3.4	4.9	6.0	5.6	6.0
YP_658634.1	5.7	6.4	3.9	6.8	7.8	6.9	7.6
ZP_08043416.1	5.9	7.7	3.9	7.5	8.2	8.2	7.8
ZP_08046055.1	9.7	12.3	4.5	10.9	11.0	10.9	9.7
ZP_08558864.1	6.8	6.9	3.1	5.9	8.8	8.8	8.2
ZP_08558865.1	4.7	4.0	3.1	3.9	6.1	6.0	6.4
ZP_08558872.1	16.8	13.1	8.0	13.8	15.3	14.5	10.2
ZP_08559182.1	7.0	7.3	4.2	5.9	9.3	9.1	9.1
ZP_08559183.1	4.8	6.0	3.2	5.2	7.2	6.3	7.0
ZP_08560032.1	12.3	2.7	7.6	9.7	1.4	16.1	1.1
ZP_08560869.1	8.9	9.7	8.7	18.4	11.4	15.1	14.1
ZP_08561299.1	12.8	10.0	12.9	12.4	17.4	12.8	14.4
ZP_08561572.1	6.9	8.4	3.3	6.4	9.4	10.6	9.6
ZP_08561573.1	5.5	5.7	2.9	5.0	8.0	6.2	7.8
ZP_08561724.1	7.1	7.1	3.7	6.7	7.6	8.9	7.3
ZP_08562196.1	14.4	14.9	6.7	12.4	15.3	17.2	16.8
ZP_08963685.1	10.9	6.5	8.5	9.1	17.4	17.1	16.5
ZP_08963686.1	5.1	5.3	3.3	4.4	5.4	6.2	5.7
ZP_08964026.1	16.3	17.3	8.4	12.7	9.0	15.5	9.4
ZP_08965165.1	5.9	6.5	4.1	7.1	7.6	9.0	7.2
ZP_08965468.1	7.9	7.4	3.9	6.3	9.4	10.2	10.7
ZP_08965469.1	6.0	5.2	1.6	5.2	6.1	6.4	7.5
ZP_08965593.1	6.5	6.5	3.1	6.5	8.4	9.0	8.0
ZP_08965594.1	6.3	5.2	3.1	3.9	6.0	6.0	5.7
ZP_08967159.1	8.9	6.1	3.6	7.8	7.8	8.7	7.5
ZP_08968424.1	15.0	13.4	18.7	13.5	8.5	13.1	8.6
ZP_08968710.1	13.7	13.5	8.3	11.4	14.7	14.0	14.5
ZP_08968711.1	5.6	5.0	3.1	5.0	6.0	6.4	5.5
ZP_09027914.1	13.5	20.9	14.3	7.8	15.4	15.7	14.8
ZP_09028183.1	14.6	17.4	11.3	7.8	10.5	7.6	11.2
ZP_09028711.1	7.2	7.1	4.0	7.0	8.1	8.7	8.1
ZP_09030271.1	5.9	4.3	3.7	4.6	6.7	6.2	7.1
ZP_09030272.1	6.0	6.5	4.3	6.2	8.8	8.5	9.1
ZP_09030564.1	16.4	8.6	10.1	6.5	13.9	12.8	14.1
ZP_09946184.1	10.3	8.0	13.2	8.0	4.1	6.8	2.8
ZP_09947326.1	5.7	5.5	2.4	4.4	5.4	5.5	5.6
ZP_09948039.1	5.4	5.7	4.2	7.0	8.2	8.4	7.8
ZP_09950156.1	7.4	7.6	3.9	8.2	9.5	9.2	9.5
ZP_09950157.1	6.4	5.9	2.7	4.7	6.4	6.4	6.1
ZP_09950187.1	11.6	14.0	21.6	12.1	12.4	10.3	11.1
ZP_09950214.1	6.9	6.8	3.1	5.8	7.5	8.5	7.8
ZP_09950215.1	4.9	3.8	2.6	4.3	7.3	6.6	4.8
ZP_09950274.1	4.1	9.2	7.7	11.2	11.5	5.5	11.5

SUPPLEMENTARY FIGURES

A



B



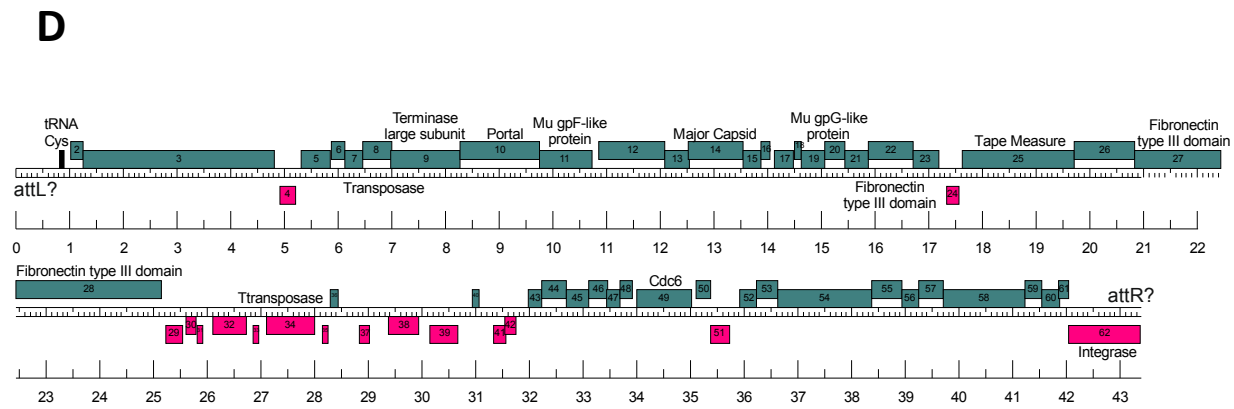
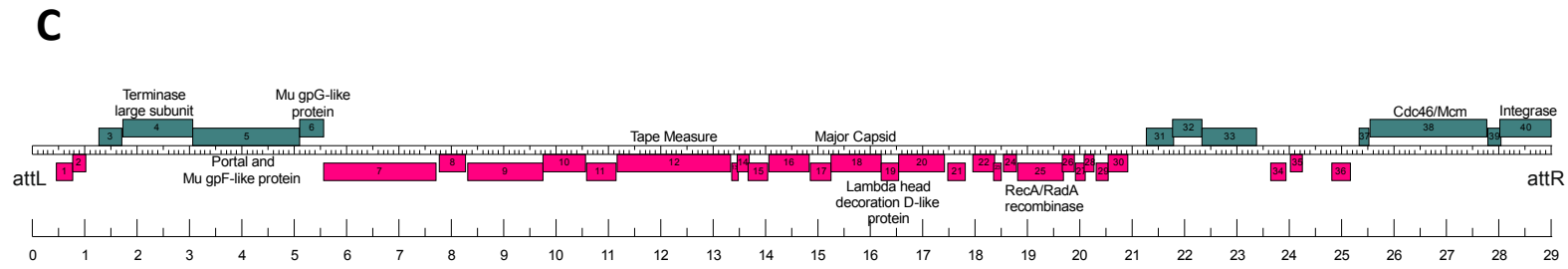
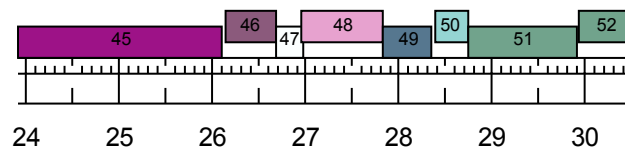
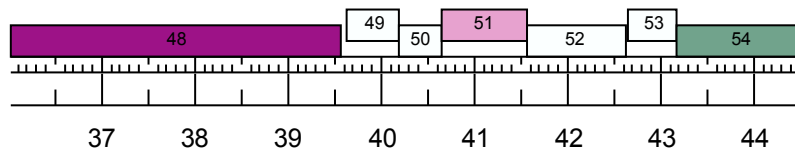


Figure S1. Genome maps of haloarchaeal proviral regions related to tailed viruses. (A) Aprof (B) Hjeo (C) Mokin (D) Ngreg

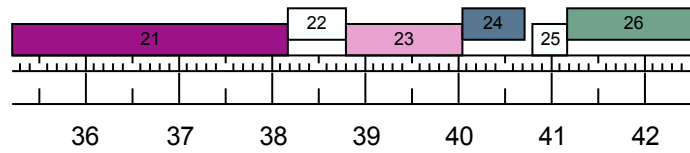
Aprof



Hjeo



HGTV-1



P2

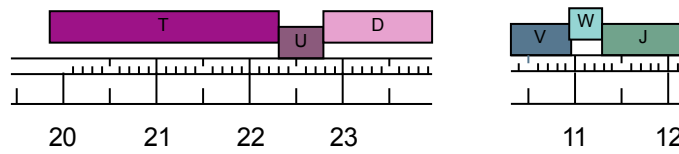


Figure S2. Maps of genome regions coding for putative homologues of P2 tail structural and assembly proteins. ORFs encoding putative homologues have the same color. Aprof and Hjeo are proviral regions identified during this study (Supplementary results 1).

>HF2_67312-67379
CGGC**GAAA**CGCCCCACCGTGGGGCG**TCGTC**GGGAGTTGGTCATCCCCGGC**TGTTGACGAG**CCCGACCAGA
1213 1415 16 17 2 2a2b3 4 5 6 7 8 9

>HF1_65540-65607
CGGC**GAAA**CGCCCCACCGTGGGGCG**TCGTC**GGGAGTTGGTCATCCCCGGC**TGTTGACGAG**CCCGACCAGA

>HRTV-5_65998-66064
TGGC**GAAA**CGCCCCACCGTGGGGCG**TCGTC**GGGAGTTGGTCGTCCCCGGC**TGTTGACGAG**CCCGACCAG

>HRTV-8_63409-63474
CGGC**GAAA**CGCCCCACTTTGGGGCG**TCGTC**GGGAGTTGGTCCCCCGGC**TGTTGACGAG**CTGACCAG

>HRTV-8_65506-65572
TGGC**GAAA**CACCCCATCGAGGGGTG**TC**ACTACCCGGTGGTTCGGGTAGT**CTGATGAT**GCCAGCCACG

Figure S3. Ribozymes in haloarchaeal tailed viruses. Header consists of virus name and ribozyme position in the genome (nt). Ribozyme catalytic site residues are in bold and are numbered below the HF2 ribozyme sequence. Complementary regions forming stems are underlined with straight lines. Regions forming pseudoknot structures are underlined with wavy lines.

I group

108_HNH YGGRWDAKR-EKAILADSEQCASCGMTRAEHKQVYGRDLHVHKKIPVRFVDD-----VSDAHFEVNLVTVCM-ECHEEFD
ZP_10771382.1_Halogramma YGSNWERQR-QEALERDDFCQHCCGLTNSEHKTERGQELVHHRVKRRAFDD-----YRDANQVENLVTLCR-ACHYDAE
YP_134093.1_Haloarcula YGSDWASYS-IDALLEWEGRCASCGLTEPTHIRTYGKSLDIHHEPPVEEFDD-----IETAHEEAELQPVCV-KCHSKVE
YP_003129161.1_Halorhabdus YGPSWSSAR-RKARERDGYTCQRCGMDDEHLDEYGHKLVHVKIKPFRTFDD-----HEDANDLGNLITVCL-DCHNRLE
YP_003482340.1_Natrialba YGEGWNKKKERVRERDGRCQHCCGRGEQEHVKLFGTKHIVHVVVPARSIED-----AQERNAMKNLVTLCRGDCHRWKVE
104_HNH FGSNWEKQR-AKALERDK-VCQRCGEDGS----D--ARLSVHHLLEPRRSFDRDDL--EEKANDLSNLVVLICM-PCHGAVE
106_HNH YGENWYAVR-RKVRSRDE-VCQKCGSDGT----D--CLLDVHHIDPIRTFEK-----PEEANTMDNLILLICR-SCHRAVE
89_HNH YGSTWDEKR-EMVLERDNYECRVCGRQ-----MVSDPLVHHLVVKKFD-----PDNAHTRDNLITLCA-THRLVE
111_HNH YGHTWSRMR-EKALDRDGYCQVCGSEE-----QLVHHLKLPVRTFDD-----PNEAHYMVNLVTCV-TCHPKLD
YP_004598316.1_Halopiger YGREWWRIR-RRALERDGYECQHCCGAGPD---ELGQNPVHVRPVSFDR-----PEDAHTMDNVVS-----HRRAE
ZP_08962642.1_Natrinema YNGRWREIR-RQALERDEHTCQHCCGLTRE----EIGHEPDVHHITPVREFND-----PQLAHTLDNVVCLICR-SCHRYAE
ZP_09029765.1_Halobacterium YGPAWWEVR-RMALERDDYECQNCGAGVA---ELGRNPVHHIEPVREFED-----PEEAHRLDNVVCLICR-PCHRRIE
AFH21819.1_halophage YGPNWEEQR-DACLERDNYRCQACGVSNEE---APIALNAHHIQPFRTFDN-----REEANRLDNLVALICL-SCHGKWE
YP_09949009.1_Halobiforma YGRQWWRIR-RRALERDDHECQHCCGATRE----EIGRNPVHHLLEPVRSFDR-----PEEAHTLSNVVALICR-RCHRAE
YP_003130543.1_Halorhabdus YGQGWNAVR-REALERDEYECVHCCGTGSD---SLGRNPVHHIIPVRAFVTSRHEKTDATIDNVVSLICP-SCHRAE
ZP_08559989.1_Halorhabdus YGQGWNAVR-REALERDEYECVHCCGIGSE---SLGRNPVHHIIPVRAFVTSRHEKTDATIDNVVSLICP-SCHRAE
YP_137220.1_Haloarcula YGPGWRAVR-EQALARDDHACVLCGTNAD---ELGRNPVHHIIPVRLFAALRALAVRDAHTLDNVVSLICP-SCHRRTE

II group

42_HNH AHRVVLML-H---GD---SPG-SQYVLEHCDNPSCVNPDELYIGTQQDNMDDWQ
49_HNH AHRFSYMLEY---GV---IE-DDFVLEHCDNKSCVNPDELYSGDHSNMQDRA
56_HNH AHRVAYRLEV---ED---PD-GSNVLEHCDNRSCVNPDELYLGTQEDNIKDRD
YP_001353918.1_Janthinobacterium AHRAAYELYV---GP---IKD-GLHVCRCNDRSCVNPDEHFLGTNHNENVADRQ
NP_521314.1_Ralstonia AHRVAYELCV---AP---IPD-GMVACRCNDRACVNPDEHFIATQRENILDMV
NP_521357.1_Ralstonia AHRLAYQLFH---GI---VLERSQHVLHRCPTACVNPDEHFLFIGTQQDNMRDMR
YP_002003969.1_phage_13a VHRLEWIKQR---GP---IPE-GLVVRHRCNPKCYNIEHLELGTQDNTRDRV
YP_002280691.1_Rhizobium AHRVSYECFH---HE---IPA-GMFVCHRCNDRKCANPDEHFLGTQDNSSDMV
YP_002328232.1_Escherichia SHRVAFELYH---SP---IPS-GKMVCHRCNDRQCCVNPDEHFLGSAQENMDDKI
YP_002964909.1_Methylobacterium AHRMAYRLAK---GS---FDE-RLFVCHRCNDRPCCVNPDEHFLWLGTHRQNMDDMF
YP_002982221.1_Ralstonia AHRLAFLVEKH---GLALSDID-GQFVVRHRCNDRPACVNPDEHLELGTQDNMRDMV
YP_003325164.1_Xylanimonas AHRVSAELA---DM---DVA-GRVVRHRCNDRPCCVNPDEHFLVPGTQAQNAVDMA
YP_003897184.1_Halomonas AHRASYLIHK---GP---IPE-GLVVCHTCNDPACVNPDEHFLWLGTHMDNTQDMM
AAP78667.1_Myxooccus_phage AHRVAYELAT---GTR---LLP-SVLVCHRCNDRRCCNPAHFLFAGTHADNADMC
NP_888749.1_Bordetella AHCASFRLV---GD---IPA-GMYVCHRCNDRPCCVNPDEHFLGTQDNQADMA
YP_005098368.1_Synechoc.phage THRLAWEFAP---GQ---IPE-GLVVRHRCNDRPCCVNPDEHLLLGTHKDNFNDKI
AFH22805.1_Halophage THRVLQTHGE--GF---TEE-KNFCCRCNDRPCCVNPDEHFLFAGSHKDNMEDAT
YP_002280661.1_Rhizobium AHRALMASTAKFNP---LP---SDFACHRCNDRPCCVNPDEHLYWGDALNLRADCM
AFH75206.1_Streptomyces AHRVAFQHAY---GY---LP---PVVRHRCNDRPCCVNPDEHLLPGTHKDNQADMA
ZP_10176254.1_Pseudomonas AHRLSYELSK---GK---IPD-GMVVHRCNDRPACINPDEHFLVLTQKDNMEDMY
ZP_10243912.1_Nitrolanceus VHRLAAMIWL---GL---DPVSSLSVCHRCNDRPACINPDEHFLVLTQKDNMEDMY

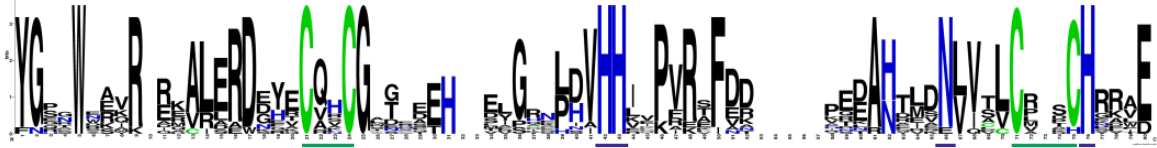
ZP_10244310.1_Nitrolancetus	SHRLAYIMAF---GP---IPD-GLHVLHHC DNPPCVNPSLFLGTARANVHDME
ZP_00209001.1_Magnetospirillum	AHRLSYRLHC---GE---VPT-GLGVLEHCDVRCCVNPDHLYAGTTQQNNADRT
YP_224185.1_phage_ES18	AHRLSYEFYSK---GI---IPE-GKVVILCDNPSCVNPKHLAVGTVSDNNADMR
YP_369478.1_Burkholderia	SHRVSYMLFN---GD---IPS-GLHVLHRC DVPCCANPNHLWLGTNDQNVQDKV
NP_041971.1_phage_T7	CHRVMS-----N---APK-GSTVLEHSCDNPLCCNPEHLSIGTPKENSTDMV
AAP33980.1_phage_T7	CHRVMS-----N---APK-GSTVLEHSCDNPLCCNPEHLSIGTPKENSIDMV
AAP34081.1_phage_T7	CHRVMS-----N---APK-DSTVLEHSCDNPLCCNPEHLSIGTPKENSTDMV
ZP_10244345.1_Nitrolancetus	AHRISWELAN---GP---IPE-GLFVCHRC DNPPCVRPDHFLFGTPADNMHDRS
EJO27443.1_Achromobacter	AHRVSYEMFV---GV---AG-EMHVLHCCDNPSCVRPDHFLFLGSKDNVDDKV
EKA33187.1_Pseudomonas	AHRASYEINC---GP---IDN-GLHVLHRC DNPRCVRPDHFLFLGTHLENMQDMV
YP_004109297.1_Rhodopseudomonas	AHRVALAL-S---GK---NVV-GRFVLEHRC DNPS CVRPDHLRRGSHKDNMADMT
YP_002982223.1_Ralstonia	AHIASYVAAN---GP---LAP-GMCMVHVC DNPLCVRPDHLEAGTDADNMNDKT

III group

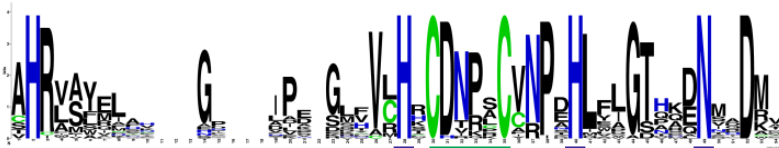
22_HNH	IHRLAVSEHGFDVAV----RDKHVHHRNGIPWDNRPGNIEVLTEEEHKSLEH
77_HNH	VHRLVAVAQFGFEAV----CGNQVHHKTGIKWDNRPSNLTLTPSDHASRH
YP_003479797.1_Natrialba	VHRLAVSEYGYEEV----GQKVVHHRNGIPWDNRPANIEVLTAEEHGRIH
YP_004598852.1_Halopiger	IHRLVMVAERGFDA----VADKHVHHRNGVRFDNRPENLELLSRSEHAERH
AFH22156.1_halophage	IHRLVAVAEYGVKAT----ENKQIHHKNGVKWDNRPENLICMSNSEHQRHE
AFH21445.1_halophage	IHRLVVVAEGADPHKVFGLSHVHHKNEIPWDNRPENLEVVTQSEHKKIHH
AFH21837.1_halophage	VHQLVAIADGANPSKIFSNGRYHVHHRNGVKWDNRPENLELQSAKTHAIEH
AFH22067.1_halophage	VHRLIAVAEGLLDPSDFRDFDTVVHHRKSEHGLDNRPNLKVMERGEHQSMH

Figure S4. Alignments of the conserved regions of HCTV-5 HNH endonucleases and related homologues from database. HNH catalytic motif residues are marked in green. Cysteine dyads are marked in yellow. Other residues conserved among the members of a given group are marked in grey.

I group



II group



III group



Figure S5. Conserved residues of aligned regions from HCTV-5 HNH homing endonucleases and homologues retrieved from the database. HNH catalytic motif residues are marked in blue. Cysteine dyads are marked in green.

```
Pyrococcus yayanosii CH1 41390 GAGGGCGGGGCTGGCGTGG-AGGATGAACGCGGTTCG 41424
HRTV-7 5775 GAGGGCGGGGCTGGCG-GGCAGGGTGAA-GCGGTTCG 5808
```

Figure S6. Alignment of *Pyrococcus yayanosii* CH1 (CRISPR ID NC_015680_2_3) and corresponding HRTV-7 genome nucleotide sequence. Matching residues are marked in yellow.

HGTV-1

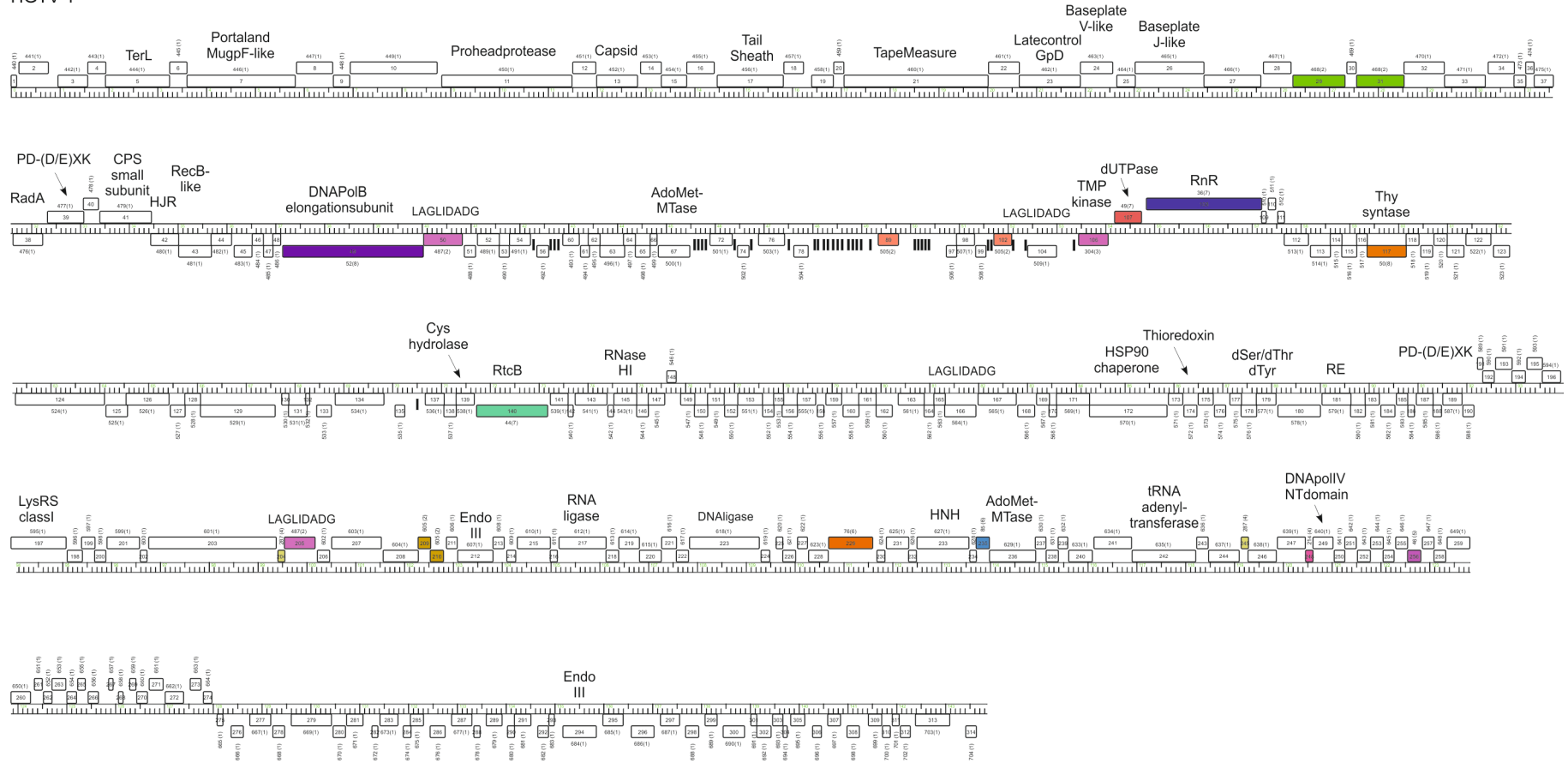


Figure S7. Genome map of HGTV-1. Pham and the number of the pham members are outlined above each ORF. TerL – terminase large subunit; AdoMet-MTase – S-adenosylmethionine-dependent methyltransferase; CPS – Carbamoyl phosphate synthase; dSer/dThr dTyr – Ser/thr and Tyr protein phosphatase; Endo III – endonuclease III; HJR – Holliday junction resolvase; HHN – Homing endonuclease; PD-(D/E)XK – nuclease; LAGLIDADG – endonuclease; LysRS classI – class I lysyl tRNA synthase; RadA – DNA dependent ATPase; RE – restriction endonuclease EcoRV; RecB-like – nuclease; RtcB – RNA ligase; RnR – ribonucleotide reductase; TMP – thymidine monophosphate.

HHTV-1

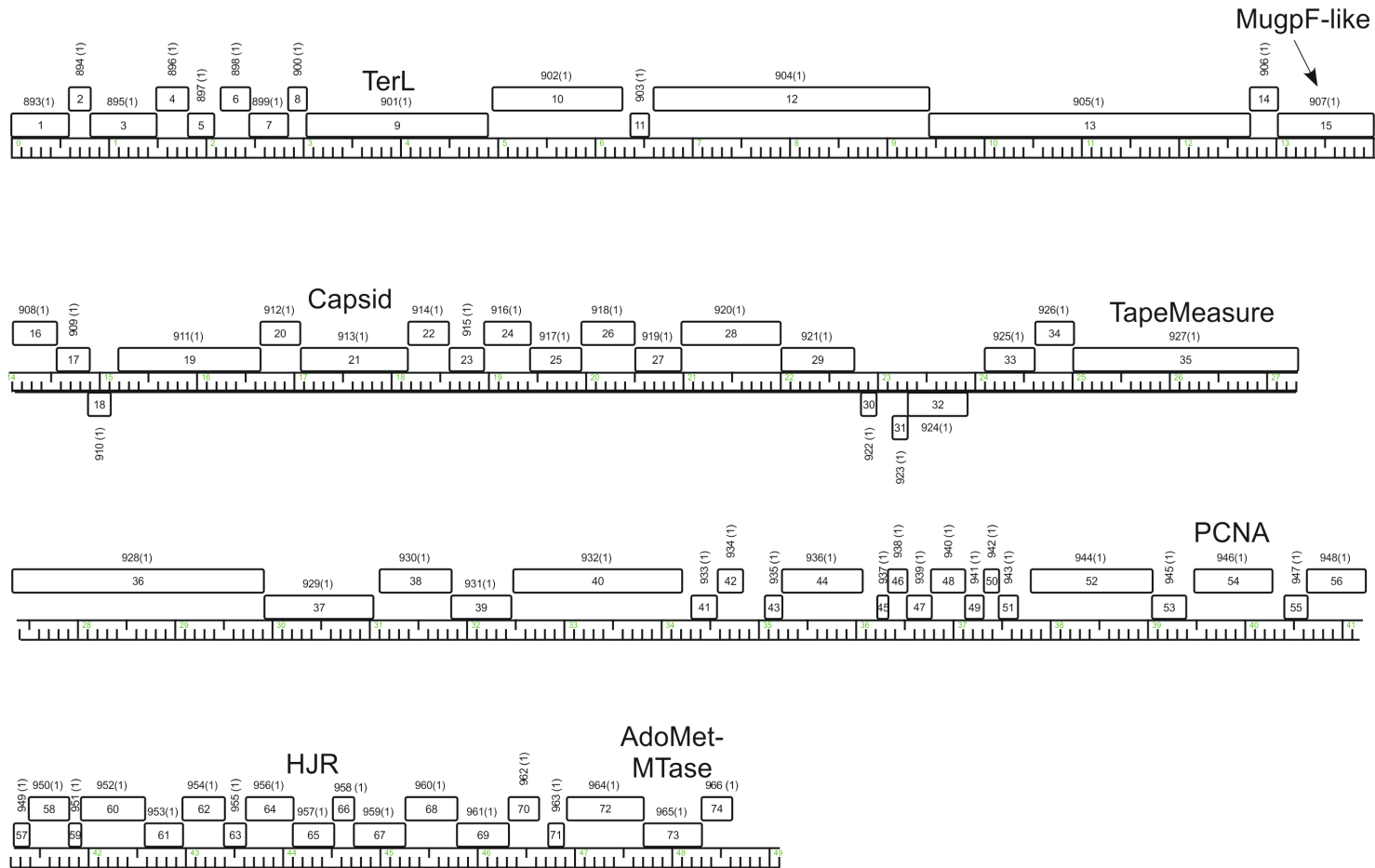


Figure S8. Genome map of HHTV-1. Pham and the number of the pham members are outlined above each ORF. TerL – terminase large subunit; AdoMet-MTase – S-adenosylmethionine-dependent methyltransferase; HJR – Holliday junction resolvase; PCNA – proliferating cell nuclear antigen.

HRTV-4

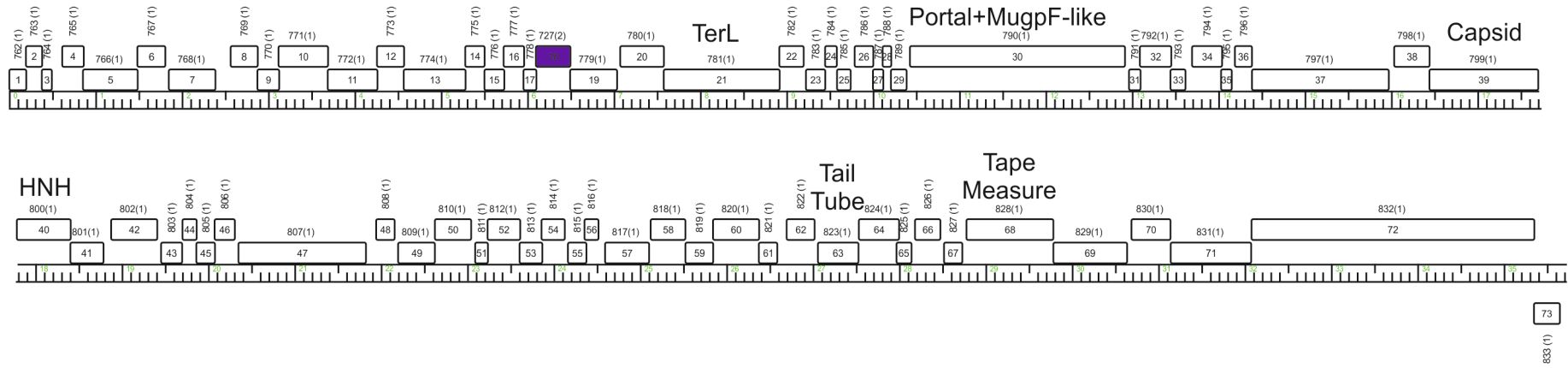


Figure S9. Genome map of HRTV-4. Pham and the number of the pham members are outlined above each ORF. TerL – terminase large subunit; AdoMet-MTase – S-adenosylmethionine-dependent methyltransferase; HNH – Homing endonuclease.


```

>ORF276 (128730-128680)
CCTTTT-TATAGTGGTATACTCCTCGTGGCACGGGGAGTGCTTAAGTAT-ATA
>ORF279 (130452-130402)
CCTTTT-TATAGTGGTATACTCCTCGTGGCACGAGGAGTCTTTTGGTAC-TTA
>ORF283 (131885-131835)
CCTTT--TATAGTGGTATACTCCTCGTGGCATGAGTATCGTTTAGGTATATAA
>ORF285 (132411-132361)
CCTTTT-TATACCCGTATACTCCTCGTGGCAAGGGGAGACTTATGGTAT-ATA
>ORF286 (132866-132816)
CCTTTT-TATAGTGATTTACTCCTCGTGGCAAGGGGAGTCTTTTGGTAC-TTA
>ORF288 (133579-133529)
TCTTTT-TATAGTGGTATACTCCTCGTG-CATGGGGAGTGTTATGGTACTTAA
>ORF289 (134022-133972)
CCTTTT-TATACCCTTACGCTCCTCGTGGCATGAGTATCGTTTAGGTAT-ATA
>ORF291 (134593-134543)
CCTTT--TATACCCGTATACTCCTCGTGGCACGGGGAGTCTTTGGTACTTAAA
>ORF293 (135102-135052)
CCTTTT-TATAGGGTTTACTCCTCGTGGCACGGGGAGGCTTTACCTAT-ATA
>ORF294 (135956-135906)
TCTTTT-TATAGGGTATACTCCTCGTGGCACGAGGAGGATTATGGTAT-ATA
>ORF295 (136549-136499)
ACTTTTATATAGGGTTTACTCCTCGTGACATGGCGAGGATTATGGTAT--AT
>ORF296 (137144-137094)
CCTTTT-TATAGTGGTATACTCCTCGTG-CAAGACGAGCGTTATGGTATATAA
>ORF297 (137668-137618)
TCTTTT-TATAGTGGTATACTTCCCGTGGCACGAGGAGCGTTATGGTAC-TTA
>ORF298 (138056-138006)
CCTTTT-TATAGTGGTATACTCCTCGTGGCACGACGAGTATTATAGTAT-ATA
>ORF299 (138435-138385)
CCTTTT-TATATGGGTTTACTCCTCGTGGCACGAGGAGTATTATGATAT-ATA
>ORF300 (139002-138952)
CTTTTA-TATAGTGTTATGCTCCGCATGGCACGAGGAGGATTATGGTAC-TTA
>ORF302 (139454-139404)
CCTTTT-TATAGGGTTTACTCCTCGTGGCACGAGGAGGATTACCCTAC-TTA
>ORF304 (139785-139735)
GGTTT-ATATAGGGGATTACTCCTCGTGGCACGACGAGCGTTATGGTAT-ATA
>ORF305 (140257-140207)
ACTTTTATATAGGGTTTACTCCTCGTGGCATGAGGAGTCTTTTGGTAT--AT
>ORF306 (140575-140525)
ACTTTTATATAGGGTTTACTCCTCGTGGCACGAGGAGGATTACCCTAC--TT
>ORF308 (141350-141300)
GGATTTATATAGGGTTTGTCTCCCGTGGCACGAGGAGTGTATGATAC--AT
>ORF309 (141808-141758)
GCTTTTATATAGGTGTTTGTCTCCCGTGGCACGAGGAGTATTATGGTAT--AT
>ORF311 (142166-142116)
AGTTTA-TATACCCGAATACTCCCCATGGCACGAGGAGGAAACACCTAT-ATA

```

Figure S10. Alignment of the repeat sequences in HGTV-1 genome. The header of the sequence indicates the ORF upstream of which the repeat is found. The nucleotide positions of the repeat are indicated in brackets. Hyphens in the sequences represent gaps in the alignment.