

Table S4. Detailed analysis of 188 Marcellivirus proteins having homologs in Redseq.

ORF	Marcellivirus self-cluster	Marcellivirus annotation	Mimi-, Marseille, <i>Acanthamoeba</i> grouped? Yes/No	Best Refseq hit			Inferred Origin**	Corresponding ML tree number in Fig. S4	Mimi- and Marcellivirus only? Yes/No	Hits to <i>Acanthamoeba</i>		
				nick	Gi number	Phyla*				Bit score, if better than to Refseq	# of hits	<i>Acanthamoeba</i> hits: Yes/No
DNA replication, recombination, repair (37 proteins, 22 trees)												
293		AP (apurinic) endonuclease family 2 - bacterial	No	Bb_Bacvu	150004972	B	NCLDV	4.1	No	0	No	
173		DNA topoisomerase II	No	El_Enccu	19074112	E	NCLDV	4.2	No	2	Yes	
156	Cluster 4	Restriction endonuclease, contains N-terminal multiple Zn-ribbon domain	No	Vn_Acapo	55819502	V	NCLDV (Iridogroup)	4.3	No	0	No	
245	Cluster 4	Restriction endonuclease, contains N-terminal multiple Zn-ribbon domain	No	Vn_Acapo	55819502	V	NCLDV (Iridogroup)	4.4	No	0	No	
327	Cluster 4	Restriction endonuclease, contains N-terminal multiple Zn-ribbon domain	No	Vn_Acapo	55819502	V	NCLDV (Iridogroup)	-	No	0	No	
349	Cluster 4	Restriction endonuclease, contains N-terminal multiple Zn-ribbon domain	No	Vn_Acapo	55819502	V	NCLDV (Iridogroup)	-	No	0	No	
289		putative nuclease linked to bacterial restriction-modification	No	Vn_Lymdi	13358473	V	NCLDV (Iridogroup)	-	No	0	No	
409	Cluster 4	Restriction endonuclease, contains N-terminal multiple Zn-ribbon domain	No	Vn_Acapo	55819502	V	NCLDV (Iridogroup)	-	No	0	No	
40	Cluster 4	Restriction endonuclease, contains N-terminal multiple Zn-ribbon domain	No	Vn_Acapo	55819502	V	NCLDV (Iridogroup)	-	No	0	No	
154	Cluster 4	Zn-ribbon like motif, a part of Restriction endonuclease	No	Vn_Acapo	55819502	V	uncertain	-	No	0	No	
347	Cluster 4	Zn-ribbon like motif, a part of Restriction endonuclease	No	Vn_Acapo	55819502	V	uncertain	-	No	0	No	
301	Cluster 10	Vsr/MutH/archaeal HJR-family nuclease (C-terminal domain, N-terminal Zn-ribbon)	No	Ba_Corgl	19552988	B	Bacteria/phages	-	No	0	No	
371	Cluster 10	Csr/MutH/archaeal HJR family nuclease	No	Vn_Invir	15078985	V	Bacteria/phages	-	No	0	No	
249	Cluster 10	MutH/Vsr/archaeal HJR-like endonuclease	No	Vn_Invir	15078985	V	Bacteria/phages	-	No	0	No	
428	Cluster 10	Vsr/MutH/archaeal HJR family endonuclease	No	Vn_Invir	15078985	V	Bacteria/phages	-	No	0	No	
248	Cluster 9	HNH family endonuclease	No	Vn_Azoph	168495129	V	uncertain	4.4	No	0	No	
318	Cluster 9	HNH endonuclease	No	Bf_Clobo	237795751	B	other Eukaryota	4.5	No	0	No	
295	Cluster 9	HNH-family endonuclease (2 tandem HNH domains)	No	Bf_Blaha	221077669	B	other Eukaryota	4.6	No	0	No	
387a	Cluster 9	HNH-family endonuclease	No	Bb_Grafo	120436764	B	other Eukaryota	4.7	No	0	No	
387b	Cluster 9	HNH-family endonuclease	No	Bb_Grafo	120436765	B	other Eukaryota	4.7	No	0	No	
212		Proliferating Cell Nuclear Antigen (DNA replication and repair), also present in phycodna, irido	No	El_Homsa	4505641	E	NCLDV (Iridogroup)	4.8	No	1	Yes	

181	Cluster 19	putative nuclease linked to bacterial restriction-modification systems	No	El_Copci	169867348	E	other Eukaryota	4.9	No	0	No	
367	Cluster 19	putative nuclease linked to bacterial restriction-modification systems	No	El_Copci	169867142	E	other Eukaryota	4.9	No	0	No	
216		RuvC-like Holiday Junction Resolvase	No	Vn_Invir	109287975	V	NCLDV (Iridogroup)	4.10	No	0	No	
315		DNA repair exonuclease SbcCD, D subunit (active nuclease of the calcineurin family)	No	Vn_Invir	15078956	V	NCLDV (Iridogroup)	4.11	No	0	No	
118		SMC1 family ATPase	No	Vn_Invir	15078763	V	NCLDV (Iridogroup)	-	No	0	No	
258		AlkB, Alkylated DNA repair protein	No	Bp_Shefr	114562122	B	Bacteria/phages	4.12	No	3	Yes	
319		XRN1, 5'-3' exonuclease	No	E9_Ostlu	145347657	E	NCLDV (Iridogroup)	4.13	No	3	Yes	
404		AAA family ATPase	No	Ec_Tetth	118378556	E	NCLDV	4.14	No	4	Yes	
365		flap-specific endonuclease (poxvirus G5 homolog)	No	El_Yarli	50556476	E	NCLDV (Iridogroup)	4.15	No	3	Yes	
152		D5 family helicase-primase	No	Vn_Invir	15078896	V	NCLDV (Iridogroup)	4.16	No	0	No	
184		helicase origin-binding-protein (UL9) like (closest to mimi)	No	Vn_Acapo	55818882	V	Bacteria/phages (Herpesviridae)	4.17	No	0	No	
36	Cluster 14	Highly derived D5-like helicase-primase	No	Ew_Triva	123194009	E	Bacteria/phages	4.18	No	0	No	
459	Cluster 14	Highly derived D5-like helicase-primase	No	Ew_Triva	123194009	E	Bacteria/phages	4.18	No	0	No	
72		ATP-dependent DNA ligase	No	Vn_Erwph	219681337	V	uncertain	4.19	No	0	No	
329		DNA polymerase elongation subunit (family B)	No	E9_Vitvi	225431691	E	NCLDV (Iridogroup)	4.20	No	3	Yes	
398		uracil-DNA glycosylase, D4 (plus additional N-term)	No	Bp_gampr	225091234	B	Bacteria/phages	4.21	No	3	Yes	
408		helicase superfamily I member (RecD, ATP-dependent exoDNAse (exonuclease V), alpha subunit), ortholog in mimi	No	Bb_Salru	83814185	B	Bacteria/phages	4.22	No	0		
Nucleotide metabolism (11 proteins, 9 trees)												
217		dNMP kinase	No	Bf_Pelth	147678482	B	Bacteria/phages	4.23	Yes	0	No	
17		Dam-like adenine-specific DNA methylase	No	Vn_Parbu	157953891	V	NCLDV (Phycodna)	4.24	No	0	No	
429		bifunctional dihydrofolate reductase-thymidylate synthase, found in some phycodna	No	E9_Vitvi	225428412	E	Amoebozoa (Acan)	4.25	No	3	No	
157		A32-like packaging ATPase	No	Vn_Invir	109287966	V	NCLDV (Iridogroup)	4.26	No	1 (short)	Yes	
211		ribonucleoside diphosphate reductase, alpha subunit	No	Vn_Parbu	157953136	V	NCLDV	4.27	No	3	Yes	
268		Ribonucleotide reductase, small chain	No	Ek_Trybr	74025554	E	uncertain	4.28	No	1	Yes	
236		Thymidine kinase	No	Vn_Ostvi	163955199	V	other Eukaryota	4.29	No	weak	No	
131		DHH family phosphohydrolase	No	Bq_Sulye	237756294	B	other Eukaryota	4.30	Yes	0	No	
326		P-loop ATPase or GTPase	Yes (grouped by tree)	Vn_Acapo	55819479	V	Amoebozoa (Acan)	4.31	Yes	145	1	Yes

423		putative P-loop kinase or ATPase	No	El_Vanpo	156846500	E	uncertain	-	No	0	No	
15		dUTPase (fragment; w/o C-terminal)	No	Bb_Capoc	229255544	B	Bacteria/phages	-	No	0	No	
Signaling (Serine/Threonine protein kinases) (15 proteins, 12 trees)												
160	Cluster 2	Serine/Threonine protein kinase	No	Ec_Parte	145516466	E	Amoebozoa (Acan)	4.33	No	3	Yes	
195	Cluster 2	Serine/Threonine protein kinase	No	Ec_Parte	145545584	E	Amoebozoa (Acan)	4.34	No	3	Yes	
199	Cluster 2	Serine/Threonine protein kinases, catalytic domain	No	El_Triad	195999760	E	other Eukaryota	4.35	No	3	Yes	
191	Cluster 2	Serine/Threonine protein kinase	No	El_Drovi	195400809	E	other Eukaryota	4.36	No	0	No	
264	Cluster 2	Serine/Threonine protein kinase catalytic domain-containing protein	No	El_Macmu	109088453	E	Amoebozoa (Acan)	4.37	No	4	Yes	
424	Cluster 2	Serine/threonine-protein kinase	No	Ea_Dicdi	66809477	E	other Eukaryota	4.38	No	5	Yes	
84	Cluster 2	serine/threonine protein kinase	No	Bp_Acier	148261089	B	Amoebozoa (Enta)	4.39	No	3	Yes	
432	Cluster 2	Serine/threonine-protein kinase	No	Bp_Myxxa	108762484	B	Amoebozoa (Acan)	4.40	No	2(both are short)	Yes	
446	Cluster 2	Serine/threonine-protein kinase	No	Bf_Eubdo	160915564	B	other Eukaryota	4.41	No	3	Yes	
427	Cluster 2	Serine/threonine-protein kinase	No	El_Caebr	157752042	E	uncertain	4.42	No	3	Yes	
352	Cluster 11	divergent Serine/Threonine protein kinase	No	El_Brafl	219454788	E	uncertain	4.43	No	0	No	
353	Cluster 11	divergent Serine/Threonine protein kinase	No	E9_Orysa	115489454	E	uncertain	4.43	No	0	No	
360	Cluster 11	divergent Serine/Threonine protein kinase	No	Eh_Heman	160331667	E	uncertain	4.43	No	0	No	
260		divergent Serine/Threonine protein kinase	No	Vn_Invir	15078921	V	NCLDV (Iridogroup)	-	No	0	No	
122	Cluster 2	serine/threonine protein kinase, duplicated kinase domain	Yes (grouped by tree; unique)	Vn_Acapo	55819780	V	Amoebozoa (Acan)	4.32	Yes	Y	Yes	
Protein and lipid synthesis, modification, and degradation (28 proteins, 10 trees)												
269		chaperone protein DnaJ (both J and Zn-finger domains, bacterial type)	No	Bp_CanBl	71891902	B	NCLDV	4.44	No	3	Yes	
310		Erv1/Alr family protein - ortholog of poxvirus E10, disulfide oxidoreductase	No	Vn_Infsp	19881448	V	NCLDV	4.45	No	58,2	1	Yes
151	Cluster 3	F-box containing protein	No	El_Chagl	116208368	E	other Eukaryota	-	No	0	No	
19	Cluster 3	F-box containing protein	No	El_Trica	91088783	E	other Eukaryota	-	No	0	No	
42	Cluster 3	F-box containing protein	No	Bh_Chlau	163849439	B	other Eukaryota	-	No	0	No	
28	Cluster 3	F-box containing protein	No	El_Aspte	115390899	E	other Eukaryota	-	No	0	No	
106	Cluster 3	F-box containing protein	No	Bu_Fusmo	237736453	B	Bacteria/phages	-	No	1(short)	Yes	
146	Cluster 3	F-box containing protein	No	El_Brafl	219484040	E	other Eukaryota	-	No	weak	No	
385	Cluster 3	F-box containing protein	No	El_Culqu	170052485	E	uncertain	-	No	weak	No	

39	Cluster 3	F-box containing protein	No	El_Bosta	194678444	E	other Eukaryota	-	No	weak	No	
285	Cluster 18	papain-like cysteine peptidase (Cathepsin B group)	No	XX_Giala	159109223	E	Amoebozoa (Acan)	4.46	No	4	Yes	
276	Cluster 18	papain-like cysteine peptidase (Cathepsin B group) unique among NCLDV	No	E9_Phypa	168026641	E	Amoebozoa (Acan)	4.46	No	3	Yes	
189		lipase class 3 (bacteria-specific)	No	El_Nemve	156304116	E	NCLDV	4.47	No	weak	No	
223		Lipase class 3	No	Ec_Tetth	118378943	E	Amoebozoa (Acan)	4.47	No	2	Yes	
334		inactivated thioredoxin/glutaredoxin, poorly conserved homologs in some irido, phycodna, mimi	No	Vn_Invir	109287908	V	uncertain	-	No	0	No	
278		metallopeptidase WLM	No	Vn_Invir	109287920	V	NCLDV (Phycodna)	4.48	No	2(short)	Yes	
228		Patatin-like phospholipase, also present in phycodna, irido, asco	No	Vn_Acapo	55819440	V	uncertain	4.49	No	0	No	
206		RING-finger-E3 Ub ligase	No	Ek_Trycr	71420237	E	other Eukaryota	-	No	short	No	
231	Cluster 21	thioredoxin	No	Bp_Psest	146280962	B	Bacteria/phages	-	No	1	Yes	
332	Cluster 21	Thioredoxin	No	Ec_Parte	145527612	E	other Eukaryota	-	No	4 hits; all are short	Yes	
313	Cluster 21	thioredoxin	No	Vn_Invir	15078908	V	Bacteria/phages	-	No	0	No	
247		Ubiquitin	No	Ec_Tetth	118399883	E	Amoebozoa (Acan)	4.50	No	133	10	Yes
416		peptidoglycan peptidase	No	Ec_Parte	145498469	E	Bacteria/phages	4.51	No	0	No	
331		ankyrin repeat protein	No	Bp_Sorce	162451350	B	Bacteria/phages	4.52	No	1(weak)	Yes	
316		ankyrin repeat-containing protein	No	Bp_Ricbe	91205812	B	Bacteria/phages	-	No	weak	No	
330		ankyrin repeat-containing protein	No	El_Aspor	169782465	E	other Eukaryota	-	No	weak	No	
445		ankyrin repeat-containing protein	No	Ea_Dicdi	66801041	E	other Eukaryota	-	No	4	Yes	
407		ubiquitin-like protein	No	El_Brafl	219418836	E	other Eukaryota	4.53	No	0	No	
Transcription and chromatin modification (31 proteins, 16 trees)												
203		D6/D11-like helicase	No	Vn_Acapo	55819789	V	NCLDV (Poxvir)	4.54	No	1 (short)	Yes	
241		D6/D11 like helicase	No	Vn_Lymdi	51870029	V	NCLDV (Iridogroup)	4.54	No	0	No	
270		helicase of superfamily II	No	Vn_Invir	15078874	V	NCLDV (Iridogroup)	4.55	No	0	No	
166	Cluster 15	Histone 2A-domain-containing protein	No	El_Danre	68442127	E	Amoebozoa (Dict)	4.56	No	1(+3 orf414)	Yes	
414	Cluster 15	histone H2B (partial) fused to histone H2A	No	El_Musmu	28316756	E	Other Eukaryota	4.56	No	3	Yes	
413		histone H3	No	Ek_Leima	157868174	E	Other Eukaryota	4.57	No	0	No	
306		Poxvirus early transcription factor (VETF), large subunit	No	Vn_Invir	15079007	V	NCLDV (Iridogroup)	-	No	0	No	

197		RNA ligase (conserved in irido and asco[previously undetected]; in baculoviruses fused to a C-terminal polynucleotide kinase)	No	Vn_Invir	15079104	V	NCLDV (Iridogroup)	4.58	No	0	No	
314		3-domain mRNA capping enzyme, apparently, not of viral origin	No	E9_Ostlu	145343780	E	Amoebozoa (Acan)	4.59	No	1	Yes	
213		Transcription initiation factor TFIIB	No	Ae_Metka	20094294	A	NCLDV (Phycodna)	4.60	No	0	No	
277		NUDIX hydrolase (D10 ortholog)	No	Vn_Acapo	55819250	V	uncertain	-	No	0	No	
242		NUDIX hydrolase (D1/D9 homologs, decapping enzyme?)	No	Bp_Dicda	242238628	B	Bacteria/phages	-	No	0	No	
291		SWIB/MDM2 domain-containing protein - chromatin condensation, irido-phycodna	No	E9_Poptr	224056218	E	uncertain (Eukaryota of NCLDV?)	-	No	0	No	
340		Transcription Factor A2-like	No	Vn_Invir	15078994	V	NCLDV	-	No	0	No	
129		ribonuclease H - bacterial	No	Bp_Meslo	13476236	B	Bacteria/phages	4.61	No	0	No	
244		ribonuclease III (conserved in irido, asco)	No	Vn_Ambti	45686033	V	NCLDV (Iridogroup)	4.62	No	0	No	
222		Transcription factor S-II (TFIIS)-domain-containing protein, RNAP subunit	No	Ew_Triva	123425416	E	NCLDV	4.63	No	0 (checked, deleted)	No	
383		RNAP subunit Rpb5 - highly divergent from other viruses	No	E9_Zeama	226530454	E	NCLDV	4.64	No	1	Yes	
193		multiple Zn fingers	Yes (unique to these three)	Vn_Acapo	55819526	V	Amoebozoa (Acan)	4.65 (without eukaryota other than Acanthamoeba)	Yes	249	4	Yes
321	Cluster 28	zinc finger protein	No	El_Brafl	219422757	E	other Eukaryota	-	No	0	No	
70	Cluster 28	Zn-finger protein	No	El_Anoga	158285292	E	other Eukaryota	-	No	0	No	
324		zinc finger	No	El_Ratno	149274619	E	other Eukaryota	-	No	0	No	
81		Zn-ribbon-like motif, derived cytidine deaminase?	No	El_Yarli	50555826	E	uncertain	4.66	No	0	No	
344		Zn-ribbon-containing protein	No	Bf_Bacce	196040681	B	Bacteria/phages	-	No	0	No	
391		multiple zinc ribbon protein	No	El_Aedae	157129122	E	other Eukaryota	-	No	0	No	
322		zinc finger protein	No	El_Musmu	110626083	E	other Eukaryota	-	No	0	No	
69		zinc finger protein	No	E9_Vitvi	225447701	E	other Eukaryota	-	No	0	No	
55		DNA-directed RNA polymerase subunit alpha; RNA polymerase Rpb1, domains 3-7	No	Ac_Stama	126465703	A	NCLDV	4.67	No	3	Yes	
52		DNA-directed RNA polymerase subunit alpha; RNA polymerase Rpb1, domains 1-2	No	El_Yarli	50557434	E	NCLDV	4.68	No	3	Yes	
23		DNA-directed RNA polymerase subunit beta	No	Ec_Plakn	221059854	E	other Eukaryota	4.69	No	1 (plus 2 short)	Yes	
237		A1L transcription factor	No	Vn_Invir	109287904	V	NCLDV	-	No	0	No	
Translation (5 proteins, 5 trees)												
163		Translation elongation factor EF-1alpha	No	Ea_Dicdi	66816687	E	Amoebozoa (Dict)	4.70	Yes	3	Yes	

79		Translation initiation factor SUI1	No	El_Yarli	50548179	E	Amoebozoa (Acan)	4.71	Yes	65,1	1	Yes
130		Eukaryotic peptide chain release (translation termination) factor 1	No	El_Neur	85084324	E	Amoebozoa (Acan)	4.72	Yes		2	Yes
400		cell division cycle 123 (affects eIF2) homolog	Yes (with four bacteria)	Vn_Acapo	55819747	V	Bacteria/phages	4.73	Yes		0	No
304		eukaryotic translation initiation factor 5 (N-terminal Zn-ribbon domain, homolog of eIF2-beta)	No	El_Ajeca	154274059	E	Amoebozoa (Acan)	4.74	No		3	Yes
Miscellaneous (23 proteins, 8 trees)												
99		Flavin-containing amine oxidoreductase	Yes (one group on the tree)	Vn_Acapo	55819068	V	Amoebozoa (Acan)	4.75	No		1	Yes
161		mannosyltransferase OCH1-like protein	No	Bb_Leeb1	86142254	B	uncertain	-	No			No
162		mannosyltransferase OCH1-like protein	No	Bp_Fraph	241668223	B	other Eukaryota	4.76	Yes		0	No
328		glycosyl transferase	No	Bb_Pardi	150009744	B	Bacteria/phages	4.76	Yes		0	No
363		calcineurin-like phosphoesterase	No	Bp_Rhisp	227822772	B	uncertain	4.77	Yes		3	Yes
210		Rossmann-fold nucleotide-binding protein widely conserved in bacteria and phages, truncated gene in mimi, functionally linked to NDR?	No	Bb_Capgi	228472699	B	Bacteria/phages	4.78	Yes		short	No
103	Cluster 1	MORN-repeat protein (bacterial)	No	Bu_Fusnu	197737419	B	Bacteria/phages	-	No		Y (repeats)	Yes
159	Cluster 1	MORN-repeat protein (bacterial)	No	Bi_Fibsu	241899983	B	Bacteria/phages	-	No		Y (repeats)	Yes
124	Cluster 1	MORN-like domain	No	Ba_Nocfa	54022764	B	Bacteria/phages	-	No		Y (repeats)	Yes
361	Cluster 1	MORN-repeat protein (bacterial)	No	Bu_Fusnu	197735567	B	Bacteria/phages	-	No		Y (repeats)	Yes
355	Cluster 1	MORN-repeat protein (bacterial)	No	Bu_Fussp	237743875	B	Bacteria/phages	-	No		Y (repeats)	Yes
366	Cluster 1	MORN-repeat protein (bacterial)	No	Bu_Fussp	237740028	B	Bacteria/phages	-	No	41,2	Y (repeats)	Yes
434	Cluster 1	MORN repeat-containing protein (a protein conserved in bacteria)	No	Bb_Capgi	228471690	B	Bacteria/phages	-	No		Y (repeats)	Yes
134	Cluster 1	MORN-repeat protein	No	Bp_Helci	224438121	B	Bacteria/phages	-	No		Y (repeats)	Yes
137	Cluster 1	MORN repeat protein	No	Bp_Burth	167836308	B	Bacteria/phages	-	No		0	No
142	Cluster 1	MORN repeat protein	No	Ba_Strsv	197785586	B	Bacteria/phages	-	No		0	No
143	Cluster 1	MORN repeat protein	No	Ec_Thean	85000317	E	Bacteria/phages	-	No		Y	Yes
144	Cluster 1	MORN repeat protein	No	El_Aspni	145253667	E	Bacteria/phages	-	No		Y	Yes
279		bacterial MORN repeats	No	B1_Elumi	187250911	B	Bacteria/phages	-	No		0	No
74		Metal dependent phosphohydrolase with conserved 'HD' motif	No	Bb_Chrg1	227368134	B	other Eukaryota	4.79	Yes		1 (short; checked, removed)	Yes
426		WD repeat protein	No	Ea_Dicdi	66811920	E	Amoebozoa (Acan)	4.80	Yes		3	Yes
451		glycosyltransferase	No	El_Monbr	167520782	E	uncertain	4.81	No		0	No

342		major capsid protein (ortholog of poxvirus D13)	No	Vn_Invir	15078986	V	NCLDV	4.82	No	2	Yes	
		unknown (conserved protein) (38 proteins, 7 trees)										
339		unknown, conserved in irido, phycodna, and mimi	No	Vn_Invir	15078999	V	NCLDV	-	No	0	No	
126		uncharacterized protein conserved in several bacteria and eukaryotes, and mimivirus (low similarity)	Yes (one group on the tree)	Bp_Rossp	225130762	B	Amoebozoa (Acan)	4.83	No	109	1	Yes
183	Cluster 7	contains an uncharacterized conserved domain found in diverse bacterial and phage proteins	No	Bp_Escoco	218703839	B	Bacteria/phages	4.84	No		0	No
180	Cluster 7	T5orf172 domain (pfam1544)-containing protein	No	Ew_Triva	123367738	E	uncertain	-	No		0	No
34	Cluster 7	contains an uncharacterized conserved domain found in diverse bacterial and phage proteins	No	Bp_Xanau	154247566	B	Bacteria/phages	4.84	No		0	No
263		conserved in only 2 iridoviruses	No	Vn_Invir	15079041	V	uncertain	-	No		0	No
104		Uncharacterized protein conserved in bacteria and mimivirus	No	Ea_Dicdi	166240195	E	uncertain	4.85	Yes		0	No
111		uncharacterized protein with homologs in some bacteria	No	Bp_Metpe	124262756	B	Amoebozoa (Acan)	4.86	No	229	1	Yes
230		uncharacterized, conserved domain in MIMI_L371	Yes (one group on the tree)	Vn_Acapo	55819246	V	Amoebozoa (Acan)	4.87	No		1	Yes
255		uncharacterized TIM-barrel protein conserved in numerous bacteria	No	Bp_Geosp	222054426	B	Bacteria/phages	4.88	No		0	No
420		uncharacterized protein conserved in bacteria and archaea	No	Ba_Micsp	238061377	B	Amoebozoa (Acan)	4.89	No		1	Yes
296		membrane protein	No	Vn_Invir	15078925	V	NCLDV	-	No		0	No
253		low compl prot, proline-rich repeats	No	Bf_Cloth	125975557	B	uncertain	-	No		Y (low complexity)	Yes
368		low complexity	No	El_Droan	194753654	E	uncertain	-	No		Y (low complexity)	Yes
114		HA-like repeats	No	Bp_Limsp	149925860	B	Bacteria/phages	-	No		Y (repeats)	Yes
26	Cluster 5	distinct repeats; similar to hypothetical protein PBCV1_A540L	No	Bi_CanKo	94967383	B	Bacteria/phages	-	No		Y (repeats)	Yes
67	Cluster 5	low complexity, HA-like repeats	No	Bi_CanKo	94967383	B	Bacteria/phages	-	No		Y (repeats)	Yes
45	Cluster 5	low complexity, HA-like repeats	No	Bi_CanKo	94967383	B	Bacteria/phages	-	No		Y (repeats)	Yes
66	Cluster 5	low complexity, HA-like repeats	No	Bi_CanKo	94967383	B	Bacteria/phages	-	No		Y (repeats)	Yes
44	Cluster 5	distinct repeats	No	Bb_Leeb1	86140522	B	Bacteria/phages	-	No		Y (repeats)	Yes
25	Cluster 5	low complexity, HA-like repeats	No	Bf_Weipa	241894892	B	Bacteria/phages	-	No		Y (repeats)	Yes
149	Cluster 16	low complexity	No	El_Brafl	219481367	E	uncertain	-	No		Y (repeats)	Yes
73	Cluster 16	low complexity, invasin-like repeats	No	Bp_Escoco	110643846	B	uncertain	-	No		Y (repeats)	Yes
24		T5orf172 domain (pfam1544)-containing protein	No	Ba_Rhoer	33867084	B	Bacteria/phages	-	No		1	Yes
403		RING-domain (cd00162) -containing protein	No	Vn_Acapo	55819461	V	uncertain	-	No		2 (short)	Yes
229		uncharacterized ,C-terminal domain conserved in irido, asco, phycodna, mimi	No	Vn_Invir	15078780	V	NCLDV	-	No		0	No

333	unknown, orthologs in all phycodna	No	Vn_Ostvi	163955082	V	NCLDV	-	No	0	No
259	uncharacterized, unique to Chilo virus	No	Vn_Invir	15078840	V	NCLDV	-	No	0	No
16	similar to hypothetical protein PBCV1_A579L	No	Vn_Parbu	9632130	V	NCLDV	-	No	0	No
338	membrane protein, also homolog in mimi	No	Vn_Acatu	155371047	V	NCLDV	-	No	0	No
369	distant homolog only in Chilo iridovirus	No	Vn_Invir	15079069	V	uncertain	-	No	0	No
219	low complexity, leucine rich repeat protein	No	El_Ratno	62638851	E	uncertain	-	No	0	No
18	conserved in Mimi-, Mama-, marseillevirus and Chordopoxvirinae. pfam00651;BTB/POZ domain	No	E9_Zeama	219362591	E	other Eukaryota	-	No	0	No
433	hypothetical protein conserved inMimi- and Marseillevirus	No	Vn_Acapo	55819715	V	NCLDV	-	Yes	0	No
430	hypothetical protein conserved inMimi- and Marseillevirus	Yes (no tree however)	Vn_Acapo	55819043	V	NCLDV	-	Yes	3	Yes
233	uncharacterized, poorly conserved inirido, phycodna	No	Vn_Invir	15079047	V	NCLDV	-	No	0	No
200	hypotetical protein found in Irido- and marseillevirus	No	Vn_Invir	15079178	V	NCLDV	-	No	weak	No
21	uncharacterized bacterial protein	No	Bd_Deira	15806713	B	Bacteria/phages	-	No	0	No

* B, *Bacteria*; E, *Eukaryota*; V, *Virus*

**Acan, *Acanthamoeba*; Dict, *Dictyostelium*; Enta, *Entamoeba*; Phycodna, *Phycodnaviruses*; Poxvir, *Poxviridae*