Supplementary figures

S1 Multiple sequence alignment (MSA) of *Yaraviridae* MCP proteins.

S2 MSA of "*Gamadviridae*" MCP proteins.

S3 Top hits of HHpred searches initiated with *Yaraviridae* and "*Gamadviridae*" MCP protein MSAs.

S4 VLTF2 protein MSA.

S5. Top hits of HHpred searches initiated with Mriyavirus VLTF2 protein MSA.

S6. Top HHpred hits for "*Gamadviridae*" (right) and *Yaraviridae* (left) RuvC MSAs.

S7. Phylogenetic tree of RuvC proteins.

S8. MSA of selected SF3_hel1 proteins.

S9. Phylogenetic tree of SF3_hel1 proteins.

S10. SF3_hel2 protein MSA.

S11. Top hits of HHpred search initiated with Mriyavirus HUH endonuclease MSA.

S12. MSA of HUH and C-terminal domain of HUH_long proteins.

S12. MSA of mriyavirus PDDEXK nuclease.

S13. Phylogenetic tree of PDDEXK nucleases.

S14. MSA of Yaravirus gene 48 and its homologs in mriyaviruses (Mriya_48) and in NCBI Genbank (nr) database.

S15. A, Phylogenetic tree of Mriya_48; B, predicted structures of two Mriya_48 proteins.

S16. MSA of Yaravirus gene 50 and its homologs in mriyaviruses (Mriya_50) and in NCBI Genbank (nr) database.

S17. Predicted structure of Mriya_50 (exemplified with Ple2@KY346835_21). Leucine site chains highlighted in magenta.

S18. MSA of conserved domain of Yaravirus gene 1 and its homologs in mriyaviruses (Mriya_1).

S19. Predicted structure of Mriya_1, exemplified with Ple2@KY346835_29.

S20. A, MSA of Yaravirus gene 51 with mriyavirus and nr homologs. B, Predicted structure of Yaravirus gene 51.

S21. A, Predicted structure of PEV_22; B, same structure overlayed with 6g41_A – the structure of minor capsid protein of Cafeteria mavirus.

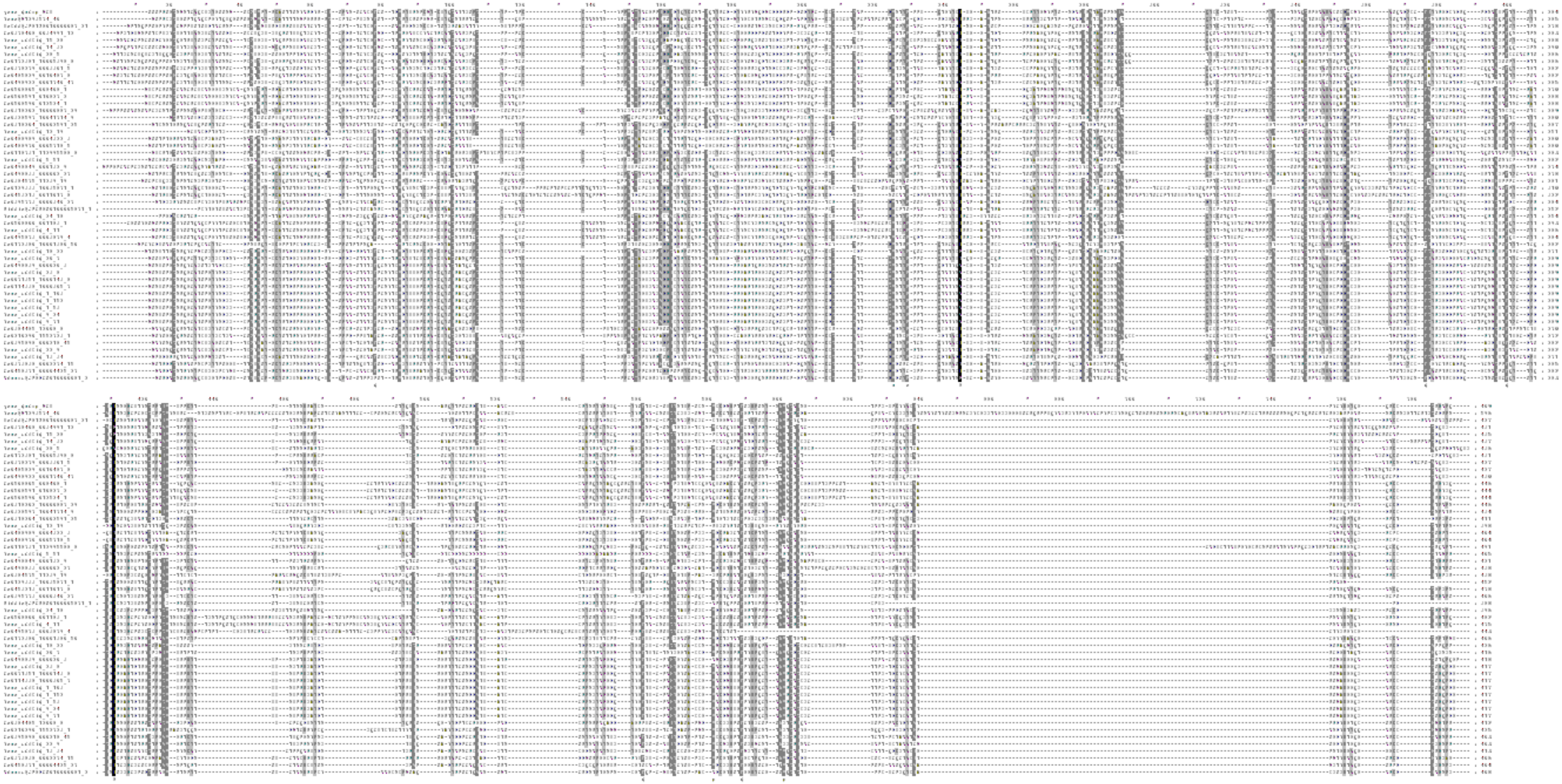
S22. Predicted structure of PEV_26 superimposed with gp2.5, a single strand DNA binding protein from bacteriophage T7 (pdb 1je5, green, z-score 13.3).

S23. Predicted structure of Yaravirus minor capsid protein

S24. MSA of Yaravirus minor capsid protein and related sequences.

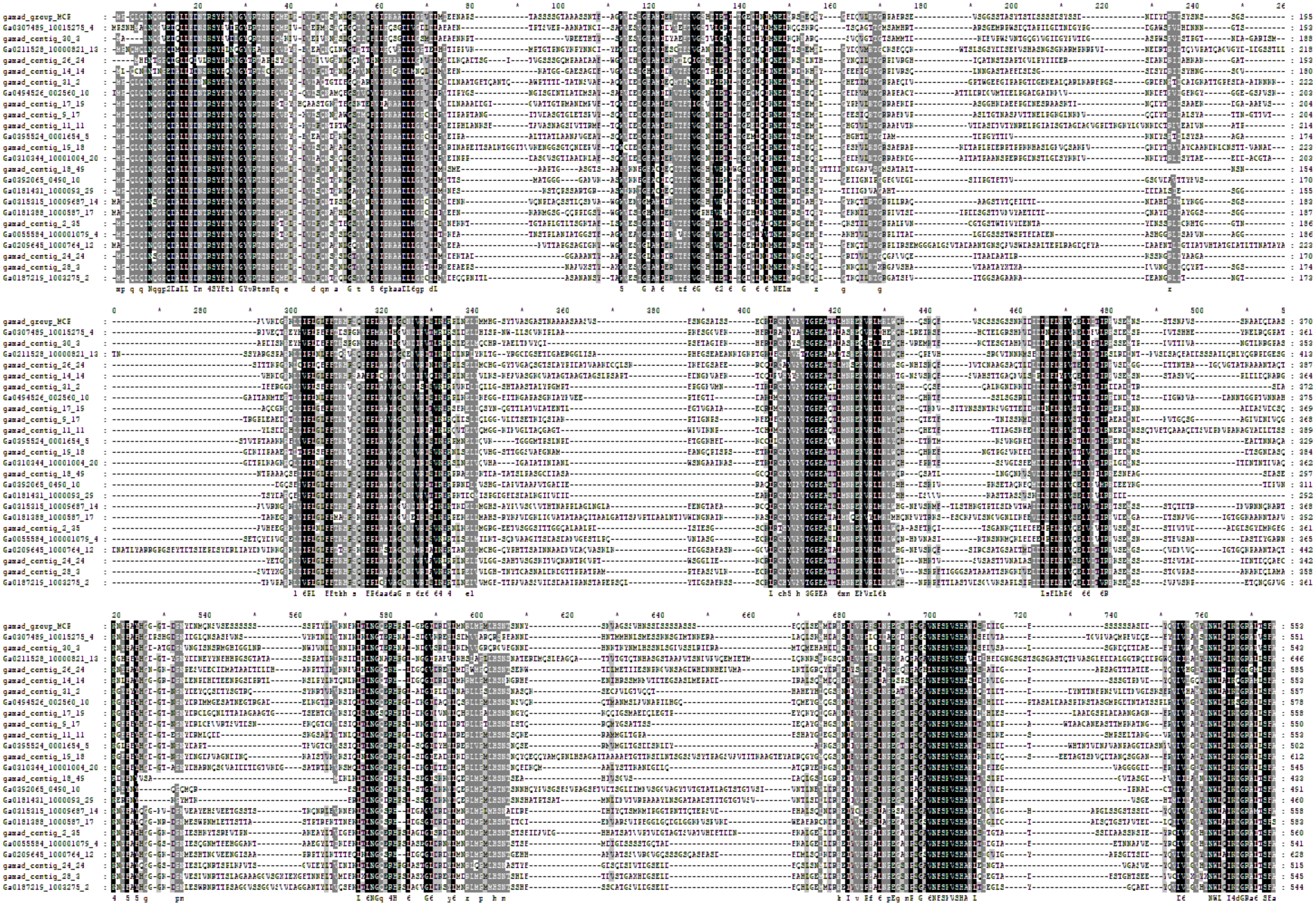
S1

Yaraviridae MCP



S2

“*Gamadviridae*” MCP



S3

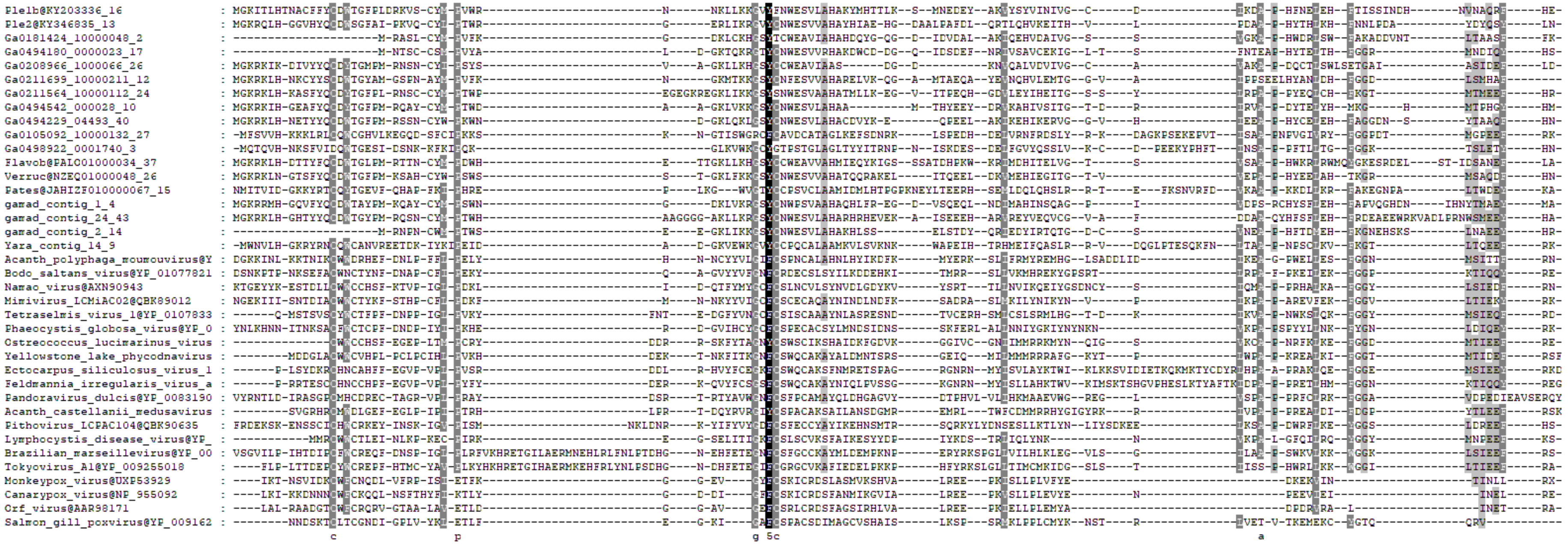
HHpred search initiated from both MCP alignments brings mimi-, irido-, asco-, and phycodnavirus MCP with high confidence.

Nr	Hit	Name	Probability	E-value	Score	SS	Aligned cols	Target Length
☐ 1	Q7T6X4	CAPS2_MIMIV Probable capsid protein 2 OS=Acanthamoeba polyphaga mimivirus OX=212035 GN=MIMI_R439 PE=3 SV=2	100	2e-74	585.41	34.7	395	591
☐ 2	8H2I_bx	MCPv1; giant virus, nucleocytoplasmic large DNA viruses (NCLDVs), viral assembly, Paramecium bursaria chlorella virus 1	100	2.6e-73	573.31	37	397	520
☐ 3	8RBT_O	Major capsid protein; cryo-EM, subtomogram averaging, EhV-201, enveloped virus, capsid, major capsid protein, VIRUS; 12.	100	1.3e-72	565.58	39.1	390	496
☐ 4	8H2I_bC	MCPv2; giant virus, nucleocytoplasmic large DNA viruses (NCLDVs), viral assembly, Paramecium bursaria chlorella virus 1	100	6.9e-73	560.36	35.9	379	486
☐ 5	Q7T6Y5	CAPS4_MIMIV Probable capsid protein 4 OS=Acanthamoeba polyphaga mimivirus OX=212035 GN=MIMI_R441 PE=3 SV=2	100	5.7e-73	567.88	34	391	512
☐ 6	O39164	MCP_IRV16 Major capsid protein OS=Costelytra zealandica iridescent virus OX=68348 GN=MCP PE=3 SV=1	100	1.1e-72	563.02	33.3	391	462
☐ 7	A4KXB1	CAPSD_HVAVE Major capsid protein OS=Heliothis virescens ascovirus 3e OX=260797 GN=MCP PE=3 SV=1	100	9.5e-73	561.7	32.3	391	455

Nr	Hit	Name	Probability	E-value	Score	SS	Aligned cols	Target Length
☐ 1	Q7T6X4	CAPS2_MIMIV Probable capsid protein 2 OS=Acanthamoeba polyphaga mimivirus OX=212035 GN=MIMI_R439 PE=3 SV=2	100	1.3e-77	669.27	40.4	419	591
☐ 2	Q7T6Y5	CAPS4_MIMIV Probable capsid protein 4 OS=Acanthamoeba polyphaga mimivirus OX=212035 GN=MIMI_R441 PE=3 SV=2	100	1.9e-76	651.33	44.5	425	512
☐ 3	O39164	MCP_IRV16 Major capsid protein OS=Costelytra zealandica iridescent virus OX=68348 GN=MCP PE=3 SV=1	100	3e-76	643.72	44.8	434	462
☐ 4	8H2I_bx	MCPv1; giant virus, nucleocytoplasmic large DNA viruses (NCLDVs), viral assembly, Paramecium bursaria chlorella virus 1	100	5.4e-76	650.01	45.8	408	520
☐ 5	5TIP_B	Major capsid protein; VIRAL PROTEIN; HET: 7CV, XYP, FUC, BGC, RM4, MAN, XXR, GLA; 2.0A {Paramecium bursaria Chlorella vi	100	7.3e-76	636.15	42.6	378	436
☐ 6	A4KXB1	CAPSD_HVAVE Major capsid protein OS=Heliothis virescens ascovirus 3e OX=260797 GN=MCP PE=3 SV=1	100	1.4e-75	637.32	44.6	411	455
☐ 7	P22176	MCP_LCDV1 Major capsid protein OS=Lymphocystis disease virus 1 (isolate Darai) OX=654922 GN=MCP PE=3 SV=2	100	9.2e-76	639.42	43.3	404	459

top HHpred hits for Yaraviridae (left) and “Gamadviridae” (right) MCP alignments

S4



```
Ple1b@KY203336_16              : MGKITLHTNACFFYCDWTGFPLDRKVS-CYM-PVWR--------------------------N------NKLLKKGVYFNWESVLAHAKYMHTTLK--S--MNEDEY--AKVYSYVINIVG--C------D--------------IKDA-P-HFNEIEH--FTISSINDH-------NVNAQRF-----HE-
Ple2@KY346835_13               : MGKRQLH-GGVHYQCDWSGFAI-PKVQ-CYL-PTWR--------------------------G------ERLIKRGVYCNWESVVAHAYHIAE-HG--DAALPAFDL--QRTLQHVKEITH--V------L--------------PDAA-P-HYTHIKH--FNNLPDA----------YDYQSY-----LN-
Ga0181424_10000048_2           : ---------------------M-RASL-CYM-PVFK--------------------------G------DKLCKHGSYTCWEAVIAHAHDQYG-QG--D--IDVDAL--AKIQEHVDAIVG--S------S--------------VGKA-P-HWDRISW--FAKADDVNT--------LTAASF-----FK-
Ga0494180_0000023_17           : ---------------------M-NTSC-CSM-PVYA--------------------------L----D-GKTQKRGTYCNWESVVRHAKDWCD-DG--Q--IDSDEF--NRIVSAVCEKIG--L-T----S--------------FNTEAP-HYTEITH--FGGR-------------MNDIQY-----HS-
Ga0208966_1000066_26           : MGKRKIK-DIVYYQCDYTGMPM-RNSN-CYI-PSYS--------------------------V----A-GKLLKHGSYCCWEAVIAAS------DG--D----------KNVQALVDVIVG--C------A--------------VAKA-P-DQCTISWLSETGAI------------ASIDEF-----LD-
Ga0211699_10000211_12          : MGKRKLH-NKCYYSCDWTGYAM-GSPN-AYM-PVFK--------------------------N------GKMTKKGSYCNFESVVAHARELVK-QG--A--MTAEQA--YEVNQHVLEMTG--G-V----A--------------IPPSEELHYANIDH--FGGD-------------LSMHAF--------
Ga0211564_10000112_24          : MGKRKLH-KASFYQCDYTGFPL-RNSC-CYM-PTWP--------------------------EGEGKREGKLIKKGSYSNWESVAAHATMLLK-EG--V--ITPEQH--GDVLEYIHEITG--S-S----Y--------------LRPA-P-PYEQICH--FKGT-------------MTMEEF-----HR-
Ga0494542_000028_10            : MGKRKIH-GEAFYQCDYTGFPM-RQAY-CYM-PTWD--------------------------A----A-GKLVKKGSYCNWESVLAHAA---------M--THYEEY--DRVKAHIVSITG--T-D----R--------------IRVA-P-DYTEIYH--MKG-----H--------MTPHGY-----HM-
Ga0494229_04493_40             : MGKRKLH-NETYYQCDWTGFPM-RSSN-CYW-PKWN-------------------------------D-GKLQKLGSYCNWESVLAHACDVYK-E-------QPEEL--AKIKEHIKERVG--G-V----H--------------IEEA-P-HYCEIEH--FAGGDN--S--------YTAAQF-----HN-
Ga0105092_10000132_27          : -MFSVVH-KKKLRLCQWCGHVLKEGQD-SFCIPKKS--------------------------K----N-GTISWGRCFCAVDCATAGLKEFSDNRK-----LSPEDH--DEIVRNFRDSLY--R-K----DAGKPSEKEPVT---ISAA-PNPVGIVRY--FGGPDT-----------MGPEEF-----RK-
Ga0498922_0001740_3            : -MQTQVH-NKSFVIDQWTGESI-DSNK-KFKIPQK----------------------------------GLKVWKGCYGTPSTGLAGLTYYITRNP--N--ISKDES--DEIFGVYQSSLV--K-C----D--PEEKYPHFT---INSA-P-PFTLITG--FGGK-------------TSLETY-----HN-
Flavob@PALC01000034_37         : MGKRKLH-DTTYFQCDWTGLPM-RTTN-CYM-PDWH--------------------------E----TTGKLLKHGSYCCWEAVVAHMIEQYKIGS--SSATDHPKW--KRIMDHITELVG--T------S--------------VSAA-P-HWKRIRWMQYGKESRDEL----ST-IDSANEF-----LA-
Verruc@NZEQ01000048_26         : MGKRKLN-GTSFYQCDWTGFPM-KSAH-CYW-PSWS--------------------------S----T-GKLFKKGSYCNWESVVAHATQQRAKEL-----ITQEEL--DKVMEHIEGITG--T-V-------------------VEPA-P-HYEEIAH--TKGR-------------MSAQDF-----HN-
Pates@JAHIZF010000067_15       : NMITVID-GKKYRTCQYTGEVF-QHAP-FKI-PHRE--------------------------P----LKG---WVGTYCCPSVCLAAMIDMLHTPGPKNEYLTEERH--SEMLDQLQHSLR--R-T----E---FKSNVRFD---VKAA-P-KKDLIKR--FAKEGNPA---------LTWDEY-----KA-
gamad_contig_1_4               : MGKRRMH-GQVFYQCDWTAYPM-KQAY-CYM-PSWN--------------------------G------DKLVKRGSYCNWPSVAAHAQHLFR-EG--D--VSQEQL--NDIMAHINSQAG--P------I--------------VDPS-RCHYSFIEH--FAPVQGHDN-----IHNYTMAEY-----MA-
gamad_contig_24_43             : MGKRKLH-GHTYYQCDWTGYPM-RQSN-CYM-PTWH--------------------------AAGGGG-AKLLKRGSYCNWESVLAHARHRHEVEK--A--ISEEEH--ARVREYVEQVCG--V-A----F--------------DDAA-QYHFSFIEH--FRDEAEEWRKVADLPRNWSMEEY-----HA-
gamad_contig_2_14              : ---------------------M-RNPN-CWM-PTWS--------------------------E----E-GKLIKKGSYCNWESVLAHAKHLSS--------ELSTDY--QRIEDYIRTQTG--D-C----S--------------VNEA-P-HFTDMEH--FKGNEHSKS--------LNAEEY-----HR-
Yara_contig_14_9               : -MWNVLH-GKRYRNCQWCANVREETDK-IYKIPEID--------------------------A----D-GKVEWKGVYCCPQCALAAMKVLSVKNK-----WAPEIH--TRHMEIFQASLR--R-V----DQGLPTESQKFN---LTAA-P-NPSCIKV--FKGT-------------LTQEEF-----RK-
Acanth_polyphaga_moumouvirus@Y : DGKKINL-KKTNIKCWWDRHEF-DNLP-FFL-PELY--------------------------H----N-NCYYVLGIFCSPNCALAHNLHYIKDFK------MYERK--SLIFRMYREMHG--LSADDLID--------------IKEA-G-PWELIES--FGGN-------------MSITTF-----RN-
Bodo_saltans_virus@YP_01077821 : DSNKPTP-NKSEFACWNCTYNF-DNAP-CFI-PEKY--------------------------Q----A-GVYYVFGNFCRDECSLSYILKDDEHKI------TMRR---SLIVKMHREKYGPSRT--------------------IRPA-F-PKEIIEK--FGGP-------------KTIQQY-----RE-
Namao_virus@AXN90943           : KTGEYYK-ESTDLLCWWCCHSF-KTVP-IGL-PDKL--------------------------I----D-QTFYMYGCFCSLNCVLSYNVDLGDYKV------YSRT---TLINVIKQEIYGSDNCY----S--------------IQMA-P-PRHAIKA--FGGY-------------LSIEDF-----RN-
Mimivirus_LCMiAC02@QBK89012    : NGEKIII-SNTDIACWWCTYKF-STHP-CFL-PDKF--------------------------M----N-NKYYVIGCFCSCECAQAYNINDLNDFK------SADRA--SLMKILYNIKYN--V------P--------------IKPA-P-AREVFEK--FGGV-------------LTIEKY-----RK-
Tetraselmis_virus_1@YP_0107833 : ------Q-MSTSVSCYWCTFPF-DNPP-IGL-PVKY------------------------FNT----E-DGFYVNGCFCSISCAAAYNLASRESND------TVCERH-SMICSLSRMLHG--T-D----K--------------IKVA-P-NWKSIQK--FGGY-------------MSIEQF-----RD-
Phaeocystis_globosa_virus@YP_0 : YNLKHNN-ITNKSACFWCTCPF-DNDP-IYI-PKHE--------------------------R----D-GVIHCYGCFCSPECACSYLMNDSIDNS------SKFERL-ALINNIYGKIYNYNKN--------------------VKPA-PSPYYLINK--FYGN-------------LDIQEY-----RK-
Ostreococcus_lucimarinus_virus : --------------CWWCCHSF-EGEP-LTM-PCRY------------------------DDR----R-SKFYTAGNYCSWSCIKSHAIDKFGDVK------GGIVC--GNIIMMRRKMYN--QIG----S--------------VKCA-P-NRFKIKE--FGGD-------------MTIEEF-----RE-
Yellowstone_lake_phycodnavirus : --------MDDGLACWWCVHPL-PCLPCIHL-PVKH------------------------DEK----T-NKFITKGNFCSWQCAKAYALDMNTSRS------GEIQ---MIIMMMRRRAFG--KYT----P--------------IWPA-P-KREAIKI--FGGT-------------MTIDEF-----RSF
Ectocarpus_siliculosus_virus_1 : ------P-LSYDKRCHNCAHFF-EGVP-VPL-PVSR------------------------DDL----R-HVYFCEGKFCSWQCAKSFNMRETSPAG------RGNRN--MYISVLAYKTWI--KLKKSVIDIETKQKMKTYCDYRIHPA-A-PRAKIQE--FGGE-------------MSIEEY-----RKD
Feldmannia_irregularis_virus_a : ------P-RRTESCCHNCCHPF-EGVP-VPL-PYFY------------------------DER----K-QVYFCSGSFCSWQCAKAYNIQLPVSSG------KGNRN--MYISLLAHKTWV--KIMSKTSHGVPHESLKTYAFTKIDPA-P-PRETIHM--FGGN-------------KTIQQY-----REG
Pandoravirus_dulcis@YP_0083190 : VYRNTLD-IRASGPCMHCDREC-TAGR-VPL-PRAY------------------------DSR----T-RTYAVWGNFCSFPCAMAYQLDHGAGVY------DTPHVL-VLIHKMAAEVWG--REG----L--------------VSPA-P-PRFCIRK--FGGA-------------VDPEDIEAVSERQY
Acanth_castellanii_medusavirus : --------SVGRHRCMWDLGEF-EGLP-IPL-PTRH------------------------LPR----T-DQYRVRGIYCSPACAKSAILANSDGMR------EMRL---TWFCDMMRRHYGIGYRK----R--------------IVPA-P-PREAIDI--FDGP-------------YTLEEF-----RSK
Pithovirus_LCPAC104@QBK90635   : FRDEKSK-ENSSCICHWCRKEY-INSK-IGV-PISM---------------------NKLDNR----K-YIFYVYGDFCSFECCYAYIKEHNSMTR------SQRKYLYDNSESLLKTLYN--LIYSDKEE--------------LKSA-P-DWRFIKE--YGGS-------------LDREEF-----HS-
Lymphocystis_disease_virus@YP_ : -----------MMRCWWCTLEI-NLKP-KEC-PIRK--------------------------E----G-SELITIGKFCSLSCVKSFAIKESYYDP------IYKDS--TRIIQLYNK------------N--------------VKPA-L-GFQIIRQ--YGGY-------------MNPEEF-----KS-
Brazilian_marseillevirus@YP_00 : VSGVILP-IHTDIPCFWCREQF-DNSP-IGL-PLRFVKHRETGILAERMNEHLRLFNLPTDHG----N-EHFETEGNFCSFGCCKAYMLDEMPKNP------ERYRKSPGLIVILHLKLEG--VLS----G--------------ILAA-P-SWKVIKK--WGGK-------------LSIEEF-----RS-
Tokyovirus_A1@YP_009255018     : ----FLP-LTTDEPCYWCREPF-HTMC-YAV-PLKYHKHRETGIHAERMKEHFRLYNLPSDHG----N-DHFETEGIFCGRGCVKAFIEDELPKKP------HFYRKSLGLITIMCMKIDG--SLS----T--------------IISS-P-HWRLIKK--WGGI-------------LTIEEF-----RA-
Monkeypox_virus@UXP53929       : ----IKT-NSVIDKCWFCNQDL-VFRP-ISI-ETFK--------------------------G----G-EV----GYFCSKICRDSLASMVKSHVA------LREE---PKISLLPLVFYE-------------------------------DKEKVIN--------------------TINLL-----RX-
Canarypox_virus@NP_955092      : ----LKI-KKDNNNCWFCKQQL-NSFTHYFI-KTLY--------------------------G----D-DI----GFFCSKICRDSFANMIKGVIA------LREE---PKVSLLPLEVYE---------N---------------------PEEVIEI---------------------INEL-----RE-
Orf_virus@AAR98171             : ----LAL-RAADGTCWFCRQRV-GTAA-LAV-ETLD--------------------------G----G-AV----GEFCSRLCRDSFAGSIRHLVA------LREE---PKIELLPLEMYA-------------------------------DPDRVRA--L------------------INET-----RA-
Salmon_gill_poxvirus@YP_009162 : --------NNDSKTCLTCGNDI-GPLV-YKI-ETLF--------------------------E----G-KI----GAFCSPACSDIMAGCVSHAIS------LKSP---SRMKLPPLCMYK--NST----R--------------IVET-V-TKEMEKC--YGTQ-------------QRV-----------
                                               c                 p                                          g 5c                                                                     a
```

S5

6\. **P20982 VLTF2_VACCC Viral late gene transcription factor 2 OS=Vaccinia virus (strain Copenhagen) OX=10249 GN=VLTF2 PE=2 SV=1**

Probability: 94.57%, E-value: 0.1, Score: 47.8, Aligned cols: 44, Identities: 11%, Similarity: 0.121, Template Neff: 3.5

```
Q ss_pred               CCCCccCCCCCCCCCCCCcceeEEEeCCEEEEEeeEcCHHHHHHHHHH
Q tmp_VLTF2       8     GKSYYQCDWTGFPMESANCYMPVWSANGKLVWKGSYCNWESVVAHAKE   55 (318)
Q Consensus       8     ~kt~~~CWWC~h~f~~~P~~IP~~~~~g~y~~~G~FCS~eCa~Ayi~~   55 (318)
                       ......||||..+....++.|=+-. +|   ..|.|||--|..++...
T Consensus      46     ~~~~~~CWfC~q~l~~~~~~iETl~-g~---~Vg~FCSkiCrDSfa~~   89 (150)
T P20982        46     NSVIDKCWFCNQDLVFRPISIETYK-GG---EVGYFCSKICRDSLASM   89 (150)
T ss_pred               CCCCCceecCCCcccCCCeEEEecC-CC---chheecCHHHHHHHHHH
```

Template alignment

7\. **O72910 VLTF2_FOWPN Viral late gene transcription factor 2 OS=Fowlpox virus (strain NVSL) OX=928301 GN=VLTF2 PE=2 SV=1**

Probability: 90.65%, E-value: 0.75, Score: 42.63, Aligned cols: 42, Identities: 14%, Similarity: 0.174, Template Neff: 3.4

```
Q ss_pred               CCCCccCCCCCCCC-CCCCcceeEEEeCCEEEEEeeEcCHHHHHHHH
Q tmp_VLTF2       8     GKSYYQCDWTGFPM-ESANCYMPVWSANGKLVWKGSYCNWESVVAHA   53 (318)
Q Consensus       8     ~kt~~~CWWC~h~f-~~~P~~IP~~~~~g~y~~~G~FCS~eCa~Ayi   53 (318)
                       ......||+|.... ...++.|=+-. +|   ..|.|||-=|..+++
T Consensus      46     ~~~~~~CWfC~q~i~~~~~~~iETl~-G~---~IG~FCS~ICrDSfA   88 (154)
T O72910        46     RKDKDNCWFCKQDMNTYNPYFIETLY-GD---HIGVFCSKICRDSFA   88 (154)
T ss_pred               cCCCCceeccCCcccCCCCeEEEEec-cc---cceeeehHHhHHHHH
```

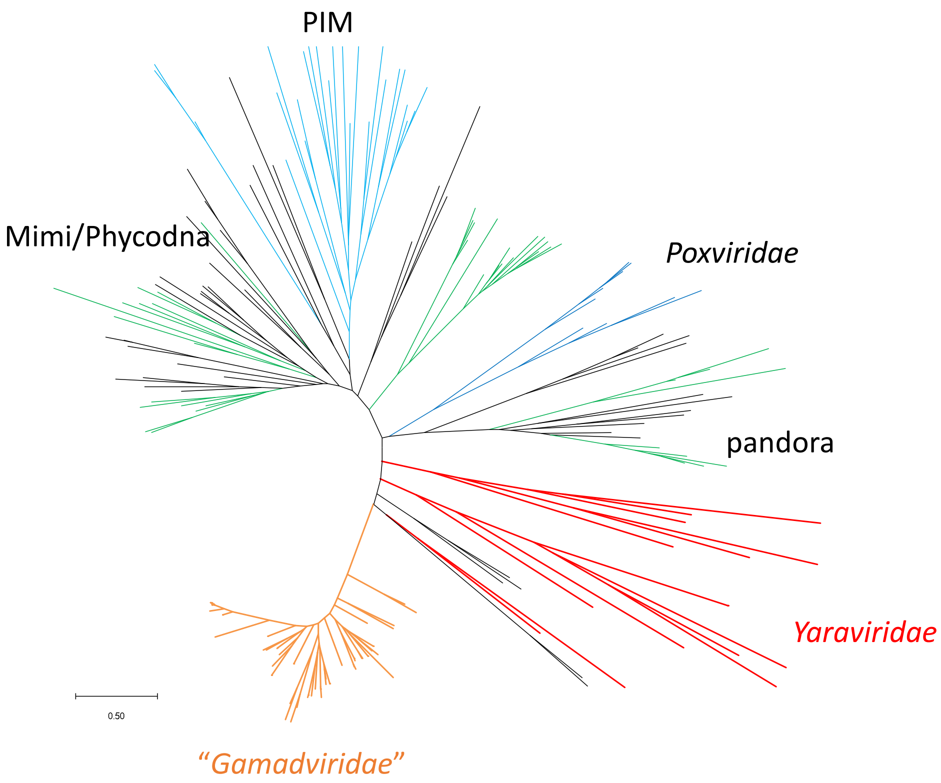
S6

Nr	Hit	Name	Probability	E-value	Score	SS	Aligned cols	Target Length
☐ 1	6P7A_A	Holliday junction resolvase; RESOLVASE, HYDROLASE; 3.081A {Fowlpox virus}	99.78	4.2e-17	115.02	14.9	138	148
☐ 2	Q9J546	RUVV_FOWPN Holliday junction resolvase OS=Fowlpox virus (strain NVSL) OX=928301 GN=FPV187 PE=1 SV=1	99.78	2.2e-17	117.98	13.3	145	156
☐ 3	4KTW_B	RuvC endonuclease; endonuclease, DNA junctions, replication, recombination, phage packaging, Holliday junction, RNase-H	99.77	5.9e-17	116.12	15.1	141	161
☐ 4	4EP4_A	Crossover junction endodeoxyribonuclease RuvC; resolvase, Hydrolase; 1.28A {Thermus thermophilus} SCOP: c.55.3.6	99.75	1.4e-16	114.2	14.3	148	166
☐ 5	7XHJ_A	Crossover junction endodeoxyribonuclease RuvC; nuclease, DNA repair, Homologous recombination, Holliday junction resolva	99.74	3e-16	114.34	13.9	140	179
☐ 6	cd16962	RuvC; Crossover junction endodeoxyribonuclease RuvC. Crossover junction endodeoxyribonuclease RuvC is also called Hollid	99.73	4e-16	110.27	14.1	138	156
☐ 7	1HJR_C	HOLLIDAY JUNCTION RESOLVASE (RUVC); SITE-SPECIFIC RECOMBINASE; 2.5A {Escherichia coli} SCOP: c.55.3.6	99.73	2.6e-16	111.81	12.8	141	158
☐ 8	5E6F_B	CNPV261 Holliday junction resolvase protein; Poxvirus, Resolvase, Holliday Junction, Metal-binding, VIRAL PROTEIN; HET:	99.73	1.7e-15	107.76	16.7	146	152

Nr	Hit	Name	Probability	E-value	Score	SS	Aligned cols	Target Length
☐ 1	6LW3_B	Crossover junction endodeoxyribonuclease RuvC; Holliday junction resolvase RuvC Endonuclease, DNA BINDING PROTEIN; 2.38A	99.68	1.4e-14	101.5	16	147	163
☐ 2	1HJR_C	HOLLIDAY JUNCTION RESOLVASE (RUVC); SITE-SPECIFIC RECOMBINASE; 2.5A {Escherichia coli} SCOP: c.55.3.6	99.68	1.4e-14	101.01	15.7	145	158
☐ 3	7XHJ_A	Crossover junction endodeoxyribonuclease RuvC; nuclease, DNA repair, Homologous recombination, Holliday junction resolva	99.67	1.9e-14	102.89	16	151	179
☐ 4	4EP4_A	Crossover junction endodeoxyribonuclease RuvC; resolvase, Hydrolase; 1.28A {Thermus thermophilus} SCOP: c.55.3.6	99.65	7.1e-14	98.46	16.8	152	166
☐ 5	cd16962	RuvC; Crossover junction endodeoxyribonuclease RuvC. Crossover junction endodeoxyribonuclease RuvC is also called Hollid	99.65	5e-14	97.77	15.4	142	156
☐ 6	6IS8_A	Monokaryotic chloroplast 1; Holliday junction resolvase-DNA complex, DNA BINDING PROTEIN-DNA complex; HET: 1PE; 1.68A {Z	99.61	1.1e-13	98.5	14.5	151	174
☐ 7	PF02075.21	; RuvC ; Crossover junction endodeoxyribonuclease RuvC	99.6	1.8e-13	94.57	14.4	137	148
☐ 8	6KVN_A	NtMOC1; chloroplast, resolvase, PLANT PROTEIN; 2.02A {Nicotiana tabacum}	99.6	2.7e-13	96.23	15.3	150	171

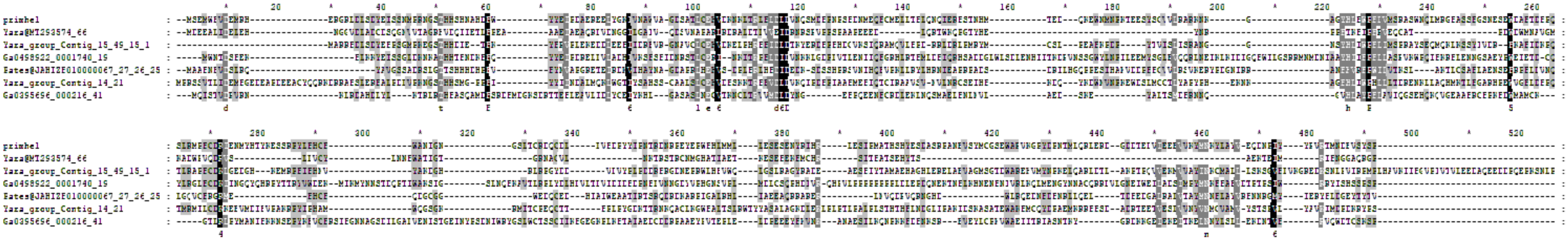
top HHpred hits for “Gamadviridae” (right) and Yaraviridae (left) RuvC alignments

S7


PIM
Mimi/Phycodna
Poxviridae
pandora
Yaraviridae
0.50
"Gamadviridae"

S8

SF3_hel1 (most of *Yaraviridae*; Yaravirus gene 69)



primase domain

```
Q ss_pred               CCCCCCccceeeEEEC---ChHHHHHHHHHHHHHHHhcCceeeccccccccceeEEeeCCCCcccccCCCccHHhHHHHH
Q hel_prim      33      STHHSHNAHDFWYYEI---PDAEREELYGKIVNAWVAGDSATLCEHVDKNKLTRLFLDLDVNQSMDFPNPSFDNMEQFCM   109 (401)
Q Consensus     33      STHH~hn~~~F~Y~~i---P~~~~eeL~~~Iv~~~V~G~s~~LCEHVDK~~~~R~FlDLD~~~~~DFP~~~~~~~~Q~~M   109 (401)
                        -||-+.+   ..-|.|   |+++.+++++.+.++...|...+|||-..+. ..+||+|+|.+...+....      +...
T Consensus     24      ~TH~~~~---~G~~~i~~~p~~~~~~F~~y~~~i~~~~~l~l~E~~~~~~-~~pl~~DiD~~~~~~~~~~~------~~~~   93 (162)
T PF18188.4     24      PTHVSQR---GGHYHIIFRSQQMYNEYLSAWKKTILDKSKMYIEEMRTKH-AFRLYVDIDLKVEIEKPCD------IVEL   93 (162)
T ss_pred               CCeEeeC---CceEEEecCCHHHHHHHHHHHHHHhhcCCCeeEEEecCCC-ceeEEEEEEEEcCCCCccc------ccch

Q ss_pred               HHHHHHHHHHHhhhhhhhccCHHHhCCCCCCCCCeEEEEEEEecc----------CCCceeceeeEeeecHHHHHH
Q hel_prim      110     ELLTFLQNQIERFSTNHMTEDQKEWNMNPKTEESYSCVVLRARKN----------KGAGYHLFWPFLVMSRASWNQ   175 (401)
Q Consensus     110     e~LtF~~~~ieR~~~~~M~~~~~~~~~~npktee~Ys~v~L~aR~N----------~G~~YHL~WPFLvMS~~s~~q   175 (401)
                        +++...++.|.++-.                +....|+++.+.+.          --.|+|+.||.++.+......
T Consensus     94      ~~i~~~~~~i~~~~~----------------~~~~~~~V~~~~~~~~~~~~~~~~~~~K~G~Hi~fP~i~v~~~~~~~   153 (162)
T PF18188.4     94      GWVDSIMQFTKQYFP----------------KKDISMVLTQCHGKWEDAVTTTAKYKSGYRLYFHEIIVDYDTFCD   153 (162)
T ss_pred               HHHHHHHHHHHHHCC----------------CCCeEEEEEecccccccCCCCcceeeceEEEEecceEECHHHHHH
```

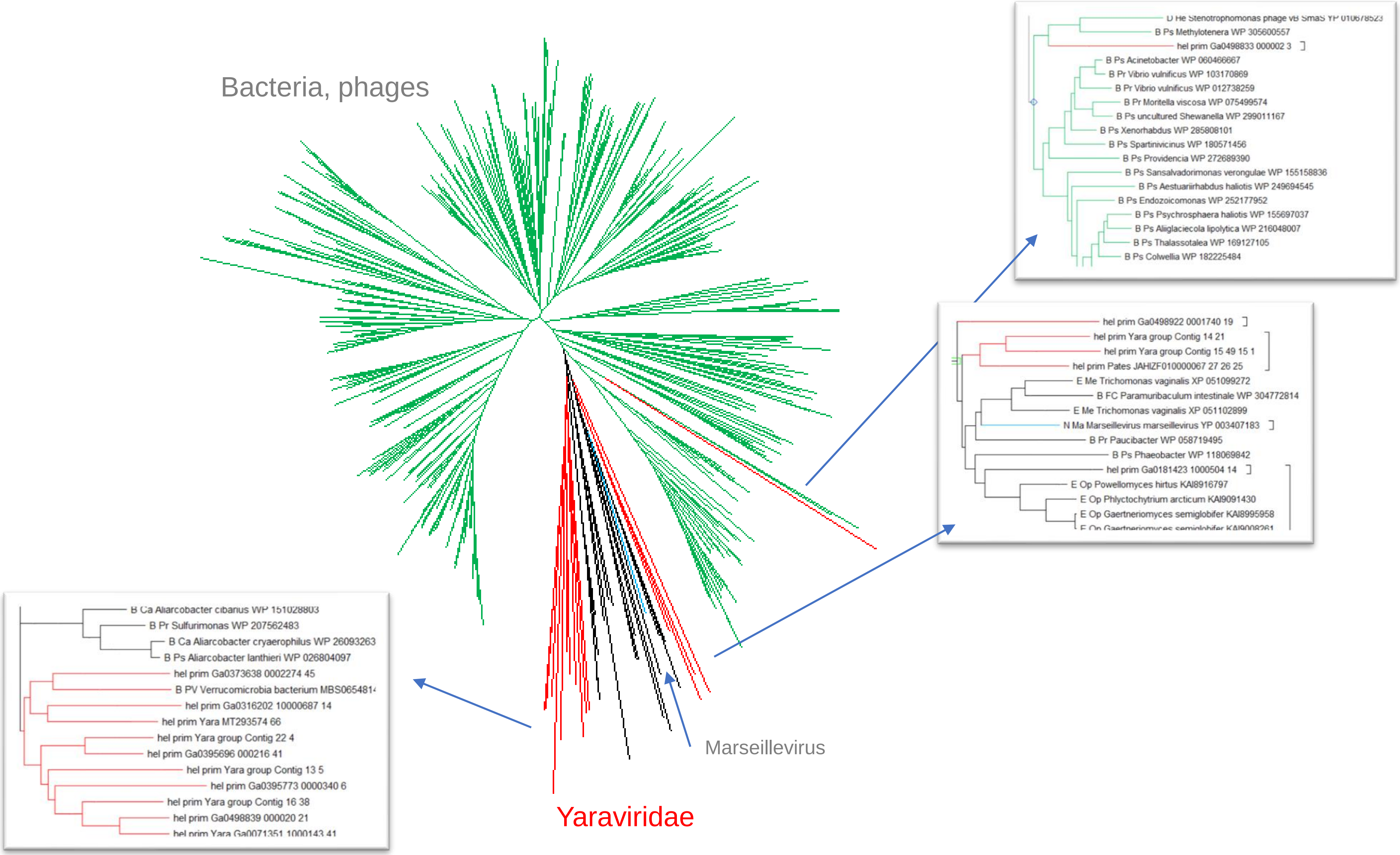
helicase domain

```
Q ss_pred               CCHHHHHHHHHHHHHHhhCchhcCCEEEEEeCCCCCCHhHHHHHHHHHHChhcEEECCHHHhcccchHHhCCCEEEEecC
Q hel_prim      486     GNQEKYNYLMNWLAYMVQHPWQKMETCMILTGPEGAGKSIFMHAFGKIFGESYCEVSDIDDILGHFNDMLEGKLIIYLDE   565 (926)
Q Consensus     486     ~d~e~~~~l~~~la~~l~~p~~~~~~~~vl~G~~g~GKSt~~~ll~~l~G~~~~~~~~~~~~l~~~f~~~l~gk~lvi~~E   565 (926)
                        ++++.+++++.|+|+++++| .+.+.+++|+|++|+||||+++++..++|..++...+...+.+.|++.+.|+++++++|
T Consensus     115     ~~~~~~~~l~~~~~~~~~~~~-~~~~~~~~l~G~~g~GKst~~~~l~~l~g~~~~~~~~~~~~~~~~~~~~~~~~~~~~dE   193 (416)
T 6K9C_A        115     YDAKSMGLFINWLAFIYQYK-ERTGVAWIFMGKQGTGKGLLVDLLKKIFEEHMSSNITDANLDSQFNPYLYNKLIVHLNE   193 (416)
T ss_dssp               TCHHHHHHHHHHHHHHHHHC-CCCSCEEEEEESTTSSHHHHHHHHHHHHGGGEEEEECHHHHTSSCCGGGTTCSEEEEEC
T ss_pred               CCHHHHHHHHHHHHHHHHCc-ccccEEEEEECCCCCCHHHHHHHHHHHHCcccceeCChHHhcccccHHHcCCeEEEEec

Q ss_pred               CCCCC----cchHhHHHHHhhCCeEEEeeCCCCceeeecceEEEEeeCCccccCCCCCCcceEEEEEcCccccCChHHHH
Q hel_prim      566     VDHPD----SDQNSALKNFVTGDSMVIEKKYQDPFYVENYAMIIMSNNDNRVINAVGPDARRWVLFDCENNARSDNDQYF   641 (926)
Q Consensus     566     ~~~~~----~~~~~~lK~l~g~d~i~~r~K~~~~~~~~~~~~~~~~i~~nN~p~~~~~~~~~~~~RR~~vi~f~~~~~~~~~~~~   641 (926)
                        .....    ....+.||+++|++.+.++.||++.+++.+.+.+.+++++|.+.....+.+. ||+.+|++...+....  .+
T Consensus     194     ~~~~~~~~~~~~~~~lk~~~~~~~~~~~~~~k~~~~~~~~~~~~~~~~~i~~tN~~~~~~~~~~~~~~-rR~~~i~~~~~~~~~~~~~--~~   270 (416)
T 6K9C_A        194     VSADNRKSRMLVKNRLKTWITDETLYINRKNMKEVEIKNFCNFIINSNETIPVDIEDSD-RRFNVIECNNVLKEQE--WW   270 (416)
T ss_dssp               SSCSSHHHHHHTHHHHHHHHHCSEEEECCTTSCCEEEECCCEEEEEECSSCCSCCCTTC-CSEEEEEECCCGGGST--TC
T ss_pred               cCcCCchhhHHHHHHHHHHhcCCeEEEeecCCCceeEEeeeEEEEEcCCCCCcCCCCCC-ccEEEEecCccccccc--cc
```

SF3_hel1 (Yaravirus gene 69)

S9



S10

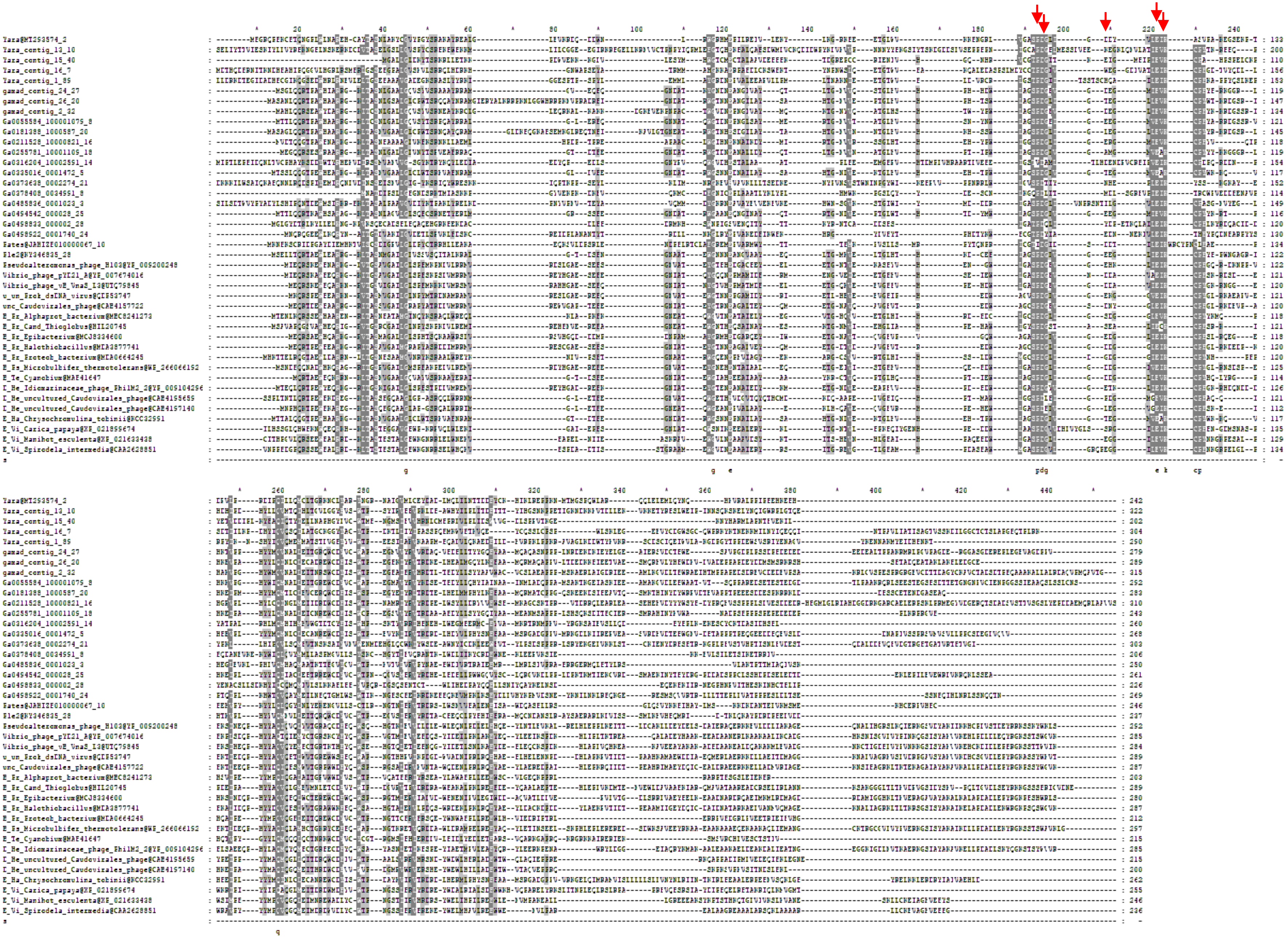
SF3_hel2

some SF3_hel2 proteins have frameshifts (marked in orange on the figure).

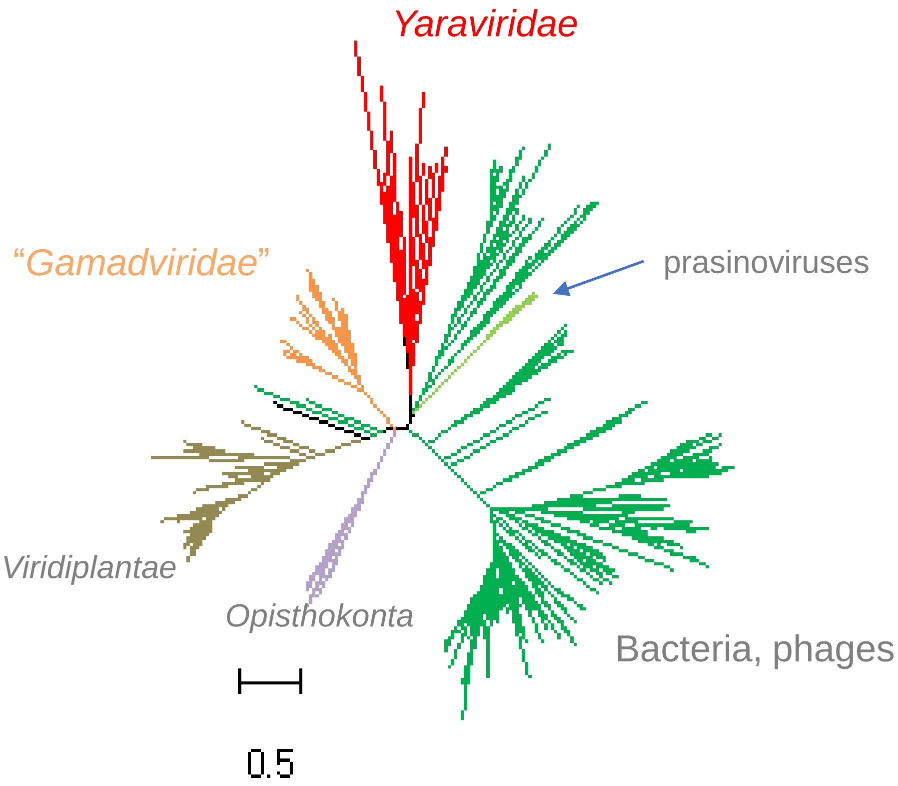
S11

Nr	Hit	Name	Probability	E-value	Score	SS	Aligned cols	Target Length
☐ 1	2X3G_A	SIRV1 HYPOTHETICAL PROTEIN ORF119; VIRAL PROTEIN, HOST MEMBRANE; HET: MSE; 1.8A {SULFOLOBUS ISLANDICUS RUDIVIRUS 1 VARIA	95.52	0.097	42.87	6	65	119
☐ 2	PF07232.14	; DUF1424 ; Putative rep protein (DUF1424)	93.77	6.8	36.39	13.8	115	316
☐ 3	P11334	REP_SPV4 Replication-associated protein ORF2 OS=Spiroplasma virus 4 OX=10855 GN=ORF2 PE=3 SV=1	92.74	3.1	39.96	10.4	107	320
☐ 4	P03627	REP_BPPF3 Replication-associated protein G2P OS=Pseudomonas phage Pf3 OX=10872 GN=II PE=3 SV=1	92.66	8.1	36.1	12.6	109	278
☐ 5	Q9G050	REP_BPPHM Replication-associated protein VP4 OS=Bdellovibrio phage phiMH2K OX=145579 GN=ORF4 PE=3 SV=1	91.61	1.4	42.09	6.8	67	315
☐ 6	PF05840.16	; Phage_GPA ; Bacteriophage replication gene A protein (GPA)	89.69	12	35.26	10.7	119	316
☐ 7	Q8QL45	Y117_SIRV1 Uncharacterized protein 117 OS=Sulfolobus islandicus rod-shaped virus 1 OX=157898 GN=117 PE=4 SV=1	88.17	2.8	35.65	5.1	63	117

S12

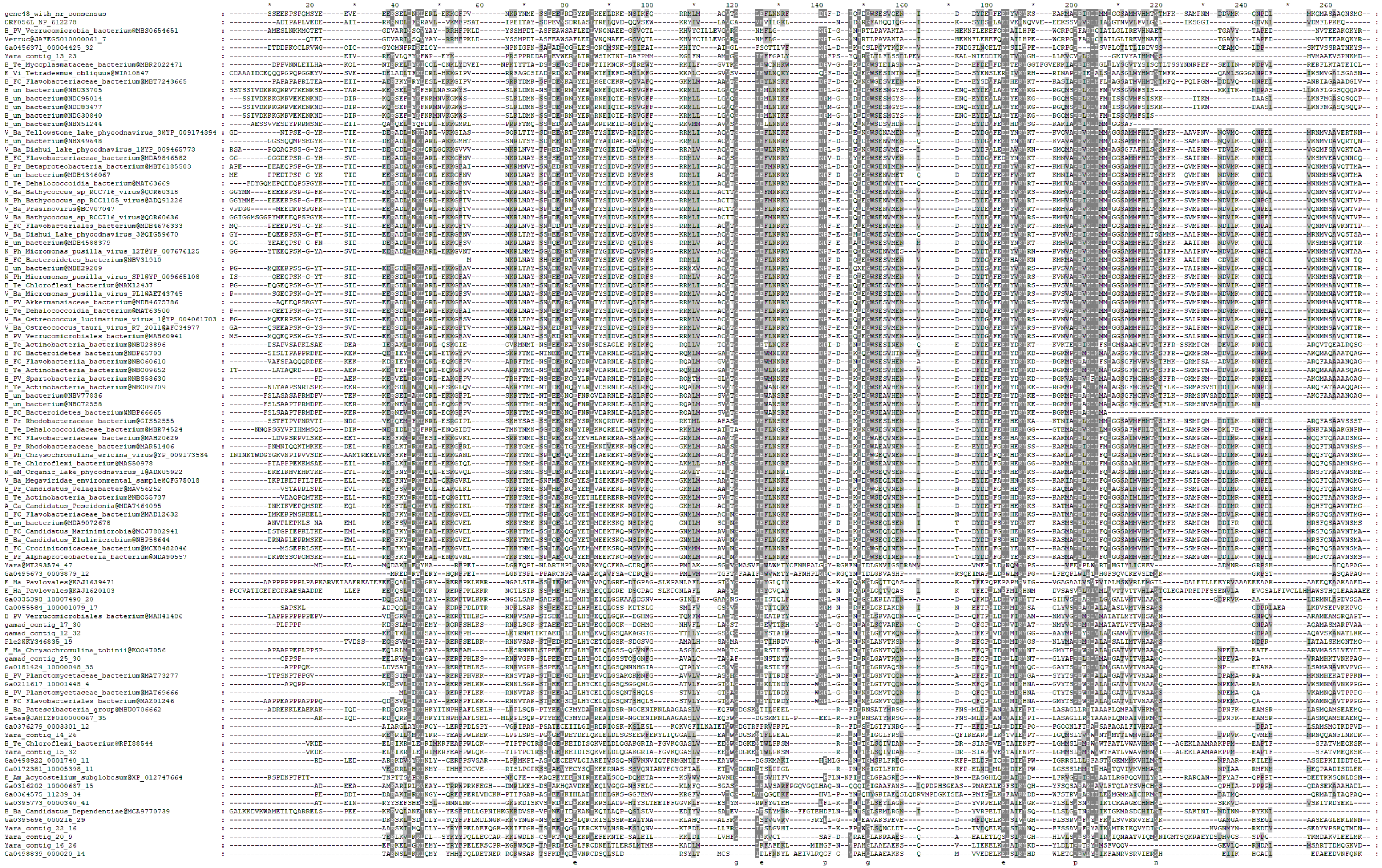

Yara@MT293574_2
Yara_contig_13_10
Yara_contig_15_40
Yara_contig_16_7
Yara_contig_1_89
gamad_contig_24_27
gamad_contig_26_20
gamad_contig_2_32
Ga0055584_100001079_8
Ga0181388_1000587_20
Ga0211528_10000821_16
Ga0255781_10001109_18
Ga0316204_10002591_14
Ga0335016_0001472_5
Ga0373638_0002274_21
Ga0378408_0034991_8
Ga0485836_0001023_3
Ga0494542_000028_25
Ga0498833_000002_28
Ga0498922_0001740_24
Pates@JAHIZE010000067_10
Ple2@NY346835_28
Pseudoalteromonas_phage_H103@YP_009200248
Vibrio_phage_pYD21_A@YP_007674016
Vibrio_phage_vB_VnaS_L3@UTQ79845
u_un_Prok_dsDNA_virus@QIE53747
unc_Caudovirales_phage@CAB4157722
B_Pr_Alphaprot_bacterium@MBC8241273
B_Pr_Cand_Thioglobus@HIL20745
B_Pr_Epibacterium@MCJ8334600
B_Pr_Halothiobacillus@MDA3877741
B_Pr_Proteob_bacterium@MDA0664245
B_Ps_Microbulbifer_thermotolerans@WP_266066192
B_Te_Cyanobium@MAP41647
I_He_Idiomarinaceae_phage_Phi1M2_2@YP_009104296
I_He_uncultured_Caudovirales_phage@CAB4195659
I_He_uncultured_Caudovirales_phage@CAB4197140
E_Ha_Chrysochromulina_tobinii@KOO32991
E_Vi_Carica_papaya@XP_021899674
E_Vi_Manihot_esculenta@XP_021633438
E_Vi_Spirodela_intermedia@CAA2628851
s
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
320
340
360
380
400
420
440
g
g e
pdg
e k
cp
q

S13

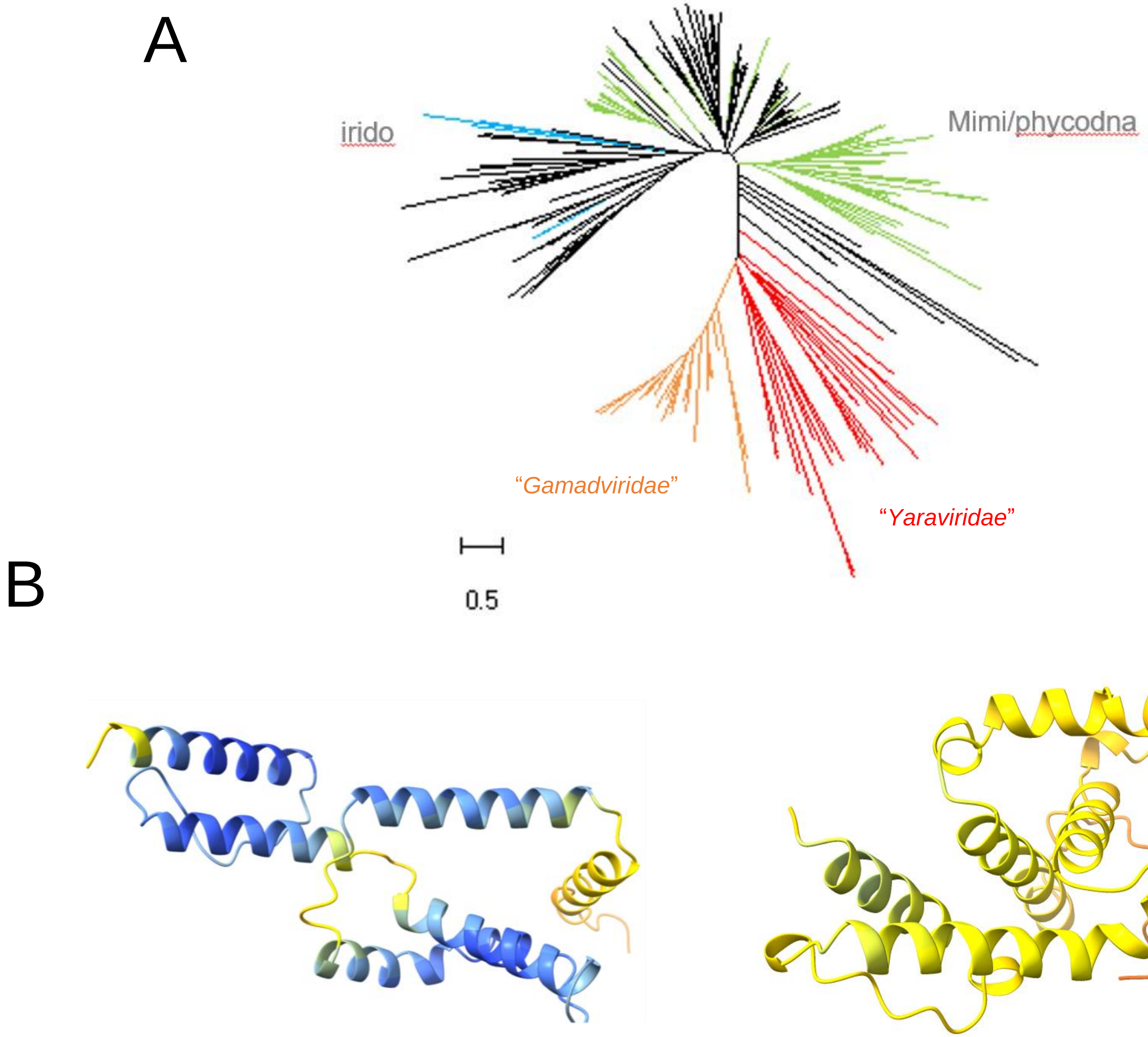

Yaraviridae
"Gamadviridae"
prasinoviruses
Viridiplantae
Opisthokonta
Bacteria, phages
0.5

S14

gene48 with nr hits



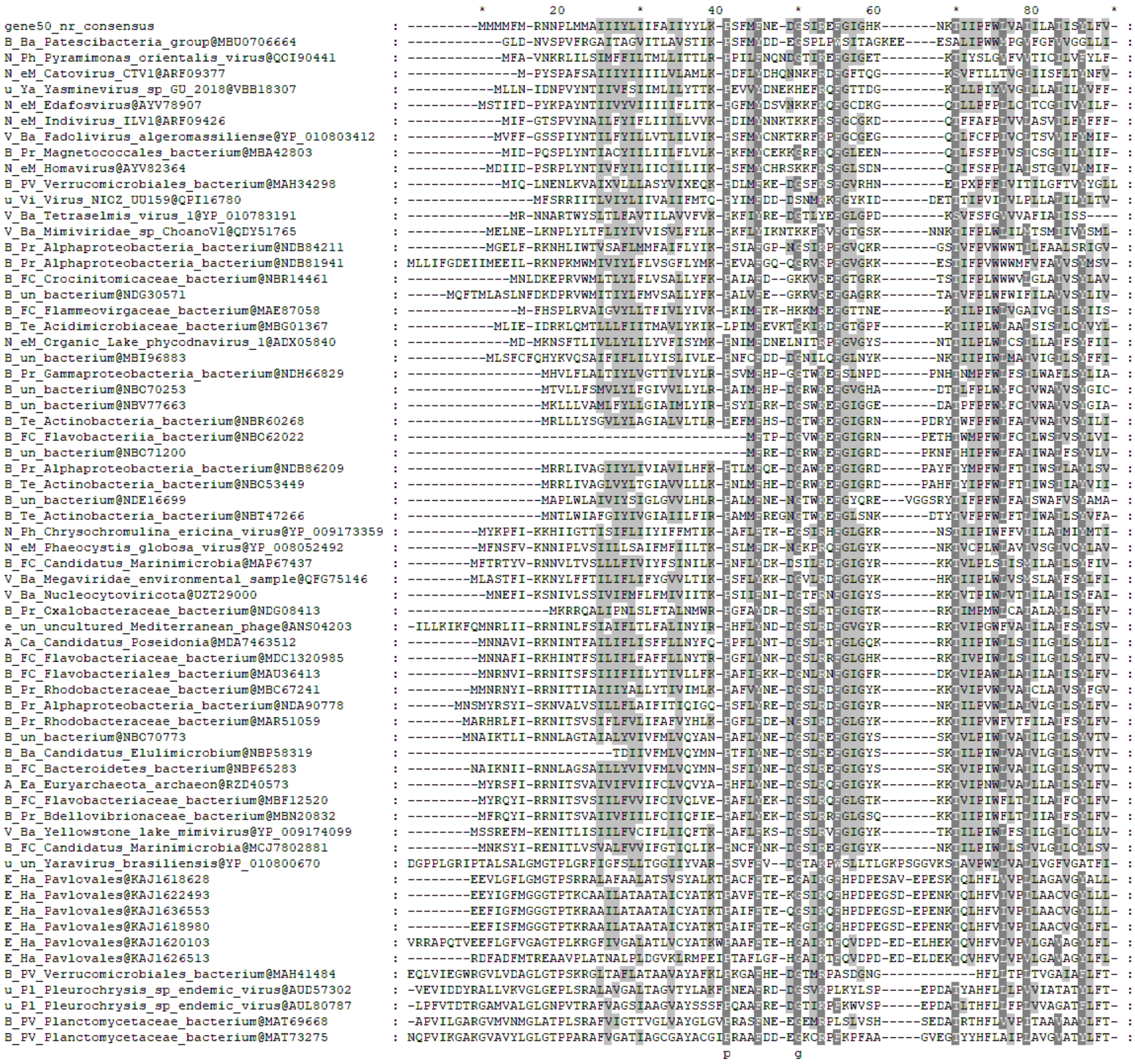
S15



gene48 (Ple2@KY346835_19)

(Yaravirus)

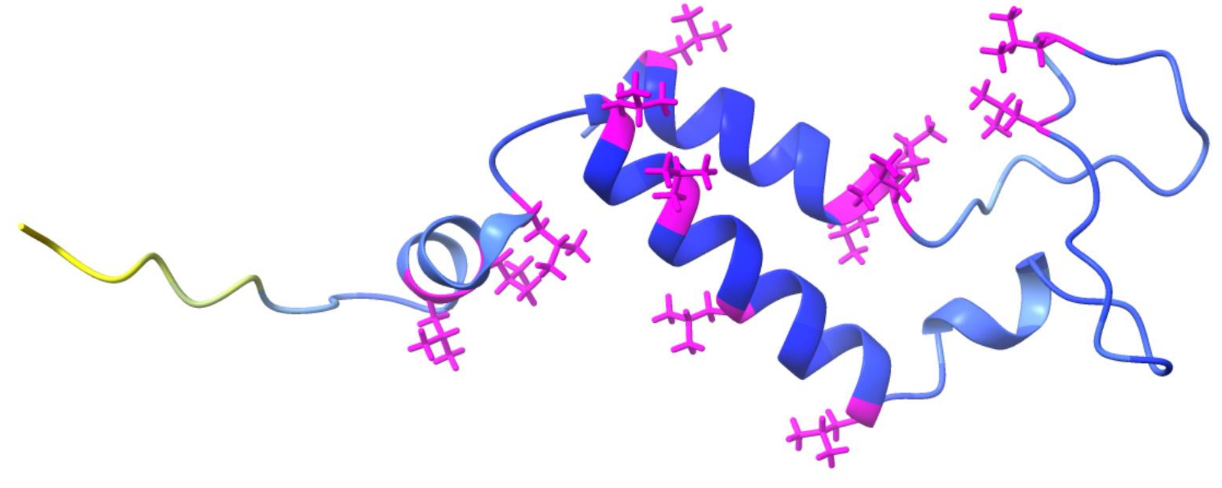
S16



```
                                                            *        20         *        40         *        60         *        80         *
gene50_nr_consensus                              : ---------MMMMFM-RNNPLMMAIIIYLIIFAIIYYLK-PSFMFNE-DGSIREFGIGHK-------NKTIIPFWLVAIILAIISYLFV- :
B_Ba_Patescibacteria_group@MBU0706664            : ------------GLD-NVSPVFRGAITAGVITLAVSTIK-PSFMYDD-EGSPLPWSITAGKEE----ESALIPWWMPGVFGFVVGGLLI- :
N_Ph_Pyramimonas_orientalis_virus@QCI90441       : ------------MFA-VNKRLILSIMFFILTMLLITTLR-PPILFNQNDGTIKEFGIGET--------KTIYSLGVFVVTICILVFYLF- :
N_eM_Catovirus_CTV1@ARF09377                     : --------------M-PYSPAFSAIIIYIIIILVLAMLK-PDFLYDHQNNKFRDFGFTQG--------KSVFTLLTVGIIISFLTYNFV- :
u_Ya_Yasminevirus_sp_GU_2018@VBB18307            : -----------MLLN-IDNPVYNTIIVFSIIMLILYTTK-PEVVYDNEKHEFRQFGTTDG--------KTLLPIYVVGILLAIILYVFF- :
N_eM_Edafosvirus@AYV78907                        : ---------MSTIFD-PYKPAYNTIIVYVIIIIIFLITK-PGFMYDSVNKKFKQFGCDKG--------QTLLPFPLLCITCGIIVYILF- :
N_eM_Indivirus_ILV1@ARF09426                     : ------------MIF-GTSPVYNAILFYIFLIIILLVVK-PDIMYNNKTKKFKSFGCGKD--------QTFFAFPLVVLASVILFYFFF- :
V_Ba_Fadolivirus_algeromassiliense@YP_010803412  : -----------MVFF-GSSPIYNTILFYILLVTLILVIK-PSFMYCNKTKRFKPFGCGEG--------QTLFCFPIVCLTSVVIFYMIF- :
B_Pr_Magnetococcales_bacterium@MBA42803          : ------------MID-PQSPLYNTIACYIILIILFLVLK-PKFMYCEKKGRFKQFGLEEN--------QTLFSFPIVSICSGIILYIIF- :
N_eM_Homavirus@AYV82364                          : ----------MDIID-PSRPLYNTIVFYILIICIILIIK-PSFMYCHRSKKFKSFGLSDN--------QTIFSFPLIAISTGIVLYMIF- :
B_PV_Verrucomicrobiales_bacterium@MAH34298       : ------------MIQ-LNENLKVAIXVLLLASYVIXEQK-PDLMFKE-DGSFKSFGVRHN--------ETPXPFFIVITILGFTVYYGLL :
u_Vi_Virus_NICZ_UU159@QPI16780                   : ----------------MFSRRIITLVIYLIIVAIIFMTQ-PYIMFDD-DSNMKKFGYKID------DETTTIPVILVLPLLALILYLTV- :
V_Ba_Tetraselmis_virus_1@YP_010783191            : -------------MR-NNARTWYSLTLFAVTILAVVFVK-PKFIYRE-DGTLYEFGLGPD--------KSVFSFGVVVAFIAIISS---- :
V_Ba_Mimiviridae_sp_ChoanoV1@QDY51765            : ----------MELNE-LKNPLYLTFLIYIVVISVLFYLK-PKFLYIKNTKKFKVFGTGSK------NNKTIFPLWLILMTSMIIVYSML- :
B_Pr_Alphaproteobacteria_bacterium@NDB84211      : ----------MGELF-RKNHLIWTVSAFLMMFAIFLYIK-PSIAFGP-NGSIKPFGVQKR-------GSTVFPVWWWTILFAALSRIGV- :
B_Pr_Alphaproteobacteria_bacterium@NDB81941      : MLLIFGDEIIMEEIL-RKNPKMWMIVIYLFLVSGFLYMK-PEVAFGQ-QGRVRPFGVGKK-------ESTIFPVWWWMFVFAVVSYMSV- :
B_FC_Crocinitomicaceae_bacterium@NBR14461        : -------------MNLDKEPRVWMLTLYLFLVSALLYFK-PAIAFD--GKKVREFGTGRK-------TSTIFPLWWWVLGLAIVSYLAV- :
B_un_bacterium@NDG30571                          : -----MQFTMLASLNFDKDPRVWMITIYLFMVSALLYFK-PALVFE--GKRVREFGAGRK-------TATVFPLWFWIFILAVVSYLIV- :
B_FC_Flammeovirgaceae_bacterium@MAE87058         : --------------M-FHSPLRVAIGVYLLTFIVLYIVK-PKIMFTK-HKKMREFGTTNE--------KTILPIWLVGAIVGILSYIIS- :
B_Te_Acidimicrobiaceae_bacterium@MBG01367        : -----------MLIE-IDRKLQMTLLLFIITMAVLYKIK-LPIMFEVKTGKIKDFGTGPF--------KTIIPLWLAALSISLLCYVYL- :
N_eM_Organic_Lake_phycodnavirus_1@ADX05840       : -------------MD-MKNSFTLIVLLYLILYVFISYMK-PNIMFDNELNITRPFGVGYS-------NTTILPLWLCSILLAIFSYFII- :
B_un_bacterium@MBI96883                          : ----------MLSFCFQHYKVQSAIFIFLILYISLIVLE-PNFCFDD-DGNILQFGLNYK-------NKTIIPIWIMAIVIGILSYFFI- :
B_Pr_Gammaproteobacteria_bacterium@NDH66829      : -----------------MHVLFLALTIYLVGTTIVLYLR-PSVMFHP-GGTWKEFSLNPD------PNHTNMPFWLFSILWAFLSYLIA- :
B_un_bacterium@NBC70253                          : -----------------MTVLLFSMVLYLFGIVVLLYLR-PAIMFHP-DGRWKEFGVGHA-------DTTLFPLWMFCIVWAVVSYGIC- :
B_un_bacterium@NBV77663                          : -----------------MKLLLVAMLFYLLGIAIMLYIR-PSYIFRK-DGSWKEFGIGGE-------DATPFPFWMFCIVWAVVSYGIA- :
B_Te_Actinobacteria_bacterium@NBR60268           : -----------------MRLLLYSGVLYLAGIALVLTLR-PEFMFHS-DGTWKEFGIGRN-----PDRYTWFPFWIFALVWAIVSYILI- :
B_FC_Flavobacteriia_bacterium@NBC62022           : -------------------------------------------MFTP-DGVWKEFGIGRN-----PETHTWMPFWLFCILWSIVSYLVI- :
B_un_bacterium@NBC71200                          : -------------------------------------------MFRE-DGRWKEFGIGRD-----PKNFTHIPFWLFAIIWALFSYVLV- :
B_Pr_Alphaproteobacteria_bacterium@NDB86209      : -----------------MRRLIVAGIIYLIVIAVILHFK-PTLMFQE-DGAWKEFGIGRD-----PAYFTYMPFWLFTIIWSLLAYLSV- :
B_Te_Actinobacteria_bacterium@NBC53449           : -----------------MRRLIVAGLVYLTGIAVVLLLK-PNLMFHE-DGRWKEFGIGRD-----PAHFTYIPFWLFTIIWSIIAYVII- :
B_un_bacterium@NDE16699                          : -----------------MAPLWLAIVIYSIGLGVVLHLR-PALMFNE-NGTWKEFGYQRE---VGGSRYTIFPFWLFAISWAFVSYAMA- :
B_Te_Actinobacteria_bacterium@NBT47266           : -----------------MNTLWIAFGIYIVGIAIILFIR-PAMMFREGNGTWKEFGLSNK------DTYTVFPFWLFTLIWAILSYVFA- :
N_Ph_Chrysochromulina_ericina_virus@YP_009173359 : ---------MYKPFI-KKHIIGTTISIFLIIYIFFMTIK-PAFLFTK-EGSIRHFGLGKR-------NSTIIPIWFFVIILAIMIYMTI- :
N_eM_Phaeocystis_globosa_virus@YP_008052492      : ---------MFNSFV-KNNIPLVSIILLSAIFMFIILTK-PSLMFDK-NGKPRQFGLGYK-------NKTVCPLWLAVIVSGIVCYLAV- :
B_FC_Candidatus_Marinimicrobia@MAP67437          : --------MFTRTYV-RNNVLTVSLLLFIVIYFSINILK-PNFLYDK-DSILRDFGIGYR-------KKTVLPLSIISMILAILSYFIV- :
V_Ba_Megaviridae_environmental_sample@QFG75146   : --------MLASTFI-KKNYLFFTILIFLIFYGVVLTIK-PSFLYKK-DGVLRDFGLGYK-------HKTIIPLWLVSMSLAVFSYLFI- :
V_Ba_Nucleocytoviricota@UZT29000                 : ----------MNEFI-KSNIVLSSIVIFMFLFMIVIITK-PSIIFNI-DGTFRNFGIGYS-------KKTVTPIWLVTIILAIISYFAI- :
B_Pr_Oxalobacteraceae_bacterium@NDG08413         : ------------------MKRRQALIPNLSLFTALNMWR-PGFAYDR-DGSLRTFGIGTK-------RKTIMPMWLCAIALAMLSYLFV- :
e_un_uncultured_Mediterranean_phage@ANS04203     : -ILLKIKFQMNRLII-RRNINLFSIAIFLTLFALINYIR-PHFLYND-DGSLRDFGVGYR-------RKTVIPGWFVAIILAIFSYLSV- :
A_Ca_Candidatus_Poseidonia@MDA7463512            : ---------MNNAVI-RKNINTFAILIFLISFFLLNYFQ-PPFLYNT-DGSLRTFGLGQK-------RKTIIPIWLLSIILGILSYLLI- :
B_FC_Flavobacteriaceae_bacterium@MDC1320985      : ---------MNNAFI-RKHINTFSILIFLFAFFLLNYTR-PGFLYNK-DGSLRRFGLGHK-------RKTIVPIWLLSIILGILSYLFV- :
B_FC_Flavobacteriales_bacterium@MAU36413         : ---------MNRNVI-RRNITSFSIIIFIILYTIVLLFK-PAFIFKK-DGNLRNFGIGFR-------DKTVIPAWLLAIILAIISYLFV- :
B_Pr_Rhodobacteraceae_bacterium@MBC67241         : --------MMNRNYI-RRNITTIAIIIYALLYTIVIMLK-PAFVYNE-DGSLRDFGIGYK-------KKTVIPVWLVAICLAIVSYFGV- :
B_Pr_Alphaproteobacteria_bacterium@NDA90778      : ------MNSMYRSYI-SKNVALVSILLFLAIFITIQIGQ-PSFLYRE-DGSLRDFGIGYR-------NKTILPVWLLAIVLGILSYLFV- :
B_Pr_Rhodobacteraceae_bacterium@MAR51059         : -------MARHRLFI-RKNITSVSIFLFVLIFAFVYHLK-PGFLFDE-NGSIRDFGLGYK-------KKTIIPVWFVTFILAIFSYLFV- :
B_un_bacterium@NBC70773                          : -------MNAIKTLI-RNNLAGTAIALYVIVFMLVQYAN-PAFLYNE-DGSLREFGIGYS-------SKTVLPIWIVAIILGILSYVTV- :
B_Ba_Candidatus_Elulimicrobium@NBP58319          : --------------------------TDIIVFMLVQYMN-PTFIYNE-DGSLREFGIGYS-------SKTILPIWLVAIVLGILSYVTV- :
B_FC_Bacteroidetes_bacterium@NBP65283            : --------NAIKNII-RNNLAGSAILLYVIVFMLVQYMN-PSFIYNE-DGSLREFGIGYS-------SKTVIPIWLVAIILGILSYVTV- :
A_Ea_Euryarchaeota_archaeon@RZD40573             : ---------MYRSFI-RRNITSVAIVIFVIIFCLVQVYA-PHFLYNE-DGSLRQFGIGYK-------KKTVIPNWLVALLLAILSYLFV- :
B_FC_Flavobacteriaceae_bacterium@MBF12520        : ---------MYRQYI-RRNITSVSIILFVVIFCIVQLVE-PAFLYEK-DGSLRQFGLGTK-------KKTVIPIWFLTLILAIFCYLFV- :
B_Pr_Bdellovibrionaceae_bacterium@MBN20832       : ---------MFRQYI-RRNITSVAIIVFIILFCIIQFIE-PAFLYEK-DGSLRRFGLGSQ-------KKTIIPIWFLTLIIAIFSYLFV- :
V_Ba_Yellowstone_lake_mimivirus@YP_009174099     : --------MSSREFM-KENITLISIILFVCIFLIIQFTK-PAFLFKS-DGSLRVFGIGYK-------TKTILPIWLFSIILGILCYLLV- :
B_FC_Candidatus_Marinimicrobia@MCJ7802881        : ---------MNKSYI-RENITLVSVALFVVIFGTIQLIK-PNCFYNK-DGSIREFGIGYK-------NKTILPIWLLSLVLGILCYLSV- :
u_un_Yaravirus_brasiliensis@YP_010800670         : DGPPLGRIPTALSALGMGTPLGRFIGFSLLTGGIIYVAR-PSVFFV--DGTAKPWSLLTLGKPSGGVKSTAVPWYLVALLVGFVGATFI- :
E_Ha_Pavlovales@KAJ1618628                       : --------EEVLGFLGMGTPSRRALAFAALATSVSYALKTPACFFTE-EGAIKGFHPDPESAV-EPESKTQLHFLVVPLLAGAVGYALL- :
E_Ha_Pavlovales@KAJ1622493                       : --------EEYIGFMGGGTPTKCAAILATAATAICYATKTPAVFFTE-KGSIKQFHPDPEGSD-EPENKTQLHFVIVPILAACVGYLLL- :
E_Ha_Pavlovales@KAJ1636553                       : --------EEFIGFMGGGTPTKRAAILATAATAICYATKTPAIFFTE-QGSIKQFHPDPEGSD-EPENKTQLHFVIVPILAACVGYLLL- :
E_Ha_Pavlovales@KAJ1618980                       : --------EEFISFMGGGTPTKRAAILATAATAICYATKTPAIFFTE-KGGIKQFHPDPEGSD-EPENKTQLHFVIVPILAACVGYLFL- :
E_Ha_Pavlovales@KAJ1620103                       : VRRAPQTVEEFLGFVGAGTPLKRGFIVGALATLVCYATKWPAAFFTE-HGAIKTFQVDPD-ED-ELHEKTQVHFVLVPVLGAVAGYLFL- :
E_Ha_Pavlovales@KAJ1626513                       : --------RDFADFMTREAAVPLATNALPLDGVKLRMPEIPTAFLGF-HGAIKTFQVDPD-ED-ELDEKTQVHFVLVPVLGAVAGYLFL- :
B_PV_Verrucomicrobiales_bacterium@MAH41484       : EQLVIEGWRGVLVDAGLGTPSKRGLTAFLATAAVAYAFKLPKGAFHE-DGTMRPASDGNG------------HFLLTPLTVGAIAFLFT- :
u_Pl_Pleurochrysis_sp_endemic_virus@AUD57302     : -VEVIDDYRALLVKVGLGEPLSRALAVGALTAGVTYLAKFPNEAFRD-DGSVKPLKYLSP-----EPDATYAHFLILPVVIATATYLFT- :
u_Pl_Pleurochrysis_sp_endemic_virus@AUL80787     : -LPFVTDTRGAMVALGLGNPVTRAFVAGSIAAGVAYSSSFPQAAFRE-DGTIKPFKWVSP-----EPDATLTHFLLFPLVVAGATFLFT- :
B_PV_Planctomycetaceae_bacterium@MAT69668        : -APVILGARGVMVNMGLATPLSRAFVIGTTVGLVAYGLGVPRASFNE-EGEMRPLSLVSH-----SEDATRTHFLVVPITAAVAAYLFT- :
B_PV_Planctomycetaceae_bacterium@MAT73275        : NQPVIKGAKGVAVYLGLGTPPARAFVGATIAGCGAYACGIPRAAFDD-EGKCRPFKPFAA-----GVEGTYYHFLAIPLAVGVATYLFT- :
                                                                                           p        g
```

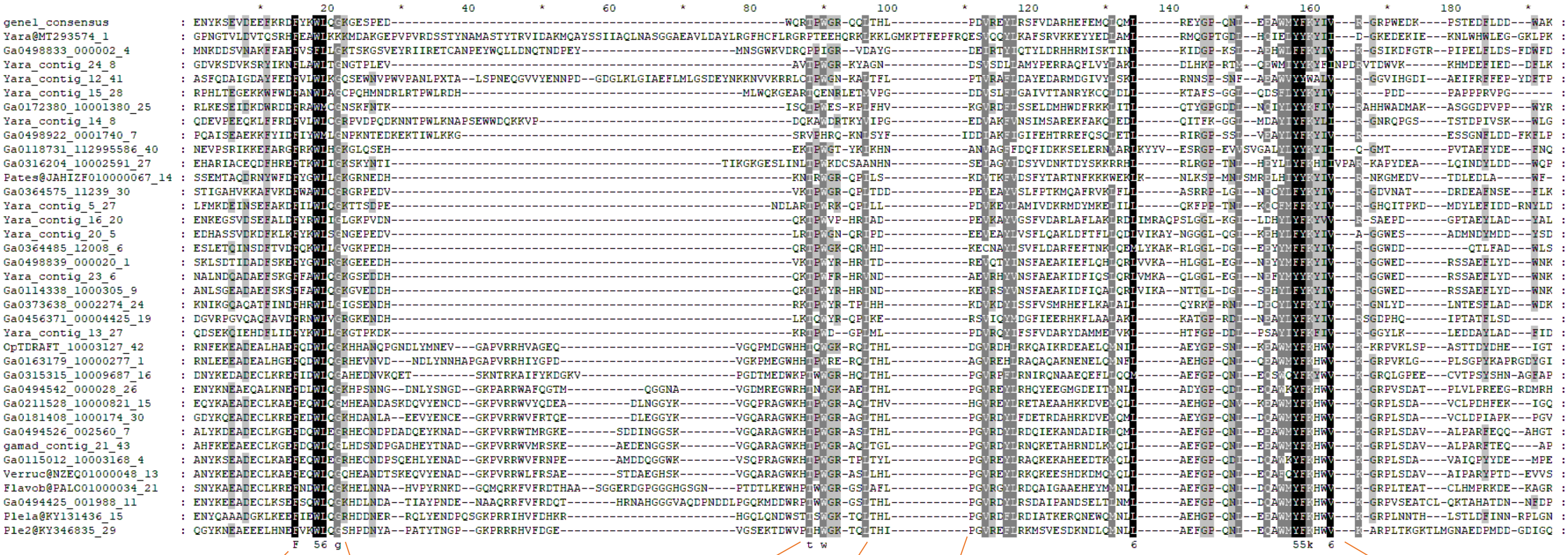
Mriya_50 with nr hits

S17

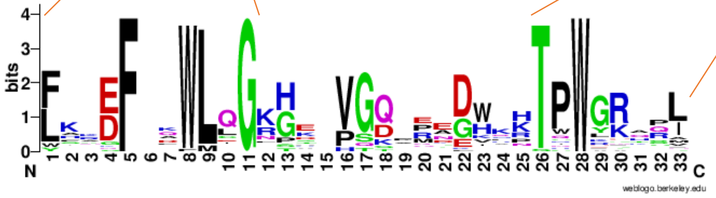


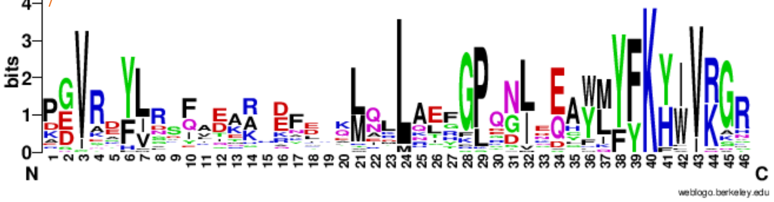
Mriya_50 (Ple2@KY346835_21)

S18

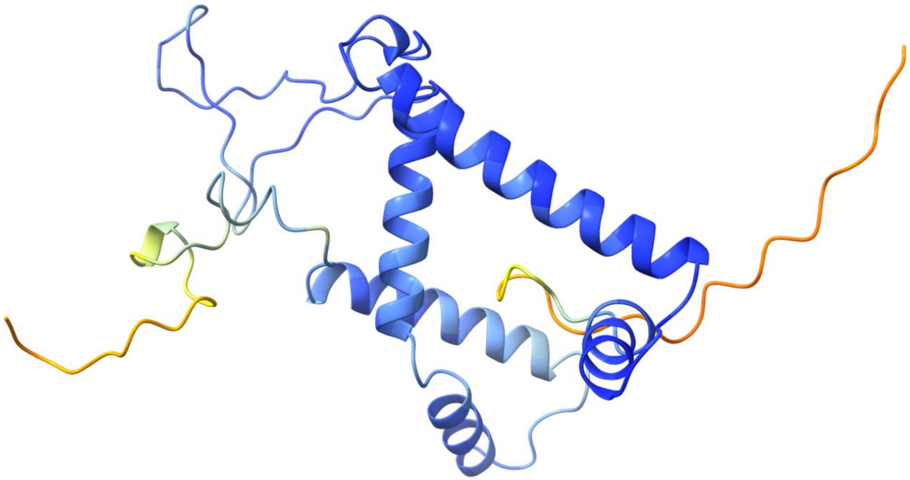


```
                                    *        20         *        40         *        60         *        80         *        100        *        120        *        140        *        160        *        180        *
gene1_consensus          : ENYKSEVDEEFKRDFYKWLQGKGESPED--------------------------------------------------------WQRTPWGR-QQLTHL-----------PDVREYLRSFVDARHEFEMQLQML------REYGP-QNL--EEAWMYFKYIV---R-GRPWEDK----PSTEDFLDD---WAK :
Yara@MT293574_1          : GPNGTVLDVTQSRHFEAWLKKKMDAKGEPVPVRDSSTYNAMASTYTRVIDAKMQAYSSIIAQLNASGGAEAVLDAYLRGFHCFLRGRPTEEHQRKLKKLGMKPTFEPFRQESVQQYLKAFSRVKKEYYEDLAML------RMQGPTGDL--HQIELYYKYII---D-GKEDEKIE---KNLWHWLEG-GKLPK :
Ga0498833_000002_4       : MNKDDSVNAKFFAEFVSFLLGKTSKGSVEYRIIRETCANPEYWQLLDNQTNDPEY----------------------MNSGWKVDRQPPIGR--VDAYG-----------DEIRTYIQTYLDRHHRMISKTINL------KIDGP-KSL--AEHWLFFKYIV---K-GSIKDFGTR--PIPELFLDS-FDWFD :
Yara_contig_24_8         : GDVKSDVKSRYIKNFLAWLTGNGTPLEV---------------------------------------------------------AVTPWGR-KYAGN------------DSVSDLLAMYPERRAQFLVYLAKL------DLHKP-RTM--QEWMIYYKYFINPDRVTDWVK------KHMDEFIED--DFLK :
Yara_contig_12_41        : ASFQDAIGDAYFEDFVLWLKGQSEWNVPWVPANLPXTA--LSPNEQGVVYENNPD--GDGLKLGIAEFLMLGSDEYNKKNVVKRRLCTPWGN-KAITFL-----------PTVRAFLDAYEDARMDGIVYLSKL------RNNSP-SNF--AEAWVYYWALV---R-GGVIHGDI---AEIFRFFEP-YDFTP :
Yara_contig_15_28        : RPHLTEGEKKWFWDFANWLAGCPQHMNDRLRTPWLRDH----------------------------------------MLWQKGEARTQENRLETMVPG-----------DDVSLFLGAIVTTANRYKCQLDLL------KTAFS-GGL--QDSFLYYKYIV---R---PDD------PAPPPRVPG------ :
Ga0172380_10001380_25    : RLKESEIDKDWRDDFRAWMCGNSKFNTK--------------------------------------------------------ISQTPWES-KPLFHV-----------KGVRDFLSSELDMHWDFRKKLITL------QTYGPGDDL--NQIYLYYKFIV---RAHHWADMAK---ASGGDPVPP---WYR :
Yara_contig_14_8         : QDEVPEEQKLFFRDFVLWLCGRPVDPQDKNNTPWLKNAPSEWWDQKKVP------------------------------------DQKAWDRTKYVIPG-----------EDVAKFVNSIMSAREKFAKQLEDL------QITFK-GGL--MDAYIYFKYLI---R-GNRQPGS----TSTDPIVSK---WLG :
Ga0498922_0001740_7      : PQAISEAEKKFYIDFIYWMLGNPKNTEDKEKTIWLKKG-----------------------------------------------SRVPHRQ-KNLSYF----------IDDIAKFIGIFEHTRREFQSQLETL------RIRGP-SSL--VEAYIYFKYIV---R------------ESSGNFLDD-FKFLP :
Ga0118731_112995586_40   : NEVPSRIKKEFARGFRKWLHGKGLQSEH---------------------------------------------------------EKTPWGT-YKIKHN-----------ANVAGFFDQFIDKKSELERNVARLKYYV--ESRGP-EVVSVGALYIYYKYII---Q-GMT--------PVTAEFYDE---FNQ :
Ga0316204_10002591_27    : EHARIACEQDFHREFTKWLIGKSKYNTI-----------------------------------------------TIKGKGESLINLTPWKDCSAANHN-----------SEIAGYIDSYVDNKTDYSKKRRHL------RLRGP-TNL--HEYLIYFKHIIVPAR-KAPYDEA----LQINDYLDD---WQP :
Pates@JAHIZF010000067_14 : SSEMTAQDRNYWFDFYGWLLGKGRNEDH---------------------------------------------------------KNIRWGR-QPILS------------KDVTKFVDSFYTARTNFKKKWEKLK-----NLKSP-MNLSMRELHIYYKYIV---R-NKGMEDV----TDLEDLA-----WF- :
Ga0364575_11239_30       : STIGAHVKKAFVKDFWAWLCGRGRPEDV---------------------------------------------------------VKTPWGR-QPLTDD-----------PEVEAYVSLFPTKMQAFRVKLFLL------ASRRP-LGI--NECYIFYKYIV---R-GDVNAT-----DRDEAFNSE---FLK :
Yara_contig_5_27         : LFMKDEINSEFAKDFILWLQGKTTSDPE------------------------------------------------------NDLARTPWRK-QPLLL------------PDVKEYLAMIVDKRMDYMKELILL------QKFPP-TNL--KQCFMFFKYIV---R-GHQITPKD---MDYLEFIDD-RNYLD :
Yara_contig_16_20        : ENKEGSVDSEFALDFYRWLIGLGKPVDN---------------------------------------------------------QKTPWVP-HRIAD------------PEVKAYVGSFVDARLAFLAKLRDLIMRAQPSLGGL-KGI--LDHYIYFKYVV---R-SAEPD------GPTAEYLAD---YAL :
Yara_contig_20_5         : EDHASSVDKDFKLKFYKWLSGNGEPEDV---------------------------------------------------------LRTPWGN-QRIPD------------EEVEAYLVSFLQAKLDFTFLLQDLVIKAY-NGGGL-QGI--KEHYIFYKYIV---A-GGWES------ADMNDYMDD---YSD :
Ga0364485_12008_6        : ESLETQINSDFTVDFQKWLLGVGKPEDH---------------------------------------------------------QRTPWGK-QRVHD------------KECNAYLSVFLDARFEFTNKLQEMLYKAK-RLGGL-DGI--EEYYMFFKYIV---R-GGWDD---------QTLFAD---WLS :
Ga0498839_000020_1       : SKLSDTIDADFSKEFYGWLRGKGEEEDH---------------------------------------------------------VKTPWYR-HRITD------------REVQTYINSFAEAKIEFLQHLQRLVVKA--HLGGL-EGL--NEYYMFFKYIV---R-GGWED------RSSAEFLYD---WNK :
Yara_contig_23_6         : NALNDQADAEFSKGFFAWLQGKGSEDDH---------------------------------------------------------QKTPWFR-HRVND------------AEVRHYVNSFAEAKIDFIQSLQRLVMKA--QLGGL-EGI--NEFYMYYKYIV---R-GGWED------RSSAEFLYD---WNK :
Ga0114338_1000305_9      : ANLSGEADAEFSKSFFAWLQGKGVEDDH---------------------------------------------------------QKTPWYR-HRIND------------KEVRSYVNSFAEAKIDFIQALQRLVIKA--NTTGL-DGI--SEHYIFYKYIV---R-GGWED------RSSAEFLYD---WNK :
Ga0373638_0002274_24     : KNIKGQAQATFINDFHRWLLGIGSENDH---------------------------------------------------------RKTPWYR-TPIHH------------KDVKDYISSFVSMRHEFLKALALL------QYRKP-RNL--DECYIYFKYIV---R-GNLYD------LNTESFLAD---WDK :
Ga0456371_00004425_19    : DGVRPGVQAQFAVDFRNWLVGRGKENDH---------------------------------------------------------LKTQWYR-QPIKE------------RSVIQYMDGFIEERHKFLAALAKL------KATGP-RDI--NEAYLYFKYIV---RSGDPHQ------IPTATFLSD------ :
Yara_contig_13_27        : QDSEKQIEHDFLIDFYKWLLGKGTPKDK---------------------------------------------------------KRTPWD--GPIML------------PDVRQYIFSFVDARYDAMMELVKL------HTFGP-DDL--PSAYIYFKFIV---R-GGYLK------LEDDAYLAD---FID :
OpTDRAFT_10003127_42     : RNFEKEADEALHAEFQDWLQGKHHANQPGNDLYMNEV---GAPVRRHVAGEQ-------------------------VGQPMDGWHHTQWGK-RQLTHL-----------DGVRDHLRKQAIKRDEAELQMNIL------AEYGP-SNL--KEAWMYFKHWV---K-KRPVKLSP---ASTTDYDHE---IGT :
Ga0163179_10000277_1     : RNLEEEADEALHGEFQDWLQGRHEVNVD---NDLYNNHAPGAPVRRHIYGPD-------------------------VGKPMEGWHHTPWRE-RQLTHL-----------AGVREHLRAQAQAKNENELQMNFL------AEHGP-QNL--QEAWMYFKHWV---K-GRPVKLG----PLSGPYKAPRGDYGI :
Ga0315315_10009687_16    : DNYKEDADECLKREFIDWLQGAHEDNVKQET---------SKNTRKAIFYKDGKV----------------------PGDTMEDWKPTWWGR-HQLTHL-----------PGVRPFLRNIRQNAAEQEFLLQQM------AEFGP-QNI--EQSWQYFKYWV---K-GRQLGPEE---CVTPSYSHN-AGFAP :
Ga0494542_000028_26      : ENYKNEAEQALKNEFDLWLQGKHPSNNG--DNLYSNGD--GKPARRWAFQGTM-----------QGGNA--------VGDMREGWRHTNWGK-AELTHL-----------PGVREYLRHQYEEGMGDEITMNLL------ADYGP-QDL--EQAWKYFKHWV---K-GRPVSDAT---PLVLPREEG-RDMRH :
Ga0211528_10000821_15    : EQYKAEADECLKAEFEQWLQGMHEANDASKDQVYENCD--GKPVRRWVYQDEA---------DLNGGYK--------VGQPRAGWKHTPWGR-AQLTHV-----------HGVREYLRETAEAAHKKDVEMQLL------AEHGP-QNV--KEAWMYFKHWV---K-GRPLSDA----VCLPDHFEK---IGQ :
Ga0181408_1000174_30     : GDYKQEADECLKREFEDWLQGKHDANLA--EEVYENCE--GKPVRRWVFRTQE---------DLEGGYK--------VGQARAGWKHTPWGR-AGLTHL-----------PGVRDYLFDETRDAHRKDVEMQML------AEYGP-QNL--DQAWMYFKHWV---K-GRPLSDA----VCLDPIAPK---PGV :
Ga0494526_002560_7       : ALYKDEADECLKGEFDQWLEGRHECNDPDADQEYKNAD--GKPVRRWTMRGKE-------SDDINGGSK--------VGQARAGWKHTPWGR-ASLTHL-----------PGVRDYLRDQIEKANDADIRLQML------AEFGP-QNI--EEAWMYFKHWV---K-GRPLSDAV---ALPARFEQQ--AHGT :
gamad_contig_21_43       : AHFKEEAEECLKGEFDQWLQGLHDSNDPGADHEYTNAD--GKPVRRWVMRSKE-------AEDENGGSK--------VGQARAGWKHTPWGR-AQLTGL-----------PGVRDYLRNQKETAHRNDLKMQLL------AEFGP-QNL--EEAWMYFKHWV---K-GRPLSDAV---ALPARFTEQ----AP :
Ga0115012_10003168_4     : ANYKSEADECLKAEFEQWLEGRHECNDPSQEHLYENAD--GKPVRRWVFRNPE-------AMDDQGGWK--------VSQPRAGWKHTPWGR-TPLTYL-----------PGVREYLRAQKEKAHEEDTKMQLL------AEFGP-QDI--DQAWKYFKHWV---K-GRPLSDA----VAIQPYYDE---MPE :
Verruc@NZEQ01000048_13   : ANYKEEADECLKAEFEQWLQGQHEANDTSKEQVYENAD--GKPVRRWLFRSAE-------STDAEGHSK--------VGQARAGWKHTPWGR-ASLLHL-----------PGVREYLRKQKEESHDKDMQMQLL------AEFGP-QNL--EQAFQYFKHWV---K-GRPLSDAV---AIPARYPTD--EVVS :
Flavob@PALC01000034_21   : SNYKAEADECLKREFNDWLQGKHELNNA--HVPYRNKD--GQMQRKFVFRDTHAA--SGGERDGPGGGHGSGN----PTDTLKEWHPTWWGR-GSLAFL-----------PGVRGYLRDQAIGAAEHEYMMNLL------AEFGP-QNI--DQAWMYFKHWV---K-GRPLTEAT---CLHMPRKDE--KAGR :
Ga0494425_001988_11      : ENYKEEADECLKSEFSQWLQGKHDLNDA--TIAYPNDE--NAAQRRFVFRDQT-------HRNAHGGGVAQDPNDDLPGQKMDDWRPTWWGR-GSLTHL-----------PGVRDYLRSDAIPANDSELTLNML------AEFGP-QNI--DQAWMYFKHWV---K-GRPVSEATCL-QKTAHATDN--NFDP :
Plela@KY131436_15        : ENYQAAADGKLKEEFIEWLQGRHDDNER--RQLYENDPQSGKPRRIHVFDHKR------------------------HGQLQNDWSTTSWGK-TQLTHL-----------PGVRDFLRDIATKERQNEWQMNLL------AEHGP-QNL--EEAWMYFKHWV---K-GRPLNNTH---LSTLDFINN-RPLGN :
Ple2@KY346835_29         : QGYKNEAEEELHNEFVKWLQGSHPDNYA--PATYTNGP--GKPRRRHVFDGE-------------------------VGSEKTDWVPTHWGK-TQLTHI-----------PGVREFLRKMSVESDKNDLQMNLL------AEFGP-QNL--EQAWMYFKHWV---K-ARPLTKGKTLMGNAEDPMDD-GDIGQ :
                                         F  56 g                                                                  t w                                           6                      55k  6
```





S19

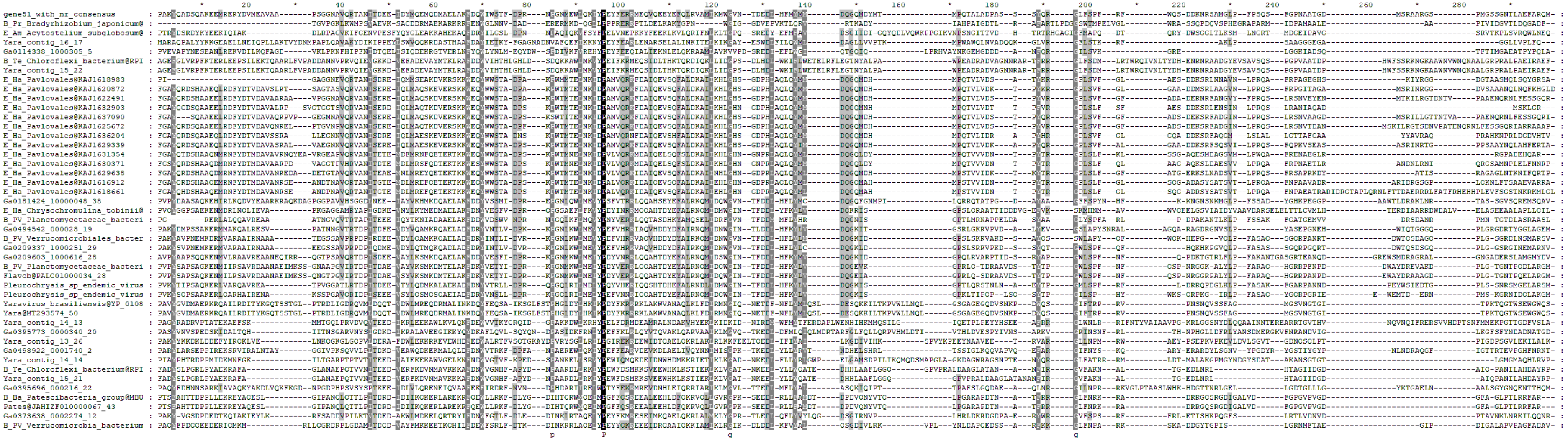


Mriya_1, exemplified with Ple2@KY346835_29

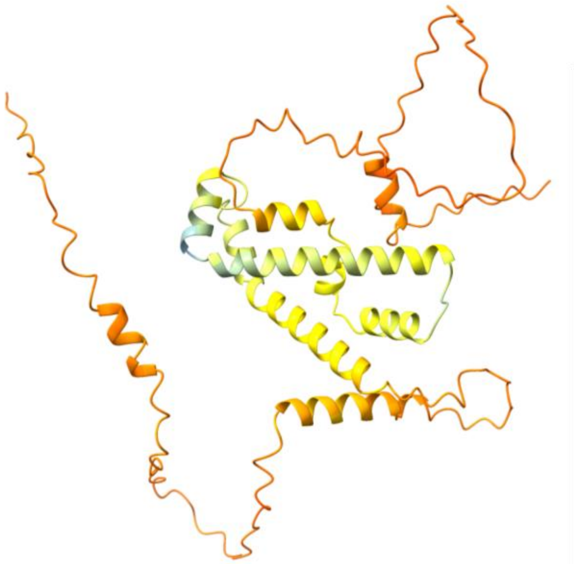
S20

Mriya_51

A



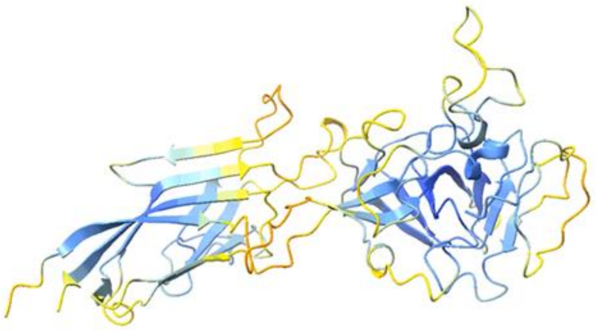
B



(Yara@MT293574_50, aka YP_010800671)

S21

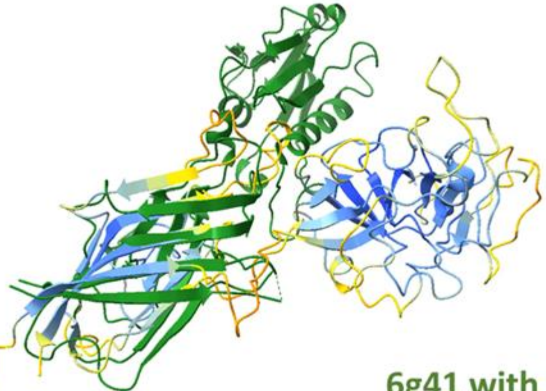
A



Foldseek:
4 uncharacterized proteins

Dali
6.4 , 6g41-A , MOLECULE: MINOR CAPSID PROTEIN;
→ Cafeteriavirus-dependent mavirus (Varidnaviria, Bamfordvirae, Preplasmiviricota)
5.4 , 3j26-N , MOLECULE: CAPSID PROTEIN V20;
→ resolution structure of the Sputnik virophage by cryo-EM

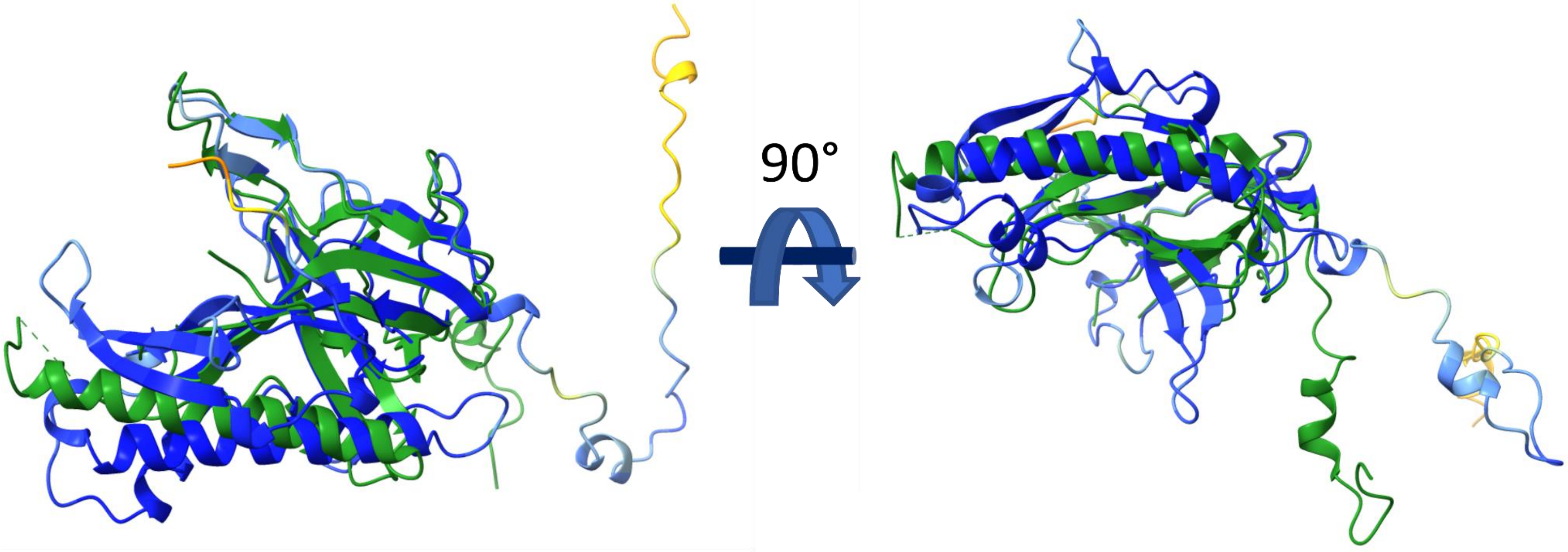
B



**6g41 with SJR (overlayed with Ple_22)
and insertion domain**

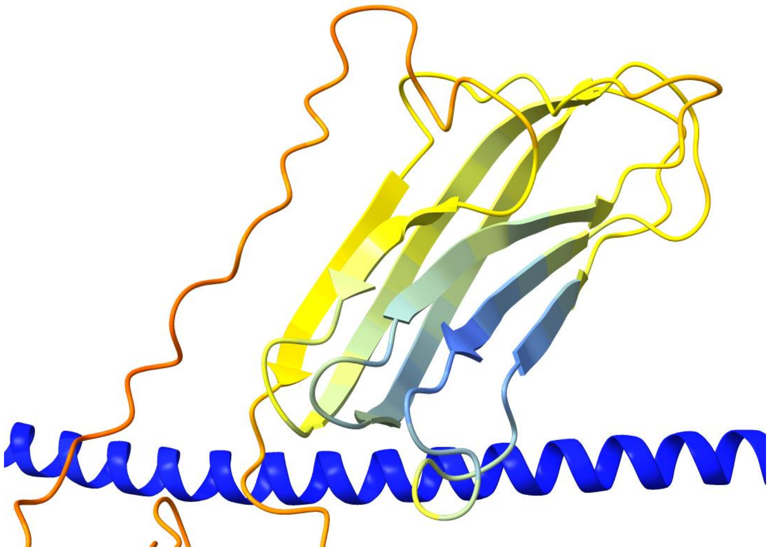
S22

PEV_26 - ssb



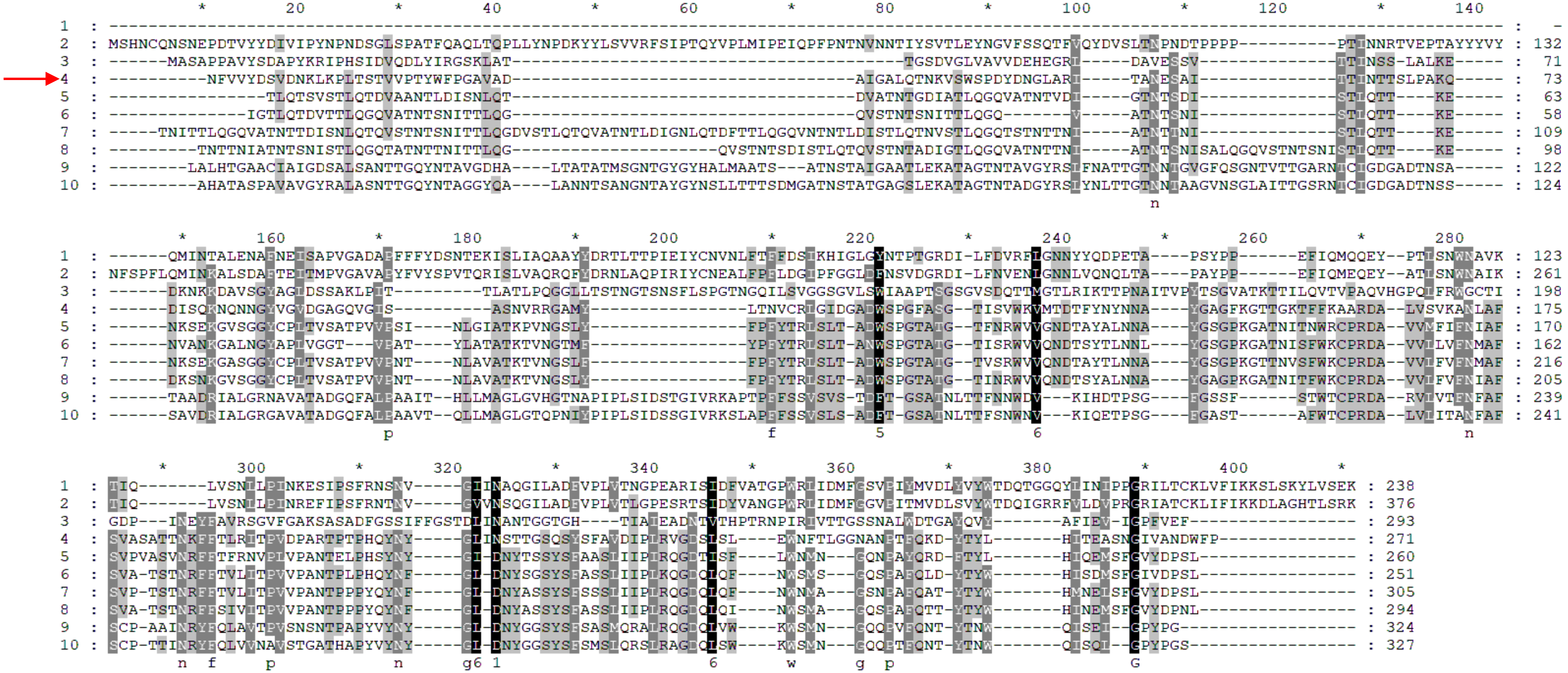
S23

Yaravirus minor capsid protein (YP_010800666.1)



S24

Yaravirus minor capsid protein (C-terminal of YP_010800666.1)



```
              *        20         *        40         *        60         *        80         *       100         *       120         *       140
1  : ---------------------------------------------------------------------------------------------------------------------------------------------- : -
2  : MSHNCQNSNEPDTVYYDIVIPYNPNDSGLSPATFQAQLTQPLLYNPDKYYLSVVRFSIPTQYVPLMIPEIQPFPNTNVNNTIYSVTLEYNGVFSSQTFVQYDVSLTNPNDTPPPP----------PTINNRTVEPTAYYYVY : 132
3  : ------MASAPPAVYSDAPYKRIPHSIDVQDLYIRGSKLAT----------------------------------------TGSDVGLVAVVDEHEGRI-----DAVESSV--------------TTINSS-LALKE----- : 71
4  : ----------NFVVYDSVDNKLKPLTSTVVPTYWFPGAVAD-----------------------------------AIGALQTNKVSWSPDYDNGLARI-----TANESAI--------------TTINTTSLPAKQ----- : 73
5  : ----------------TLQTSVSTLQTDVAANTLDISNLQT-----------------------------------DVATNTGDIATLQGQVATNTVDI-----GTNTSDI--------------STLQTT----KE----- : 63
6  : --------------IGTLQTDVTTLQGQVATNTSNITTLQG-----------------------------------QVSTNTSNITTLQGQ-------V-----ATNTSNI--------------STLQTT----KE----- : 58
7  : -----TNITTLQGQVATNTTDISNLQTQVSTNTSNITTLQGDVSTLQTQVATNTLDIGNLQTDFTTLQGQVNTNTLDISTLQTNVSTLQGQTSTNTTNI-----ATNTTNI--------------STLQTT----KE----- : 109
8  : ---------TNTTNIATNTSNISTLQGQTATNTTNITTLQG---------------------QVSTNTSDISTLQTQVSTNTADIGTLQGQVATNTTNI-----ATNTSNISALQGQVSTNTSNISTLQTT----KE----- : 98
9  : --------LALHTGAACIAIGDSALSANTTGQYNTAVGDHA----LTATATMSGNTGYGYHALMAATS---ATNSTAIGAATLEKATAGTNTAVGYRSIFNATTGTNNTGVGFQSGNTVTTGARNTCIGDGADTNSA----- : 122
10 : ---------AHATASPAVAVGYRALASNTTGQYNTAGGYQA----LANNTSANGNTAYGYNSLLTTTSDMGATNSTATGAGSLEKATAGTNTADGYRSIYNLTTGTNNTAAGVNSGLAITTGSRNTCIGDGADTNSS----- : 124
                                                                                                                   n

            *       160         *       180         *       200         *       220         *       240         *       260         *       280
1  : ------QMINTALENAFNEISAPVGADAPFFFYDSNTEKISLIAQAAYYDRTLTTPIEIYCNVNLFTFFDSIKHIGLGYNTPTGRDI-LFDVRFIGNNYYQDPETA----PSYPP------EFIQMQQEY--PTISNWNAVK : 123
2  : NFSPFLQMINKALSDAFTEITMPVGAVAPYFVYSPVTQRISLVAQRQFYDRNLAQPIRIYCNEALFPFLDGIPFGGLDFNSVDGRDI-LFNVENIGNNLVQNQLTA----PAYPP------EFIQMEQEY--ATISNWNAIK : 261
3  : ------DKNKKDAVSGYAGIDSSAKLPIT---------TLATLPQGGLLTSTNGTSNSFLSPGTNGQILSVGGSGVLSWIAAPTSGSGVSDQTTMGTLRIKTTPNAITVPYTSGVATKTTILQVTVPAQVHGPQLFRWGCTI : 198
4  : ------DISQKNQNNGYVGVDGAGQVGIS----------ASNVRRGAMY----------------LTNVCRLGIDGADWSPGFASG--TISVWKVMTDTFYNYNNA----YGAGFKGTTGKTFFKAARDA--LVSVKANLAF : 175
5  : ------NKSEKGVSGGYCPLTVSATPVVPSI----NLGIATKPVNGSLY----------------FPFYTRLSLT-ADWSPGTATG--TFNRWVVGNDTAYALNNA----YGSGPKGATNITNWRCPRDA--VVMFIFNIAF : 170
6  : ------NVANKGALNGYAPLVGGT---VPAT----YLATATKTVNGTMF----------------YPFYTRLSLT-ANWSPGTATG--TISRWVVQNDTSYTLNNL----YGSGPKGATNISFWKCPRDA--VVLLVFNMAF : 162
7  : ------NKSEKGASGGYCPLTVSATPVVPNT----NLAVATKTVNGSLF----------------FPFYTRLSLT-ADWSPGTATG--TVSRWVVQNDTAYTLNNA----YGSGPKGTTNVSFWKCPRDA--VVLFVFNMAF : 216
8  : ------DKSNKGVSGGYCPLTVSATPVVPNT----NLAVATKTVNGSLY----------------FPFYTRLSLT-ADWSPGTATG--TINRWVVQNDTSYALNNA----YGAGPKGATNITFWKCPRDA--VVLFVFNIAF : 205
9  : ------TAADRIALGRNAVATADGQFALPAAIT--HLLMAGLGVHGTNAPIPLSIDSTGIVRKAPTPFFSSVSVS-TDFT-GSATNLTTFNNWDV---KIHDTPSG----FGSSF------STWTCPRDA--RVLVTFNFAF : 239
10 : ------SAVDRIALGRGAVATADGQFALPAAVT--QLLMAGLGTQPNIYPIPLSIDSSGIVRKSLAPFFSSVSLS-ADFT-GSATNLTTFSNWNV---KIQETPSG----FGAST------AFWTCPRDA--LVLITANFAF : 241
                                 p                                      f          5               6                                           n

          *       300         *       320         *       340         *       360         *       380         *       400         *
1  : TIQ-------LVSNILPINKESIPSFRNSNV-----GIINAQGILADFVPLVTNGPEARISIDFVATGPWRLIDMFGSVEIYMVDLYVYWTDQTGGQYLINIPPGRILTCKLVFIKKSLSKYLVSEK : 238
2  : TIQ-------LVSNILPINREFIPSFRNTNV-----GVVNSQGILADFVPLVTLGPESRTSIDYVANGPWRLIDMFGGVEITMVDLSVYWTDQIGRRFVLDVPRGRIATCKLIFIKKDLAGHTLSRK : 376
3  : GDP---INEYFAVRSGVFGAKSASADFGSSIFFGSTDLINANTGGTGH----TIAIEADNTVTHPTRNPIRIVTTGSSNALWDTGAYQVY-------AFIEV-IGPFVEF----------------- : 293
4  : SVASATTNKFFTLRITPVDPARTPTPHQYNY-----GLINSTTGSQSYSFAVDIPLRVGDSLSL----EWNFTLGGNANPTFQKD-YTYL-------HITEASNGIVANDWFP-------------- : 271
5  : SVPVASVNRFFTFRNVPLVPANTELPHSYNY-----GI-DNYTSSYSFAASLIIPIRQGDTISF----LWNMN---GQNPAYQRD-FTYL-------HIQEMSFGVYDPSL---------------- : 260
6  : SVA-TSTNRFFTVLITPVVPANTPLPHQYNF-----GI-DNYSGSYSFASSLIIPIKQGDQIQF----NWSMS---GQSPAFQLD-YTYW-------HISDMSFGIVDPSL---------------- : 251
7  : SVP-TSTNRFFTVLITPVVPANTPPPYQYNF-----GI-DNYASSYSFSSSLIIPIRQGDQIQF----NWNMA---GSNPAFQAT-YTYW-------HMNEISFGVYDPSL---------------- : 305
8  : SVA-TSTNRFFSIVITPVVPANTPPPYQYNF-----GI-DNYASSYSFASSLIIPIRQGDQIQI----NWSMA---GQSPAFQTT-YTYW-------HINEMSFGVYDPNL---------------- : 294
9  : SCP-AAINRYFQLAVTPVSNSNTPAPYVYNY-----GI-DNYGGSYSFSASMQRAIRQGDQIVW----KWSMN---GQQPVFQNT-YTNW-------QISEI--GPYPG------------------ : 324
10 : SCP-TTINRYFQLVVNAVSTGATHAPYVYNY-----GI-DNYGGSYSFSMSLQRSIRAGDQISW----KWSMN---GQQPTFQNT-YTNW-------QISQI--GPYPGS----------------- : 327
            n  f     p            n      g6 1                     6       w      g  p                        G
```

1 Sputnik_UMZ08530 .1 minor virion protein [Mimivirus-dependent virus Sputnik]
2 Zamilon_YP_008859634 .1 minor virion protein [Zamilon virus]
3 YP_010800667.1 hypothetical protein QKS73_gp57 [Yaravirus brasiliensis]
4 YP_010800666.1:200-470 tail needle DNA stabilization protein [Yaravirus brasiliensis]
5 YP_009507121.1:265-524 chromosome segregation protein (SMC) [Tunisvirus fontaine2]

6 YP_009094634.1:305-555 hypothetical protein MEL_133 [Melbournevirus]
7 YP_009345180.1:241-545 hypothetical protein B1750_gp035 [Noumeavirus]
8 YP_009238991.1:321-614 hypothetical protein A3303_gp486 [Brazilian marseillevirus]
9 YP_010804828.1:961-1284 hypothetical protein QKU73_gp304 [Acanthamoeba castellanii medusavirus]
10 YP_010802084.1:451-777 hypothetical protein QKT49_gp315 [Acanthamoeba castellanii medusavirus]