

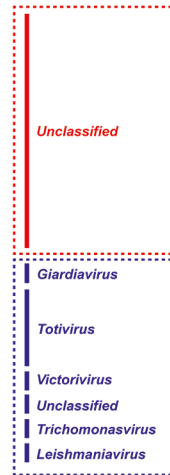
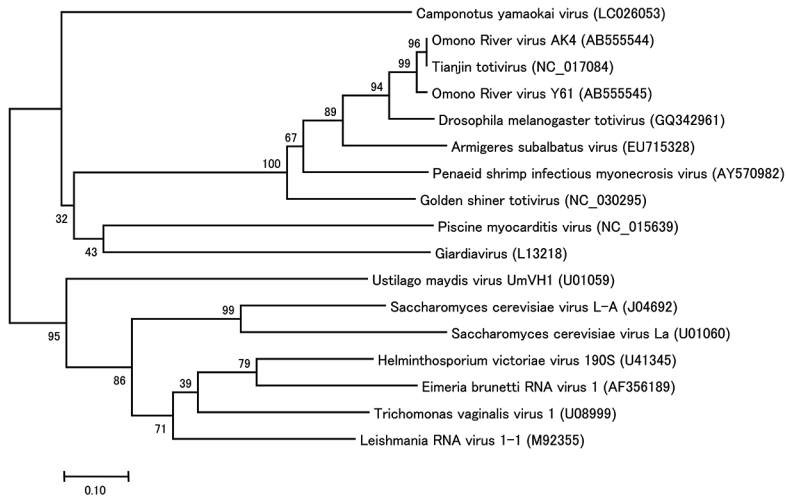
# **The infectious particle of insect-borne totivirus-like Omono River virus has raised ridges and lacks fibre complexes**

\*Kenta Okamoto<sup>1</sup>, Naoyuki Miyazaki<sup>2</sup>, Daniel S. D. Larsson<sup>1</sup>, Daisuke Kobayashi<sup>3</sup>,  
Martin Svenda<sup>1</sup>, Kerstin Mühlig<sup>1</sup>, Filipe R.N.C. Maia<sup>1</sup>, Laura H. Gunn<sup>1</sup>, Haruhiko  
Isawa<sup>3</sup>, Mutsuo Kobayashi<sup>3</sup>, Kyoko Sawabe<sup>3</sup>, \*Kazuyoshi Murata<sup>2</sup>, and Janos Hajdu<sup>1</sup>

1. The Laboratory of Molecular Biophysics, Department of Cell and Molecular biology, Uppsala University, Sweden
2. National Institute for Physiological Sciences (NIPS), Japan
3. Department of Medical Entomology, National Institute of Infectious Diseases (NIID), Japan

\*Corresponding author1: Kenta Okamoto e-mail: kenta.okamoto@icm.uu.se

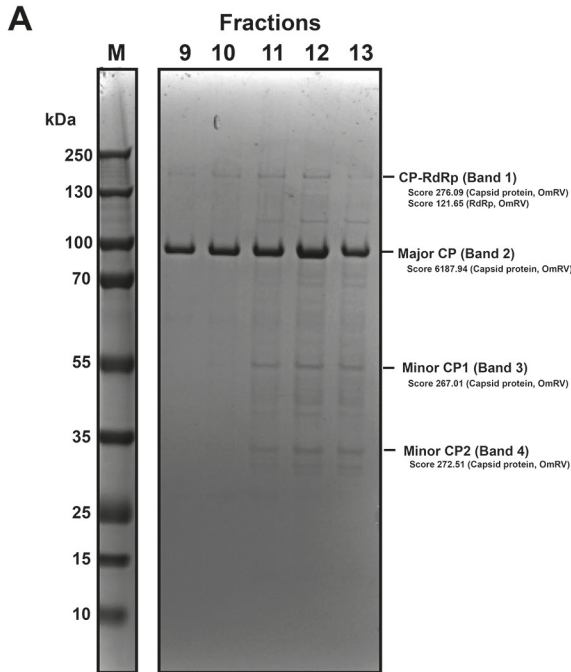
\*Corresponding author2: Kazuyoshi Murata e-mail: kazum@nips.ac.jp



**Metazoan hosts**  
Totivirus-like viruses

**Fungal or protozoan hosts**  
5 genera of Totiviridae family

Fig. S1 Phylogenetic relationships among five genera of *totiviridae* viruses and unclassified *totiviridae* viruses inferred from RdRp amino acid sequences. The GenBank or the RefSeq accession numbers are shown in the figure. The unclassified *totiviridae* viruses above the Giardia virus are called totivirus-like viruses including OmRV, IMNV, PMCV, AsMV, and DTV and are infectious to metazoan hosts such as mosquitoes, shrimps, and fish. The five Totivirus genera and the unclassified viruses below Giardia virus are infectious to fungal or protozoan hosts except plant *Ustilago maydis* virus (UmVH1). The sequences were aligned using ClustalW<sup>1</sup>. The phylogenetic tree was built using a neighbour-joining algorithm in ClustalW.



**B**

N-term

MPGYSANYREGPVI VPSGDQVVTSPTSHGRSKGVSELLTVGGQTYKAIASNEVV  
KSGQAKGEQKFFDENAKIVLLGTFPFKKDLTKEGVEPNPGPVVSDPYANTIYGPL  
PTHKEAAINLGEQPSAPVKRQRKHAVVDRVYKTLTKNGILVTKVVIAGVATYI  
VYKVGSEVWSWYAAITKPLLGVQTAVKEEYHTLQHKYDYQWYDFCPMCLLSKYI  
ANSVNGTDAEHHNHVMHALNGNIDCDSVFGNNFNITTSPTQLTMSGGLPAGKY  
QTTLTVQALIGGTGVVVGTVTFAGKTVAYQVDFDSFASFDLGTVTVSASTAPSVI  
WTGSTGATLTMVNIICKPI TPTSVAISGQPIWTPYAPAQAVMTVPAVAKALKN  
TQRAADLSTRNKTRHGSNNAISVLNQLATQIDVEDQRLPIYKIEGEREGTHEPSF  
SCQLTFRGMTKSAFGRKTRKEAKTAAASLILEELTRKSNSTWPAADAKLTLTSGWVR  
DLTEEGIEPNPGPIANEDGAECWDPEELKMLQAMMAREQQKVLPMIDDPKLVLD  
LDEMQLLAPQESCNKTVSVDEYQEHVRHMSNYSGAVGEAFVVRWYLLNETTNR  
APLVVMDSDMDEPIATPVDVSPAKRPSGVSRRSDKEEEKPKPKPKKEKIVPAPQP  
PKTQEQLDDEERKFAMQRIADRIKSDPIKLLTWIKRPTSTFKREVLSLAFDIPHI  
NSKPSLDQYETIVCYFNDIPDELITMLLQYDRKWTLSLTDCKLSILSACNVRA  
SLEAAKAHNKLMHAYNGNPI SADFSEVENAPSFSLAENTDEVLPKPTGLEIQT  
ITNIVGDANPNQSRIFDQDRLRGNQYSAGGLVTQNAVAI PFTNLI PRTIRVGN  
LVNSANRLQITETNVSEYYSNP IATKLESEMISDQVKNQFSTWRDNTSLQGFN  
AFDIATINTAILPNGLSLESMLLKL SLLHSIKAMNVDAASINRSQYQVIDHNTVP  
TIGAPAVVGVNNSPVFGECCGNNPVYFPGGDTGAI AFHVTLQTVTPDERKSYAIF  
VPPAILQATSDANEALALFALSMEWPHALYTVTKQTTDLAGANAGQVFIPTQS  
THHIGRRVLDLIPRREIAPNPTTLVAANAMCMVRPQAGPDATAGAIPLAAGQL  
FNMNF IGAPAFEWPLTSLYLSWAGRFDITTRIQYMGRLATMVGVKDAYNAAEHL  
NVALSQVAPKMTTAAAGWAAQAANSQQSDVCYSLLTVTRSAANFLANQPAAD  
MRVYDTPATWKNVALGLATAANLVPEQSMDFVFGDARTSFWERLQAI PMCIA  
WTMYHSGRITTLAWDNAYTDNTNKLQKMRNTFTTQSVGTIIPARYGKIVCN  
LYKNMFHRAPAYVATSVGGKELHITFERWLPGGTYANVYSGAGAVVNCPSPLVI  
PDINWCQYFTAKLPLFAGAFPPAQGQNSKGFNSKQGLMIHRNQNNLVAPYLEKF  
ADNSSYFPVGGQPEINDMATWNGRLWMTTGNVQYLDYSGAAIVEAVPPAGELPVG  
KQIPLLAGENAPIELTNAATTCVPRYSNDGRRIFTYLLTQAQSVIPVQACNRAANL  
ARSCWLLSNVYAEFALQALGDEVEDAFDILTNSFLDVAKSVAESAGEVPATKAL  
TDLQAVDVSSLPSTSDPSNVLSQPAPLMSPTSSS C-term

**C**

ORF1 (CP)

N-term

MPGYSANYREGPVI VPSGDQVVTSPTSHGRSKGVSELLTVGGQTYKAIASNEVV  
KSGQAKGEQKFFDENAKIVLLGTFPFKKDLTKEGVEPNPGPVVSDPYANTIYGPL  
PTHKEAAINLGEQPSAPVKRQRKHAVVDRVYKTLTKNGILVTKVVIAGVATYI  
VYKVGSEVWSWYAAITKPLLGVQTAVKEEYHTLQHKYDYQWYDFCPMCLLSKYI  
ANSVNGTDAEHHNHVMHALNGNIDCDSVFGNNFNITTSPTQLTMSGGLPAGKY  
QTTLTVQALIGGTGVVVGTVTFAGKTVAYQVDFDSFASFDLGTVTVSASTAPSVI  
WTGSTGATLTMVNIICKPI TPTSVAISGQPIWTPYAPAQAVMTVPAVAKALKN  
TQRAADLSTRNKTRHGSNNAISVLNQLATQIDVEDQRLPIYKIEGEREGTHEPSF  
SCQLTFRGMTKSAFGRKTRKEAKTAAASLILEELTRKSNSTWPAADAKLTLTSGWVR  
DLTEEGIEPNPGPIANEDGAECWDPEELKMLQAMMAREQQKVLPMIDDPKLVLD  
LDEMQLLAPQESCNKTVSVDEYQEHVRHMSNYSGAVGEAFVVRWYLLNETTNR  
APLVVMDSDMDEPIATPVDVSPAKRPSGVSRRSDKEEEKPKPKPKKEKIVPAPQP  
PKTQEQLDDEERKFAMQRIADRIKSDPIKLLTWIKRPTSTFKREVLSLAFDIPHI  
NSKPSLDQYETIVCYFNDIPDELITMLLQYDRKWTLSLTDCKLSILSACNVRA  
SLEAAKAHNKLMHAYNGNPI SADFSEVENAPSFSLAENTDEVLPKPTGLEIQT  
ITNIVGDANPNQSRIFDQDRLRGNQYSAGGLVTQNAVAI PFTNLI PRTIRVGN  
LVNSANRLQITETNVSEYYSNP IATKLESEMISDQVKNQFSTWRDNTSLQGFN  
AFDIATINTAILPNGLSLESMLLKL SLLHSIKAMNVDAASINRSQYQVIDHNTVP  
TIGAPAVVGVNNSPVFGECCGNNPVYFPGGDTGAI AFHVTLQTVTPDERKSYAIF  
VPPAILQATSDANEALALFALSMEWPHALYTVTKQTTDLAGANAGQVFIPTQS  
THHIGRRVLDLIPRREIAPNPTTLVAANAMCMVRPQAGPDATAGAIPLAAGQL  
FNMNF IGAPAFEWPLTSLYLSWAGRFDITTRIQYMGRLATMVGVKDAYNAAEHL  
NVALSQVAPKMTTAAAGWAAQAANSQQSDVCYSLLTVTRSAANFLANQPAAD  
MRVYDTPATWKNVALGLATAANLVPEQSMDFVFGDARTSFWERLQAI PMCIA  
WTMYHSGRITTLAWDNAYTDNTNKLQKMRNTFTTQSVGTIIPARYGKIVCN  
LYKNMFHRAPAYVATSVGGKELHITFERWLPGGTYANVYSGAGAVVNCPSPLVI  
PDINWCQYFTAKLPLFAGAFPPAQGQNSKGFNSKQGLMIHRNQNNLVAPYLEKF  
ADNSSYFPVGGQPEINDMATWNGRLWMTTGNVQYLDYSGAAIVEAVPPAGELPVG  
KQIPLLAGENAPIELTNAATTCVPRYSNDGRRIFTYLLTQAQSVIPVQACNRAANL  
ARSCWLLSNVYAEFALQALGDEVEDAFDILTNSFLDVAKSVAESAGEVPATKAL  
TDLQAVDVSSLPSTSDPSNVLSQPAPLMSPTSSS C-term

ORF2 (RdRp)

N-term

MSTPGDILLTDYR V KSDLA VTRLR I KDFIARIDPNI V KFMNQHLC HLDQTILANL  
CIWQLNGLRHLVTLHSLGMLNDIDTYATKMAK V SFAKRFPFETDNARQRWCEV  
NTLTGYM QNDIGTFDYDKEFESLATGGNEHPWWRKRFEEKVRDLMTFQEAPFV  
SFETYVKDGYWLTSGSSSVGK V EWSYDGD T GFKRARKNMLVDLYTKEEYQIAID  
WDGELMNRVFIKDELAKRRLAVASNIEAYLNQAYLLYLFHGFGKNYKI TLDEKP  
NETHKRNCRILKLLK EGSYALPFDFKGFDRQPTTDEIKTI I KRVIDLVS PRVPAS  
HRRLFNTI AFKNIACYDKNYLSPLTKRTRVQRTGGLPSGIRPSTLIGNLWNMIAT  
DIARDVTKIDLGQDYIQEIALKGDDTYIVAPNFVCLVFRYAYXSINAVGENSKF  
GIMQNAEFLRTEISSSGVGRWTRAI PSVTRQRPWNPEPWTEQQVQTIANNIY  
LLERRCKQDCTLLHMANKIKWSKMMHQSYLWHLPKHHGGFGIYEWSGWLPDCKL  
PLTQPPVFDVANLNPSSVDLSWYSLSDSEKRAYQKVEFSKIAANDIPGPAKHVM  
RSYIVALRACKPEWSKTLVNFKPIPLISGPTYTNPFWPHRIKTDTSNTNGFPQ  
LSEFVRQHQIAKRAKIPVPPLRDMIAKHYPSAYKVMNRCELQGWHRDTSINIACG  
KIPTEPLKILNPI LAAWVQIVL NAGLLKWSGRKNIANRLYAATTLAVHHIQCEG  
GASMYAF C-term

Fig. S2 Identification of OmRV proteins in the purified fractions. A) The four protein bands that were analysed using tandem-MS spectrometry. The proteins were identified and scores calculated by SEQUEST<sup>2</sup>. Each lane number represents a 1 mL fraction from a 14 mL 5-50% sucrose density gradient numbered from top to bottom. B) Detected peptides from the major and minor OmRV proteins. The full amino acid sequence of ORF1 is shown. Peptides only detected in the minor CP1 and minor CP2 bands (Band 3 and Band 4), and not detected in the major CP band (Band 2), are shown in red. Peptides detected in the major CP band (Band 2) are shown in blue. C) Detected peptides of the CP (ORF1) and the RdRp (ORF2) from Band 1. The detected peptides are shown in red.

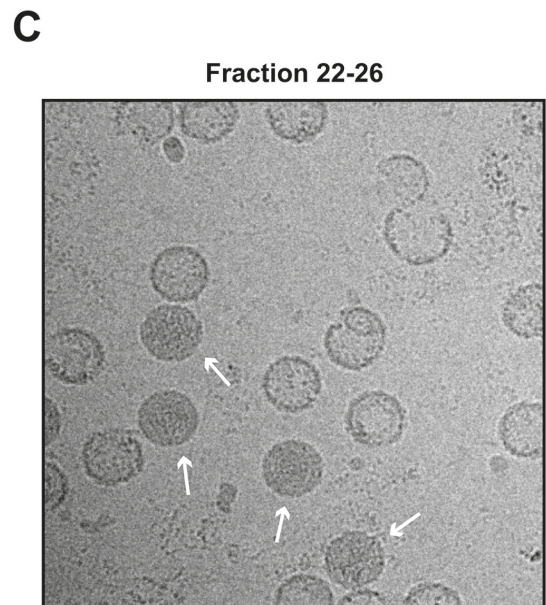
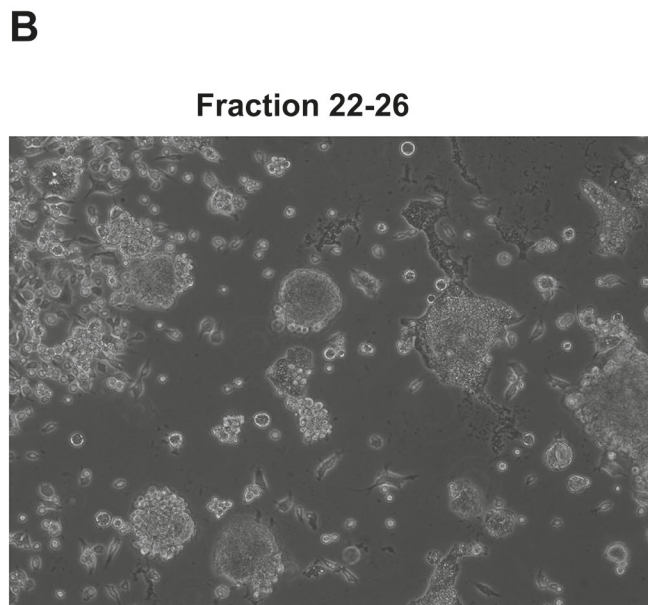
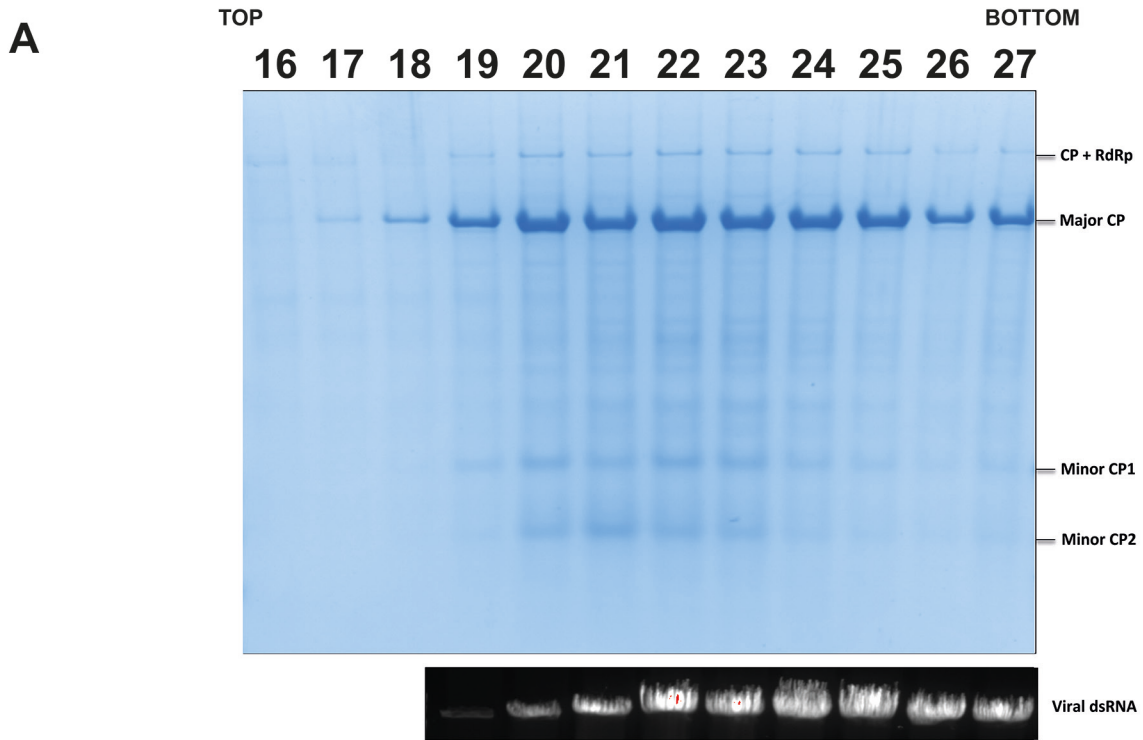


Fig. S3 Protein and dsRNA compositions and the infectivity titre in purified OmRV fractions. A) SDS-PAGE and agarose gel electrophoresis of the protein- and dsRNA-containing fractions from the 14 mL 5-50% sucrose density gradient. Each lane number represents a 500  $\mu$ L fraction numbered from top to bottom. B) Cytopathic effects on C6/36 mosquito cells three days post inoculation of pooled fraction 22-26. The sample was diluted  $10^5$  times before inoculation. C) Raw cryo-EM images of pooled fraction 22-26 that was used for the plaque assay. The dsRNA-full particles are indicated by arrows.

A

```

D8VY-MCP 1 FID-ADPVERA98FFLLA99HTD100VE101LP102FT103LE104Q105IT106TH107VD108AN109QR110IT111QD112RI113ND114Q115Y116AG117LV118Q119AV120AI121FP122FL123PI124TR125---VGR126IV127NS 96
D7V-MCP 1 FIC-QD98VERA99FFLLA100HTD101VE102LP103FT104LE105Q106IT107TH108VD109AN110QR111IT112QD113RI114ND115Q116Y117AG118LV119Q120AV121AI122FP123FL124PI125TR126---SDG127VER128E 97
A8TV-MCP 1 ILE-N98TL99LD100VE101ED102RA103Q104IV105Q106AL107LP108Q109IE110Q111Q112VI113TV114AG115DF116FF117TI118PD119QR120IR121ND122V123AN124IT125Q126AV127LP128Q129IE130Q131Q132VI133TV134AG135DF136FF137TI138PD139QR140IR141ND142V143AN144IT145Q146AV147LP148Q149IE150Q151Q152VI153TV154AG155DF156FF157TI158PD159QR160IR161ND162V163AN164IT165Q166AV167LP168Q169IE170Q171Q172VI173TV174AG175DF176FF177TI178PD179QR180IR181ND182V183AN184IT185Q186AV187LP188Q189IE190Q191Q192VI193TV194AG195DF196FF197TI198PD199QR200IR201ND202V203AN204IT205Q206AV207LP208Q209IE210Q211Q212VI213TV214AG215DF216FF217TI218PD219QR220IR221ND222V223AN224IT225Q226AV227LP228Q229IE230Q231Q232VI233TV234AG235DF236FF237TI238PD239QR240IR241ND242V243AN244IT245Q246AV247LP248Q249IE250Q251Q252VI253TV254AG255DF256FF257TI258PD259QR260IR261ND262V263AN264IT265Q266AV267LP268Q269IE270Q271Q272VI273TV274AG275DF276FF277TI278PD279QR280IR281ND282V283AN284IT285Q286AV287LP288Q289IE290Q291Q292VI293TV294AG295DF296FF297TI298PD299QR300IR301ND302V303AN304IT305Q306AV307LP308Q309IE310Q311Q312VI313TV314AG315DF316FF317TI318PD319QR320IR321ND322V323AN324IT325Q326AV327LP328Q329IE330Q331Q332VI333TV334AG335DF336FF337TI338PD339QR340IR341ND342V343AN344IT345Q346AV347LP348Q349IE350Q351Q352VI353TV354AG355DF356FF357TI358PD359QR360IR361ND362V363AN364IT365Q366AV367LP368Q369IE370Q371Q372VI373TV374AG375DF376FF377TI378PD379QR380IR381ND382V383AN384IT385Q386AV387LP388Q389IE390Q391Q392VI393TV394AG395DF396FF397TI398PD399QR400IR401ND402V403AN404IT405Q406AV407LP408Q409IE410Q411Q412VI413TV414AG415DF416FF417TI418PD419QR420IR421ND422V423AN424IT425Q426AV427LP428Q429IE430Q431Q432VI433TV434AG435DF436FF437TI438PD439QR440IR441ND442V443AN444IT445Q446AV447LP448Q449IE450Q451Q452VI453TV454AG455DF456FF457TI458PD459QR460IR461ND462V463AN464IT465Q466AV467LP468Q469IE470Q471Q472VI473TV474AG475DF476FF477TI478PD479QR480IR481ND482V483AN484IT485Q486AV487LP488Q489IE490Q491Q492VI493TV494AG495DF496FF497TI498PD499QR500IR501ND502V503AN504IT505Q506AV507LP508Q509IE510Q511Q512VI513TV514AG515DF516FF517TI518PD519QR520IR521ND522V523AN524IT525Q526AV527LP528Q529IE530Q531Q532VI533TV534AG535DF536FF537TI538PD539QR540IR541ND542V543AN544IT545Q546AV547LP548Q549IE550Q551Q552VI553TV554AG555DF556FF557TI558PD559QR560IR561ND562V563AN564IT565Q566AV567LP568Q569IE570Q571Q572VI573TV574AG575DF576FF577TI578PD579QR580IR581ND582V583AN584IT585Q586AV587LP588Q589IE590Q591Q592VI593TV594AG595DF596FF597TI598PD599QR600IR601ND602V603AN604IT605Q606AV607LP608Q609IE610Q611Q612VI613TV614AG615DF616FF617TI618PD619QR620IR621ND622V623AN624IT625Q626AV627LP628Q629IE630Q631Q632VI633TV634AG635DF636FF637TI638PD639QR640IR641ND642V643AN644IT645Q646AV647LP648Q649IE650Q651Q652VI653TV654AG655DF656FF657TI658PD659QR660IR661ND662V663AN664IT665Q666AV667LP668Q669IE670Q671Q672VI673TV674AG675DF676FF677TI678PD679QR680IR681ND682V683AN684IT685Q686AV687LP688Q689IE690Q691Q692VI693TV694AG695DF696FF697TI698PD699QR700IR701ND702V703AN704IT705Q706AV707LP708Q709IE710Q711Q712VI713TV714AG715DF716FF717TI718PD719QR720IR721ND722V723AN724IT725Q726AV727LP728Q729IE730Q731Q732VI733TV734AG735DF736FF737TI738PD739QR740IR741ND742V743AN744IT745Q746AV747LP748Q749IE750Q751Q752VI753TV754AG755DF756FF757TI758PD759QR760IR761ND762V763AN764IT765Q766AV767LP768Q769IE770Q771Q772VI773TV774AG775DF776FF777TI778PD779QR780IR781ND782V783AN784IT785Q786AV787LP788Q789IE790Q791Q792VI793TV794AG795DF796FF797TI798PD799QR800IR801ND802V803AN804IT805Q806AV807LP808Q809IE810Q811Q812VI813TV814AG815DF816FF817TI818PD819QR820IR821ND822V823AN824IT825Q826AV827LP828Q829IE830Q831Q832VI833TV834AG835DF836FF837TI838PD839QR840IR841ND842V843AN844IT845Q846AV847LP848Q849IE850Q851Q852VI853TV854AG855DF856FF857TI858PD859QR860IR861ND862V863AN864IT865Q866AV867LP868Q869IE870Q871Q872VI873TV874AG875DF876FF877TI878PD879QR880IR881ND882V883AN884IT885Q886AV887LP888Q889IE890Q891Q892VI893TV894AG895DF896FF897TI898PD899QR900IR901ND902V903AN904IT905Q906AV907LP908Q909IE910Q911Q912VI913TV914AG915DF916FF917TI918PD919QR920IR921ND922V923AN924IT925Q926AV927LP928Q929IE930Q931Q932VI933TV934AG935DF936FF937TI938PD939QR940IR941ND942V943AN944IT945Q946AV947LP948Q949IE950Q951Q952VI953TV954AG955DF956FF957TI958PD959QR960IR961ND962V963AN964IT965Q966AV967LP968Q969IE970Q971Q972VI973TV974AG975DF976FF977TI978PD979QR980IR981ND982V983AN984IT985Q986AV987LP988Q989IE990Q991Q992VI993TV994AG995DF996FF997TI998PD999QR1000IR1001ND1002V1003AN1004IT1005Q1006AV1007LP1008Q1009IE1010Q1011Q1012VI1013TV1014AG1015DF1016FF1017TI1018PD1019QR1020IR1021ND1022V1023AN1024IT1025Q1026AV1027LP1028Q1029IE1030Q1031Q1032VI1033TV1034AG1035DF1036FF1037TI1038PD1039QR1040IR1041ND1042V1043AN1044IT1045Q1046AV1047LP1048Q1049IE1050Q1051Q1052VI1053TV1054AG1055DF1056FF1057TI1058PD1059QR1060IR1061ND1062V1063AN1064IT1065Q1066AV1067LP1068Q1069IE1070Q1071Q1072VI1073TV1074AG1075DF1076FF1077TI1078PD1079QR1080IR1081ND1082V1083AN1084IT1085Q1086AV1087LP1088Q1089IE1090Q1091Q1092VI1093TV1094AG1095DF1096FF1097TI1098PD1099QR1100IR1101ND1102V1103AN1104IT1105Q1106AV1107LP1108Q1109IE1110Q1111Q1112VI1113TV1114AG1115DF1116FF1117TI1118PD1119QR1120IR1121ND1122V1123AN1124IT1125Q1126AV1127LP1128Q1129IE1130Q1131Q1132VI1133TV1134AG1135DF1136FF1137TI1138PD1139QR1140IR1141ND1142V1143AN1144IT1145Q1146AV1147LP1148Q1149IE1150Q1151Q1152VI1153TV1154AG1155DF1156FF1157TI1158PD1159QR1160IR1161ND1162V1163AN1164IT1165Q1166AV1167LP1168Q1169IE1170Q1171Q1172VI1173TV1174AG1175DF1176FF1177TI1178PD1179QR1180IR1181ND1182V1183AN1184IT1185Q1186AV1187LP1188Q1189IE1190Q1191Q1192VI1193TV1194AG1195DF1196FF1197TI1198PD1199QR1200IR1201ND1202V1203AN1204IT1205Q1206AV1207LP1208Q1209IE1210Q1211Q1212VI1213TV1214AG1215DF1216FF1217TI1218PD1219QR1220IR1221ND1222V1223AN1224IT1225Q1226AV1227LP1228Q1229IE1230Q1231Q1232VI1233TV1234AG1235DF1236FF1237TI1238PD1239QR1240IR1241ND1242V1243AN1244IT<
```



Fig. S4 PSIPRED-based secondary structure prediction and sequence alignment. A) PROMALS3D sequence alignment of arthropod totivirus-like virus (OmRV, DTV, AsTV, IMNV) major CPs (MCP) using PSIPRED-based secondary structure prediction<sup>3</sup>. The red and blue sequences indicate the  $\alpha$ -helices and  $\beta$ -sheets predicted by PSIPRED protein sequence analyses. The predicted consensus  $\alpha$ -helices (h) and  $\beta$ -sheets (e) are indicated. All arthropod totivirus-like viruses have a conserved C-terminal  $\alpha$ -helix-rich domain not found in fungal and protozoan totiviruses. B) PSIPRED-based secondary structure prediction of TVV1. The red and blue sequences indicate  $\alpha$ -helices and  $\beta$ -sheets, respectively, as predicted by PSIPRED protein sequence analyses. TVV1 does not have the C-terminal  $\alpha$ -helix-rich domain found in arthropod totivirus-like viruses. C) Superimposition of ScV-L-A capsid protein (green cartoon) on the segmented density of the TVV1 CP (grey surface). The ScV-L-A capsid forms an  $\alpha + \beta$  and a  $\beta$ -rich domain. The orientation of the TVV1 CP is same as those of the MCPs in Fig. 8A and 8B (right). A 90-degree rotated view (left). A yellow pentagon indicates a 5-fold axis.

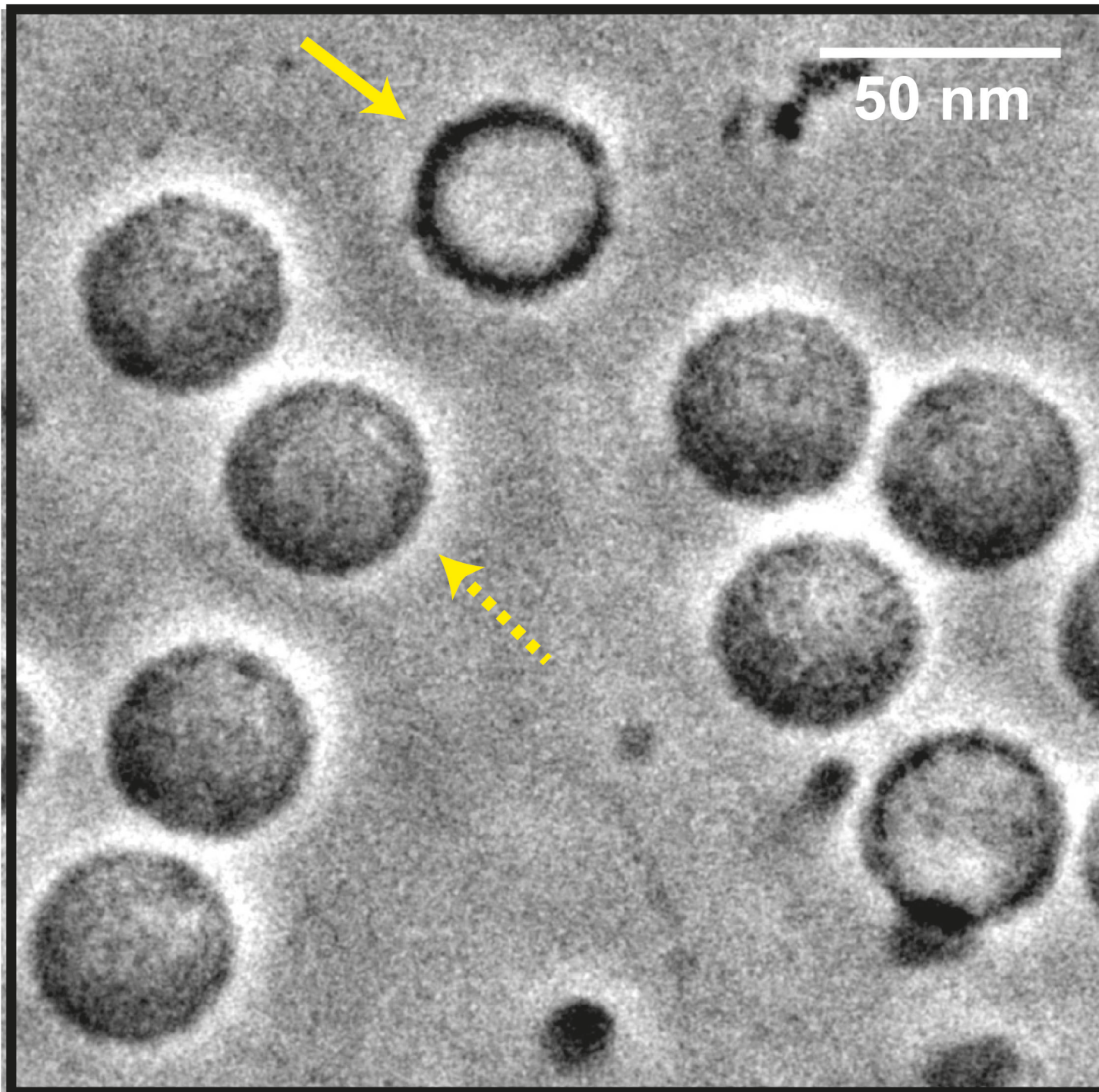


Fig. S5 Zernike phase contrast cryo-EM image of ice-embedded OmRV. Apart from being observed in focus using the Zernike phase plate<sup>4</sup>, the image was taken under the same conditions as the defocus contrast image (Fig. 2A). Solid and dotted arrows indicate empty and full virus particles, respectively.

## References

- 1 Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
- 2 Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* **5**, 976-989 (1994).
- 3 Pei, J. & Grishin, N. V. PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. *Methods in molecular biology* **1079**, 263-271 (2014).
- 4 Murata, K. *et al.* Zernike phase contrast cryo-electron microscopy and tomography for structure determination at nanometer and subnanometer resolutions. *Structure* **18**, 903-912 (2010).