Improving the visualization of viruses in soil

Amar D. Parvate[1], Trinidad Alfaro[1], Regan McDearis[1], Amy Zimmerman[1], Kirsten Hofmockel[1], William C. Nelson[1], James E. Evans[1,3]*

[1]Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99354, USA
[2]School of Biological Sciences, Washington State University, Pullman, WA 99354, USA
*Corresponding Author: James E. Evans (james.evans@pnnl.gov)

KEYWORDS: Soil sample, cryo-EM, virus diversity

ABSTRACT

Viruses are numerically the most abundant forms on Earth, and most are present in soil. Even though viruses are highly abundant in soil and critical to rhizosphere function, visualizing the diverse morphotypes within soil has been challenging. The difficulty is primarily due to the heterogenous nature of isolated suspensions that typically contain nanometer to micron scale debris which renders protein crystallography for structural studies unfeasible and hinders cryo-electron microscopy due to ice thickness and contrast issues. Here we employed and compared a simple spin filtration method to cleanup solutions of extracted viruses for direct observation with cryo-electron microscopy. Although relatively simple, the method employs common physical biochemical separation steps to remove large and small debris which dramatically improves image quality and preservation of structural features to permit visualizing morphotypes not typically seen with conventional negative stain approaches. In addition to tailed and non-tailed polyhedral phages, several under reported or novel morphotypes of soil viruses are directly visualized as a particle library with both 2D and 3D information.

INTRODUCTION

Viruses are numerically the most abundant forms on Earth with estimates of up to 10^{31} total virus particles worldwide [1, 2] and are being explored as alternatives to pesticides, antibiotics or as source of genes encoding molecules of biotechnological importance for commercialization purposes [3-5]. Interestingly, the majority of viruses are believed to be present in soil where they are estimated to range between 10^5 to 10^9/ g and outnumber bacteria by 5-1,000 fold [6]. Viruses are known to integrate themselves into complex soil ecosystems and interact with microbial communities [7]. In the case of agricultural land, multiple factors impact the soil virome including the use of fertilizers, the type of fertilizer, as well as type of crops grown [8].

Furthermore, the soil type, moisture content, season of sampling, and depth of sampling can all impact the number and types of viruses observed from sample to sample [9]. This has largely limited scientific study to viruses or phages which are known to either positively or negatively impact human activities, livestock, or crops, and to a limited number of viruses which have been successfully isolated and propagated through a host [10, 11]. As of 2023, soil viromes constituted <1% of the total viromes available in public databases, making them some of the least studied forms of life on the planet [12].

"Soil" is a deceptively simple term applied to an extremely complex layer of matter covering the entire planet which harbors its own ecosystems consisting of everything from viruses to unicellular microbes, fungi as well as the rhizosphere. Soil contains solid, liquid and gaseous phases with stoichiometries varying over time and space. For soil virome studies, typically soil in the top few inches is utilized [2]. Viruses regulate microbial communities in soil environments in cycles, seasonally and spatiotemporally and may influence soil nutrient cycles in lands under cultivation [2, 13]. Negative stain TEM is the dominant approach used in already documented studies highlighting the morphological diversity of soil viruses with nanoscale imaging. Major benefits for negative stain TEM are that it is cheap and easy to perform, the resulting heavy metal stained grids are stable at room temperature for years and can be reimaged multiple times, and in general this method requires ~10-100x less sample than other approaches like cryo-electron microscopy. One of the most comprehensive examples of high-quality negative stain TEM analysis of the soil virome was reported recently and focused on the discovery and characterization of giant virus-like particles in forest soil[14]. While they presented stunning images, the authors also acknowledged that their method of negative staining with heavy metal and high salt content in conjunction with a chemical fixation step and dehydration could perturb the viral ultrastructure resulting in shrinking and other artifacts. They also mention that "it is not possible to prove the viral nature of any particles analyzed only by negative staining TEM, regardless of how similar to published virus structures they may look," which is partly due to the inability of negative staining to capture the full structural details and whole 3D structure of large particles. This is a widely recognized limitation of negative staining which generally fails to preserve high resolution features beyond ~16 Angstrom for non-crystalline samples, fails to stain sugar or glycan surfaces, and fails to capture the full height and context of the virus components depending on stain depth and local area reproducibility. Several researchers have also reported that the method of extraction influences the virus morphotypes observed using negative stain electron microscopy [6]. The difference in extraction methods includes but is not limited to factors

like buffer composition, type of soil, and whether a de-clumping step has been utilized, as these can interfere with the uniformity of the negative stain.

Due to these historic challenges for direct morphological analysis of soil viromes in native-like and hydrated (not chemically fixed or dehydrated) conditions, many studies are currently exploring the soil virome with only genomic methods[15, 16]. However, both genomic and morphological characterization remain critical aspects for taxonomic classification of viral Families, and both are needed to fully understand the diversity of viruses present in soil [17]. Cryo-electron microscopy (cryo-EM) is a well-established technique for imaging biological samples in near-native suspension conditions (i.e. monodisperse suspensions without any use of fixatives or heavy metal salt stains) and allows the preservation of atomic details while permitting direct visualization of material up to 0.8-1 micrometer in thickness which is an ideal range for most virus particles. Here we present a sample preparation approach that quickly and simply converts native soil isolates into a suspension compatible with high-resolution cryo-EM workflows for visualizing the soil virome with 2D and 3D analyses. Using this workflow, we observed several different morphotypes of tail-less and tailed phages with symmetrical heads and tail components, as well as a diverse array of other morphotypes including round, elongated and polymorphic viruses, and some morphotypes not previously reported using direct visualization.

METHODS

Field Site Description

Soil was collected on 7 March 2023 from the Tall Wheatgrass Irrigation Field Trial at the Washington State University Irrigated Agriculture Research and Extension Center (WSU IAREC) in Prosser, WA. As described previously [18], the site has aridisol soils with low organic matter content and alkaline pH and was planted with Tall Wheatgrass in 2018. Irrigation treatments from 25% to 100% Field Water Capacity (FWC) have been ongoing since spring 2019. Soil was collected from a random location within plot 23, which is planted with the Alkar cultivar of Tall Wheatgrass and irrigated at 25% FWC, using a 5 cm diameter hammer corer. The soil from 5-15 cm depth was used in this study. Soil was transported to the laboratory on blue ice, where it was homogenized and sieved (<2 mm) prior to further sample processing.

Enrichment of virus-like particles from soil

Virus-sized particles were enriched from freshly collected soils (maintained at 4ºC after homogenization) following methods adapted from Santos-Medellín et al. [13]. For each sample, 20 g of soil was weighed into 50 mL tubes. Each tube received 18 mL of protein supplemented

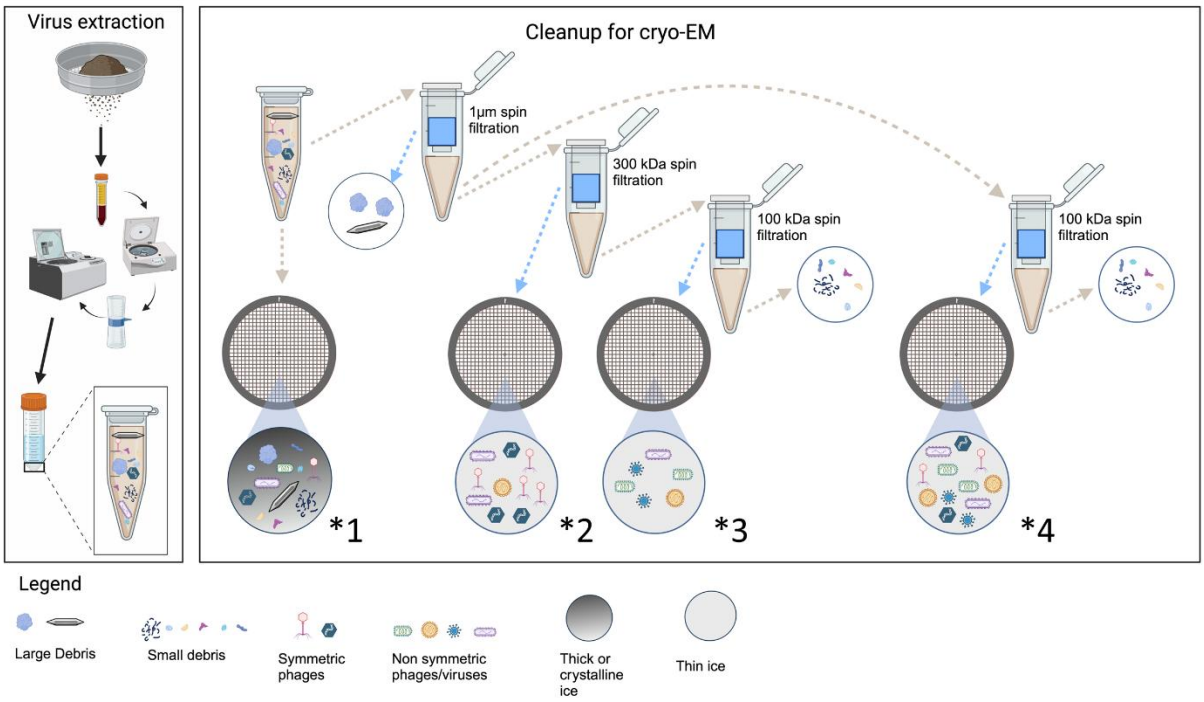
104 phosphate-buffered saline solution (PPBS: 2% bovine serum albumin, 10% phosphate-buffered
105 saline, 1% potassium citrate, and 150 mM MgSO4) and was mixed at 300 rpm for 10 minutes at
106 4°C. After mixing, tubes were centrifuged at 4,000 x g for 10 minutes at 4°C. The supernatant
107 was collected, and the soil pellet was resuspended in 18 mL PPBS, mixed as above, and
108 centrifuged. These steps were repeated once more. The pooled supernatant was then centrifuged
109 twice at 10,000 x g for 8 minutes at 4°C for removal of remaining soil particles. The supernatant
110 was then filtered through a 0.22 µm filter for removal of bacterial cells. To concentrate the virus-
111 like particles (VLP), the filtrate was ultracentrifuged in a 45 Ti rotor at 32,000 x g for 3 hours at
112 4°C, and the supernatant discarded. To resuspend the VLP pellet, 300 µL of cell culture grade
113 water was added directly to the surface of the pellet, incubated at 4°C for 1 hour at 150 rpm, and
114 gently resuspended via pipetting.
115
116 **Clean-up step before TEM imaging**
117 The resulting VLP suspensions were subsampled (40 µL) for direct visualization by cryo-EM. Each
118 suspension was diluted ~9-10X in TBD buffer (350-400 µL) and loaded on a Pleuriselect 1 µm
119 membrane filter (Cat # SKU 46-10001-15) to separate large micron size debris. Samples were
120 centrifuged and the filtrate was loaded on a 300 kDa MWCO Amicon spin concentrator (Millipore)
121 to further concentrate the retentate to ~30 µL by centrifugation. The filtrate was then loaded on a
122 100 kDa MWCO spin concentrator and centrifuged. The retentate from this step was concentrated
123 to 30 µL. Alternatively, the filtrate from the Pleuriselect filter was loaded directly onto a 100 kDa
124 MWCO filter and the retentate was concentrated to 30 µL. All centrifugation steps were performed
125 at 5000-10,000 x g at 4°C for 15 min.
126
127 **Grid preparation for TEM imaging**
128 Aliquots of soil samples before the Pleuriselect step, and retentates from the 300 and 100 kDa
129 MWCO filters were used for preparing cryo-EM grids. Briefly, 3 µL of sample was loaded onto
130 glow discharged holey carbon grids (Quantifoil, 300 mesh, R1/2). Samples were blotted for 3s
131 and plunge frozen in liquid ethane using a Leica EM GP2 to obtain vitrified ice for imaging. For
132 negative staining, 2 µL suspension from the 300 kDa MWCO retentate was loaded on holey
133 carbon grids and excess sample was blotted away. Staining was performed using Nano-W™ (Ted
134 Pella Inc.), following conventional procedures of negative stained sample preparation [18]
135
136 **TEM, cryo-EM and cryo-ET**

137 Samples were loaded on a 300 keV (Thermo Fisher Scientific) Titan Krios and all screening and
138 data collection was performed using standard EPU software. Images were collected either at
139 nominal magnifications ranging from 135x, 580x, 11500x for low dose screening and 42000x for
140 data collection using a K3 direct electron detector (Gatan Inc.) fitted with a Bioquantum energy
141 filter and a slit width of 20 eV. Dose was not optimized for low dose screening images, but
142 exposure was kept to 1 s. For data collection, movies were collected at a pixel size of 1.1 Å in
143 counting mode, using a defocus of -1 to -3 µm, at an exposure of 2-2.5 s exposure across 32
144 frames resulting in a total dose of 25-28 $e^-/Å^2/s$. Motion correction was performed using standard
145 parameters in CryoSparc v4.2 and micrographs were Fourier cropped to 0.25 x [19]. Motion
146 corrected micrographs were inspected and sorted manually in Fiji [20] to visualize different virus
147 morphotypes. For cryo-ET, tilt series were collected from +/- 60° in 3° increments with a total dose
148 of 100 $e^-/Å^2$ using a dose symmetric scheme. Each tilt was collected as a movie with 3 subframes,
149 and an exposure of 0.36s per tilt. Tilt series were collected at a defocus of – 4 µm. Motion
150 correction and restacking of the tilt series was performed using MotionCorr2 using scripts written
151 in house [21]. Individual tilt images were then [22]denoised using TOPAZ and tilt-series were
152 imported in IMOD 4.12.11, binned 4-8x and tomograms were reconstructed using standard
153 reconstruction workflows using the ETomo suite [23]. Missing wedge correction was performed
154 using Isonet [24]. and final visualization of the reconstructed volumes was done in 3dmod and
155 ChimeraX [25].
156
157 **RESULTS**
158 **Conventional negative stain TEM introduces artifacts and occludes some native details**
159 Using a soil sample from a managed agricultural field site in Prosser, WA which was
160 enriched for viruses using an approached optimized for downstream genomics analysis, we
161 performed Negative Stain TEM to get an initial image of the sample. While the typical fields of
162 view for Negative Stain TEM can allow researchers to pick and choose the best local area and
163 particle of interest, we chose random locations to provide more reproducible examples of best
164 imaging conditions. As seen in Figure S1, two random locations within the same ~50 micron x 50
165 micron grid square show vastly different backgrounds. While similar elongated symmetrical
166 particles with intact tails can be identified in both images with the arrowheads, the two particles
167 show a putatively intact virus with genomic content still inside the capsid (red arrowhead) and a
168 particle that clearly has stain inside the capsid indicating the DNA/RNA content was already
169 released or the particle itself is damaged and not intact (white arrowhead). Other particle
170 morphologies can be seen but the images also show aggregation, stain artifacts and particle

171 spatial overlap. For example, while a bottle-shaped phage is seen in **Figure S1A** (white asterisk)
172 the stain prevented observing any high-resolution features like surface glycoproteins which are
173 expected to be present on that morphotype[26, 27].
174

175 **Filtration improves cryo-EM imaging of viruses**

176 Suspensions of native viruses isolated from soil generally have low abundance compared
177 to the amount of material needed for protein crystallography and they contain nanometer to
178 micron scale debris which complicate cryo-EM. Debris that is usually not a problem in
179 conventional negative stain electron microscopy or multi-omics becomes a hindrance in cryo-EM
180 for observing any viruses or virus like particles (VLPs) as the largest structures tend to set the
181 local ice thickness gradients (**Figure 1, grid *1**).



182

183 **Figure 1 Workflow cartoon of methods to prepare soil samples for cryo-EM.** Virus extraction
184 - Soil samples were processed per Santos-Medellin et al to extract viruses. Clean-up for Cryo-
185 EM – Suspensions of viruses natively isolated from soil contain viruses as well as nanometer to
186 micrometer size debris. Soil suspensions frozen on grids as-is and imaged with cryo-EM
187 contained lots of debris and thick ice which hindered visualization of virus particles (*1). However,
188 soil suspensions passed through a 1 µm pore size Pleuriselect filter got rid of larger debris in the
189 retentate and allowed the filtrate containing the viruses and smaller debris to be subsequently
190 further concentrated and fractionated using 2 separate approaches. In the first approach,
191 sequential filtering steps help separate different size viruses. Here the filtrate from the Pleuriselect
192 filter can be concentrated through a 300 kDa MWCO filter to concentrate viruses and get rid of
193 nanometer sized debris. This sample resulted in thin ice and facilitated visualization of viruses.

Cryo-EM images collected on the retentate (*2) capture most of the virus diversity. Filtrate from the 300 kDa MWCO filter can be further processed through a 100 kDa MWCO filter. Imaging the retentate (*3) showed fewer total viruses that were mostly smaller virus like particles, while imaging the filtrate showed mostly junk protein. If desired the sequential approach allows for visualizing the differently sized viruses. Alternatively filtrate from the Pleuriselect membrane can be directly loaded on a 100 kDa MWCO filter. The retentate (*4) from this approach captures the entire mixture of small and large viruses while and most small protein debris is removed to provide quality thin ice to directly visualize the viral diversity. (Created using BioRender.com).

Using direct freezing of the as-delivered initial suspension, we observed thick ice and the presence of large aggregates and mineral debris despite the original preparation for viral enrichment including filtration through a 0.22 µm pore size membrane filter (**Figure 2** and **S2**). We therefore employed a cleanup step to get rid of debris and aggregates that might occlude or hinder visualization of viruses and VLPs in cryo-EM imaging. In principle, cryo-EM imaging with a phase plate and energy filter can image specimens up to 1 micron in thickness. Our cleanup step therefore consisted of centrifuging the suspension through a 1 µm pore size membrane Pleuriselect filter to separate out the larger micron-sized debris. We then subject the cleaned filtrate to simultaneous concentration and size selection filtration using 300 or 100 kDa MWCO spin columns to further improve the sampling efficiency and image quality.

Following this simple cleanup, we observed that at lower magnifications (135x-580x) the atlases and montages appeared roughly similar between the as-delivered and final cleaned up samples **(Figure S2 A-C, S3A-C)**. However, images and montages collected at 11,500 x **(Figure S2D, S3D)** showed clear differences where the as-delivered sample depicted high amounts of debris which was absent from the cleaned samples **(Figure S2E, S3E).**

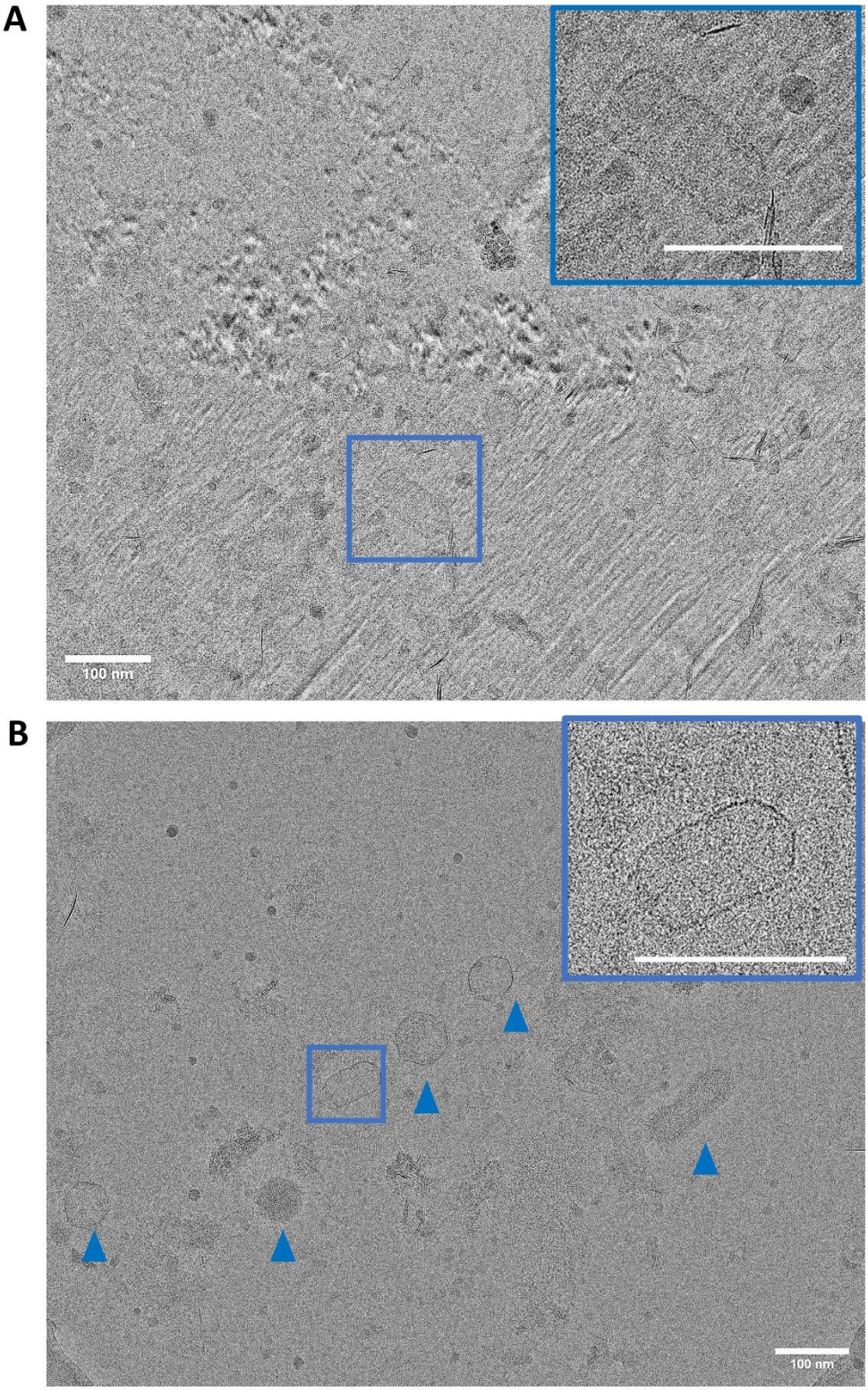


Figure 2 – Ice quality comparison of grids frozen before and after the cleanup step. (A) A cryo-EM image at 42000 x nominal magnification showing overall devitrified crystalline ice. The

221 inset is a zoomed in view of a phage (blue boundary) which shows the crystalline ice hindering
222 visualization of the phage morphology. (B) A cryo-EM image from the sample post cleanup which
223 shows vitreous ice that allows visualization of multiple phage and virus particles (blue triangles).
224 The inset is zoomed in view of a phage in blue boundary showing glycoprotein spikes on an
225 unidentified elongated virus. Scale bar = 100 nm
226

227 Any protein macromolecules like viruses frozen for cryo-EM should be ideally present in
228 a layer of uniform vitreous ice only slightly thicker than the target macromolecule to maximize
229 contrast and resolution. To compare ice quality on the grids with virus suspensions frozen as-
230 delivered and after filtration, we collected images at 42,000x where the presence of ice
231 crystallization artifacts from uneven vitrification would be obvious. Crystalline ice has unit cell with
232 dimensions around 3.4 Angstroms which results in prominent spots visible in the Fast Fourier
233 Transform of the image or in severe cases the lattice is visible in the direct image. Due to this,
234 images which have crystalline ice cannot be used for high resolution structural studies. As
235 expected, images of the as-delivered virus suspension showed aggregates, debris as well as
236 devitrified ice **(Figure 2A)** with multiple microdomains showing tell-tale signs of crystalline ice.
237 While a bacilliform phage like particle was visible in the field of view, no high-resolution features
238 could be clearly seen even in magnified views **(Figure 2A inset)**. However, after filtration and
239 removal of the larger debris, the cryo-EM image (**Figure 2B**) of the virus suspension showed
240 uniform vitrified ice (confirmed by absence of any detectable lattice in the direct image nor any
241 diffraction spots in the FFT of the image, data not shown)**.** The vitreous ice, with minimal debris
242 from particle aggregation, allowed visualization of several phages and one elongated virus in the
243 field of view. A zoomed in view of the virus particle **(Figure 2B, inset)** showed similar morphology
244 as that in Figure 2A but here the vitreous ice allowed better visualization of the capsid and other
245 features.

246 To estimate how many viruses may have escaped observation in the 300 kDa MWCO
247 retentate, we further concentrated the 300 kDa MWCO filtrate over a 100 kDa MWCO spin
248 concentrator and imaged the resulting retentate (Figure 1, *3). Unsurprisingly, we mostly
249 observed vlps with diameters between 20-50 nm and an occasional virus. By quantifying the
250 number of observed phages seen in roughly equivalent datasets of around 1100 captured
251 micrographs from each of the 300 kDa MWCO and 100 kDa MWCO retentates, we concluded
252 that the 300 kDa spin concentration step is efficient to capture most (>90%) of the viral diversity
253 from our current sample and the 100 kDa spin concentration post 300 kDa spin concentration
254 may not be absolutely necessary beyond getting rid of small debris that can obscure high-
255 resolution imaging of smaller viruses. Thus, if the main goal is to evaluate the diversity of the

256 viromes, a combination of Pleuriselect 1 micrometer filtration followed directly by 100 kDa MWCO
257 filtration (Figure 1, *4) provides the best option for capturing all viruses and vlps present in solution
258 while removing both large and small debris that hinder high-quality imaging.
259
260 **Direct visualization of viral diversity in soil extracts**
261 While previous negative stain TEM studies [28-32] of soil extracts and high-resolution
262 cryo-EM structural characterization of purified viruses have provided detailed classification of
263 symmetric viruses, we decided to classify any detected viruses based on known and unknown
264 morphotypes to fully capture the diversity. The most readily identifiable morphology consisted of
265 tail-less polyhedral phages and phages with tails of varying length **(Figure 3A-B)**. Figure 3B.1
266 shows a bacilliform phage while 3B.2 shows a myovirus-like phage with relatively thicker tail, and
267 Figure 3B.3 a podovirus-like phage with a very short tail. Figure 3B.5 shows a siphovirus-like
268 phage where the head, tail and tail fibers are clearly visible. Our workflow also clearly enables
269 visualization of features like contractile and non-contractile tails **(Figure 3B.2, 3B.1,5)** as well as
270 glycoproteins on several symmetrical phages **(Figure 3B.4).** Diameters for the icosahedral phage
271 heads ranged from 50-80 nm whereas the bacilliform phages had 100-120 nm long heads with
272 40 nm diameter. Several images showed the presence of long thin tube-like structures with 10-
273 20 nm diameter and of varying lengths (200 nm – 1 µm) which could not be visually assigned to
274 any group of viruses **(Figure 3C.1,2, Figure S4 C)** and could be remnants of [31]broken tails or
275 contractile sheaths or something else currently undetermined. Another major morphotype
276 observed had an overall round shape, diameter <50 nm and no visible glycoproteins on the
277 surfaces **(Figure 3C.3,4,5, Figure S5 C)**.
278 In addition to the tailed and polyhedral viruses which are readily seen with negative stain
279 TEM and above with the cryo-EM workflow described here, we also identified several virus
280 morphotypes which have only be occasionally reported [28] and some novel ones which have
281 never been directly visualized before. These novel morphotypes could belong to any of the
282 multiple genera reported from soil viromes which currently have no morphology associated with
283 them. These virus morphotypes were broadly divided into groups based on either shape, size or
284 other visual similarities **(Figure 4)**. The presence of an intact membrane and visible glycoprotein
285 spikes of any shape and size were considered as pre-requisites for a visual “feature” observed in
286 an image to be counted as a virus. Round viruses with diameters between 50-200 nm and visually
287 different surface glycoprotein spikes made up the numerically second largest group (**Figure 4.1).**
288 A diverse group of morphotypes showed the presence of a very thick coat surrounding the viral
289 capsid/core **(Figure 4.2-6)**. The sizes of these viruses ranged from 20-157 nm with coat thickness

290 from 15-35 nm. We observed an inovirus like morphotype (caterpillar like) that had a cylindrical
291 central core dotted with putative glycoprotein spikes orthogonal to the axial virus core **(Figure**
292 **4.7,8).** Bottle shaped viruses with diverse morphologies **(Figure 4.9, 4.10**) and filamentous
293 viruses **(Figure 4.11)** were also occasionally observed. Interestingly, the images of bottle shaped
294 viruses by cryo-EM clearly show a glycoprotein-like coat that is not visualized in a similar negative
295 stain image (compare Figure 4.10 with Figure S1A white asterisk). Another group of viruses with
296 diameters ranging from 30 to 85 nm showed the presence of fiber like glycoprotein like projections
297 between 31-47 nm in length (here called LG viruses) **(Figure 4.12)**. The LG viruses are visually
298 similar to morphotypes described by Fischer *et al* as viruses with capsid fibers[14]. Of note, the
299 size/dimension of virus morphotypes did not have a linear correlation with either the measured
300 thickness of the external coat or length of the glycoprotein projections. Some polymorphic viruses
301 with irregular shapes were also observed which have not been reported before in literature
302 **(Figure 4.13,14).** While Figures 3 and 4 document the library of 2D morphotypes identified by this
303 imaging approach, we also summarize the results in Figure 5 to illustrate the diversity more
304 graphically.

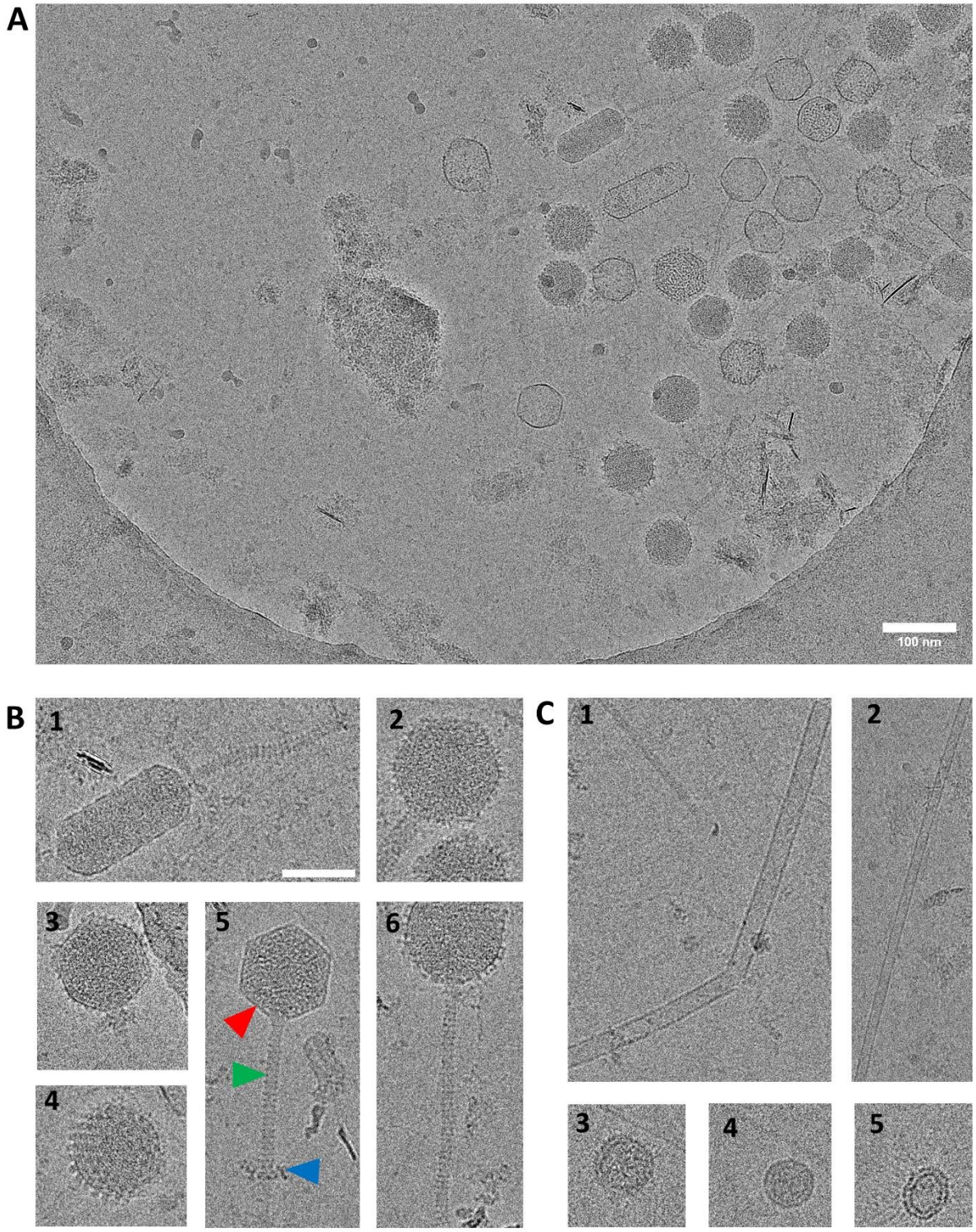


Figure 3 Cryo-EM images of symmetric viruses in filtrate of the sample 100 kDa MWCO filter – A) Cryo-EM image showing multiple phages with symmetric shapes. Tailed and tail-less phages along with other icosahedral viruses. Phages display both icosahedral and bullet shaped heads. B) Zoomed in view of 3 select phages from (A) shown. 1. A bacilliform phage showing a

long non contractile tail and siphovirus like morphology. 2. Phage showing a thick contractile tail having myovirus like morphology. 3. Podovirus like phage showing an extremely short tail (green triangle) (Phages in 1,2,4 are zoomed in views of phages from 3A). 4) Tail-less phage with glycoproteins visible on the virus surface. 5. Phage showing intact head (red triangle), contractile tail (green triangle) as well as tail fibers (blue triangle). 6. Phage showing a slightly rounded head relative to the phage shown in 3A.6. C) 1 and 2 are tube like VLPs of differing diameters abundantly observed in images. 3, 4, and 5 are VLPs which are <50 nm in diameter and have no visible glycoproteins on the surface. Scale bar = 50 nm for B and C.

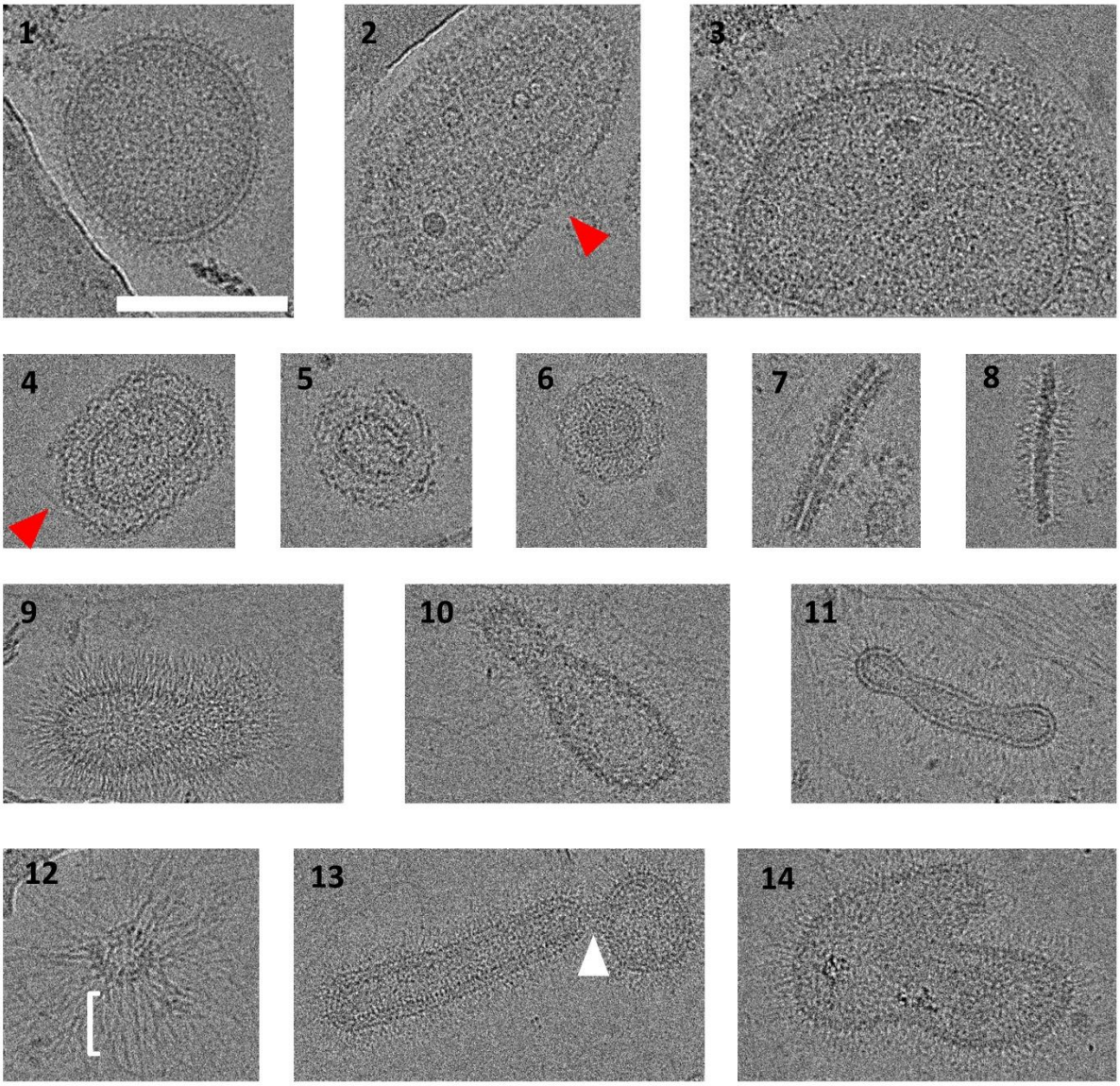
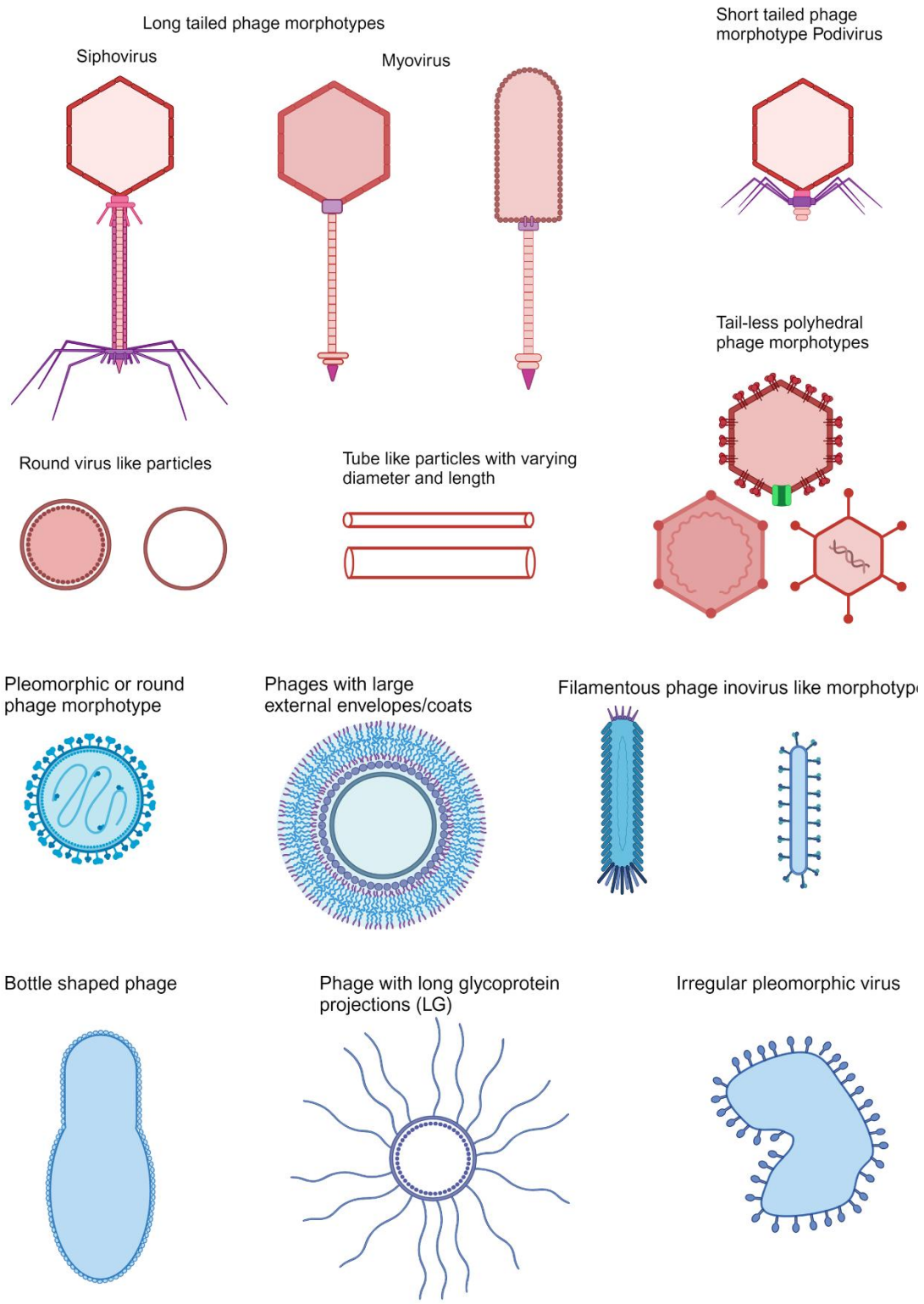


Figure 4- Cryo-EM images of polymorphic viruses in filtrate of the sample 100 kDa MWCO filter. 1) Example of a round phage. 2-6) Diverse uncharacterized phages displaying a wide range of diameters and a thick outer coating/layer external to the virus core (red triangle). 7-8) Caterpillar like appearance of a virus with long thin central core and glycoproteins on the surface putatively belonging to the Inoviridae Family. 9-10) Tube shaped viruses with diverse morphologies 11) Elongated or filamentous phage. 12) Phage showing unique long projections which could be

325 glycoproteins or glycans (LG, red arrows) with the white parenthesis indicating the length of the
326 glycoproteins. 12) A polymorphic vlp showing a head and body morphology on either side of a
327 narrow neck (white triangle) 13) Polymorphic phage having plasmavirus like irregular morphology.
328 Scale bar 100 nm.

329

330 One of the main benefits of cryo-EM, besides imaging samples in near-native conditions
331 without crystallization, is that it can be used to both produce a 2D library of images as above and
332 obtain 3D volumetric information using tomography. Since the polymorphic viruses have not been
333 widely studied and their overall shape can be hard to recognize in 3D from a single 2D projection,
334 we performed cryo-electron tomography on a small-subset of the viruses **(Figures S4 and S5)**.
335 This approach allows tracing the contiguous and intact membranes for the polymorphic phages
336 as well as detecting glycoprotein extensions and in some cases density inside the viruses that
337 could correspond to packed DNA/RNA.



338

339 ***Figure 5 - Cartoon representation of virus morphotypes in current study.*** *The brown/red*
340 *morphotypes have been reported in previous studies and were also observed in the current*

341 *cryo-EM analyses. The blue morphotypes are mostly uncharacterized, diverse phages some of*
342 *which have not yet been reported in literature.* (Created using Biorender.com)

343

344 **Discussion**

345 Negative staining and conventional TEM has been widely used to visualize viruses and
346 VLPs from soil environments, but a number of technical and operational problems have plagued
347 this technique and hindered viral research. While negative stained TEM is a quick and easy way
348 to visualize viruses, a recent report delved into the gradual decline in quality of electron
349 microscopy images seen over the years by evaluating hundreds of publications where the loss
350 was largely ascribed to a loss of expertise and training [33]. Additionally, previous methods for
351 extracting viruses from soil have sometimes been too harsh or biased towards tailed phages [34].
352 Most studies of TEM imaging of tailed viruses isolated from soil have reported broken tails or loss
353 of tails, where the frequency of loss of tail is inversely proportional to the tail length [6, 35]. Here
354 we have demonstrated an alternate technique for visualizing viruses from challenging samples
355 that employs highly selective but gentle filtration and cryo-EM to mitigate or eliminate many of the
356 problems that plague negative stain TEM images.

357 So far, cryo-EM of soil viruses has been restricted to studies where a single virus isolated
358 from soil sample can be propagated and purified in quantities amenable for cryo-EM single
359 particle studies [36]. Here we overcome this challenge/bottleneck and demonstrate the ability to
360 visualize the native viral community isolated directly from soil using cryo-EM by using common
361 and simple physical separation techniques. We used microcentrifuge spin columns to
362 simultaneously remove unwanted debris and to concentrate the sample to increase likelihood of
363 detection on the grid which overcame two of the main bottlenecks. The resulting grids were then
364 imaged using semi-automated cryo-EM screening and data collection where we able to detect at
365 least one virus or VLP in ~70% of the images collected on the 300 kDa MWCO retentate. Since
366 most of the virus suspension loaded on a grid gets blotted away prior to vitrification, the virus
367 library we have obtained **(Figure 5)** cannot yet be truly representative and it is possible that
368 several morphotypes could not be imaged. However, this drawback also extends to libraries
369 obtained using negative stained TEM and our work sets the foundation for being able to visualize
370 the viral diversity with cryo-EM while future work can focus on optimizing the loading and sampling
371 efficiency to capture more of the diversity. For example, multiblot methods are known to increase
372 the local concentration of particles trapped in cryo-EM grid holes even when using wicking, while
373 lossless sample deposition and vitrification can be accomplished using the Vitrojet [37] and other
374 similar preparative instruments.

375 Another way to validate the diversity is linking the observed cryo-EM 2D and 3D data with
376 omics analysis on the same samples. While beyond the focus of the current study, orthogonal
377 multi-omics would be needed to provide more information about the morphotypes imaged in this
378 study to confirm that a novel virus morphotype is truly an infectious virus. For example, gene
379 transfer agents (GTAs) appear and function like phages but carry fragments of cellular genome
380 resulting in horizontal gene transfer. It would be very challenging to distinguish between a
381 bacteriophage and a GTA purely based on 2D images [38]. Additionally, some of the round virus
382 morphotypes we have observed may be a result of cellular blebbing. Without further orthogonal
383 studies, it is not possible to unambiguously assign all phage morphotypes to the Class
384 *Caudoviricetes* [17].
385 Given estimates of the diversity of viruses in soil, there is ample opportunity for this work
386 to be replicated on soil samples from numerous diverse locations, during different times of the
387 year to obtain a broader representation of viruses resident in soil. Further, for higher resolution
388 studies, soil virus extractions could be processed using the virus reduction approaches used in
389 aquatic virology to selectively propagate a few viruses in high copy number [39, 40]. Complete
390 purification and isolation of these viruses would not be needed; their selective amplification would
391 provide suitable samples for high-resolution cryo-EM/ET studies using sub-tomogram averaging,
392 while also allowing verification of host-virus infectivity. Hypothetically, a dataset with just a few
393 hundred symmetric virus particles could be used to obtain 2D classes and sub-nanometer
394 resolution 3D maps. The selectively amplified viruses could be used to infect target host cells and
395 samples can be frozen at various stages of infection to perform cryo-ET, which will provide
396 structural details of interaction of host-virus systems from soil in unprecedented resolution.
397
398
399 **Acknowledgement and Funding:**
400 The research was performed using EMSL (grid.436923.9), a DOE Office of Science User Facility
401 sponsored by the Biological and Environmental Research program located at PNNL. A portion of
402 this research was performed on project award (10.46936/fics.proj.2022.60449/60008585) under
403 the FICUS program and used resources at the DOE Joint Genome Institute and the
404 Environmental Molecular Sciences Laboratory, which are DOE Office of Science User Facilities.
405 Both facilities are sponsored by the Biological and Environmental Research program and
406 operated under Contract Nos. DE-AC02-05CH11231 (JGI) and DE-AC05-76RL01830 (EMSL). A
407 portion of this research was supported by the DOE Office of Biological and Environmental
408 Research FWP 74915 and D7G33.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

JEE, ADP conceived this project. ADP, RWM, TA, AEZ performed all the experiments and ADP performed all the data analyses. JEE, WCN and KSH provided supervision and research support (funding). All the authors contributed to drafting and editing the manuscript.

References

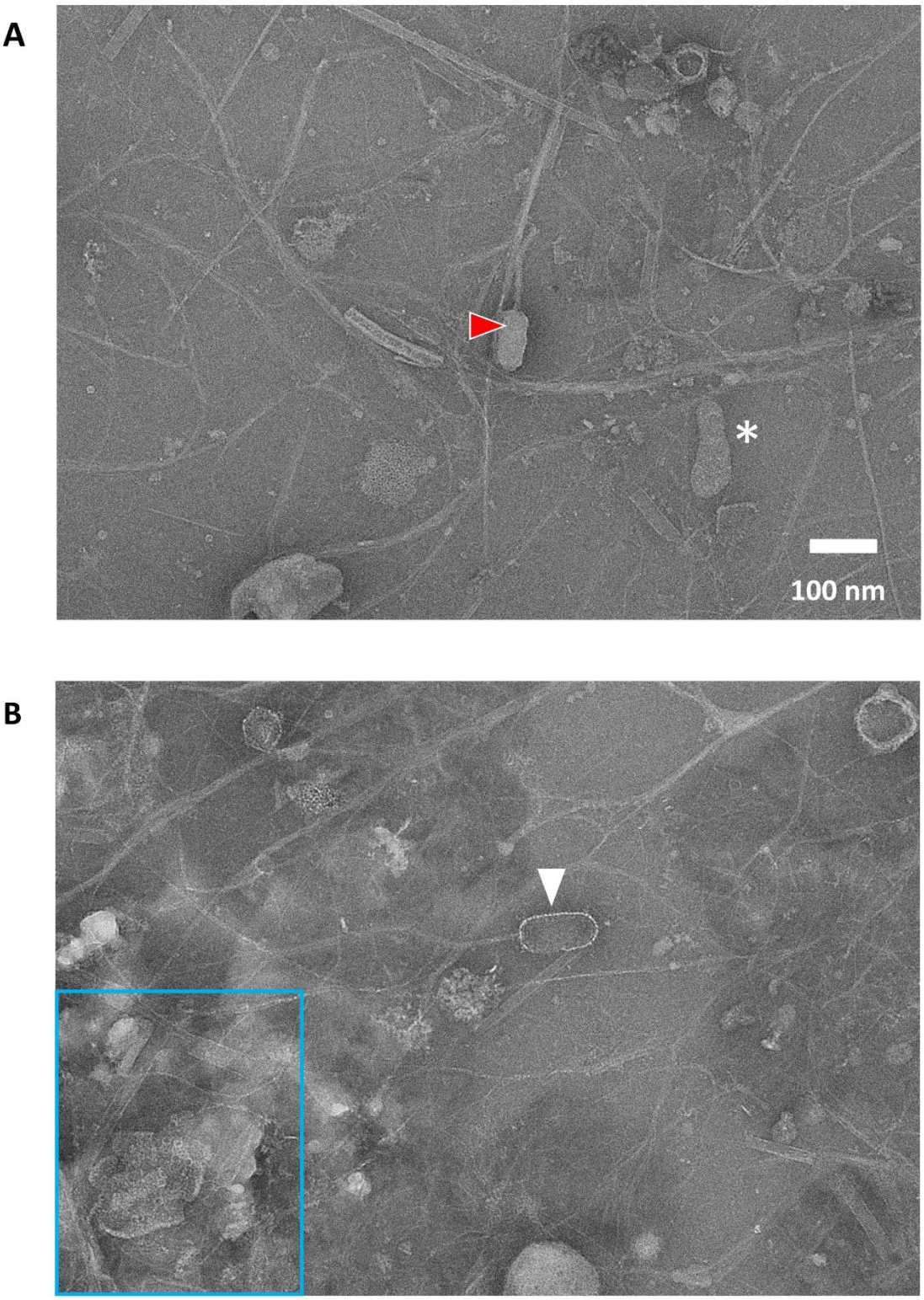
1. Bar-On, Y.M., R. Phillips, and R. Milo, *The biomass distribution on Earth.* Proc Natl Acad Sci U S A, 2018. **115**(25): p. 6506-6511.
2. Williamson, K.E., et al., *Viruses in Soil Ecosystems: An Unknown Quantity Within an Unexplored Territory.* Annu Rev Virol, 2017. **4**(1): p. 201-219.
3. Duan, Y., et al., *Bacteriophage targeting of gut bacterium attenuates alcoholic liver disease.* Nature, 2019. **575**(7783): p. 505-511.
4. Adriaenssens, E.M., et al., *T4-related bacteriophage LIMEstone isolates for the control of soft rot on potato caused by 'Dickeya solani'.* PLoS One, 2012. **7**(3): p. e33227.
5. Jansson, J.K. and R. Wu, *Soil viral diversity, ecology and climate change.* Nat Rev Microbiol, 2023. **21**(5): p. 296-311.
6. Williamson, K.E., et al., *Estimates of viral abundance in soils are strongly influenced by extraction and enumeration methods.* Biology and Fertility of Soils, 2013. **49**(7): p. 857-869.
7. Santos-Medellín, C., et al., *Viral but not bacterial community succession is characterized by extreme turnover shortly after rewetting dry soils.* bioRxiv, 2023.
8. Chen, L., et al., *Effect of different long-term fertilization regimes on the viral community in an agricultural soil of Southern China.* European Journal of Soil Biology, 2014. **62**: p. 121-126.
9. Narr, A., et al., *Soil Viral Communities Vary Temporally and along a Land Use Transect as Revealed by Virus-Like Particle Counting and a Modified Community Fingerprinting Approach (fRAPD).* Front Microbiol, 2017. **8**: p. 1975.
10. Noteborn, W.E.M., et al., *Using cryogenic electron microscopy methods to gain insight into structure and initial host attachment of the flagellotropic bacteriophage 7-7-1.* bioRxiv, 2025.
11. Almeida, G.M.d.F., et al., *Genomic and structural insights into Jyvaskylavirus, the first giant virus isolated from Finland.* eLife, 2025. **13**.
12. Graham, E.B., et al., *A global atlas of soil viruses reveals unexplored biodiversity and potential biogeochemical impacts.* Nat Microbiol, 2024. **9**(7): p. 1873-1883.
13. Santos-Medellin, C., et al., *Spatial turnover of soil viral populations and genotypes overlain by cohesive responses to moisture in grasslands.* Proc Natl Acad Sci U S A, 2022. **119**(45): p. e2209132119.
14. Fischer, M.G., U. Mersdorf, and J.L. Blanchard, *Amazing structural diversity of giant virus-like particles in forest soil.* bioRxiv, 2023.
15. Hazard, C., et al., *Beneath the surface: Unsolved questions in soil virus ecology.* Soil Biology & Biochemistry, 2025. **205**.

16. Roux, S. and J.B. Emerson, *Diversity in the soil virosphere: to infinity and beyond?* Trends Microbiol, 2022. **30**(11): p. 1025-1035.
17. Turner, D., et al., *Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee.* Arch Virol, 2023. **168**(2): p. 74.
18. Naylor, D., et al., *Interactive effects of depth and differential irrigation on soil microbiome composition and functioning.* Frontiers in Microbiomes, 2023. **2**.
19. Punjani, A., et al., *cryoSPARC- algorithms for rapid unsupervised cryo-EM structure determination.* Nature Methods, 2017. **14**: p. 290-297.
20. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.
21. Zheng, S.Q., et al., *MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy.* Nat Methods, 2017. **14**(4): p. 331-332.
22. Bepler, T., et al., *Topaz-Denoise: general deep denoising models for cryoEM and cryoET.* Nat Commun, 2020. **11**(1): p. 5208.
23. Kremer, J.R., D.N. Mastronarde, and J.R. McIntosh, *Computer visualization of three-dimensional image data using IMOD.* J Struct Biol, 1996. **116**(1): p. 71-6.
24. Liu, Y.T., et al., *Isotropic reconstruction for electron tomography with deep learning.* Nat Commun, 2022. **13**(1): p. 6482.
25. Pettersen, E.F., et al., *UCSF ChimeraX: Structure visualization for researchers, educators, and developers.* Protein Sci, 2021. **30**(1): p. 70-82.
26. Prangishvili, D., P. Forterre, and R.A. Garrett, *Viruses of the Archaea: a unifying view.* Nat Rev Microbiol, 2006. **4**(11): p. 837-48.
27. Wang, F., et al., *Spindle-shaped archaeal viruses evolved from rod-shaped ancestors to package a larger genome.* Cell, 2022. **185**(8): p. 1297-1307 e11.
28. Ackermann, H.W., *5500 Phages examined in the electron microscope.* Arch Virol, 2007. **152**(2): p. 227-43.
29. Swanson, M.M., et al., *Viruses in soils: morphological diversity and abundance in the rhizosphere.* Annals of Applied Biology, 2009. **155**(1): p. 51-60.
30. Reavy, B., et al., *Distinct circular single-stranded DNA viruses exist in different soil types.* Appl Environ Microbiol, 2015. **81**(12): p. 3934-45.
31. Zheng, J., et al., *A Capsid Structure of Ralstonia solanacearum podoviridae GP4 with a Triangulation Number T = 9.* Viruses, 2022. **14**(11).
32. Effantin, G., et al., *Cryo-electron microscopy three-dimensional structure of the jumbo phage PhiRSL1 infecting the phytopathogen Ralstonia solanacearum.* Structure, 2013. **21**(2): p. 298-305.
33. Ackermann, H.W., *Sad State of Phage Electron Microscopy. Please Shoot the Messenger.* Microorganisms, 2013. **2**(1): p. 1-10.
34. Williamson, K.E., K.E. Wommack, and M. Radosevich, *Sampling natural viral communities from soil for culture-independent analyses.* Appl Environ Microbiol, 2003. **69**(11): p. 6628-33.
35. Williamson, K.E., R.R. Helton, and K.E. Wommack, *Bias in bacteriophage morphological classification by transmission electron microscopy due to breakage or loss of tail structures.* Microsc Res Tech, 2012. **75**(4): p. 452-7.
36. Okamoto, K., et al., *Structural variability and complexity of the giant Pithovirus sibericum particle revealed by high-voltage electron cryo-tomography and energy-filtered electron cryo-microscopy.* Sci Rep, 2017. **7**(1): p. 13291.
37. Ravelli, R.B.G., et al., *Cryo-EM structures from sub-nl volumes using pin-printing and jet vitrification.* Nat Commun, 2020. **11**(1): p. 2563.
38. Bardy, P., et al., *Structure and mechanism of DNA delivery of a gene transfer agent.* Nat Commun, 2020. **11**(1): p. 3034.

39. Wei, W., et al., *Reduced bacterial mortality and enhanced viral productivity during sinking in the ocean.* ISME J, 2022. **16**(6): p. 1668-1675.
40. Wilhelm, S.W., S.M. Brigden, and C.A. Suttle, *A dilution technique for the direct measurement of viral production: a comparison in stratified and tidally mixed coastal waters.* Microb Ecol, 2002. **43**(1): p. 168-73.

Supplementary Figures

514



515

Figure S1 Negative stained image of viruses natively isolated after the Pleuriselect step – Conventional negatively stained images of viruses collected at 42000 x nominal magnification. A) Example of a good negative staining where only the background has been stained and phage shape is clearly visible (white triangle) B) Example of an overstained image where stain seems to have penetrated inside the phage head (read triangle). The images show a lot of background junk and loss of high-resolution features. Area in blue boundary shows aggregates in the field of view which prevents identifying any VLPs due to spatial overlap.

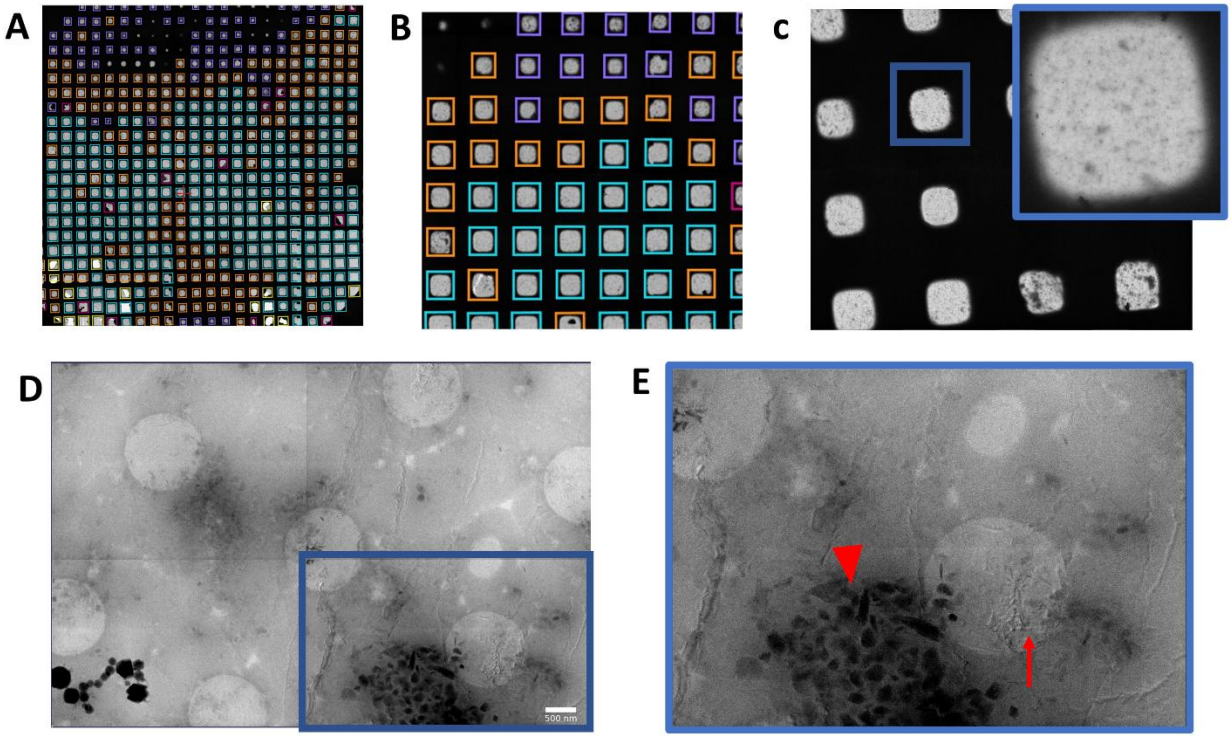


Figure S2 – Comparison of scales of imaging for virus samples before purification. A) A stitched atlas images collected at very low magnification showing the entire cryo-EM grid showing. Squares on the grid are colored from relatively thick (purple) to thin (blue) with the blue colored squares having ice thickness ideal for cryo-EM imaging to B) A zoomed in area of the atlas level magnification shows no visible difference in the grid with samples before or after purification. C) A higher mag view of the grid with an individual square (blue border and inset) that shows no visible difference in ice quality. D) A 2x2 montage of individual images collected at 11500x magnification for the sample before Pleuriselect cleanup. At this magnification, the ice begins to show irregularities. E) Zoomed in image of the individual image in the montage shows presence of debris (red triangle) as well as crystalline ice (red arrow). Scale bar for D and I – 500 nm. Each circular hole is 1 µm in diameter.

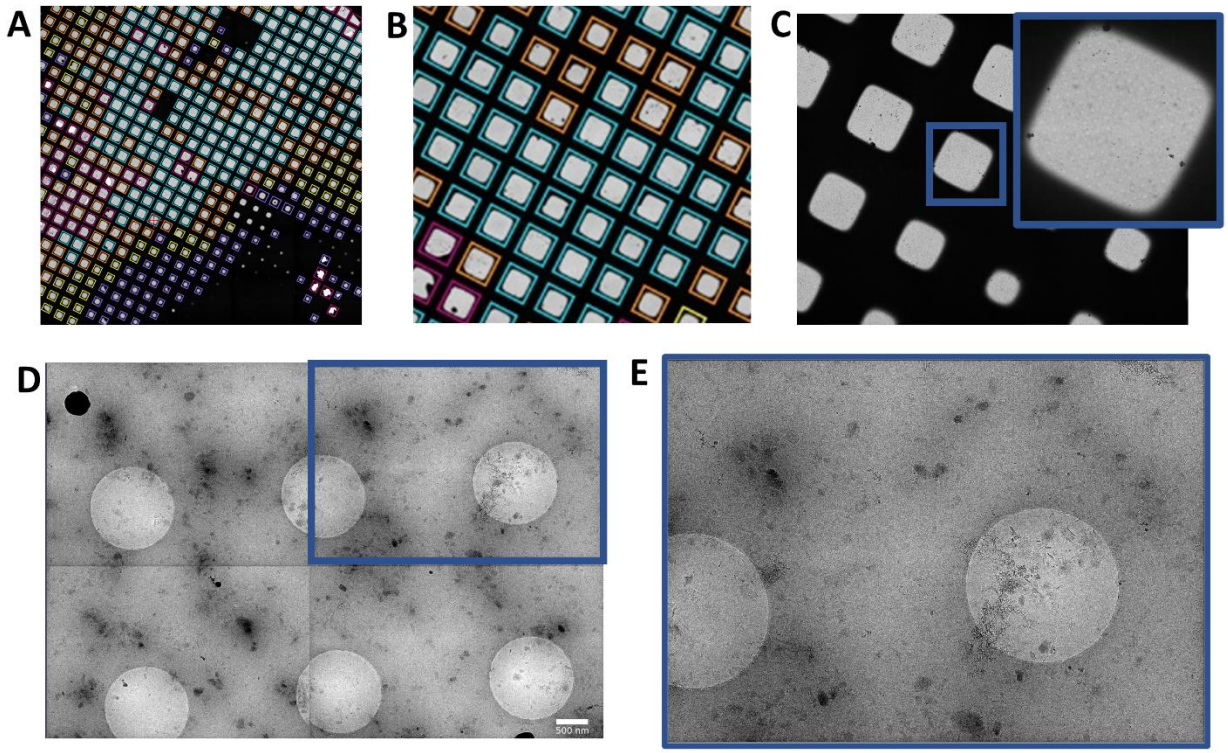


Figure S3 - Comparison of scales of imaging for virus samples after purification. A) stitched atlas images collected at very low magnification showing the entire cryo-EM grid showing. Squares on the grid are colored from relatively thick (purple) to thin (blue) with the blue colored squares having ice thickness ideal for cryo-EM imaging to B) A zoomed in area of the atlas level magnification shows no visible difference in the grid with samples before or after purification. C) A higher mag view of the grid with an individual square (blue border and inset) that shows no visible difference in ice quality. D) A 2x2 montage of individual images collected at 11500x magnification for the sample after Pleuriselect cleanup. E) Zoomed in image of the individual image in the montage shows no detectable. No visible large debris or crystalline ice seen in the montage or the zoomed in image. Scale bar for D and I – 500 nm. Each circular hole is 1 μm in diameter.

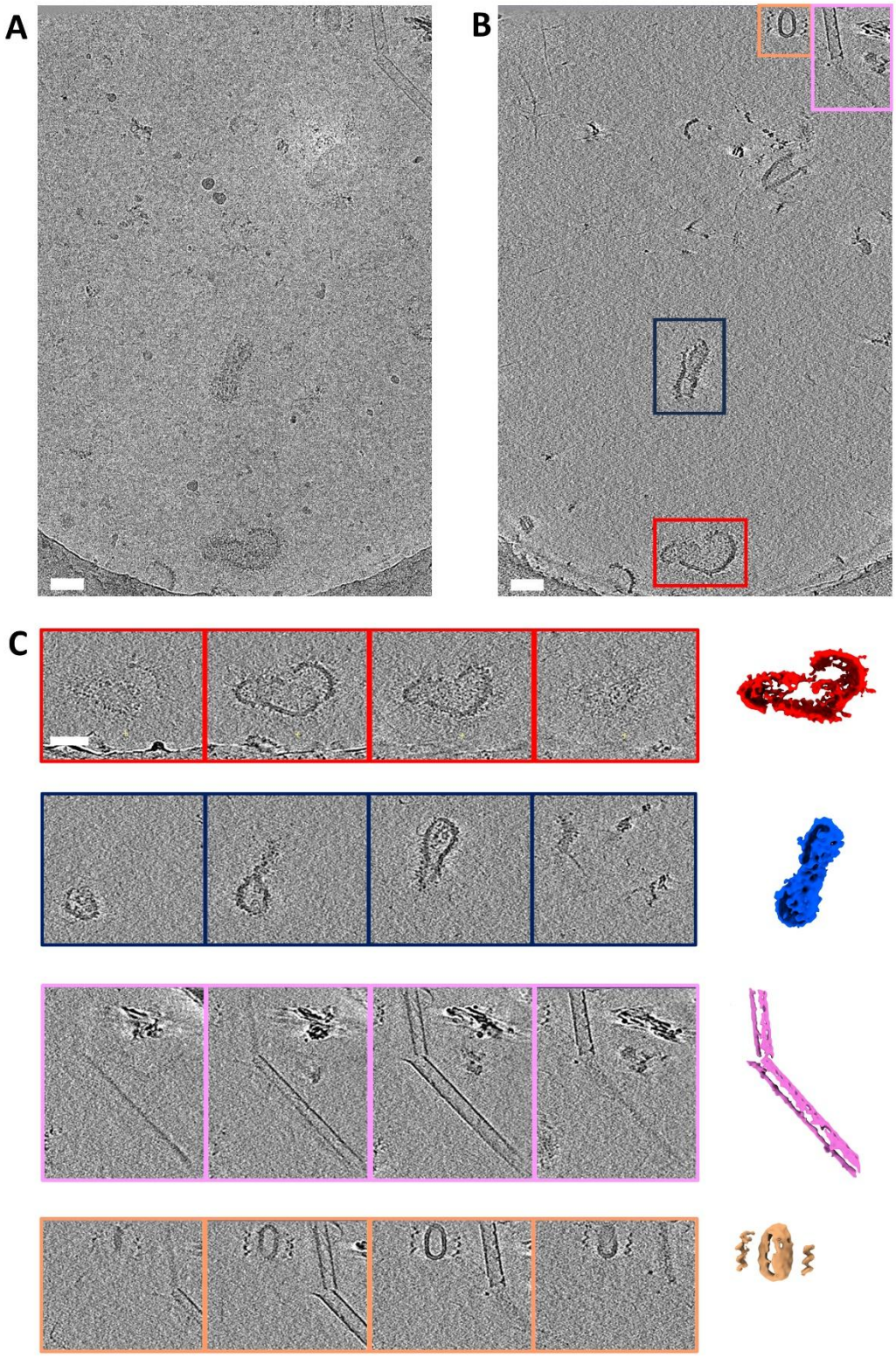


Figure S4 - Cryo-tomographic imaging of viruses natively isolated from soil. A)Zero-degree tilt image of the tomogram with diverse morphotypes. B) Slice through the reconstructed

551 tomogram with 3 morphotypes in in the red, blue and yellow boundary. C) Successive Z slices
552 through the tomogram for each of the individual morphotypes tracing appearance and
553 disappearance of the viruses/vlps. False color rendering of the virus/vlps are shown as visualized
554 using ChimeraX. Red – a comma or bean shaped vlp showing intact membranes and glycoprotein
555 spikes visible on its surface. Blue – An elongated virus morphotype with blebs internal to its
556 membrane and glycoprotein spikes on the surface. Pink – A tube like morphotype observed with
557 uncharacterized material internal to the tube. Beige – Morphotype with a thick coat surrounding
558 the central core. Slices in pink and beige boundaries have overlapping regions. All scale bars =
559 25 nm

560

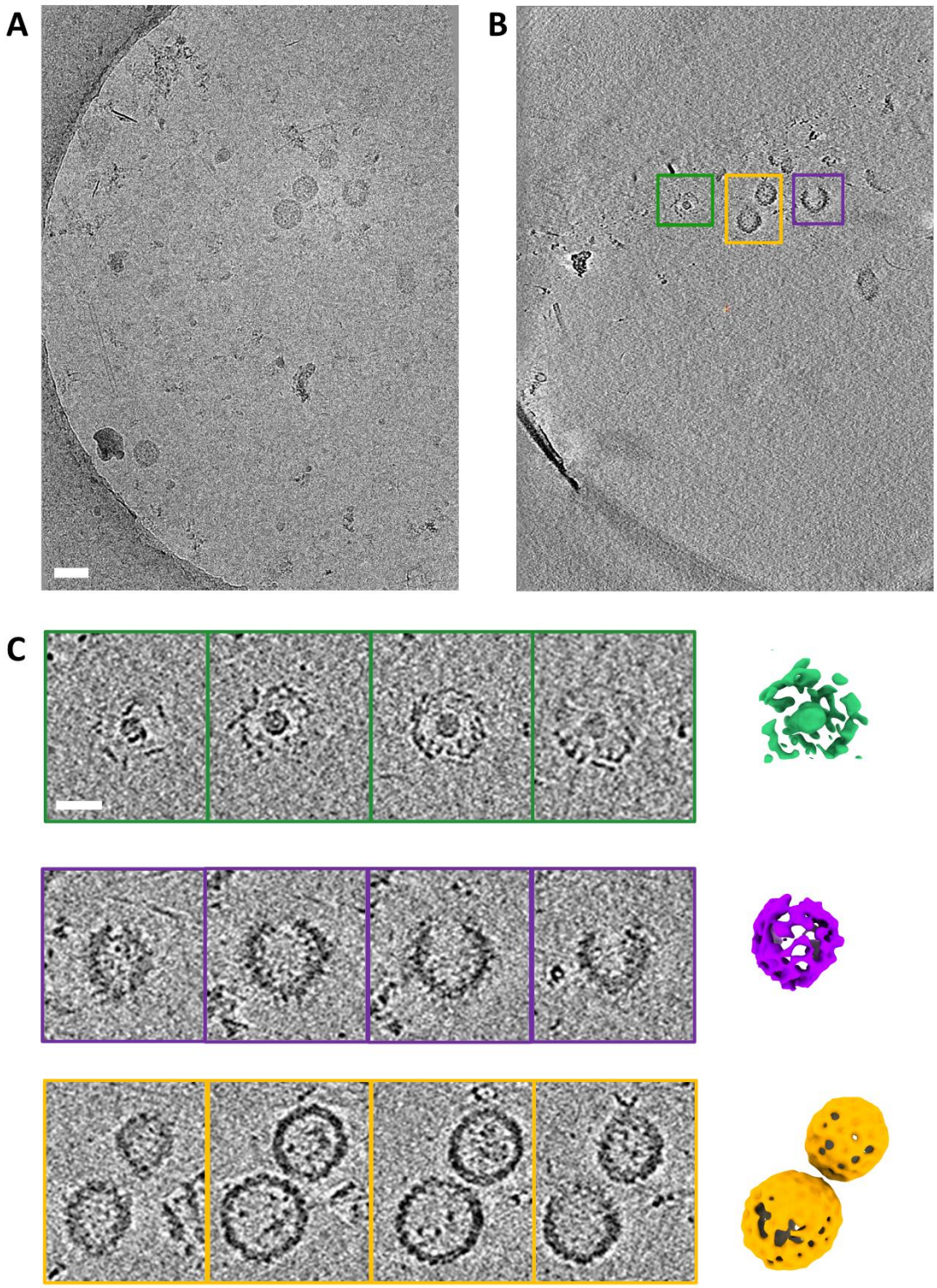


Figure S5 Cryo-tomographic imaging of vlps natively isolated from soil. . A)Zero-degree tilt image of a tomogram with diverse morphotypes/vlps B) Slice through the reconstructed tomogram with 3 morphotypes in in the green, violet and gold. C) Successive Z slices through the tomogram

565 for each of the individual morphotypes tracing appearance and disappearance of the viruses/vlps.
566 False color rendering of the virus/vlps are shown as visualized using ChimeraX. Green – A round
567 morphotype with a small core and a thick coat. Violet– A curved morphotype. Gold – Two round
568 vlps very close to each other. No glycoproteins visible on their surface. All scale bars = 50 nm

569