PNAS RESEARCH ARTICLE | EVOLUTION

Evolutionarily conserved grammar rules viral factories of amoeba-infecting members of the hyperdiverse *Nucleocytoviricota* phylum

Sofia Rigou[a,1,2], Alain Schmitt[a,1], Anaísa B. Moreno[a,1], Audrey Lartigue[a], Lucile Danner[a], Lotte Mayer[b], Claire Giry[a], Feres Trabelsi[a], Lucid Belmudes[c], Natalia Olivero-Deibe[d], Hugo Le Guenno[e], Yohann Couté[c], Mabel Berois[f], Matthieu Legendre[a], Sandra Jeudy[a], Chantal Abergel[a,3], and Hugo Bisio[a,3]

Affiliations are included on p. 11.

Edited by Eugene Koonin, NIH, Bethesda, MD; received June 11, 2025; accepted August 4, 2025

Despite sharing fewer than 10 core genes, the hyperdiverse *Nucleocytoviricota* phylum (ranging from poxviruses to giant viruses) universally assembles viral factories (VFs) resembling biomolecular condensates. Regardless, it is unclear how these viruses achieve such a level of functional conservation without clear conserved genetic information. We demonstrate that the VFs produced by amoeba-infecting viruses have liquid-like properties and identify a conserved molecular grammar governing viral factory scaffold protein: charge-patterned intrinsically disordered regions that drive phase separation independently of sequence homology. This grammar predicts functional scaffold proteins across the 15 viral families, revealing evolutionary constraints invisible to sequence or structural analysis. Strikingly, VFs exhibit subcompartmentalization analogous to nuclei, segregating transcription and mRNA processing (inner condensates) from replication (interphase zones) and translation (host cytoplasm). Our work establishes phase separation as a fundamental organizational principle bridging extreme genomic diversity, explaining how biological complexity emerges without gene conservation. This grammar is likely also conserved in non-amoeba-infecting members of the phylum and thus may represent a primordial solution for organelle-like organization, with broad implications for antiviral targeting.

Nucleocytoviricota | mimivirus | phase separation | biomolecular condensate | giant virus

Significance

Viral factories (VFs) are organelle-like compartments that are conserved across the *Nucleocytoviricota* phylum despite extensive genomic diversity. How such phenotypic conservation is maintained remains unknown. Here, we show that VFs of amoeba-infecting viruses emerge from a conserved organizational molecular grammar, condensed within intrinsically disordered regions, thereby efficiently bypassing the need for gene conservation. This compact mode of functional inheritance likely provides a robust platform for genome diversification.

The phylum *Nucleocytoviricota* exemplifies the remarkable diversity of the virosphere, encompassing viruses with a wide range of morphologies, genome sizes, and hundreds of unique genes. This diversity spans from Yaravirus, which forms 80-nm icosahedral particles and carries a 45-kbp dsDNA genome encoding 74 predicted proteins (1), to Pandoravirus, which produces 1-µm ovoid particles and harbors a 2.5-Mbp genome encoding over 1,500 genes (2, 3). Despite this diversity, these viruses are unified by only ten core genes, most of which are not universally conserved among all members (4). How, then, do these viruses maintain a coherent replication strategy? Virtually all members of *Nucleocytoviricota* [including Poxviridae and the giant viruses (5, 6)] form viral factories (VFs): cytoplasmic organelle-like compartments that spatially coordinate genome replication and/or virion assembly (7). Importantly, while some viruses have acquired nuclear genome replication, cytoplasmic VFs are still visualized by electron microscopy (8, 9). However, the molecular mechanisms underlying VF biogenesis remain poorly understood (10). To fill this gap, mimivirus was used as a primary model since it displays a biphasic nature when imaged by electron microscopy and a highly synchronized biogenesis and maturation (10). Moreover, mimivirus is a well-studied member of the family Mimiviridae, which is ubiquitous in the environment (11, 12). Mimivirus is a fully cytoplasmic giant virus for which virions are internalized into the host cell by phagocytosis (13). The acidic compartment of the phagosome, in combination with oxidative stress through Fenton reactions, triggers the opening of the stargate structure located at an apex of the icosahedral capsid (14–16). Once opened, the fusion of the membrane of the phagosome with the virion's internal membrane leads to the release of the core in the cytoplasm (13, 17, 18). The core contains the genome of the virus and the machinery needed to establish infection (19). The core consequently develops to establish a viral factory, a highly dynamic and complex organelle-like structure putatively formed by over 300 proteins (20). The VF not only protects the viral DNA from cellular insults (21) but also segregates DNA replication and transcription from translation (20).

Author contributions: S.R., A.S., A.B.M., C.A., and H.B. designed research; S.R., A.S., A.B.M., A.L., L.D., L.M., C.G., F.T., L.B., N.O.-D., H.L.D., S.J., and H.B. performed research; M.B. and M.L. contributed new reagents/analytic tools; S.R., A.S., A.B.M., L.M., H.L.G., Y.C., S.J., and H.B. analyzed data; M.B., M.L., S.J., C.A., and H.B. supervision; and S.R., A.S., A.B.M., C.A., and H.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]S.R., A.S., and A.B.M. contributed equally to this work.

[2]Present address: Division of Microbial Ecology, Centre for Microbiology and Environmental Systems Science, University of Vienna, Vienna 1030, Austria.

[3]To whom correspondence may be addressed. Email: chantal.abergel@igs.cnrs-mrs.fr or hugo.bisio@igs.cnrs-mrs.fr.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2515074122/-/DCSupplemental.

Published August 27, 2025.

Biomolecular condensates, formed through phase separation, are increasingly recognized as key structures in viral replication, host manipulation, and morphogenesis (22, 23). Phase separation is the process by which biomolecules demix into membrane-less compartments (22–25). These structures typically contain scaffold proteins, which drive condensate formation, and client proteins, which are selectively recruited (22–25). VFs formed by phase separation have been described in several RNA and DNA viruses, including members of the Herpesviridae, Adenoviridae, Reoviridae, and the order Mononegavirales (reviewed in ref. 22).

Here, we hypothesize that phase separation provides a unifying mechanism for assembling viral replication centers across the genetically diverse *Nucleocytoviricota*, despite minimal sequence conservation. To test this, we leveraged available genetic tools for giant viruses to investigate the nature of VFs across the phylum (26–28).

Results

Universal Liquid-Like Properties of Amoeba-Infecting Nucleocytoviricota VFs. To gather information associated with the fine ultrastructure of VFs, we first fluorescently labeled proteins (described later in the manuscript) enriched in a previous proteome of mimivirus [order: *Imitervirales* (Fig. 1*A*)] VFs (20) and identified two subcompartments recapitulating the ultrastructure visualized by electron microscopy (Fig. 1 *B* and *C*). Multiple inner layers can be seen in a single VF early during infection, while a single and continuous outer layer englobes all inner layers (Fig. 1 *B* and *C* and *SI Appendix*, Fig. S1*A*). The number of inner layers per VF linearly increases with the multiplicity of infection, suggesting that each genome unit induces the formation of individual inner layers (*SI Appendix*, Fig. S1*B*).

Live-cell imaging of mimivirus-infected cells revealed that upon contact, two VFs juxtaposed and never separated during the remaining recording time (*SI Appendix*, Fig. S1*C*), suggesting that the outer layer of two independent VFs fused as observed by fluorescence and electron microscopy (Fig. 1 *B* and *C*). Importantly, imaging mimivirus-infected cells early during infection to record fusion events is challenging for two main reasons: 1) The high mobility of the amoeba precludes the use of higher magnification since the cell quickly exits the area being recorded. 2) It is extremely infrequent to find individual mimivirus VFs: independent inner layers without a shared outer layer within the same cell. Indeed, once VFs are sufficiently big to be visualized, the inner layers of the viral factory almost always share the same outer layer. Recovery of fluorescence after photobleaching (FRAP) was observed when an area of the outer layer of the VF was bleached (Fig. 1*D*). Fluorescence gain was accompanied by a partial decrease in fluorescence in the nonbleached area, but this decrease was incomplete, indicating some diffusion along with a fraction of nonmobile molecules (Fig. 1*D*). Fixed VFs did not show any recovery, consistent with crosslinking (*SI Appendix*, Fig. S1*D*). Thus, mimivirus VFs exhibit some liquid-like properties but do not behave as a classical condensate formed by liquid–liquid phase separation.

To compare with other viruses from the phylum, we labeled noumeavirus (29) [order: *Pimascovirales* (Fig. 1*A*)] VFs using cells expressing the C-terminal GFP-tagged mel_H2B-H2A (28). Noumeavirus is a good model for this experiment since 1) it contains a single detectable layer; 2) multiple VFs can be easily observed early in infection, and 3) noumeavirus infection rapidly disrupts the movement of the amoeba (29). Noumeavirus VFs were observed as dynamic organelles subjected to deformation, and the encounter of two VFs yielded coalescence (Fig. 1*E*). To specifically address the liquid-like properties of VFs on the class *Pokkesviricetes*, we fluorescently tagged PN-99 with GFP and performed FRAP experiments on pacmanvirus lostcity (30) [order: *Asfuvirales* (Fig. 1*A*)] VFs (Fig. 1*F*). Fast recovery was observed upon photobleaching and diffusion from the nonbleached side of the VF led to a homogenous fluorescent intensity. Thus, both noumeavirus and pacmanvirus show classical behavior of condensates formed by liquid–liquid phase separation.

To facilitate comparisons, we investigated whether these membrane-less organelles would dissolve by treating infected cells with 10% 1,6-hexanediol (23) or hypotonic stress (31). Mimivirus VF only partially dissolved during the treatment, further reinforcing nonclassical liquid-like properties of the factories (*SI Appendix*, Fig. S2 *A–C*). On the other hand, treatment with 10% 1,6-hexanediol efficiently dissolved the VFs of noumeavirus (*Marseilleviridae*), pithovirus (*Pithoviridae*), pacmanvirus, and the mammalian-infecting vaccinia virus (*Poxviridae*) (*SI Appendix*, Fig. S2 *A–C*). All VFs were also partially dissolved upon hypotonic stress (*SI Appendix*, Fig. S2*A*), reinforcing some liquid-like phase separation behavior of these organelles.

Additionally, a subgroup of members of the phylum transfer their genome to their host nucleus, where it replicates. Mollivirus is one of those viruses (32). Regardless, mollivirus major capsid protein (virion protein) accumulates in a cytoplasmic, uncharacterized subcompartment (likely the VFs) in infected cells before loading into the particles (21). This localization could also be disrupted by 1,6-hexanediol treatment (*SI Appendix*, Fig. S2*D*). Importantly, both 1,6-hexanediol and hypotonic shock treatment have limitations worth mentioning: 1,6-hexanediol has several known off-target effects, including components of the cytoskeleton (33) and the inhibition of kinases and phosphatases (34), all of which could affect VFs independently of liquid-like phase separation. Hypotonic stress induces swelling of the cells, effectively reducing the concentration of proteins within them. Regardless, VFs are big in size (reaching up to 20 μm diameter) and occupy half of the cell cytoplasm volume. This likely diminishes the impact of the cell swelling effect, potentially explaining the partial dissolution of VFs with this treatment. Nevertheless, the combination of these experiments with the live imaging (*SI Appendix*, Table S1) strongly suggests that the VFs of mimivirus, noumeavirus, pacmanvirus, and probably all amoeba-infecting members of the phylum *Nucleocytoviricota* display characteristics of biomolecular condensates ranging from prototypical compartments made by liquid–liquid phase separation (like noumeavirus or pacmanvirus) to condensates with some liquid-like properties (like mimivirus).

A Two-Scaffold System Creates Biphasic Condensates in Mimivirus. Having identified a hallmark of the amoeba-infecting *Nucleocytoviricota*, we aimed to uncover the common components of these organelles between the extreme viral diversity of the phylum. Phase separation is driven by the multivalency of proteins termed scaffold proteins (35). Scaffolding proteins are abundant in condensates and tend to contain intrinsically disordered regions (IDRs) as compacted means to achieve multivalency (35). To identify VF scaffold proteins, we performed immunoprecipitations using three previously identified VF proteins as bait (R562, R505, and R336/R337) (20). Two prey proteins (R561 and R252) containing predicted IDRs (*SI Appendix*, Fig. S3*A*), were consistently enriched, indicating interaction with all three client proteins (Dataset S1).

Expression of R561 [termed Outer Layer Scaffold 1 (OLS1)] in *Acanthamoeba castellanii* demonstrated its ability to undergo phase separation in the amoeba cytoplasm (*SI Appendix*, Fig. S3*B*). In contrast, R252 [termed Inner Layer Scaffold 1

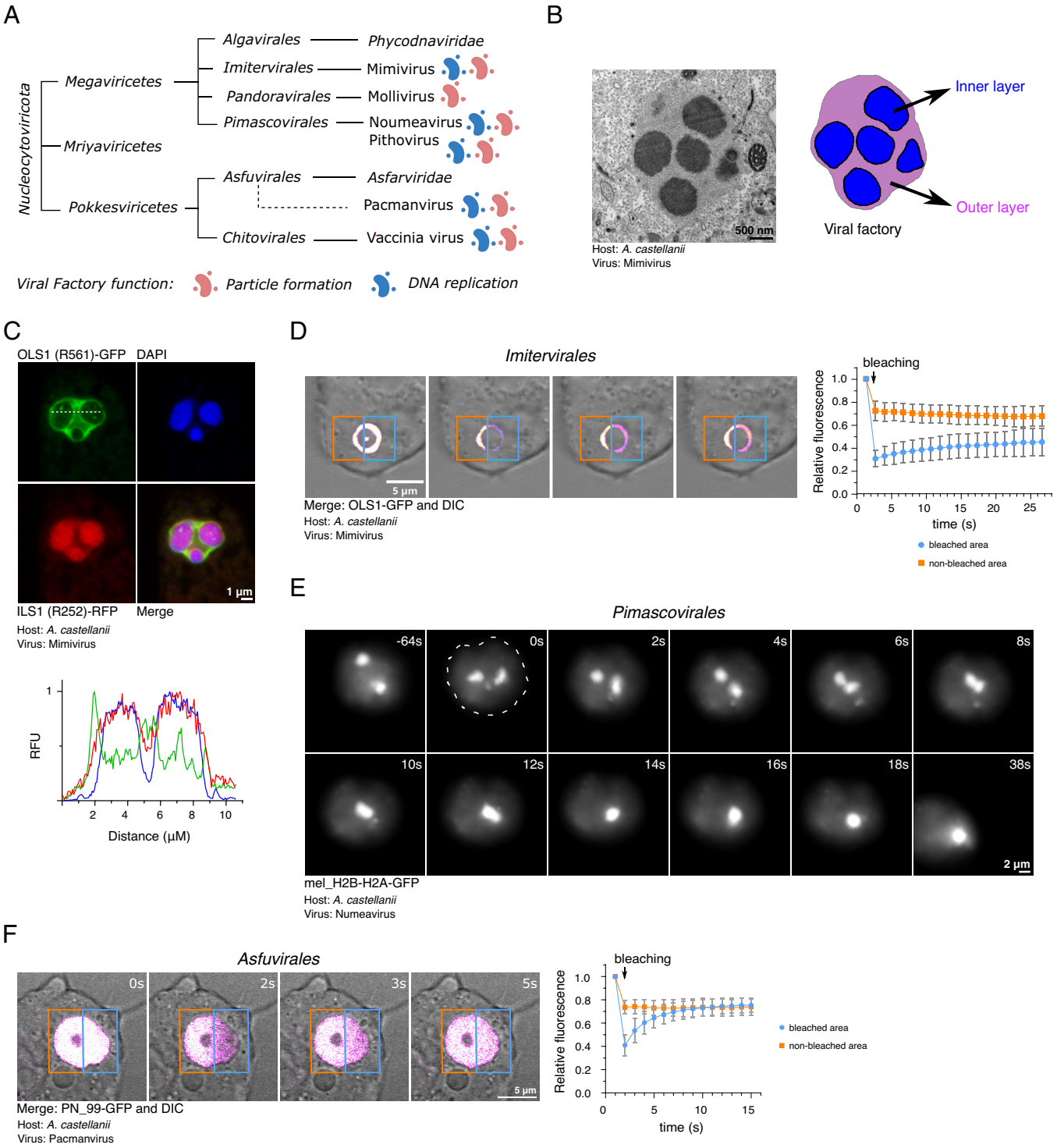

Fig. 1. *Nucleocytoviricota* VFs display liquid-like properties. (*A*) Taxonomy of viruses belonging to the *Nucleocytoviricota* phylum (adapted from ref. 6). (*B*) Transmission electron microscopy imaging (and cartoon) of an ultrathin section of an infected *Acanthamoeba castellanii* cell with mimivirus VF formed in the cytoplasm. 6 h postinfection at an MOI of 20. (*C*) *A. castellanii* cells expressing C-terminally tagged OLS1-green fluorescent protein (GFP)(R561, outer layer) and ILS1-RFP (R252, inner layer) infected with mimivirus. Line profiles (below) corresponding to the white dashed line show fluorescence patterns. 6 h postinfection at an MOI of 10. (*D*) Representative images of *Acanthamoeba* cells expressing OLS1-GFP were infected with mimivirus at an MOI of 5. Five hours postinfection, OLS1-GFP was photobleached, and fluorescence recovery was imaged by confocal microscopy. FRAP data were normalized, corrected for background, and represented in the recovery curve. Data represent the mean ± SD of seven experiments. (*E*) Live-cell imaging of noumeavirus-infected *A. castellanii* expressing mel_H2B-H2A-GFP as a marker of the VF. Recording was performed 2 h postinfection. Cell boundaries are indicated by a dashed line. (*F*) Representative images of Acanthamoeba cells expressing PN99-GFP were infected with pacmanvirus Lost City at an MOI of 10. Three hours postinfection, PN99-GFP was photobleached and fluorescence recovery was imaged by confocal microscopy. FRAP data were normalized, corrected for background, and represented in the recovery curve. Data represent the mean ± SD of 10 experiments.

(ILS1)] exhibited a diffuse cytoplasmic localization (*SI Appendix*, Fig. S3*B*). Upon mimivirus infection of cells expressing both fluorescently tagged proteins, OLS1 and ILS1 relocalized to the VFs, with OLS1 associating with the outer layer and ILS1 with the inner layer (*SI Appendix*, Fig. S3*B*). Control cells expressing only GFP or RFP did not display similar relocalizations

(*SI Appendix*, Fig. S3*C*). Notably, VF client proteins, such as R336/R337, localized not only to the VF outer layer but also to biomolecular condensates formed by the overexpressed OLS1, strongly supporting its role as a scaffolding protein forming the VF outer layer (*SI Appendix*, Fig. S3*D*). Furthermore, when OLS1 and ILS1 were coexpressed, ILS1 acted as a client protein and was recruited to OLS1 biomolecular condensate (*SI Appendix*, Fig. S3*E*).

Considering that ILS1 binds DNA (36), we reasoned that its recruitment to the VF outer layer would allow contact with viral DNA, thereby enabling phase separation. To test this hypothesis, recombinantly expressed ILS1 (*SI Appendix*, Fig. S4*A*) was analyzed for phase separation in the presence of DNA (Fig. 2*A* and *SI Appendix*, Fig. S4B). Phase separation was observed only in the presence of DNA (Fig. 2*A*). Moreover, DNA concentration significantly influenced the nature of phase separation (*SI Appendix*, Fig. S4*C*), being at lower DNA concentrations (2.5 to 20 μg/mL), networks and droplets predominated, whereas higher DNA concentrations (>20 μg/mL) prompted the formation of large gels. ILS1 concentration did not affect the nature of phase separation (*SI Appendix*, Fig. S4*D*). Both linear and circular DNA equally triggered phase separation, but mimivirus genomic DNA induced larger gel formation at lower DNA concentrations (*SI Appendix*, Fig. S4*E*). Similar results were obtained with the recombinant mCherry fused ILS1 (*SI Appendix*, Fig. S4*F*). RNA did not trigger ILS1 phase separation (*SI Appendix*, Fig. S4*B*).

Recombinant OLS1 (*SI Appendix*, Fig. S4*A*) made phase separation with classical liquid–liquid phase separation morphology, independently of other macromolecules (Fig. 2*B*). Phase separation depended on OLS1 concentrations (*SI Appendix*, Fig. S5*A*) and salt concentrations (*SI Appendix*, Fig. S5*B*). Similar to in cellula conditions, in vitro OLS1 biomolecular condensate recruited ILS1 as a client protein (*SI Appendix*, Fig. S5*C*). Finally, adding DNA to the mixture of both proteins triggered the segregation into two phases, independently of the order of component incorporation (Fig. 2*C*). Altogether, OLS1, ILS1, and DNA are sufficient to trigger VF-like biphasic phase separation in vitro.

To test whether these proteins were responsible for the VFs formation, we attempted gene knockout of both genes in host cells expressing a codon-optimized version of each protein for trans-complementation (28). Recombinant OLS1 knockout viruses were obtained (*SI Appendix*, Fig. S6*A*) with a 100-fold reduction in burst size compared to wild-type viruses (Fig. 2*D*). Moreover, a lower viral DNA accumulation was observed during infection (Fig. 2*E*), consistent with a delay in VF growth, as shown by DAPI staining of infected cells (Fig. 2 *F* and *G*). Depletion of OLS1 also reduced fusion events of VFs upon superinfection, as shown by the presence of multiple independent VFs in the same cells (Fig. 2 *F* and *H*). Visualization of mutant infectious cycle by electron microscopy demonstrated that VFs of *ols1* KO viruses lack any visible outer layer, confirming that OLS1 is the scaffold protein for this subcompartment of the VF (Fig. 2*I* and *SI Appendix*, Fig. S6*B*). Moreover, endogenous tagging of client proteins of the outer layer of the VF (like R336/R337 or R322) displayed a cytoplasmic localization upon deletion of *ols1* (Fig. 2 *J* and *K* and *SI Appendix*, Fig. S6 *C* and *D*), indicating that the role of the outer layer of the VF of mimivirus is to concentrate important components for VF functions. All phenotypes were fully restored by trans-complementation (Fig. 2 *D*–*K*). On the other hand, we were unable to obtain clonal *ils1* KO viruses. Using the trans-complementing line, we demonstrated that *ils1* is likely an essential gene (*SI Appendix*, Fig. S6*E*).

A Conserved Molecular Grammar Predicts Scaffold Protein throughout the Phylum. The molecular grammar of an IDR refers to compositional bias and sequence patterns in its primary structure (24, 25). This grammar allows condensation and specific recruitment of molecules, including client proteins (37–42). We thus reasoned that due to the plethora of client proteins recruited to the VFs, despite major changes in the primary sequence of their scaffold proteins, the molecular grammar of the condensate would not easily change. Concordantly, host-expressed mimivirus (*Imitervirales*) OLS1-GFP and ILS1-RFP relocalized to noumeavirus (*Pimascovirales*) VF upon infection (*SI Appendix*, Fig. S7 *A* and *B*).

Thus, to identify scaffold proteins in all *Nucleocytoviricota*, we built a pipeline to analyze the genomes of these viruses. First, we predicted the "IDRome" encoded in representative genomes of isolated viruses and extended the analysis to metagenomes from the giant virus database (4), the permafrost (43), and Egovirales (44) (Dataset S2). We also capitalized on methods previously described to study the biochemistry of nucleolar phase separation, relying on Nardini and CIDER (42) and adjusted two parameters for faster (*SI Appendix*, Fig. S8*A*) and more precise (*SI Appendix*, Fig. S8*B*) discrimination of particular proteins. Since the inner layer of the VF is only present in members of the *Mimiviridae*, we excluded ILS1 from the following analysis and focused bioinformatic computations on OLS1. Using 98 features (36 from Nardini and 62 from CIDER), we created a preliminary machine learning classifier that identified four candidate scaffold proteins in noumeavirus (*SI Appendix*, Fig. S7*C*).

To experimentally challenge these predictions, the four genes were codon-optimized and expressed in the amoeba. Two showed a localization coherent with biomolecular condensates (Fig. 3*A*) and relocalized to the noumeavirus VF upon infection (Fig. 3*B*). Moreover, both proteins also relocalized to the mimivirus VFs upon infection, confirming a shared molecular grammar for phase separation for these two viruses' VFs (*SI Appendix*, Fig. S7*D*). In contrast, the other two protein candidates did not spontaneously form biomolecular condensates in the amoeba cytoplasm (*SI Appendix*, Fig. S7*E*).

Using these confirmed noumeavirus scaffold proteins (NMV_095 and NMV_238) and detected homologs, we retrained the predictive machine learning classifier (*SI Appendix*, Fig. S7*F*). Features for the classifier were selected by comparing the feature values of the three scaffold proteins' IDRs and homologs to the rest of the combined IDRome of mimivirus and noumeavirus (*SI Appendix*, Fig. S8 *C* and *D*). Final predictions of scaffold proteins in all orders of *Nucleocytoviricota* were then generated with this optimized classifier (Dataset S2).

Clear segregation of resulting candidate IDRs can be observed in the unsupervised UMAP representation of the *Nucleocytoviricota* IDRome features (Fig. 3*C* and *SI Appendix*, Fig. S9*A*). As expected, scaffold IDRs have a low Anchor2 score (45) and are not predicted to become structured upon changes in the environment (i.e., binding to an interacting protein) (*SI Appendix*, Fig. S9*B*). The most discriminative features in our classification, and thus, the features associated with the molecular grammar of the VFs, were related to charge segregation (Fig. 3 *E* and *F* and *SI Appendix*, Fig. S9 *A* and *C*). Within the negative charges, glutamate seems more important, as confirmed scaffold IDRs and candidates have 15% of this residue (±6), while the rest of the IDRs have 6% (±8). Scaffold IDRs and candidates have also a relatively higher E/D ratio (0.25 ± 0.26 vs. 0 ± 0.45) and a higher K/R ratio (0.48 ± 0.36 vs. 0.07 ± 0.53).

Putative scaffold proteins could be identified to be encoded by members of all *Nucleocytoviricota* (15 families and unclassified

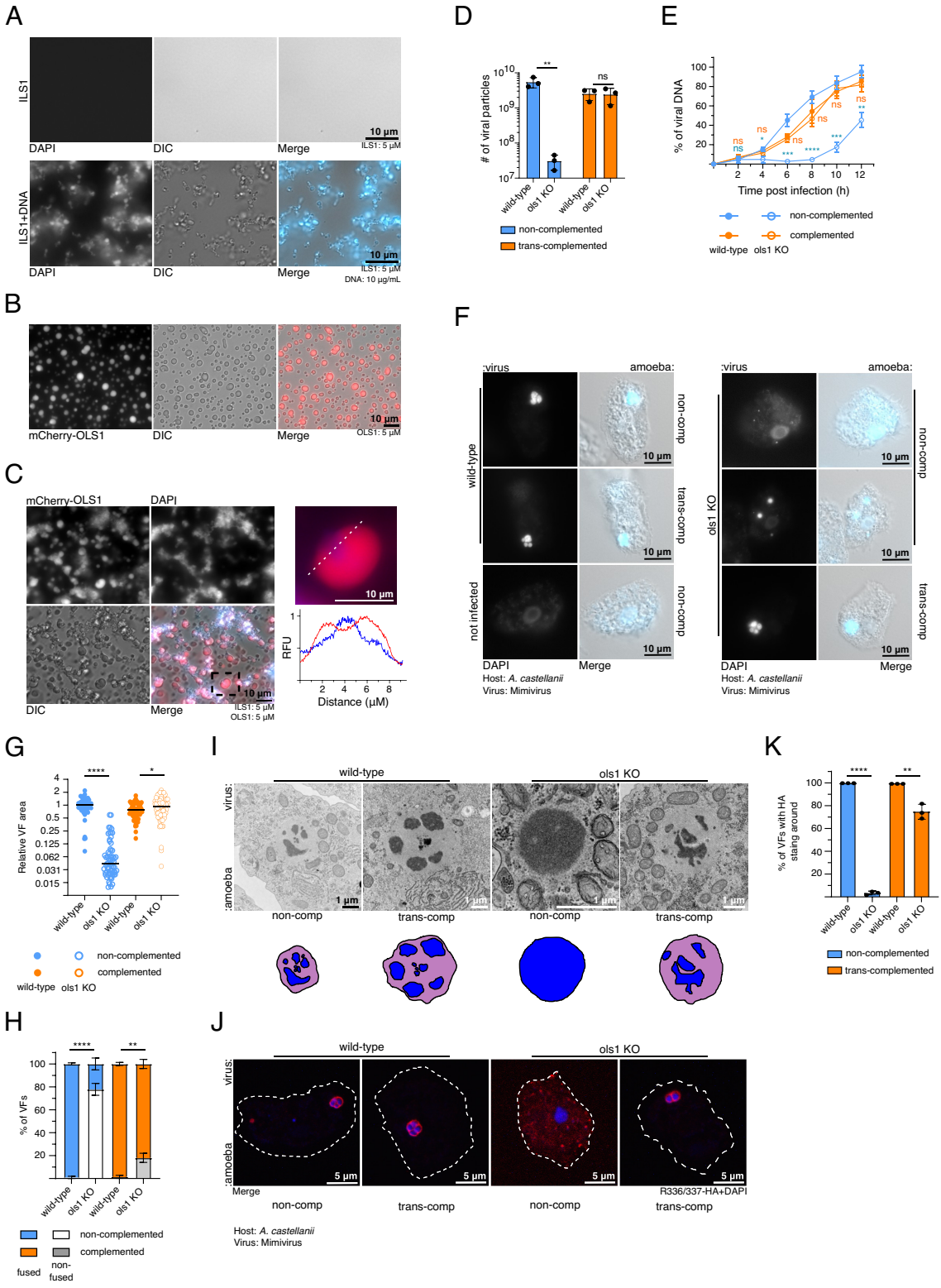

Fig. 2. Two scaffold proteins are involved in mimivirus viral factory assembly. (*A*) In vitro phase separation of ILS1 in the presence or absence of DNA. DAPI was used to confirm nucleic acid presence in the condensate. (*B*) In vitro phase separation of mCherry-OLS1 at 50 mM NaCl. (*C*) In vitro phase separation of mCherry-OLS1 and ILS1 in the presence of 10 μg/mL DNA. DAPI was used to confirm nucleic acid presence in the condensate. A plot profile is shown demonstrating a biphasic interaction between the two phase-separated compartments. (*D*) Quantification of the number of viruses produced upon knock-out of *ols1* in mimivirus. Data correspond to the mean ± SD of three independent experiments. (*E*) DNA replication was analyzed by qPCR. Viral DNA is represented as a percentage of total DNA in the sample. Data correspond to the mean ± SD of three independent experiments. (*F*) Representative light fluorescence microscopy images of *A. castellanii* cells infected with wild-type mimivirus or *ols1* KO viruses 6 h pi. DAPI: VFs. (*G*) Quantification of the size of VF generated, as shown in (*F*). At least 50 VF were recorded per condition during three independent experiments. The area of DAPI staining was measured using ImageJ. (*H*) Quantification of the number of VF either fused or separated in infected cells as shown in (*F*). At least 100 infected cells were recorded per condition. Data correspond to the mean ± SD of three independent experiments. (*I*) Transmission electron microscopy imaging (and representative cartoon) of the mimivirus replication cycle in *A. castellanii*. Images were acquired 4 to 6 h pi (MOI=20). (*J*) Localization of client protein (R336/337) from the outer layer of the VF endogenously tagged with C-terminal 3xHA. 6 hpi. DAPI: VFs. (*K*) Quantification of the number of (*I*). At least 100 infected cells were quantified per condition. Data correspond to the mean ± SD of three independent experiments. ns, $P > 0.05$, *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, and ****$P \leq 0.0001$.

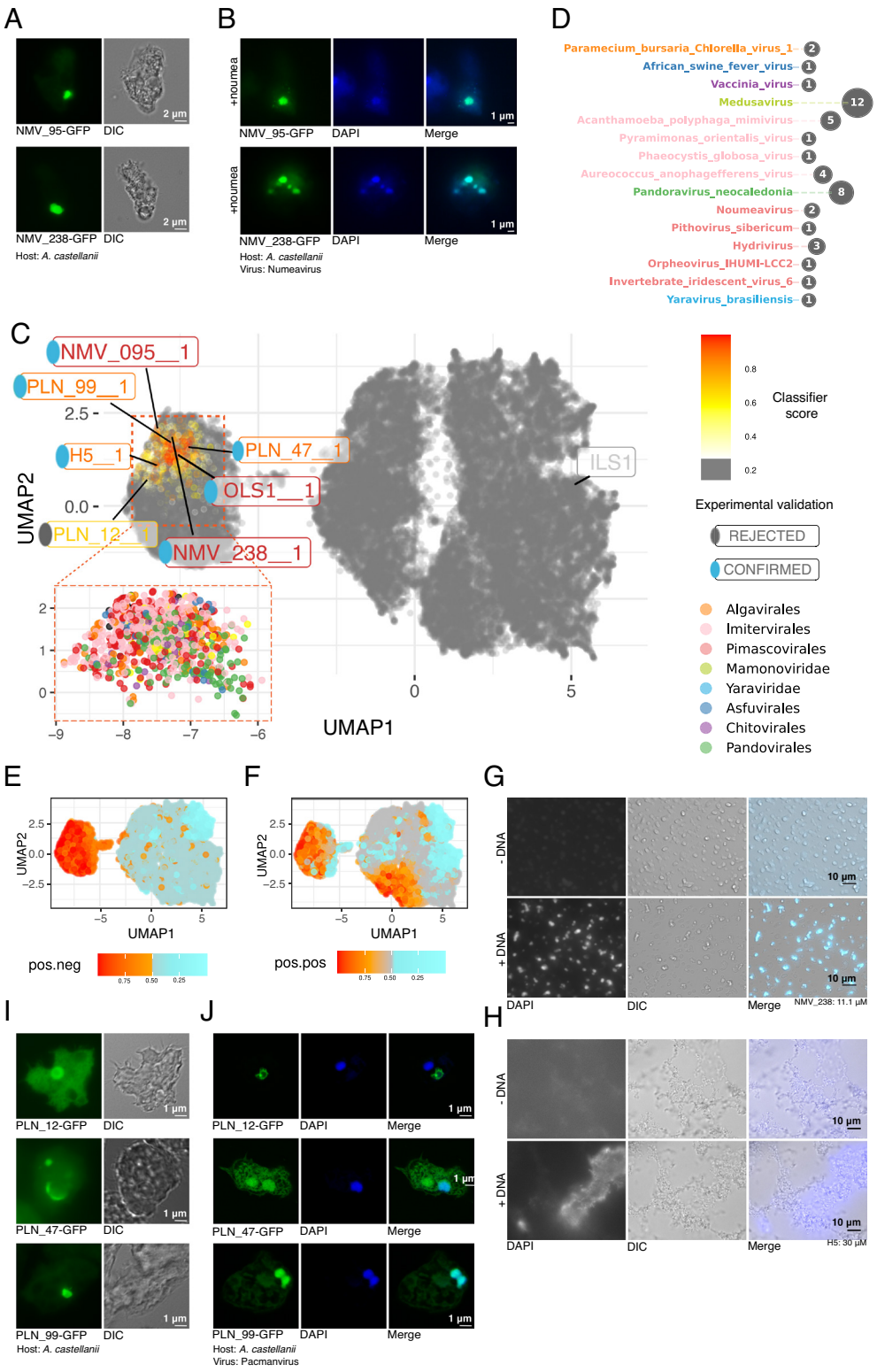

Fig. 3. A common grammar allows the identification scaffold proteins across the entire *Nucleocytoviricota* phylum. (*A*) *A. castellanii* cells expressing C-terminally GFP-tagged NMV_095 or NMV238. (*B*) *A. castellanii* cells expressing C-terminally GFP-tagged NMV_095 or NMV238 were infected with noumeavirus. VFs were labeled using DAPI 2 to 3 h pi. (*C*) Uniform Manifold Approximation and Projection (UMAP) representation of the IDRs in representative genomes and metagenomics giant viruses, based on the 11 features selected for the classifier. Each point represents one predicted IDR. IDRs corresponding to negative predictions are shown in gray. (*Inset*) Scaffold candidates were labeled according to their Order classification (or higher-level classification available). (*D* and *E*) Details of the most discriminant features on the UMAP. Pos.neg (positive from negative segregation) and pos.pos (positive to all other residues) are plotted. (*F*) Count of candidate proteins for VF phase separation predicted by the classifier in representatives of each *Nucleocytoviricota* family. (*G*) NMV_238 and (*H*) H5 form biomolecular condensates in vitro in the absence and presence of DNA. The solution contains 25 mM NaCl and 0.6 µg/mL DNA, where indicated. (*I* and *J*) *A. castellanii* cells expressing C-terminally tagged scaffold candidates of pacmanvirus lostcity were either noninfected (*I*) or infected with pacmanvirus (*J*). VFs were labeled using DAPI 6 h pi.

viruses) (Fig. 3 *C*, *D*, and *F*). In all these orders, between 4.4% (*Algalvirales*) and 5.4% (*Chitovirales*) of all predicted IDRs were classified as candidate scaffold proteins for phase separation and VFs generation. The only exception was the *Egovirales*, scoring 8.7% of positive proteins in their respective IDRomes. In *Pimascovirales*, homologous proteins were identified as scaffolds in pithovirus sibericum and cedratvirus kamchatka (pv_12 and ck125, 39% identity/63% similarity), highlighting the consistency of the method. Similarly, orthopoxvirus' H5 protein was identified as the only candidate in the vaccinia and monkeypox viruses (93% identity). On the other hand, in *Pandoravirales* [which present a nuclear DNA replication (12)], eight candidate proteins were identified. Regardless, all candidate VF scaffold IDRs predicted in all *Nucleocytoviricota* are distributed on the UMAP without order segregation, further reinforcing the shared molecular grammar for the cytoplasmic VFs (Fig. 3*C*).

To challenge scaffold protein predictions across the *Nucleocytoviricota*, we expressed recombinantly vaccinia virus (order: *Chitovirales*) H5 [as previously described (46)] and noumeavirus (order: *Pimascovirales*) NMV_238 (*SI Appendix*, Fig. S10*A*). Similar to OLS1, both H5 and NMV_238 formed biomolecular condensates in the absence of other macromolecules (Fig. 3 *G* and *H*). Phase separation depended on protein concentration (*SI Appendix*, Fig. S10 *B* and *C*) and salt concentrations (*SI Appendix*, Fig. S10 *D* and *E*). Interestingly, in contrast to OLS1, which does not bind dsDNA, both H5 and NMV_238 were capable of binding and concentrating DNA, as shown by DAPI staining (Fig. 3 *G* and *H*). Moreover, condensation was slightly favored in the presence of dsDNA (*SI Appendix*, Fig. S10 *B*–*E*). Overall, both scaffold proteins display the capacity to phase-separate independently of other macromolecules upon diminishing ionic strength (similar to OLS1), but display DNA binding and seclusion similarly to ILS1.

While H5 shows a propensity to form biomolecular condensates, the recombinant protein did not show the expected liquid–liquid phase separation behavior (however, H5 is heavily phosphorylated during the infectious cycle (47), potentially affecting phase separation). To further address the validity of scaffold predictions in the *Pokkesviricetes*, we expressed the three candidates identified as potential scaffold proteins from pacmanvirus lostcity (order: *Asfuvirales*) in *A. castellanii*. PN-99-GFP showed the classical pattern of phase separation similar to mimivirus OLS1 (Fig. 3*I*). PN_47-GFP displayed similar characteristics but also strongly interacted with the contractile vacuole (Fig. 3*I*). Finally, PN-12-GFP showed a nuclear localization (Fig. 3*I*). Moreover, upon infection with pacmanvirus, both PN-99-GFP and PN_47-GFP relocalized to the VFs while PN-12-GFP remained nuclear (Fig. 3*J*). Interestingly, homologues of PN-99-GFP and PN_47-GFP were consistently identified as scaffolds in all pacmanvirus isolates, while PN-12-GFP was only identified in pacmanvirus lostcity and A23 (Dataset S2). Taking these data together, both PN-99-GFP and PN_47-GFP are scaffold proteins encoded by pacmanvirus.

Importantly, while candidate scaffold proteins were found in all orders of *Nucleocytoviricota* irrespective of their host, the experimental validation of this grammar was mostly centered on amoeba-infecting viruses. Thus, further analysis on other host-pathogen models will be necessary to confirm its universality.

Functional Subcompartmentalization of VFs. While a previous study reported proteins tentatively localized at the VFs of mimivirus (20), it did not differentiate proteins associated with each subcompartments of the VFs and displayed high rates of false positives (*SI Appendix*, Fig. S11 *A* and *B*). To fill this gap, we performed immunoprecipitations of the two scaffold proteins and identified copurified proteins by mass spectrometry (MS)-based proteomics (Fig. 4*A*, *SI Appendix*, Fig. S11*C*, and Dataset S3). Several proteins enriched with OLS1 and/or ILS1 were endogenously tagged to confirm their localization (Fig. 4*B* and *SI Appendix*, Fig. S11*B*). None of the false-positive examples identified in the previous proteome study were detected in these immunoprecipitations (Dataset S3). Moreover, enrichment of proteins localized to the outer layer of the VF copurified with OLS1, while inner layer proteins or virion proteins were coimmunoprecipitated majorly with ILS1 (Fig. 4*C*). Importantly, while this study shows a low rate of false positive identifications, other proteins not detected in the immunoprecipitations are still localized at the VF (Fig. 4*C*).

Several host ribosomal proteins were enriched in the immunoprecipitations of OLS1-GFP (but not with ILS1-RFP), suggesting some proximity between the outer layer of the VF and the ribosomes (Fig. 5*A* and Dataset S3). Regardless, VFs are thought to segregate replication and transcription from translation (20). To test the current consensus, we generated recombinant viruses or amoebae encoding tagged versions of several proteins involved in the central dogma (Fig. 5 *A*–*D* and *SI Appendix*, Fig. S12*A*). As previously suggested (20), host ribosomes (Fig. 5*C* and *SI Appendix*, Fig. S12*A*) and the viral translation-associated protein SUI1 (Fig. 5*D*) were excluded from the VFs. On the other hand, the virally encoded eIF4e localized both at the cytoplasm of the host and the outer layer of the VF (Fig. 5*D*). In eukaryotes, besides its classical cytoplasmic function in translation initiation, eIF4E also localizes at the nucleus, where it participates in the export of a subset of mRNA (48). If such export function is conserved in the virally encoded eIF4E, it remains to be explored, but it would explain the dual localization of the protein.

Proteins associated with replication and transcription are incorporated into the VFs but subcompartmentalized. While replication proteins localized to the outer layer of the VF, transcription proteins accumulated at the inner layer (Figs. 4*A* and 5 *A* and *B*). Pulse-labeling of mRNAs using 5-ethynyluridine (EU) for 15 min strongly suggests that mRNA production site is located at the inner layer of the VF (Fig. 5*E*), colocalizing with the RNA polymerase (Fig. 5*B*). On the other hand, we were unable to efficiently label DNA by 5-Ethynyl-2-deoxyuridine (EdU) in wild-type viruses (*SI Appendix*, Fig. S12*B*). We reasoned that due to the high AT richness of mimivirus genome (49), de novo synthesis of dTMP would be particularly efficient in infected cells, allowing thymidine to outcompete EdU for its incorporation into the DNA (*SI Appendix*, Fig. S12*C*). Thus, we generated knockout recombinant viruses on the virally encoded thymidylate synthase (TS), which concordantly blocked viral de novo synthesis of dTMP (*SI Appendix*, Fig. S12 *D* and *E*). EdU labeling significantly improved in these viruses, allowing the visualization of DNA replication with pulses of EdU labeling as short as 5 min (Fig. 5*F* and *SI Appendix*, Fig. S12*F*). Similarly to what was observed in the vaccinia virus (50), pulse labeling of DNA replication in late VF resulted in a lower fluorescence intensity than labeling in early VF (Fig. 5 *F* and *G*), showing that DNA replication decreases during the late stages of infection. 5-min labeling with EdU showed an enrichment of newly synthesized DNA at the periphery of the inner layer of the VF (Fig. 5*F*). Concordantly, DNA replication occurs at the interface between the inner layer and the outer layer, where DNA and the DNA polymerase get in contact. Finally, mRNA processing (including capping and Poly(A) synthesis) localized at the inner layer of the VF, indicating that maturation of pre-mRNA occurs at the same subcompartment as transcription (Fig. 5*H*). Overall, a significant subcompartmentalization of functions is observed at mimivirus VFs.

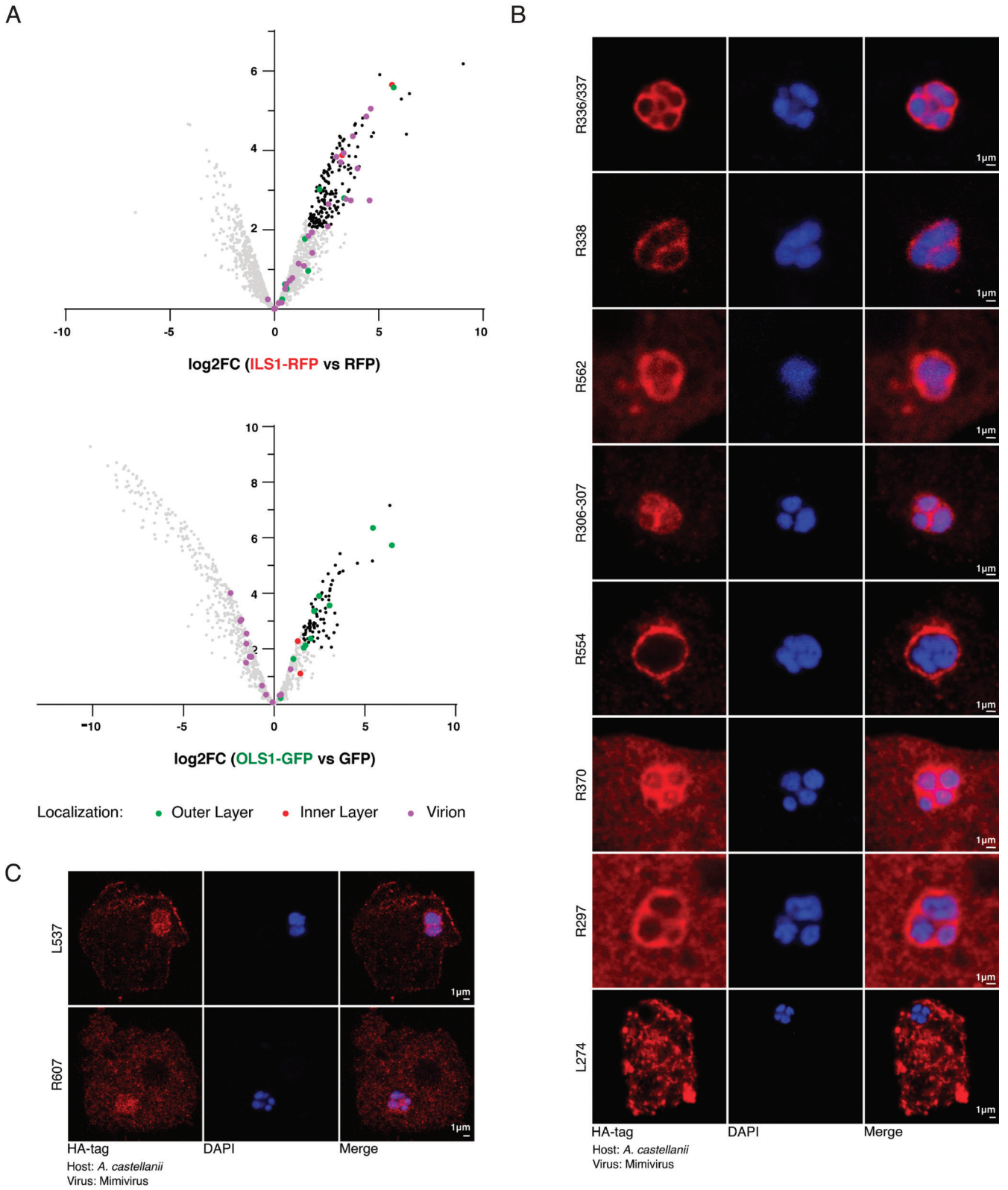

Fig. 4. Viral factory proteins are subcompartmentalized. (*A*) Immunoprecipitation experiments in *A. castellanii* cells expressing ILS1-RFP or OLS1-GFP and infected by mimivirus. Cells expressing RFP or GFP were utilized as controls. Crosslinking was performed prior to cell lysis. Immunoprecipitated proteins were analyzed through MS-based label-free quantitative proteomics (three replicates per condition). The volcano plots represent the −log10 (limma *P*-value) on the *y* axis plotted against the log2(FoldChange bait vs. control) on the *x* axis for each quantified protein (*Upper* panel: OLS1-GFP vs. GFP, *Bottom* panel: ILS1-RFP vs. RFP). Each dot represents a protein. Proteins with log2 (FoldChange) ≥ 1.6 (FoldChange > 3) and −log10(*P*-value) ≥ 2 (*P*-value = 0.01) compared to controls, were considered significant (Benjamini–Hochberg FDR < 2%) and are shown in black. Detailed data are presented in Dataset S3. (*B*) Immunofluorescence demonstrating localization of proteins enriched in *A.* Proteins were endogenously tagged with 3xHA at the C-terminal of each mimivirus gene, and infection was carried out for 6 h before fixation. VFs were labeled using DAPI. (*C*) Immunofluorescence demonstrating localization of proteins not enriched in A but still displaying a VF localization. Proteins were endogenously tagged with 3xHA at the C-terminal, and infection was carried out for 6 h before fixation. VFs were labeled using DAPI.

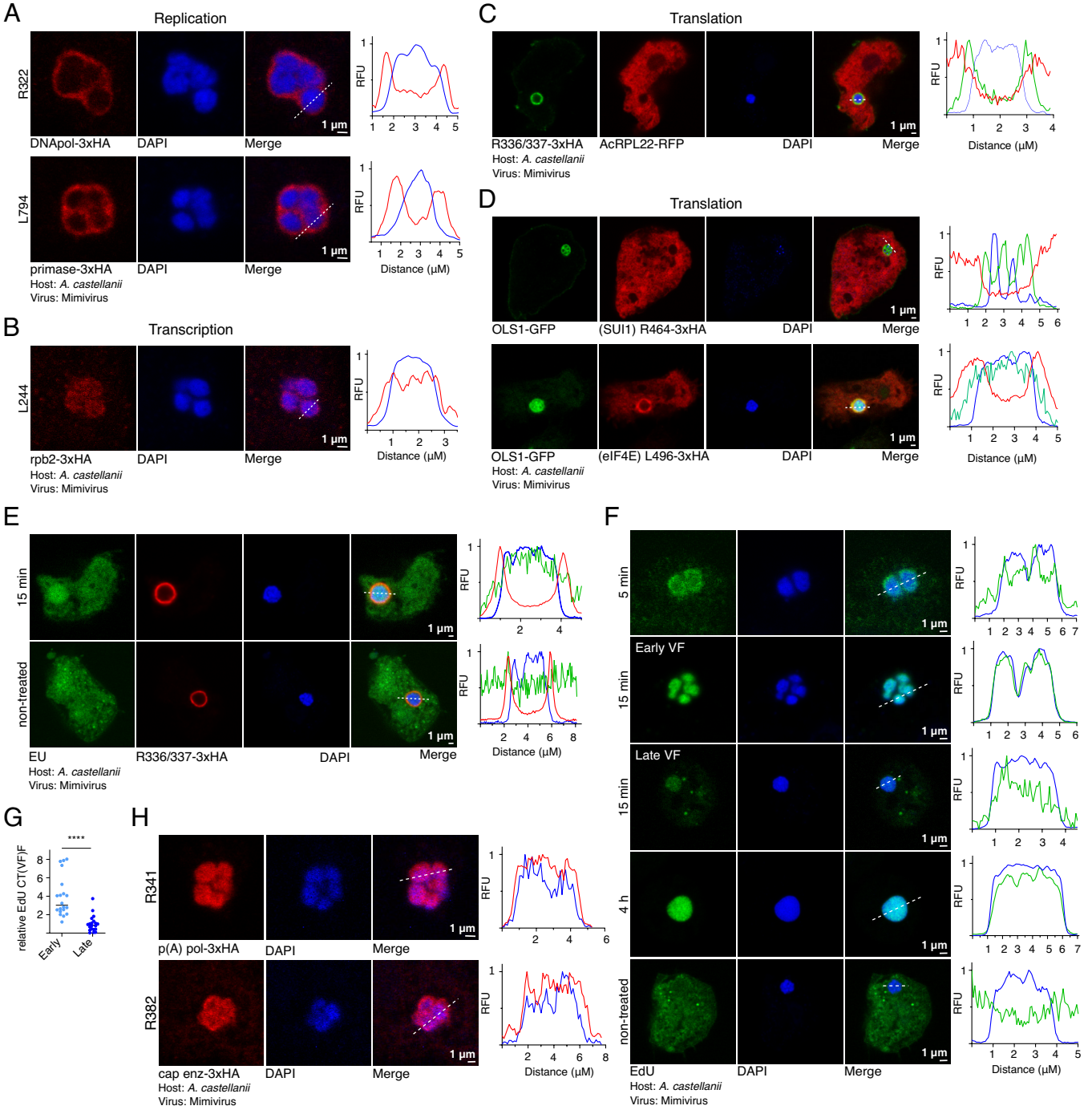

Fig. 5. Central dogma functions are subcompartmentalized at the mimivirus VFs. (*A*) Localization of proteins associated with DNA replication endogenously tagged with C-terminal 3xHA. 6 hpi. (*B*) Localization of the RNA polymerase subunit 2 (rpb2), associated with transcription, endogenously tagged with C-terminal 3xHA. 6 hpi. (*C*) Localization of host ribosomal protein RPL22-RFP, associated with translation. Protein is expressed from a second copy plasmid encoding *rpl22-rfp*. 6 hpi. (*D*) Localization of viral proteins associated with translation endogenously tagged with C-terminal 3xHA. 6 hpi. (*E*) Detection of RNA synthesis by EU labeling. Viral infection was allowed to proceed for 4 to 6 h, and labeling time is indicated. (*F*) Detection of DNA synthesis by EdU labeling. Viral infection by *thymidylate synthase* KO viruses was allowed to proceed for 4 to 6 h, and labeling time is indicated. (*G*) Quantification of the corrected total VF fluorescence (CT(VF)F) of EdU labeling as shown in *F*. Twenty VFs were reordered during three independent experiments, and the intensity of EdU staining was measured using ImageJ. ****$P \leq 0.0001$. (*H*) Localization of proteins associated with mRNA maturation endogenously tagged with C-terminal 3xHA. 6 hpi. In all cases: Line profiles correspond to the white dashed lines. DAPI: VF inner layer. R366/337-3xHA or OLS1-GFP: VF outer layer.

Discussion

The molecular grammar of an IDR in a scaffold protein determines the nature of the interaction to achieve phase separation and the selective recruitment of client proteins (37–42, 51). Here, we demonstrate that amoeba-infecting members of the *Nucleocytoviricota* phylum share a common molecular grammar governing the formation of their VFs. This observation likely extrapolates to non-amoeba-infecting viruses but further experiments would need to be performed with different host-pathogen models in order to confirm this prediction. Such compact means of functional maintenance allow its inheritance with little gene conservation, solving a long-standing paradigm of how such a diverse phylum can conserve a replication strategy. Importantly, since the common ancestor of all *Nucleocytoviricota* likely already possessed a VF with this

molecular grammar, and as the origin of *Nucleocytoviricota* is thought to predate the origin of eukaryotes (and thus, the origin of organelles) (52), this organelle-like structure might represent a primordial solution for compartmentalization. Such conservation of the grammar is likely driven by the complex number of client proteins recruited to the VFs, which would need to change their biochemical/biophysical properties simultaneously if the molecular grammar of the VF suddenly changes. Such a scenario is parsimoniously unlikely. Regardless, it is unclear whether all scaffold protein IDRs share a common origin and diversification occurred by shuffling protein fragments (53) or whether the same molecular grammar emerged in multiple IDRs on different occasions by convergent evolution (54). Which advantages the virus gains by modifying the scaffold proteins that form their VFs remains to be addressed but might be associated with emergent traits of different VFs [including the appearance of an inner layer in mimivirus allowing subcompartmentalization of replication and transcription, or the endoplasmic reticulum wrapping in poxviruses (55)].

Such grammar allowed us to identify previously neglected but well-characterized scaffold proteins in other members of the phylum. In *Poxviridae*, H5 was predicted as the sole candidate for a scaffold protein for phase separation. H5 is an essential protein for the infectious cycle of the vaccinia virus (56) and was coined as a hub protein due to its importance in DNA replication, transcription, and virion morphogenesis (46). All those phenotypes correlate with a scaffolding function for phase separation. Moreover, H5 binds DNA (46) [a function modulated by phosphorylation (47)] and localizes as puncta upon heterologous expression in mammalian cells (indicating spontaneous phase separation of the protein) (56). Interestingly, when vaccinia virus uncoating occurs, early viral proteins associated with DNA replication localize to cytoplasmic puncta, including H5 (46, 57). We propose that these puncta (known as prereplication foci) are likely formed by phase separation using H5 as a scaffold protein. Upon maturation of the prereplication foci, VF gets surrounded by the ER membranes (55). In addition, H5 strongly accumulates inside the VF (58, 59), and the VFs dissolve in the presence of 1,6-hexanediol. Moreover, VFs of the vaccinia virus have previously been shown to coalesce when present in the same cell (60). These data support the idea that phase separation is a major driver of the VF formation regardless of the ER wrapping. It has been proposed that H5 would be a key component for VF enlargement and wrapping by the ER (57). Overall, previously published data on H5 strongly support its role as an unrecognized scaffold protein for phase separation. In *Marseilleviridae*, at least two scaffold proteins localize to the VFs with differential transcriptional expression patterns (61). This allows us to hypothesize that NMV_095 might initiate VF formation while NMV_238 would allow its expansion and maturation. Further experiments will be needed to corroborate this hypothesis. In *Pandoravirales*, multiple IDRs containing a similar molecular grammar to VF scaffold proteins were identified. Regardless, these viruses transfer their DNA into the nucleus of their host. In evolutionary terms, it is parsimonious to assume that a fully cytoplasmic infectious cycle style of the majority of the members of the phylum originated prior to the nuclear one (29), as the origin of *Nucleocytoviricota* predated the origin of eukaryotes (and thus, the nucleus) (52). Thus, during the transition from cytoplasmic to nuclear viruses, the transfer of the DNA into the nucleus would generate a VF, which is no longer needed for replication and transcription but would still retain the functions associated with virion morphogenesis. Nonetheless, only imaging data [electron microscopy (8, 9) and immunofluorescence (21)] support the presence of these structures and further experiments would be needed to characterize such "Virion Factories" potentially assembled by nuclear giant viruses and their functions.

Mimivirus VFs are highly compartmentalized organelle-like structures (*SI Appendix*, Fig. S13). Biogenesis of the VFs starts by utilizing the cores as nucleating points. Similar observations were previously made on vaccinia virus, indicating some conserved mechanisms of nucleation (59). Mimivirus VFs then develop into multilayered structures that contain at least two distinctive phases. The outer layer of the VF, formed by OLS1, acts as a selective barrier and recruits VF proteins. Importantly, while we hypothesize that the outer layer of mimivirus VFs is the phylogenetically conserved phase between different *Nucleocytoviricota*, OLS1 is dispensable in mimivirus. We theorize that the presence of the inner layer allows the protection of the genome of the virus in the absence of the outer layer and, despite losing the ability to selectively recruit proteins to the VF and considering the permissive conditions of the laboratory, the inner layer is sufficient to achieve successful infection. The DNA replication machinery localizes to the outer layer of the VF and maximizes DNA replication only at the interphase between outer layer and inner layer. This interphase is also the site of virion assembly, and the competition between these two processes may at least partially explain the decrease in DNA synthesis during the late stages of the infection cycle. The inner layer of the VF contains proteins associated with transcription, which are also packaged into the virions to establish a new cycle of infection (including the RNA polymerase, transcription factor, RNA processing, etc.). Thus, the inner layer appears to be a compartment analogous to the internal content of the virion core, which is sufficient to restart RNA transcription of early genes upon infection of a new cell (13). Moreover, ILS1 has recently been proposed to work on mimivirus DNA condensation for its incorporation into the viral particle (36). Finally, translation occurs outside of the VF, a feature that differentiates mimivirus from vaccinia virus (62). Importantly, how mRNAs are exported from the VFs remains unknown. Nevertheless, since neither ILS1 nor OLS1 requires RNA for phase separation, a simple model can be envisioned where RNA is not retained by either the inner layer or outer layer of the VF. In such a case, diffusion would be sufficient to deliver mRNAs into the cytoplasm of the infected cell.

Overall, this finding raises questions, including how cytoplasmic viruses interact upon infection of the same host cell (can VFs from different viruses fuse?), and how biomolecular condensates accommodate the various functions required for VFs' multiple roles spatiotemporally. This work opens the door for developing generalist drugs to inhibit *Nucleocytiviricota* viral infections, including monkeypox and African swine fever viruses.

Materials and Methods

***A. castellanii* Growth and Virus Production.** The following viral strains have been used in this study: Acanthamoeba polyphaga mimivirus (49), noumeavirus (29), pithovirus sibericum (63), pacmanvirus Lost City (30), mollivirus sibericum (32), and Modified Vaccinia virus Ankara strain. See *SI Appendix, Supplementary Materials and Methods*. Viral purification was performed as shown in ref. 64.

Generation of DNA Constructs and Viral Lines. See *SI Appendix, Supplementary Materials and Methods*.

Protein Expression and Purification and Phase Separation Assays. See *SI Appendix, Supplementary Materials and Methods*.

Immunofluorescence, Fluorescence Microscopy, and Electron Microscopy Imaging. See *SI Appendix, Supplementary Materials and Methods*.

1,6-Hexanediol Treatment, EU and EdU Labeling. See *SI Appendix, Supplementary Materials and Methods*.

Immunoprecipitation and MS-Based Proteomic Analyses. Immunoprecipitation was performed as previously described (65). Proteins eluted from co-IP experiments were either separated by SDS-PAGE or stacked as previously described (66). Proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (67) (PXD054803). Peptides and proteins were identified by Mascot (version 2.8.3, Matrix Science). The Proline software (68) (version 2.3) was used for the compilation, grouping, and filtering of the results. Statistical analysis was performed using the ProStaR software (69). See *SI Appendix, Supplementary Materials and Methods* for the extended protocol.

Prediction of the VF Scaffold Proteins. A database was constituted with genomes of isolated viruses, predicted proteins from the Giant virus database (4) (PRJEB47746), and *Egovirales* (44). IDRs were predicted using MobiDB-lite v3.10.0 (70). For the residue binary patterns, the python package Nardini v1.1.1 was used. Compositional data and physical and chemical properties of IDRs were predicted with localCIDER v0.1.21 (71). Nardini and CIDER features were normalized. All Z-scores and normalized features were then saturated by the sigmoid function (72). Scaffold protein homologues were identified by MMseqs2 v.12 (73) with an e-value cutoff of 1e-5 and further confirmed by alignment with t-coffee v13.41.0 (74). HMM models were constructed with HMMER v3.3.2 (75). Sequences were considered as homologs only for alignments with e-values $<1e^{-10}$. The final SVM classifier was built on an rbf kernel considering 11 features whose Spearman correlation coefficient is under 0.55. These features included two Nardini features and nine CIDER features. MolPhase (76) and ParSe v2 (77) were also tested. See *SI Appendix, Supplementary Materials and Methods* for an extended method.

Data, Materials, and Software Availability. All the codes used for the bioinformatic analysis and for the VFCpredict tool developed in this study are available at https://src.koda.cnrs.fr/igs/vfcpredict (78). The mass spectrometry proteomics raw data have been deposited to the ProteomeXchange Consortium via the PRIDE (79) partner repository with the dataset identifier PXD054803 (80). All other data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank Matias Estaras Hermosel, Christian Betzel, Gamaleldin I. Harisa, Chunfu Zheng, Anouk Willemsen, Andrea Maria Guarino, Yves Gaudin, Juliana Cortines, and Kekungu-u Puro for the comments raised on the preprint version of the manuscript. We also thank the members of the Shared Imaging Platform of Luminy Campus-France BioImaging core facility (Nicolas Brouilly, Fabrice Richard, and Aïcha Aouane), Institut de Biologie du Développement de Marseille, AMU-Marseille, and the Institut de Microbiologie de la Méditerranée imaging platform (Artemis Kosta). N.O.-D. and M.B. gratefully acknowledge the Advanced Bioimaging Unit at the Institut Pasteur Montevideo & Universidad de la República. This work was funded by European Research Council grant agreement No. 101160452, ERC-2024-STG (H.B.), and European Research Council grant agreement No 832601 ERC-2018-ADG (C.A.). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council (ERC). Neither the European Union nor the granting authority can be held responsible for them. This work was also funded by Agence Nationale de la Recherche grant ProFI, ANR-10-INBS-08 (Y.C.), and Chemistry Biology Health Graduate School of University Grenoble Alpes grant ANR-17-EURE-0003 (Y.C.).

Author affiliations: [a]Information Génomique and Structurale, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 7256 (Institut de Microbiologie de la Méditerranée, FR3479), Institut microbiologie, bioénergies et biotechnologie, Origins Institute Marseille, Aix-Marseille University, Marseille Cedex 9 13288, France; [b]Division of Microbial Ecology, Centre for Microbiology and Environmental Systems Science, University of Vienna, Vienna 1030, Austria; [c]Biosciences et Bioingénierie pour la Santé, Université Grenoble Alpes, INSERM, Commissariat à l'énergie atomique et aux énergies alternatives, CNRS, CEA, FR2048, Grenoble 38000, France; [d]Immunovirology Lab, Institut Pasteur de Montevideo, Montevideo 11400, Uruguay; [e]Microscopy Core Facility, CNRS, Institut de Microbiologie de la Méditerranée, Aix-Marseille University, Marseille 13009, France; and [f]Virology department, Instituto de Química Biológica, Facultad de Ciencias, Universidad de la República, Montevideo 11600, Uruguay

1. P. V. M. Boratto *et al.*, Yaravirus: A novel 80-nm virus infecting Acanthamoeba castellanii. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 16579–16586 (2020).
2. N. Philippe *et al.*, Pandoraviruses: Amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* **341**, 281–286 (2013).
3. M. Legendre *et al.*, Diversity and evolution of the emerging Pandoraviridae family. *Nat. Commun.* **9**, 2285 (2018).
4. F. O. Aylward, M. Moniruzzaman, A. D. Ha, E. V. Koonin, A phylogenomic framework for charting the diversity and evolution of giant viruses. *PLoS Biol.* **19**, e3001430 (2021).
5. E. V. Koonin *et al.*, Global organization and proposed megataxonomy of the virus world. *Microbiol. Mol. Biol. Rev.* **84**, https://doi.org/10.1128/mmbr.00061-19 (2020).
6. E. V. Koonin, J. H. Kuhn, V. V. Dolja, M. Krupovic, Megataxonomy and global ecology of the virosphere. *ISME J.* **18**, wrad042 (2024).
7. T. Bosmon, C. Abergel, J. M. Claverie, 20 years of research on giant viruses. *NPJ Viruses* **3**, 9 (2025).
8. A. Pereira Andrade *et al.*, New isolates of pandoraviruses: Contribution to the study of replication cycle steps. *J. Virol.* **93**, e01942-18 (2019).
9. E. Milrot *et al.*, Virus-host interactions: Insights from the replication cycle of the large Paramecium bursaria chlorella virus. *Cell Microbiol.* **18**, 3–16 (2016).
10. Y. G. Kuznetsov, T. Klose, M. Rossmann, A. McPherson, Morphogenesis of mimivirus and its viral factories: An atomic force microscopy study of infected cells. *J. Virol.* **87**, 11200–11213 (2013).
11. F. Schulz *et al.*, Giant virus diversity and host interactions through global metagenomics. *Nature* **578**, 432–436 (2020).
12. F. Schulz, C. Abergel, T. Woyke, Giant virus biology and diversity in the era of genome-resolved metagenomics. *Nat. Rev. Microbiol.* **20**, 721–736 (2022).
13. Y. Mutsafi, N. Zauberman, I. Sabanay, A. Minsky, Vaccinia-like cytoplasmic replication of the giant Mimivirus. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5978–5982 (2010).
14. J. R. Cortines *et al.*, Transition metals and oxidation reactions trigger stargate opening during the initial stages of the replicative cycle of the giant Tupanvirus. *mBio* **15**, e0219224 (2024).
15. J. R. Schrad, J. S. Abrahao, J. R. Cortines, K. N. Parent, Structural and proteomic characterization of the initiation of giant virus infection. *Cell* **181**, 1046–1061.e6 (2020).
16. N. Zauberman *et al.*, Distinct DNA exit and packaging portals in the virus *Acanthamoeba polyphaga* mimivirus. *PLoS Biol.* **6**, e114 (2008).
17. J. M. Claverie, C. Abergel, Mimivirus and its virophage. *Annu. Rev. Genet.* **43**, 49–66 (2009).
18. D. Arslan, M. Legendre, V. Seltzer, C. Abergel, J. M. Claverie, Distant mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17486–17491 (2011).
19. A. Villalta *et al.*, The giant mimivirus 1.2 Mb genome is elegantly organized into a 30-nm diameter helical protein shield. *Elife* **11**, e77607 (2022).
20. Y. Fridmann-Sirkis *et al.*, Efficiency in complexity: Composition and dynamic nature of Mimivirus replication factories. *J. Virol.* **90**, 10039–10047 (2016).
21. H. Bisio *et al.*, Evolution of giant pandoravirus revealed by CRISPR/Cas9. *Nat. Commun.* **14**, 428 (2023).
22. X. Zhang, R. Zheng, Z. Li, J. Ma, Liquid-liquid phase separation in viral function. *J. Mol. Biol.* **435**, 167955 (2023).
23. M. Charman *et al.*, A viral biomolecular condensate coordinates assembly of progeny particles. *Nature* **616**, 332–338 (2023).
24. T. Zarin *et al.*, Identifying molecular features that are associated with biological function of intrinsically disordered protein regions. *Elife* **10**, e60220 (2021).
25. M. C. Cohan, M. K. Shinn, J. M. Lalmansingh, R. Pappu, Uncovering non-random binary patterns within sequences of intrinsically disordered proteins. *J. Mol. Biol.* **434**, 167373 (2022).
26. N. Philippe, A. Shukla, C. Abergel, H. Bisio, Genetic manipulation of giant viruses and their host. *Nat. Protoc.* **19**, 3–29 (2024).
27. J. M. Alempic *et al.*, Functional redundancy revealed by the deletion of the mimivirus GMC-oxidoreductase genes. *Microlife* **5**, uqae006 (2024).
28. Y. Liu *et al.*, Virus-encoded histone doublets are essential and form nucleosome-like structures. *Cell* **184**, 4237–4250.e19 (2021).
29. E. Fabre *et al.*, Noumeavirus replication relies on a transient remote control of the host nucleus. *Nat. Commun.* **8**, 15087 (2017).
30. S. Santini *et al.*, Pacmanvirus isolated from the Lost City hydrothermal field extends the concept of transpoviron beyond the family Mimiviridae. *ISME J.* **19**, wraf002 (2025).
31. R. Hancock, A role for macromolecular crowding effects in the assembly and function of compartments in the nucleus. *J. Struct. Biol.* **146**, 281–290 (2004).
32. M. Legendre *et al.*, In-depth study of *Mollivirus sibericum*, a new 30, 000-y-old giant virus infecting *Acanthamoeba*. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5327–E5335 (2015).
33. F. Yang *et al.*, Organization of microtubule plus-end dynamics by phase separation in mitosis. *J. Mol. Cell Biol.* **16**, mjae006 (2024).
34. R. Duster, I. H. Kaltheuner, M. Schmitz, M. Geyer, 1, 6-hexanediol, commonly used to dissolve liquid-liquid phase separated condensates, directly impairs kinase and phosphatase activities. *J Biol Chem* **296**, 100260 (2021).
35. D. Bracha, M. T. Walls, C. P. Brangwynne, Probing and engineering liquid-phase organelles. *Nat. Biotechnol.* **37**, 1435–1445 (2019).
36. D. Sharma, F. Coulibaly, K. Kondabagil, Mimivirus encodes an essential MC1-like non-histone architectural protein involved in DNA condensation. bioRxiv [Preprint] (2024). https://doi.org/10.1101/2024.02.22.580433 (Accessed 13 August 2025).
37. A. Boija *et al.*, Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855.e16 (2018).
38. J. Wang *et al.*, A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699.e16 (2018).
39. J. A. Greig *et al.*, Arginine-enriched mixed-charge domains provide cohesion for nuclear speckle condensation. *Mol. Cell* **77**, 1237–1250.e4 (2020).
40. H. Lyons *et al.*, Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell* **186**, 327–345.e28 (2023).
41. A. Patil *et al.*, A disordered region controls cBAF activity via condensation and partner recruitment. *Cell* **186**, 4936–4955.e26 (2023).

42. M. R. King *et al.*, Macromolecular condensation organizes nucleolar sub-phases to set up a pH gradient. *Cell* **187**, 1889–1906.e24 (2024).
43. S. Rigou, S. Santini, C. Abergel, J. M. Claverie, M. Legendre, Past and present giant viruses diversity explored through permafrost metagenomics. *Nat. Commun.* **13**, 5853 (2022).
44. M. Gaïa *et al.*, Closest relatives of poxviruses are spread in the gut of humans and animals worldwide: The egoviruses. biorxiv [Preprint] (2025). https://doi.org/10.1101/2024.03.23.586382 (Accessed 13 August 2025).
45. G. Erdos, Z. Dosztanyi, Analyzing protein disorder with IUPred2A. *Curr. Protoc. Bioinformatics* **70**, e99 (2020).
46. N. E. Kay *et al.*, Biochemical and biophysical properties of a putative hub protein expressed by vaccinia virus. *J. Biol. Chem.* **288**, 11470–11481 (2013).
47. Y. Huang *et al.*, Multi-omics characterization of the monkeypox virus infection. *Nat. Commun.* **15**, 6778 (2024).
48. K. L. B. Borden, L. Volpon, The diversity, plasticity, and adaptability of cap-dependent translation initiation and the associated machinery. *RNA Biol.* **17**, 1239–1251 (2020).
49. D. Raoult *et al.*, The 1.2-megabase genome sequence of Mimivirus. *Science* **306**, 1344–1350 (2004).
50. W. K. Joklik, Y. Becker, The replication and coating of vaccinia DNA. *J. Mol. Biol.* **10**, 452–474 (1964).
51. K. Kamagata *et al.*, Molecular principles of recruitment and dynamics of guest proteins in liquid droplets. *Sci. Rep.* **11**, 19323 (2021).
52. J. Guglielmini, A. C. Woo, M. Krupovic, P. Forterre, M. Gaia, Diversification of giant and large eukaryotic dsDNA viruses predated the origin of modern eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19585–19592 (2019).
53. B. J. Smug, K. Szczepaniak, E. P. C. Rocha, S. Dunin-Horkawicz, R. J. Mostowy, Ongoing shuffling of protein fragments diversifies core viral functions linked to interactions with bacterial hosts. *Nat. Commun.* **14**, 7460 (2023).
54. J. F. Storz, Causes of molecular convergence and parallelism in protein evolution. *Nat. Rev. Genet.* **17**, 239–250 (2016).
55. N. Tolonen, L. Doglio, S. Schleich, J. Krijnse Locker, Vaccinia virus DNA replication occurs in endoplasmic reticulum-enclosed cytoplasmic mini-nuclei. *Mol. Biol. Cell* **12**, 2031–2046 (2001).
56. K. A. Boyle, M. D. Greseth, P. Traktman, Genetic confirmation that the H5 protein is required for vaccinia virus DNA replication. *J. Virol.* **89**, 6312–6327 (2015).
57. G. Beaud, R. Beaud, Preferential virosomal location of underphosphorylated H5R protein synthesized in vaccinia virus-infected cells. *J. Gen. Virol.* **78**, 3297–3302 (1997).
58. M. D. Greseth, P. Traktman, The life cycle of the vaccinia virus genome. *Annu. Rev. Virol.* **9**, 239–259 (2022).
59. M. Mallardo *et al.*, Relationship between vaccinia virus intracellular cores, early mRNAs, and DNA replication sites. *J. Virol.* **76**, 5167–5183 (2002).
60. Y. C. J. Lin, D. H. Evans, Vaccinia virus particles mix inefficiently, and in a way that would restrict viral recombination, coinfected cells. *J. Virol.* **84**, 2432–2443 (2010).
61. R. A. L. Rodrigues *et al.*, Analysis of a Marseillevirus transcriptome reveals temporal gene expression profile and host transcriptional shift. *Front. Microbiol.* **11**, 651 (2020).
62. G. C. Katsafanas, B. Moss, Colocalization of transcription and translation within cytoplasmic poxvirus factories coordinates viral expression and subjugates host functions. *Cell Host Microbe* **2**, 221–228 (2007).
63. M. Legendre *et al.*, Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4274–4279 (2014).
64. L. Bertaux, A. Lartigue, S. Jeudy, Giant mimiviridae CsCl purification protocol. *Bio-protocol* **10**, e3827 (2020).
65. H. Bisio, A. Krishnan, J. B. Marq, D. Soldati-Favre, *Toxoplasma gondii* phosphatidylserine flippase complex ATP2B-CDC50.4 critically participates in microneme exocytosis. *PLoS Pathog.* **18**, e1010438 (2022).
66. M. G. Casabona, Y. Vandenbrouck, I. Attree, Y. Coute, Proteomic characterization of Pseudomonas aeruginosa PAO1 inner membrane. *Proteomics* **13**, 2419–2423 (2013).
67. Y. Perez-Riverol *et al.*, The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).
68. D. Bouyssie *et al.*, Proline: An efficient and user-friendly software suite for large-scale proteomics. *Bioinformatics* **36**, 3148–3155 (2020).
69. Y. Coute, C. Bruley, T. Burger, Beyond target-decoy competition: Stable validation of peptide and protein identifications in mass spectrometry-based discovery proteomics. *Anal. Chem.* **92**, 14898–14906 (2020).
70. M. Necci, D. Piovesan, Z. Dosztanyi, S. C. E. Tosatto, MobiDB-lite: Fast and highly specific consensus prediction of intrinsic disorder in proteins. *Bioinformatics* **33**, 1402–1404 (2017).
71. A. S. Holehouse, R. K. Das, J. N. Ahad, M. O. Richardson, R. V. Pappu, CIDER: Resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophys. J.* **112**, 16–21 (2017).
72. Y. LeCun, L. Bottou, G. B. Orr, K. R. Müller, Efficient backprop. *Lect. Notes Comput. Sci.* **1524**, 9–50 (1998).
73. M. Steinegger, J. Soding, MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
74. C. Notredame, D. G. Higgins, J. Heringa, T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).
75. S. R. Eddy, A new generation of homology search tools based on probabilistic inference. *Genome Inform.* **23**, 205–211 (2009).
76. Q. Liang *et al.*, MolPhase, an advanced prediction algorithm for protein phase separation. *EMBO J.* **43**, 1898–1918 (2024).
77. A. Y. Ibrahim *et al.*, Intrinsically disordered regions that drive phase separation form a robustly distinct protein class. *J. Biol. Chem.* **299** (2023).
78. S. Rigou *et al.*, VFCPredict. GitHub. https://src.koda.cnrs.fr/igs/vfcpredict. Deposited 24 July 2024.
79. Y. Perez-Riverol *et al.*, The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552. (2022).
80. S. Rigou *et al.*, Data from "Evolutionarily conserved grammar rules viral factories of amoeba-infecting members of the hyperdiverse Nucleocytoviricota phylum." ProteomeXchange Consortium. https://www.ebi.ac.uk/pride/archive/projects/PXD054803. Deposited 12 August 2024.