# Giant Viruses Encode Actin-Related Proteins

Violette Da Cunha,*[†,1] Morgan Gaia,[†,2] Hiroyuki Ogata ⓘ,[3] Olivier Jaillon,[2,4] Tom O. Delmont,[2] and Patrick Forterre* [,1,5]

[1]CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Université Paris-Saclay, Gif-sur-Yvette, France
[2]Génomique Métabolique, Génoscope, Institut François Jacob, CEA, CNRS, Université d'Évry, Université Paris-Saclay, Evry, France
[3]Institute for Chemical Research, Kyoto University, Gokasho, Uji, Japan
[4]Research Federation for the Study of Global Ocean Systems Ecology and Evolution, Tara Oceans, France
[5]Département de Microbiologie, Institut Pasteur, Paris, France
[†]These authors contributed equally to this work.
*Corresponding authors: E-mails: violette.da.cunha.vdc@gmail.com; patrick.forterre@pasteur.fr.
Associate editor: Emma Teeling

## Abstract

The emergence of the eukaryotic cytoskeleton is a critical yet puzzling step of eukaryogenesis. Actin and actin-related proteins (ARPs) are ubiquitous components of this cytoskeleton. The gene repertoire of the Last Eukaryotic Common Ancestor (LECA) would have therefore harbored both actin and various ARPs. Here, we report the presence and expression of actin-related genes in viral genomes (viractins) of some *Imitervirales*, a viral order encompassing the giant *Mimiviridae*. Phylogenetic analyses suggest an early recruitment of an actin-related gene by viruses from ancient protoeukaryotic hosts before the emergence of modern eukaryotes, possibly followed by a back transfer that gave rise to eukaryotic actins. This supports a coevolutionary scenario between pre-LECA lineages and their viruses, which could have contributed to the emergence of the modern eukaryotic cytoskeleton.

*Key words:* actin and actin-related proteins, NucleoCytoplasmic Large DNA virus, viral eukaryogenesis.

The actin family is a diverse and evolutionarily ancient group of proteins containing the conventional actin and many paralogs named ARPs (actin-related proteins) (Goodson and Hawse 2002). Different ARP families have been defined based on their degree of similarity to actin (Schroer et al. 1994). Eight ARPs (ARP1–ARP8) are conserved in most or all eukaryotes (Muller et al. 2005), suggesting that, besides conventional actin, the Last Eukaryotic Common Ancestor (LECA) already had a set of several ARPs. Actin and most ARP sequences are usually around 400 aa long, even if some ARPs are larger. Conventional actin is present in both the cytoplasm and the nucleus of eukaryotes, whereas some ARPs are primarily localized in the cytoplasm (ARP1-3, ARP10) or the nucleus (ARP4-6, ARP8) (Muller et al. 2005; Bajusz et al. 2018). Cytoplasmic ARPs are involved in key functions such as the spatiotemporal control of actin assembly and movement of vesicles along microtubules. Nuclear ARPs are implicated in processes like chromatin modulation, transcription regulation, and DNA repair (Kristó et al. 2016). It has been proposed that these different ARPs paralogs originated in the protoeukaryotic lineage, that is, in the cellular ancestors of eukaryotes that throve before the emergence of LECA, and/or evolved from distant archaeal actin homologs (crenactins) (Stairs and Ettema 2020). Recently, it has been hypothesized that eukaryotic-like actins detected in Asgard archaea (hereafter dubbed asgardactins) could have been at the origin of

eukaryotic actin and ARPs (Lindås et al. 2017; Zaremba-Niedzwiedzka et al. 2017). Some Bacteria also encode actin-like proteins (Miller et al. 2007; Shiratori et al. 2019) corresponding to xenologs originating from eukaryotes or protoeukaryotes (Miller et al. 2007).

Here, we report the discovery of actin genes (∼400 aa long) in the genomes of several viruses belonging to the *Imitervirales*, a viral order that is part of the *Nucleocytoviricota* (Koonin et al. 2020) phylum, previously known as the NucleoCytoplasmic Large DNA virus assemblage (NCLDV) (Koonin and Yutin 2019) (table 1). We notably detected one of these viral actin-like genes (hereafter dubbed viractin) in the genome of the giant Yasminevirus, recently isolated from sewage water through amoeba coculture (Bajrai et al. 2019). Yasminevirus belongs to the proposed *Klosneuvirinae* subfamily (Schulz et al. 2017; Bajrai et al. 2019) within the *Mimiviridae* family (Koonin and Yutin 2019; Guglielmini et al. 2019). We detected additional viractins in 18 metagenome-assembled genomes (MAGs) of marine and freshwater *Nucleocytoviricota* (16 published; Schulz et al. 2017, 2020; Moniruzzaman, Martinez-Gutierrez, et al. 2020; and 2 that we reconstructed from the sunlit ocean survey of the *Tara* Oceans expeditions; Endo et al. 2020; see Supplementary Material online). The genomic statistics for these 19 viractin-encoding viruses (18 MAGs and one isolate) are summarized in supplementary table S1, Supplementary Material online.

**Open Access**

**Table 1.** The Newly Identified Viractins.

| Viractin | Name | Lineage | Identity % to Human Actin | Source |
|---|---|---|---|---|
| Viractin 01 | Schulz_GVMAG_M_3300003404_4 | Clade MVGL55 MAG | 71.8 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_M_3300023184_120 | Clade MVGL55 MAG | 73.6 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_M_3300027707_43 | Clade MVGL55 MAG | 73.4 | Schulz et al. (2020) |
| Viractin 01 | Moniruzzaman_MM01_SRX802077_122_dc | Clade MVGL55 MAG | 71.5 | Moniruzzaman, Martinez-Gutierrez, et al. (2020) |
| Viractin 01 | Schulz_GVMAG_M_3300009441_23 | Clade MVGL55 MAG | 67.3 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_M_3300020185_80 | Clade MVGL55 MAG | 71.4 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_M_3300020187_94 | Clade MVGL55 MAG | 69.2 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_M_3300021375_17 | Clade MVGL55 MAG | 68.6 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_M_3300024261_25 | Clade MVGL55 MAG | 71.5 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_M_3300025767_21 | Clade MVGL55 MAG | 64.7 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_M_3300025880_30 | Clade MVGL55 MAG | 71.5 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_S_3300010368_67 | Clade MVGL55 MAG | 71.2 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_S_3300010368_72 | Clade MVGL55 MAG | 69 | Schulz et al. (2020) |
| Viractin 01 | Schulz_GVMAG_S_ERX556045_64 | Clade MVGL55 MAG | 71 | Schulz et al. (2020) |
| Viractin 01 | TARA_MED_NCLDV_Bin_121_14_c | Clade MVGL55 MAG | 72 | This study |
| Viractin 02 | Moniruzzaman_MM15_ERX556043_42_dc | Clade MVGL15 MAG | 66.9 | Moniruzzaman, Martinez-Gutierrez, et al. (2020) |
| Viractin 03 | Yasminevirus_GU_2018 | Yasminevirus | 59.2 | Bajrai et al. (2019) |
| Viractin 03 | TARA_PSW_NCLDV_Bin_158_1_c | Yasminevirus MAG | 55.4 | This study |
| Viractin 04 | Hokovirus_KY684103 | Hokovirus MAG | 27[a] | Schulz et al. (2017) |

NOTE.—For all viractins, the identity percentage is given for a coverage ranging from 99% to 100% to the human actin, except for viractin 04 ([a]) where the coverage is 73%. The summary of the 19 genomes of Imitervirales containing a viractin is provided in supplementary table S1, Supplementary Material online.

These viruses were confirmed to belong to the *Imitervirales* order through phylogenetic analyses of conserved markers from five representative viruses (supplementary fig. S1, Supplementary Material online). All the published viractin-encoding *Nucleocytoviricota* MAGs had already been described as related to *Mimiviridae*; 15 belong to a lineage dubbed MVGL55 and comprising viruses identified in various aquatic environments (supplementary table S1, Supplementary Material online), and one to a lineage dubbed MM15. These clades correspond to the newly revealed diversity of *Mimiviridae* relatives (Moniruzzaman, Martinez-Gutierrez, et al. 2020; Schulz et al. 2020) that have recently been incorporated into the *Imitervirales* (Koonin et al. 2020) order. We also detected one viractin-encoding MAG related to Yasminevirusand one to Hokovirus (table 1). Interestingly, the two viractin-encoding Yasmineviruses were characterized from very different environments: sewage water in Saudi Arabia and the Pacific Ocean. Thus, although the detection of viractin in the isolated Yasminevirus confirms its genuine presence in the viral genome, environmental surveys reveal its distribution in multiple clades of *Mimiviridae*-related viruses, always in single copy.

In addition, we detected viractin mRNA in polyA metatranscriptomes of plankton obtained from the same seawater samples used to identify viractin-encoding MAGs (notably in the Pacific Ocean and Mediterranean Sea) (Pesant et al. 2015) (supplementary table S2, Supplementary Material online). This observation strongly suggests that viractins are expressed during infection, as suggested by a separate analysis conducted on one virus of the MVGL55 clade (Ha et al. 2021). Predicted protein structure comparison (see Supplementary Material online) supports that these viractins are close structural homologs of eukaryotic actin and ARPs (fig. 1 and

supplementary table S3, Supplementary Material online), suggesting similar biochemical functions and a common ancestry.

To investigate the origins of viractins, notably with respect to the frequent assumption of recurrent gene recruitment by viruses from their hosts, we performed a global phylogenetic analysis of actin-related sequences that included crenactin and asgardactin sequences, as well as actin-like sequences that we unexpectedly identified in some *Bathyarchaea* (hereafter dubbed bathyactins) (Dombrowski et al. 2018; Zhou et al. 2018) (see Supplementary Material online). The actin, crenactin, bathyactins, and eukaryotic ARPs all formed clear monophyletic clades (fig. 1 and supplementary fig. S2, Supplementary Material online), whereas the asgardactins were distributed into five clades, suggesting ancient transfer events between their ancestors and protoeukaryotes. The 19 viractins were structured in four clades, viractins 01-04, corresponding to the four different groups of viractin-encoding *Imitervirales* (table 1 and fig. 1; supplementary fig. S1, Supplementary Material online). Surprisingly, viractins branched neither within eukaryotic actins nor in any of the eukaryotic clades of ARPs, as would be expected in the case of recent transfers from modern eukaryotes. Instead, the eukaryotic actin clade was positioned within the viractins 01 to 03, altogether forming a monophyletic sister clade to ARP1 and ARP2. The viractin 04 single protein sequence was located separately, at the base of a clade grouping asgardactins 04-05 with ARP10, suggesting a separate evolutionary scenario. A structure-based alignment (supplementary fig. S3, Supplementary Material online) and protein similarity network analyses (supplementary fig. S4, Supplementary Material online) corroborated this pattern of relationships, with viractins 01 to 03 more closely related to actin, ARP1 and ARP2
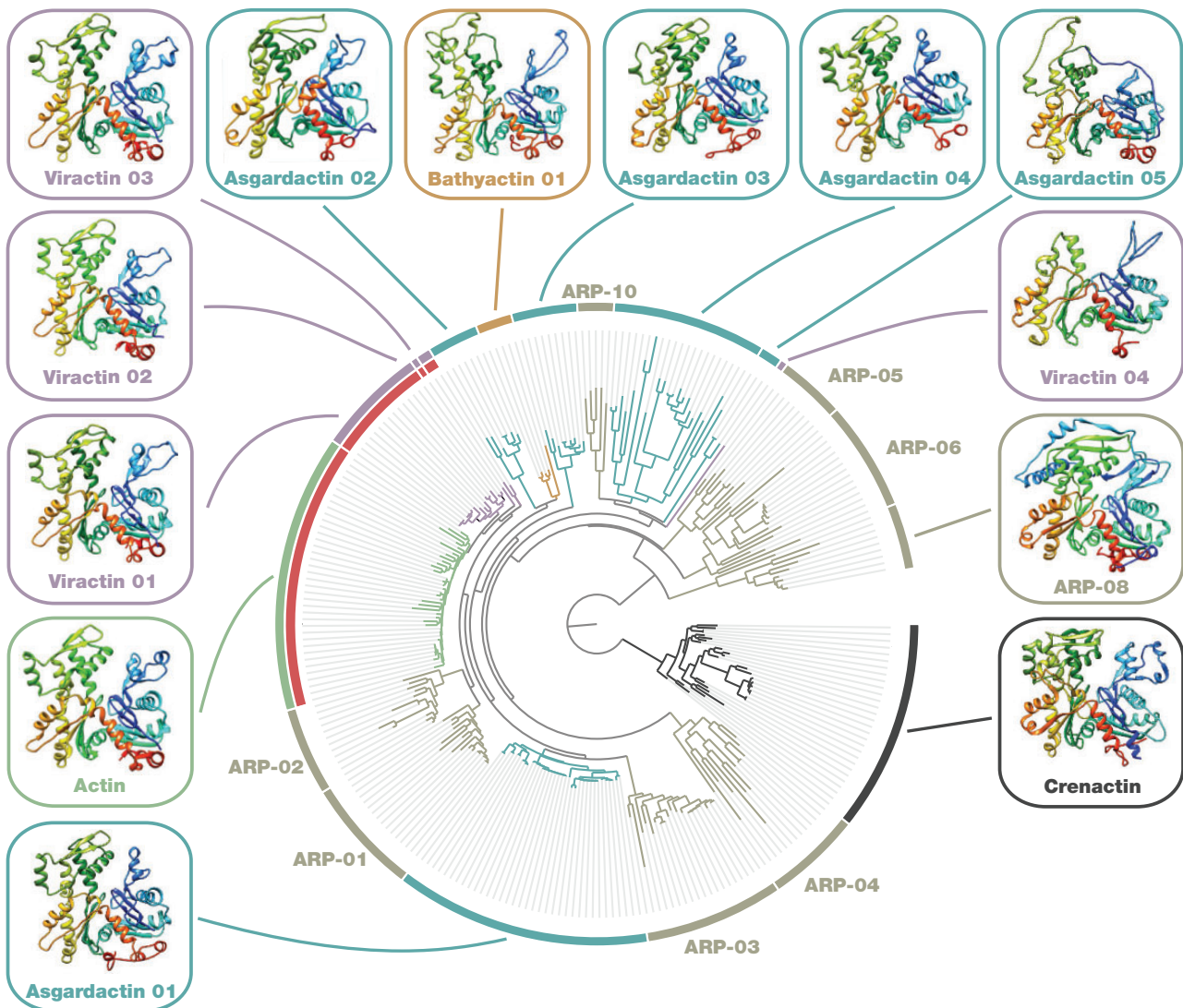
**Fig. 1.** Viractins within the different ARPs and actin clades of our initial protein sequence data set. Maximum-likelihood phylogenetic tree of our initial protein data set (GD), which included eukaryotic actin (green) and ARP (yucky green), asgardactin (light blue-green), newly identified bathyactin (orange), viractin (purple), and crenactin (black) sequences. These latter were used as outgroup. The predicted structures obtained for representatives of different clades (supplementary table S2, Supplementary Material online) are shown, as well as the structures for actin and ARP-8. The positions of the conventional actin and of the closely related viractin 01, 02, and 03 sequences are indicated in red.

proteins, but also to a lesser extent to ARP3, ARP4, ARP6, and possibly asgardactin and bathyactin sequences.

All these analyses pointed to a particularly complex evolutionary history of the actin and ARP families, reminiscent of previous conclusions from an analysis addressing the origin of LECA's proteins (Vosseberg et al. 2020). Considering the challenges of reconstructing such a complex evolutionary history, we decided to focus our subsequent analyses on the relationships between eukaryotic conventional actins and related viractins to limit potential long-branch artifacts. A subset of sequences excluding the most distant clades was selected based on a similarity network analysis (supplementary fig. S4, Supplementary Material online). The phylogenetic tree obtained displayed the same pattern of relationships (supplementary fig. S5, Supplementary Material online). In order to check if the position of viractins 01 to 03 at the base of eukaryotic actins could be due to a putative attraction artifact, we performed a clustering analysis of the associated distance matrix (see Supplementary Material online), which showed that viractins do not display larger distances relative to conventional actin sequences (supplementary fig. S6, Supplementary Material online). To further improve the resolution of the phylogenetic tree, we focused on actins and viractins using first the ARP1/ARP2 clade as outgroup (supplementary fig. S7, Supplementary Material online), and removed the few unstable actin sequences with relatively long branches and high transfer index values according to a Transfer Bootstrap Expectation (TBE) analysis (supplementary figs. S6 and S7, Supplementary Material online). These sequences could indeed correspond to paralogs of actin genes (Goodson and Hawse 2002; Schroeder et al. 2020). Finally, we then decided to remove ARP2 sequences, which had relatively long branches, and only used ARP1 sequences as outgroup (fig. 2 and supplementary fig. S8, Supplementary Material online).
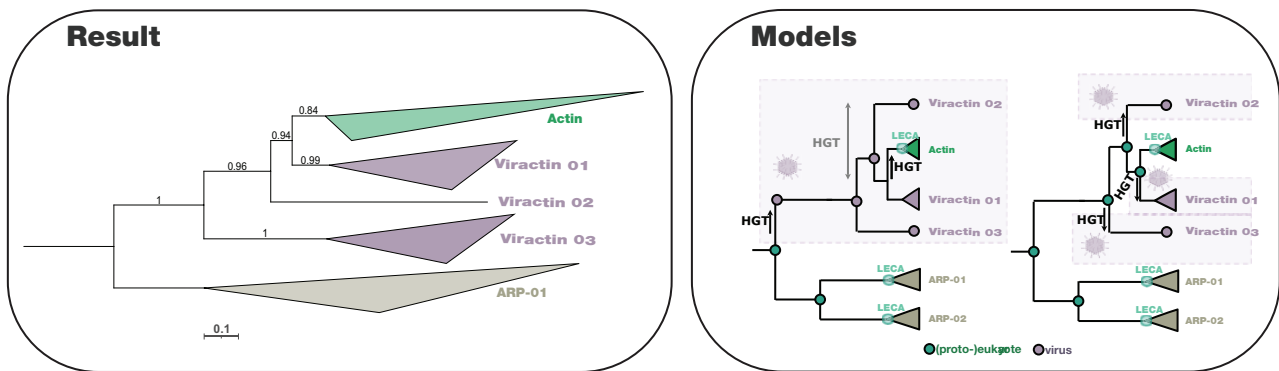
**Fig. 2.** Evolutionary relationships between conventional actin and viractins, and putative scenarios. The left panel displays a phylogenetic tree of eukaryotic actin (green) and viractin 01–03 (purple) using eukaryotic ARP-01 (yucky green) sequences as the outgroup (377 positions). Values at nodes represent their support computed by transfer bootstrap estimation (detailed tree in supplementary fig. S4, Supplementary Material online; aLRT and UFBoot support for the same tree are provided in supplementary fig. S8, Supplementary Material online). The scale bar indicates the average number of substitutions per site. The right panel displays two hypothetical scenarios of evolution for the actin based on new insights provided by the viractins.

This phylogenetic tree confirmed our initial results (fig. 2 and supplementary fig. S8, Supplementary Material online). Conventional actins confidently branched within paraphyletic viractins 01 to 03, as sister clade to viractin 01 with high TBE support. The vast majority of branches were highly supported, with minor exceptions within the conventional actin clade, highlighting the importance of dealing with unstable taxa. This tree still displayed the monophyly of the conventional actin clade representing the diversity of modern eukaryotes. Indeed, the eukaryotic actin, but also the ARP clades have a wide distribution among protists and pluricellular eukaryotes from different supergroups, including Amorphae, Archaeaplastida, TSAR, and Excavates (Burki et al. 2020), and were hence most likely acquired before LECA. Consequently, most nodes at the base of the eukaryotic clades in our trees correspond to the relative position of LECA. Since viractins, and particularly the large viractin 01 clade, branched between rather than within the eukaryotic clades, a scenario of multiple and independent recent acquisitions of the viractins from modern actins or ARPs of their hosts seems unlikely. Instead, ancestral recruitments by viruses of proto-ARP genes from protoeukaryotes appear more plausible.

Considering the position of conventional actins within viractins 01 to 03, a tempting hypothesis is that the ancestor of conventional actins was recruited by a protoeukaryotic ancestor of LECA from an *Imitervirales* (fig. 2, left model of the right panel), possibly following the integration of a viractin-encoding viral genome in the chromosome of its protoeukaryotic host. Recent observations have indicated that the integration of *Imitervirales* genomes is indeed a rather frequent phenomenon (Moniruzzaman, Weinheimer, et al. 2020), which have occurred from before the emergence of modern eukaryotes (Guglielmini et al. 2019). However, although it is possible that an increased representation eventually leads to a single viractin clade instead of three, one cannot exclude three independent transfers from different protoeukaryotic ancestors to different lineages of *Imitervirales* that took place after the divergence between

the ARP1/2 clade and conventional actin (fig. 2, right model of the right panel). In both cases, the distribution of viractin-encoding genomes among *Imitervirales*, with one relatively large subclade (MVGL55) and more distant individual genomes, suggests either several gene losses in all other related subclades, or, more likely, horizontal transfer between these viruses. Further exploration of the actual diversity of *Nucleocytoviricota* from various environments will certainly help to clarify these scenarios.

Actin and ARPs are paramount features of the eukaryotic cytoskeleton involved in various cellular processes (Goodson and Hawse 2002). Here, we report the discovery of xenolog actin genes in the viral world, which we dubbed viractins. Besides their origin, the identification of viractins raises many questions regarding the extent of the viral-driven regulation of their hosts' cytoskeleton dynamics. Actin-regulating activities are elicited by the calcium-regulated multidomain gelsolin family; LECA likely possessed a protein comprising three gelsolin domains (Archer et al. 2005; Ghoshdastider et al. 2013). We did not find proteins with three gelsolin domains in viractin-encoding *Nucleocytoviricota*. This suggests that viractins interact with gelsolin from their hosts. However, we detected one such protein that we called virgelsolin in the MAG of Klosneuvirus KNV1 (ARF12218) (supplementary table S3, Supplementary Material online), which does not possess a viractin gene, and that could interact with the actin of their hosts. This new discovery further suggests that manipulation of the host cytoskeleton is part of the infection strategy of *Imitervirales*. Recently, it has been shown that some viruses of the class *Caudoviricetes* produce a nucleus-like structure within infected bacterial cells (Chaikeeratisak et al. 2017). These viruses encode a distant homolog of the eukaryotic tubulin that localizes this nucleus-like structure in the middle of the virocell (i.e., the virus-infected cell) (Forterre 2017) and treadmills toward itself the viral capsids that were assembled on the membrane (Chaikeeratisak et al. 2019). It is possible that viractins play a similar role during viral infection by controlling the localization of the viral factory close to the host

nucleus (e.g., as seen in Yasminevirus; Bajrai et al. 2019), or could be involved in the formation of an actin cocoon around the viral factories, similar to those recently detected around *Shigella*-containing vacuoles (Kühn et al. 2020). Deciphering the role of viractins and virgelsolin during viral infection will be an exciting challenge.

The origin of the eukaryotic cytoskeleton, with actin and many ARPs already present in LECA (Muller et al. 2005), is a major step of eukaryogenesis and yet remains obscure (Akıl et al. 2021). The identification of viral actin-like genes sharing a common ancestry with eukaryotic conventional actins now brings a new piece to the puzzle. A few studies have previously emphasized a possible evolutionary relationship between the eukaryotic nucleus and *Nucleocytoviricota*' viral factories (Forterre and Gaïa 2016 and herein; Bell 2020), but no connection had been made between these viruses and the eukaryotic cytoskeleton until very recently (Kijima et al. 2021). In this context, it is interesting to note that actin is present in both the cytoplasm and the nucleus of eukaryotic cells and with nuclear ARPs seems to be involved in several nuclear-related processes (Bajusz et al. 2018).

The discovery of viractins closely related to the eukaryotic actin echoes the close association between *Imitervirales* and several eukaryotic signature features (DNA-dependent RNA polymerase II [Guglielmini et al. 2019], histones [Boyer et al. 2009; Yoshikawa et al. 2019], DNA polymerase [Takemura et al. 2015], kinesin [Subramaniam et al. 2020], myosin [Ha et al. 2021; Kijima et al. 2021]), supporting the hypothesis of a coevolution between *Nucleocytoviricota* and protoeukaryotes that could have contributed in shaping the molecular components and functions of the modern eukaryotic cell.

## Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

## Acknowledgments

## Author Contributions

Conceptualization: V.D.C., M.G., and P.F. Methodology: V.D.C., M.G., and T.O.D. Investigation: V.D.C., M.G., T.O.D., and P.F. Supervision: V.D.C., M.G., and P.F. Writing—original draft: V.D.C., M.G., T.O.D., and P.F. Writing—review and editing: V.D.C., M.G., H.O., O.J., T.O.D., and P.F.

## Data Availability

All data are available in the main text or Supplementary Material online, and on a public repository at https://figshare.com/projects/Giant_viruses_encode_novel_types_of_actins_the_viractin/82262.

## References

Akıl C, Kitaoku Y, Tran LT, Liebl D, Choe H, Muengsaen D, Suginta W, Schulte A, Robinson RC. 2021. Mythical origins of the actin cytoskeleton. *Curr Opin Cell Biol*. 68:55–63.

Archer SK, Claudianos C, Campbell HD. 2005. Evolution of the gelsolin family of actin-binding proteins as novel transcriptional coactivators. *Bioessays* 27(4):388–396.

Bajrai LH, Mougari S, Andreani J, Baptiste E, Delerce J, Raoult D, Azhar EI, La Scola B, Levasseur A. 2019. Isolation of Yasminevirus, the first member of Klosneuvirinae isolated in coculture with *Vermamoeba vermiformis*, demonstrates an extended arsenal of translational apparatus components. *J Virol*. 94(1):e01534-19.

Bajusz C, Borkúti P, Kristó I, Kovács Z, Abonyi C, Vilmos P. 2018. Nuclear actin: ancient clue to evolution in eukaryotes? *Histochem Cell Biol*. 150(3):235–244.

Bell PJL. 2020. Evidence supporting a viral origin of the eukaryotic nucleus. *Virus Res*. 289:198168.

Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG, et al. 2009. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc Natl Acad Sci U S A*. 106(51):21848–21853.

Burki F, Roger AJ, Brown MW, Simpson AGB. 2020. The New Tree of Eukaryotes. *Trends Ecol Evol*. 35(1):43–55.

Chaikeeratisak V, Khanna K, Nguyen KT, Sugie J, Egan ME, Erb ML, Vavilina A, Nonejuie P, Nieweglowska E, Pogliano K, et al. 2019. Viral capsid trafficking along treadmilling tubulin filaments in bacteria. *Cell* 177(7):1771–1780.e12.

Chaikeeratisak V, Nguyen K, Khanna K, Brilot AF, Erb ML, Coker JKC, Vavilina A, Newton GL, Buschauer R, Pogliano K, et al. 2017. Assembly of a nucleus-like structure during viral replication in bacteria. *Science* 355(6321):194–197.

Dombrowski N, Teske AP, Baker BJ. 2018. Expansive microbial metabolic versatility and biodiversity in dynamic Guaymas Basin hydrothermal sediments. *Nat Commun*. 9:4999.

Endo H, Blanc-Mathieu R, Li Y, Salazar G, Henry N, Labadie K, De Vargas C, Sullivan MB, Bowler C, Wincker P, et al. 2020. Biogeography of marine giant viruses reveals the interplay with eukaryotes and ecological functions. *Nat Ecol Evol*. 4:1639–1649.

Forterre P. 2017. Viruses in the 21st century: from the curiosity-driven discovery of giant viruses to new concepts and definition of life. *Clin Infect Dis*. 65(Suppl 1):S74–S79.

Forterre P, Gaïa M. 2016. Giant viruses and the origin of modern eukaryotes. *Curr Opin Microbiol*. 31:44–49.

Ghoshdastider U, Popp D, Burtnick LD, Robinson RC. 2013. The expanding superfamily of gelsolin homology domain proteins. *Cytoskeleton* 70(11):775–795.

Goodson HV, Hawse WF. 2002. Molecular evolution of the actin family. *J Cell Sci*. 115(13):2619–2622.

Guglielmini J, Woo AC, Krupovic M, Forterre P, Gaia M, Gaïa M. 2019. Diversification of giant and large eukaryotic dsDNA viruses predated the origin of modern eukaryotes. *Proc Natl Acad Sci U S A*. 116(39):19585–19592.

Ha AD, Moniruzzaman M, Aylward FO. 2021. High transcriptional activity and diverse functional repertoires of hundreds of giant viruses in a coastal marine system. *mSystems* 6(4):e00293-21.

Kijima S, Delmont TO, Miyazaki U, Gaia M, Endo H, Ogata H. 2021. Discovery of viral myosin genes with complex evolutionary history within plankton. *Front Microbiol*. 12:683294.

Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, Yutin N, Zerbini FM, Kuhn JH. 2020. Global organization and proposed megataxonomy of the virus world. *Microbiol Mol Biol. Rev*. 84:1–33.

Koonin EV, Yutin N. 2019. Evolution of the large nucleocytoplasmic DNA viruses of eukaryotes and convergent origins of viral gigantism. *Adv Virus Res*. 103:167-202.

Kristó I, Bajusz I, Bajusz C, Borkúti P, Vilmos P. 2016. Actin, actin-binding proteins, and actin-related proteins in the nucleus. *Histochem Cell Biol*. 145(4):373–388.

Kühn S, Bergqvist J, Gil M, Valenzuela C, Barrio L, Lebreton S, Zurzolo C, Enninga J. 2020. Actin assembly around the Shigella-containing vacuole promotes successful infection. *Cell Rep*. 31:107638.

Lindås AC, Valegård K, Ettema TJG. 2017. Archaeal actin-family filament systems. *Subcell Biochem*. 84:379–392.

Miller A, Guljamow A, Jenke-Kodama H, Saumweber H, Quillardet P, Frangeul L, Castets AM, Bouchier C, De Marsac NT. 2007. Horizontal gene transfer of two cytoskeletal elements from a eukaryote to a cyanobacterium. *Curr Biol*. 17:757–759.

Moniruzzaman M, Martinez-Gutierrez CA, Weinheimer AR, Aylward FO. 2020. Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses. *Nat Commun*. 11:1–11.

Moniruzzaman M, Weinheimer AR, Martinez-Gutierrez CA, Aylward FO. 2020. Widespread endogenization of giant viruses shapes genomes of green algae. *Nature* 588(7836):141–145.

Muller J, Oma Y, Vallar L, Friederich E, Poch O, Winsor B. 2005. Sequence and comparative genomic analysis of actin-related proteins. *Mol Biol Cell*. 16(12):5736–5748.

Pesant S, Not F, Picheral M, Kandels-Lewis S, Le Bescot N, Gorsky G, Iudicone D, Karsenti E, Speich S, Trouble R, et al. 2015. Open science resources for the discovery and analysis of Tara Oceans data. *Sci Data*. 2:1–16.

Schroeder CM, Valenzuela JR, Mejia Natividad I, Hocky GM, Malik HS. 2020. A burst of genetic innovation in Drosophila actin-related proteins for testis-specific function. *Mol Biol Evol*. 37(3):757–772.

Schroer TA, Fyrberg E, Cooper JA, Waterston RH, Helfman D, Pollard TD, Meyer DI. 1994. Actin-related protein nomenclature and classification. *J Cell Biol*. 127(6):1777–1778.

Schulz F, Roux S, Paez-Espino D, Jungbluth S, Walsh DA, Denef VJ, McMahon KD, Konstantinidis KT, Eloe-Fadrosh EA, Kyrpides NC, et al. 2020. Giant virus diversity and host interactions through global metagenomics. *Nature* 578(7795):432–436.

Schulz F, Yutin N, Ivanova NN, Ortega DR, Lee TK, Vierheilig J, Daims H, Horn M, Wagner M, Jensen GJ, et al. 2017. Giant viruses with an expanded complement of translation system components. *Science* 356(6333):82–85.

Shiratori T, Suzuki S, Kakizawa Y, Ishida K. 2019. Phagocytosis-like cell engulfment by a planctomycete bacterium. *Nat Commun*. 10:1–11.

Stairs CW, Ettema TJG. 2020. The archaeal roots of the eukaryotic dynamic actin cytoskeleton. *Curr Biol*. 30(10):R521–R526.

Subramaniam K, Behringer DC, Bojko J, Yutin N, Clark AS, Bateman KS, van Aerle R, Bass D, Kerr RC, Koonin EV, et al. 2020. A new family of DNA viruses causing disease in crustaceans from diverse aquatic biomes. *mBio* 11(1):e02938-19.

Takemura M, Yokobori SI, Ogata H. 2015. Evolution of eukaryotic DNA polymerases via interaction between cells and large DNA viruses. *J Mol Evol*. 81(1–2):24–33.

Vosseberg J, van Hooff JJE, Marcet-Houben M, van Vlimmeren A, van Wijk LM, Gabaldón T, Snel B. 2020. Timing the origin of eukaryotic cellular complexity with ancient duplications. *Nat Ecol Evol*. 5(1):92–100.

Yoshikawa G, Blanc-Mathieu R, Song C, Kayama Y, Mochizuki T, Murata K, Ogata H, Takemura M. 2019. Medusavirus, a novel large DNA virus discovered from hot spring water. *J Virol*. 93:1–25.

Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström D, Juzokaite L, Vancaester E, Seitz KW, Anantharaman K, Starnawski P, Kjeldsen KU, et al. 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541(7637):353–358.

Zhou Z, Pan J, Wang F, Gu J-D, Li M. 2018. Bathyarchaeota: globally distributed metabolic generalists in anoxic environments. *FEMS Microbiol Rev*. 42(5):639–655.