VIRUS EVOLUTION

Virus Evolution, 2025, **11(1)**, veaf063
DOI: https://doi.org/10.1093/ve/veaf063
Advance Access Publication 22 August 2025
Research Article

Convergent evolution of the G3BP1-binding motif in betacoronavirus nucleocapsid proteins

Cinzia Borgogna[1,*], Ilaria Cislaghi[1], Sarah Turati[1], Alessandra Mozzi[2], Diego Forni[2], Rachele Cagliani[2], Manuela Sironi[3,†], Marisa Gariglio[1,†]

[1]Virology Unit, Department of Translational Medicine, University of Piemonte Orientale, Via Generale Solaroli 17, Novara (NO), 28100, Italy
[2]Scientific Institute IRCCS E. MEDEA, Department of Computational Biology Unit, Via Don Luigi Monza 20, Bosisio Parini (LC), 23842, Italy
[3]School of Medicine and Surgery, University of Milano-Bicocca, Via Cadore 48, Monza (MB), 20900, Italy

*Corresponding author. Virology Unit, Department of Translational Medicine, University of Piemonte Orientale, Novara (NO), Italy.
E-mail: cinzia.borgogna@med.uniupo.it
†These authors contributed equally.

Abstract

Betacoronaviruses (β-CoVs) display divergent mechanisms to evade host antiviral responses, yet the evolutionary origin and functional relevance of their strategies remain unclear. Stress granules (SGs), central to host defenses, are disrupted by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) nucleocapsid (N) protein *via* its interaction with G3BP1 mediated by an ^{15}ITFG^{18} motif. This interaction inhibits SG assembly, enhancing viral replication and immune evasion. Here, we analyzed 179 N protein sequences across β-CoVs subgenera and identified the ITFG motif in sarbecoviruses but not in MERS-CoV or embecoviruses. Among tested CoVs, only SARS-CoV-2 N showed robust G3BP1 binding *in vitro* and in infected cells. Phylogenetic analyses revealed that the ΦxFG motif emerged independently in sarbecoviruses and a bat-infecting merbecovirus clade. The VGTF motif in these merbecoviruses also binds to G3BP1, suggesting convergent evolution in viral evasion strategies. The emergence of this motif was unrelated to 4a protein inactivation, another viral protein that inhibits SG formation.

Keywords: G3BP1; betacoronavirus; ITFG motif; nucleocapsid protein

Introduction

Coronaviruses (CoVs) are a large family of enveloped viruses characterized by positive-sense RNA genomes, primarily implicated in respiratory infections across humans and animals (Salata *et al.* 2019, Ye *et al.* 2020). They are taxonomically divided into four genera: *Alphacoronavirus, Betacoronavirus, Gammacoronavirus*, and *Deltacoronavirus*. Among these, betacoronaviruses (β-CoVs) have had the most profound impact on human health over the past two decades, exemplified by the zoonotic transmission of SARS-CoV, MERS-CoV, and SARS-CoV-2, leading to outbreaks marked by severe morbidity and substantial mortality. By contrast, endemic β-CoVs, such as HCoV-HKU1 and HCoV-OC43, are typically associated with mild respiratory infections (Deng and Baker 2021), suggesting divergent evolutionary pressure and host-virus interactions. Thus, characterizing the interplay between β-CoVs and host antiviral defenses is critical for understanding the evolutionary adaptations that sustain their persistence and pathogenicity.

A crucial aspect of the host antiviral strategy involves the assembly of stress granules (SGs), which are cytoplasmic aggregates primarily composed of RNA and RNA-binding proteins. These structures sequester key viral components and host factors, thereby hindering viral replication (McCormick and Khaperskyy 2017). Among the core mediators of SG formation are the paralogous proteins G3BP1 and G3BP2 (Ras-GTPase-activating protein SH3-domain-binding proteins), which are activated in response to cellular stresses, including viral infections (Yang *et al.* 2024). In several viral infections, SG formation is PKR (protein kinase R)-dependent (White and Lloyd 2012). PKR expression is induced by interferon, and the protein can detect viral RNA in the cytoplasm, leading to phosphorylation of the alpha subunit of eukaryotic translation initiation factor 2 (eIF2α). This, in turn, blocks protein synthesis, and the stalled translation initiation complexes, together with nucleating factors (e.g. G3BP1 and G3BP2) promote the formation of SGs (White and Lloyd 2012).

Previous studies have shown that the SARS-CoV-2 nucleocapsid (N) protein can inhibit SG formation through its interaction with G3BP1/2 *via* an ^{15}ITFG^{18} motif located in the intrinsically disordered region (IDR) at the N-terminal domain (NTD) (Huang *et al.* 2021, Kruse *et al.* 2021, Yang *et al.* 2024). Of note, an F17A point mutation in the viral N protein, which disrupts this motif, specifically abolishes the interaction with G3BP1/2. As a result, a mutant SARS-CoV-2 N-F17A displayed decreased viral replication and attenuated pathology *in vivo* (Yang *et al.* 2024). Similarly, deletion of the G3BP-binding motif in the SARS-CoV-2 N protein was shown to impair viral replication across multiple cell lines and to reduce disease severity in a mouse model of severe COVID-19. These findings support a model in which the N–G3BP interaction plays a crucial role in antagonizing host antiviral responses, while simultaneously promoting efficient viral RNA replication and gene expression (Long *et al.* 2024). However, the evolutionary

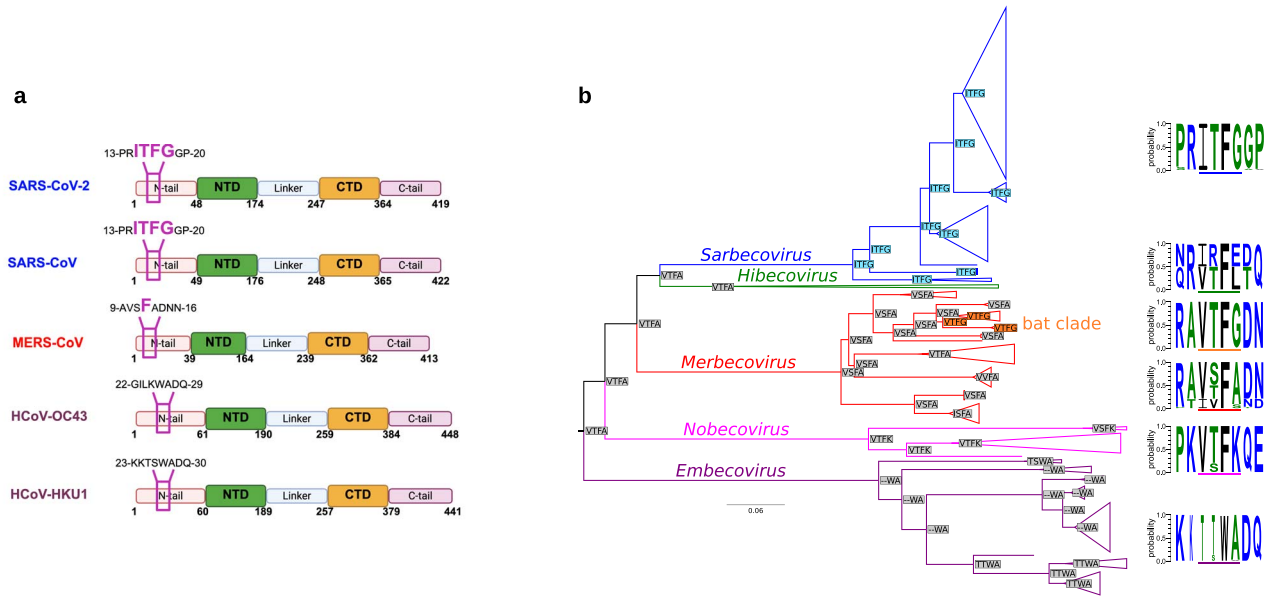


Figure 1. Origin and evolution of the ΦxFG in the N protein of betacoronaviruses. (a) Schematic representation of the N proteins of the five β-HCoVs. The NTD and C-terminal domain are highlighted. Created with BioRender.com. (b) Phylogenetic tree of the non-recombining region of N protein from 179 β-CoVs, generated using IQ-TREE. The five subgenera are indicated along with the merbecovirus bat clade carrying a VTFG motif (BtVs-BetaCoV/SC2013, KJ473821; *Merbecovirus sp.* Ej-CoV-3, LC663783; *Merbecovirus sp.* VsCoV-kj15, LC706863; *Merbecovirus sp.* VsCoV-a7, LC706864; *Merbecovirus sp.* EjCoV-3, LC706865; MERS-related NL13845, MG021451; MERS-related NL140422, MG021452). Sequence logos display amino acid status—With letter size proportional to frequency—at the positions corresponding to the ^{15}ITFG^{18} motif. Ancestral state reconstruction—Using GRASP—Of the ^{15}ITFG^{18} motif is shown for all tree nodes.

origins and functional significance of the G3BP1-binding motif changes in the nucleocapsid protein of β-CoVs remain poorly understood.

Interestingly, the ^{15}ITFG^{18} motif is conserved in the N protein of SARS-CoV but absent in closely related β-CoVs, including MERS-CoV, HCoV-OC43, or HCoV-HKU1 (Fig. 1a), raising the possibility that these viruses rely on alternative mechanisms to interact with host factors, driven by their specific evolutionary contexts. Notably, other viral and cellular proteins are known to exploit similar motifs (ΦxFG, where Φ is an amino acid residue with a hydrophobic side chain and x is any amino acid) for G3BP1/2 binding (Kruse *et al.* 2021), suggesting a broader functional relevance. In this study, we thus aimed to investigate the evolutionary history and G3BP1-binding potential of the ΦxFG motif across an extended β-CoV phylogeny.

Material and methods

Cell lines

Rhesus monkey kidney epithelial cells (LLC-MK2, ATCC: CCL-7) and Human lung fibroblast MRC-5 cells (ATCC: CCL-171) were cultured in high glucose Dulbecco's modified Eagle's medium (DMEM; Euroclone, Milano, Italy) supplemented with 10% fetal bovine serum (FBS; Euroclone, Milano, Italy) and 1% penicillin–streptomycin. Human epithelial adenocarcinoma HeLa (ATCC: CCL-2) cells were cultured in DMEM supplemented with 10% FBS, 2 mM L-glutamine and 100 U/ml penicillin/streptomycin (Invitrogen, Carlsbad, CA, USA; Thermo Fisher Scientific, Waltham, MA, USA). Cells were maintained at 37°C with 5% CO_2.

Biosafety statements and facility

All experiments with live HCoV-OC43 were performed in a biosafety level 2 (BSL2) facility at the University of Piemonte Orientale, Novara, Italy. Experiments involving live SARS-CoV-2 were performed in biosafety levels 3 (BSL3) facility at the University of Milan, Milan, Italy. Both facilities adhered to approved standard operating procedures in Italy. All personnel underwent comprehensive biosafety training before initiating work in these facilities.

Viruses

The HCoV strain OC43 (ATCC: VR-1558) was kindly provided by David Lembo (Department of Clinical and Biological Sciences, University of Turin, Turin, Italy) and propagated and maintained in MRC-5 cells. All experiments involving SARS-CoV-2 were conducted in collaboration with Serena Delbue (University of Milan, Milan, Italy). The SARS-CoV-2 strain, isolated from a nasal-pharyngeal swab, belongs to the B.1 lineage, with the spike D614G mutation, which corresponds to the predominant European lineage linked to the Northern Italian outbreak in early 2020. Its complete nucleotide sequence is available in GenBank and GISAID (accession Nos. MT748758.1 and EPI_ISL 584051, respectively). SARS-CoV-2 was propagated and maintained in Vero E6 cells at the Laboratory of Molecular Virology, University of Milan, Italy.

Plasmids

Complementary DNA (cDNA) encoding the SARS-CoV-2 ORF N (NC_045512.2, nucleotide position: 28274–29 533) were synthesized by Origene Technologies, Inc. (Rockville, MD, USA). The sequence was cloned into the pCMV6-entry mammalian expression vector (Origene Technologies, PS100001) with a C-terminal Myc-DDK tag (pCMV6-N plasmid). Mutant constructs (pCMV6-N-I15V, pCMV6-N-F17A, and pCMV6-N-G18A) were generated by site-direct mutagenesis using Pfu DNA polymerase (Promega, Madison, WI, USA) and the pCMV6-N plasmid as a template. Following site-directed mutagenesis, the template DNA was digested with DpnI restriction endonuclease, and PCR

products were used to transform TOP10 *Escherichia coli* competent cells (Invitrogen). Mutations were verified *via* Sanger sequencing.

Surface plasmon resonance

The Biacore X100 (GE Healthcare, Chicago, IL, USA) instrument was used for real-time binding interaction experiments. Recombinant G3BP1 (cat. # GTX68122-pro, GeneTex, Irvine, CA, USA) was covalently immobilized onto the surface of a sensor CM5 chip (cat # BR100012, GE Healthcare) *via* amine coupling. G3BP1 was diluted to a concentration of 4 μg/ml in 10 mM sodium acetate (pH 4.0) and injected onto the sensor chip at a flow rate of 10 μl/min. Activation of the carboxyl groups on the sensor surface was achieved by injecting for 7 min a mixture of 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) and 0.05 M N-hydroxysuccinimide (NHS). The remaining esters were blocked with a 7-min injection of ethanolamine. Taking into account molecular weights of the ligand (G3BP1) and analytes [N of SARS-CoV-2 (52 kDa, Z03488, GenScript, Piscataway, NJ, USA), HCoV-OC43 (46 kDa, 40 643-V07E-100, Sino Biological, Benjing, China), and MERS (50.43 kDa, 40 068-V08B-100, Sino Biological)], the appropriate ligand density (RL) on the chip was calculated using the following equation: RL = (ligand MW/analyte MW) × Rmax × (1/Sm), where Rmax is the maximum binding signal, and Sm corresponds to the binding stoichiometry of the interaction. The target capture level for G3BP1 was 314.5 response units (RUs). A reference flow cell was immediately blocked after activation for control purposes. Increasing concentrations (i.e. 500, 250, 125, 62.5, 31.25, and 15.625 nM) of recombinant N of SARS-CoV-2, OC43, and MERS-CoV were flowed over the G3BP1-coated CM5 sensor chip coated at 30 μl/min and 25°C. The association and dissociation times were set at 120 s and 180 s, respectively. A single regeneration step with 50 mM NaOH was performed following each analytic cycle. All the analytes tested were diluted in PBSP plus buffer (GE Healthcare). The KDs were calculated using the BIAcore evaluation software (GE Healthcare). Their reliability was calculated by assuming a 1:1 binding model, validated by quality assessment indicators.

Viral infection, immunoprecipitation, and immunoblotting

Confluent LLC-MK2 cells were infected with an MOI of 3 for SARS-CoV-2 and 0.1 for HCoV-OC43. Virus absorption was allowed for 2–8 h before replacing the media. Infected cells were incubated at 34°C in a 5% CO_2 atmosphere. Cells were collected at 16 hpi for SARS-CoV-2 and 24 hpi for HCoV-OC43. Immunoprecipitation of SARS-CoV-2 or HCoV-OC43 nucleoproteins and G3BP1, along with their interacting proteins, was performed on cytoplasmic cell extracts obtained with the NE-PER Nuclear and Cytoplasmic Extraction Reagents kit (Thermo Fisher Scientific) according to the manufacturer's instructions. 15 μg of total cell extracts were kept as input controls, while 300 μg of cytoplasmic cell extracts were incubated overnight at 4°C with the following antibodies: anti-SARS-CoV-2 N polyclonal antibody (3 μg) (40143-R001; SinoBiological), anti-HCoV-OC43 N monoclonal antibody (5 μg) (MAB9012; Millipore, Burlington, MA, USA), and anti-G3BP1 monoclonal antibody (1 μg) (ab181149; Abcam, Cambridge, UK). These antibodies were incubated with Pierce Protein G Agarose (Thermo Fisher Scientific). The resulting complexes were then washed, eluted, denatured, and subjected to immunoblotting as previously described (Lo Cigno *et al.* 2020). Immunocomplexes were detected using HRP-conjugated sheep anti-mouse immunoglobulin antibodies (GE Healthcare Europe GmbH) or TidyBlot Western Blot Detection Reagent: HRP (STAR209PA; Bio-Rad Laboratories, Hercules, CA, USA). Detection was performed using enhanced chemiluminescence (Super Signal West Pico; Thermo Fisher Scientific) and visualized on a ChemiDoc Touch Imaging System (Bio-Rad Laboratories).

Mutant analysis: Cell line transfection, immunoprecipitation, and immunoblotting

HeLa cells were seeded at a density of 1.2×10^6 cells/well into p100 plates 24 h before transfection. Transient transfections were performed using Lipofectamine™ 3 000 Transfection Reagent (Thermo Fisher Scientific) with 10 μg of plasmid DNA/plate (pCMV6-N [Origene Technologies, Rockville, MD, USA], pCMV6-N-I15V, pCMV6-N-F17A, or pCMV6-N-G18A [our laboratory, Supplementary Table S1]), according to the manufacturer's instruction. Co-immunoprecipitation assays were performed with the Pierce™ MS-Compatible Magnetic IP Kit, protein A/G (Thermo Fisher Scientific). Briefly, at 24 h post transfection, HeLa cells were rinsed twice with ice-cold PBS and lysed on ice in IP-MS Cell Lysis Buffer supplemented with EDTA-free Halt™ protease inhibitor cocktail (Thermo Fisher Scientific) for 10 min with periodic mixing. Extracts were clarified by centrifugation (13 000 × g for 10 min) and quantified using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific). A total of 500 μg of cell lysate was combined with 5 μg of the immunoprecipitation antibody and incubated overnight at 4°C, with continuous mixing to form immune complexes. Immunoprecipitation reactions were performed for 1 h at RT by incubating the sample/antibody mixture with 0.25 mg of pre-washed Pierce Protein A/G Magnetic Beads. After washes, target antigen samples were eluted in IP-MS elution buffer and dried in a speed vacuum concentrator. Samples were reconstituted in sample buffer for SDS-PAGE/WB analyses. SDS-PAGE and Western-blot were carried out following standard procedures. Samples were loaded and resolved on an 8% or 10% acrylamide/bis-acrylamide gel and, blotted onto a nitrocellulose membrane (Amersham, Cytiva, Marlborough, MA, USA). Horseradish peroxidase-conjugated secondary antibodies were used for detection. Enhanced chemiluminescence (ECL, GE Healthcare) signals were visualized and captured using ECL (GE Healthcare) and acquired with iBrightFL1000 imaging system (Thermo Fisher Scientific).

Viral sequence selection and recombination analysis

All annotated coding sequences for 179 complete β-CoV genomes were retrieved from the NCBI database (https://www.ncbi.nlm.nih.gov/, Supplementary Table S2). Orthologs of the nucleocapsid protein and the MERS-CoV accessory protein 4a were identified with OrthoFinder version 2.5.4 (Katoh and Standley 2013, Emms and Kelly 2019). Alignments were generated using MAFFT v. 7.475 (Katoh and Standley 2013). A phylogenetic maximum likelihood (ML) tree for the N protein was generated by the IQ-TREE software v.1.6.12 (Minh *et al.* 2020), which selects the best-fitting evolutionary model and generates the corresponding ML tree. The N protein alignment, along with its tree, were screened for recombination using the GARD tool (Kosakovsky Pond *et al.* 2006) from the hypothesis testing using phylogenies (HYPHY) package (Pond *et al.* 2005). This was motivated by the fact that, in the presence of recombination it is impossible to reconstruct a single phylogeny and, therefore, to infer ancestral sequence states. GARD identifies recombination breakpoints based on phylogenetic incongruence among alignment segments. A recombination event was detected in the alignment with a P-value < 0.01. Subsequent analyses used

the tree output by GARD for the non-recombinant region at the 5′ end of the breakpoint, which includes the N-terminal IDR of the N protein and is 39% of the total sequence length.

Intrinsically disordered regions

Intrinsically disordered regions were predicted using the Metapredict tool (Emenecker *et al.* 2021) with default parameters. This tool applies a deep-learning algorithm to calculate a consensus score from eight different disorder predictors (Emenecker *et al.* 2021).

Ancestral sequence reconstruction

To study the evolution of the ΦxFG motif in β-CoVs, ancestral sequence reconstruction was performed for all internal nodes of the β-CoV phylogeny. The GRASP tool (Foley *et al.* 2022) was used to infer ancestral sequences, focusing on the predicted N-terminal IDR containing the ΦxFG motif. GRASP takes as inputs a multiple sequence alignment and a phylogenetic tree. We used GARD to generate a tree of the N protein that takes into account the position of recombination breakpoints; the resulting tree was then midpoint rooted before running GRASP. Thus, we obtained inferred motifs for all nodes in the β-CoV phylogeny.

In silico prediction of motif substitutions

The 3D structure of a peptide carrying the ΦxFG motif in complex with human G3BP1 was retrieved from the Protein Data Bank (PDB ID: 7SUO). Amino acid substitutions in the ΦxFG motif, identified in β-CoV N proteins, were analyzed for their effects on binding affinity using the MutaBind2 tool (Zhang *et al.* 2020). The detrimental F17A mutation was also evaluated for confirmation alongside the observed changes.

Results

Independent evolution of a G3BP-binding motif in sarbecoviruses and a subset of bat merbecoviruses

To investigate the evolutionary significance of the ΦxFG motif within βCoVs, we first assembled a comprehensive list of 179 β-CoV N protein sequences representing the five recognized subgenera: *Sarbecovirus, Hibecovirus, Merbecovirus, Nobecovirus*, and *Embecovirus* (Supplementary Table S2). This dataset provided a robust foundation for comparative analysis of motif conservation and divergence across the phylogenetic tree.

Using Metapredict V2, we found that all β-CoVs possess an IDR at the N-terminus of the N protein. Alignment and visual inspection of N sequences confirmed the presence of the ITFG motif in all sarbecoviruses, including SARS-CoV and SARS-CoV-2, and its absence in embecoviruses, which include the low-pathogenic HCoV-OC43 and HCoV-HKU1 (Fig. 1b). Most nobecoviruses displayed the VTFK sequence at the corresponding position, which is unlikely to be a functional binding motif due to the presence of a charged Lys (K) residue instead of the canonical Gly (G) residue (Fig. 1b). The *Merbecovirus* subgenus displayed greater heterogeneity, with a Phe (F) residue universally conserved at position 12—MERS-CoV numbering—frequently followed by an Ala. However, a small clade of bat-infecting merbecoviruses carried a VTFG motif, which may retain the ability to bind to G3BP1/2 (Fig. 1b, Supplementary Table S2).

To further understand the evolution of the ΦxFG motif, we used the genetic algorithm for recombination detection (GARD) method to identify a recombination-free alignment of the N protein region where the motif is located. We then employed the Graphical representation of ancestral sequence predictions (GRASP) tool, which takes a phylogenetic tree and an alignment as inputs to infer ancestral sequences at internal nodes of the tree. The confidence of the reconstruction was very good at relevant nodes, including the β-CoV ancestor (Supplementary Table S3). This analysis indicated that the ancestor of β-CoVs carried a VTFA motif, most likely unable to affect SG assembly due to the absence of the G residue and other key structural features necessary for binding (Fig. 1b). Importantly, the ITFG motif seems to have evolved in the most recent common ancestor of sarbecoviruses and to have remained conserved (Fig. 1b). In most other subgenera, amino acid configurations are likely preventing N protein binding to G3BP1/2. Nonetheless, a VTFG motif appears to have emerged in the ancestor of a small clade of bat-infecting merbecoviruses (Fig. 1b).

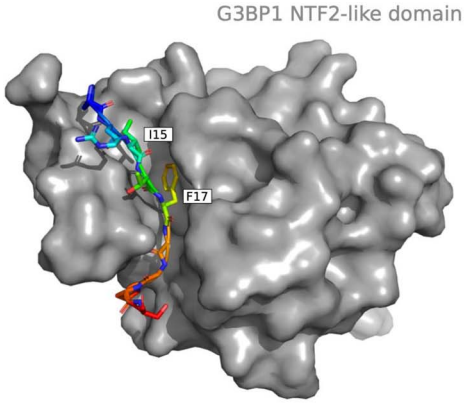
Convergent evolution of functional ΦxFG motifs enables selective G3BP1 interaction in sarbecoviruses and a clade of bat merbecoviruses

Next, we sought to predict how specific mutations in the ITFG motif observed across the β-CoV phylogeny would impact G3BP1/2 binding. For this purpose, we analyzed the resolved crystal structure of the NTD of G3BP1 bound to the ITFG motif. Using the MutaBind2 tool, we then evaluated how these mutations influenced binding affinity. We found that substitutions of F17 or G18—both relative to SARS-CoV-2—are predicted to disrupt binding (Fig. 2). Conversely, replacing Thr (T) at position 16 with a non-charged amino acid residue or Ile (I) at position 15 with another non-polar amino acid, such as Val (V), is predicted to maintain binding capacity (Fig. 2).

Overall, these findings suggest that, aside from the ITFG motif in sarbecoviruses, only the VTFG motif present in certain merbecoviruses is likely to interact with G3BP1/2, underscoring the evolutionary specificity of the ΦxFG motif in modulating host interactions among β-CoVs.

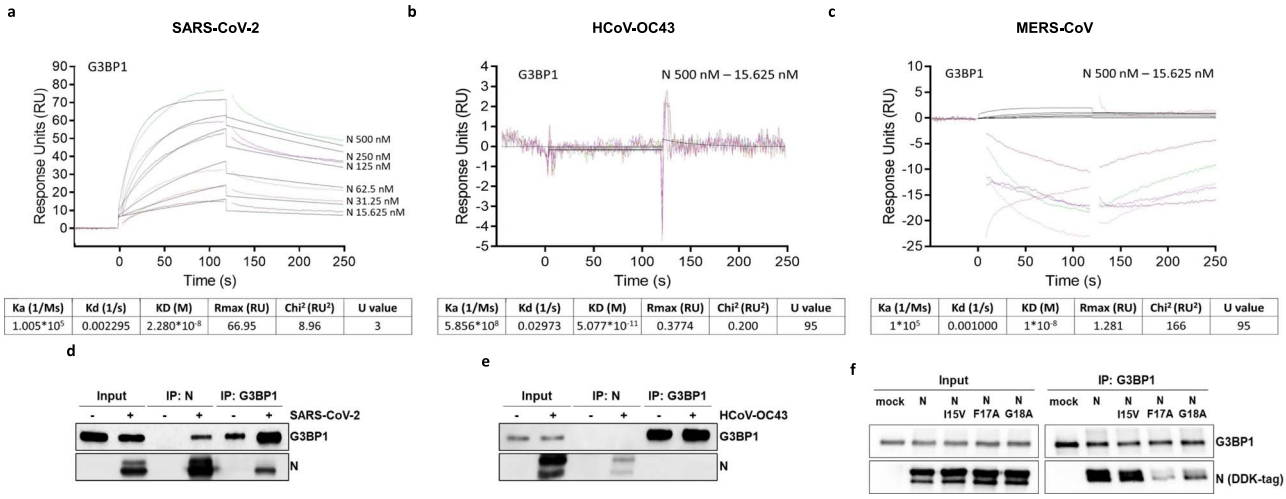
To experimentally validate the computational predictions described above, we next assessed the binding affinity of the G3BP1 protein to the N proteins from representative human-infecting β-CoVs comprising different subgenera using surface plasmon resonance (SPR) analysis. Specifically, recombinant G3BP1 protein was immobilized on a CM5 sensor chip and then probed with increasing concentrations of recombinant N proteins from SARS-CoV-2 (*Sarbecovirus*), MERS-CoV (*Merbecovirus*), or HCoV-OC43 (*Embecovirus*). The resulting SPR sensorgrams revealed a strong interaction between G3BP1 and the N protein from SARS-CoV-2, with an equilibrium dissociation constant (Kd) of 2.280×10^{-8} (Fig. 3a). In contrast, no binding was detected between G3BP1 and the N proteins from MERS-CoV (Fig. 3b) or HCoV-OC43 (Fig. 3c). These *in vitro* results were fully consistent with the computational predictions shown in Fig. 2, demonstrating the specificity of the ITFG motif in facilitating this interaction.

To further validate these observations in a cellular context, we performed co-immunoprecipitation (co-IP) assays using LLC-MK2 cells, a rhesus macaque epithelial kidney-derived cell line. Cells were infected with either SARS-CoV-2 or HCoV-OC43, as experiments with MERS-CoV were not feasible due to the unavailability of the virus. The SARS-CoV-2 N protein successfully co-immunoprecipitated with G3BP1 (Fig. 3d), whereas no interaction was detected for the N protein from HCoV-OC43-infected cells (Fig. 3e). As these cellular findings align with the SPR analysis (Fig. 3a and c), they point to the distinct role of the ITFG motif in mediating the interaction between the SARS-CoV-2 N protein and G3BP1, a characteristic absent in other representative β-CoVs.



Mutation	$\Delta\Delta G_{bind}$(Kcal/mol)	Deleterious
I15T	3.09	yes
I15V	1.38	no
T16S	1.36	no
T16R	1.51	yes
T16V	1.29	no
F17W	2.98	yes
F17R	3.85	yes
G18A	2.25	yes
G18V	2.9	yes
G18E	2.55	yes
G18L	2.39	yes
G18K	2.74	yes
G18S	2.61	yes
G18Q	2.53	yes

Figure 2. G3BP1-N protein interaction. Crystal structure of G3BP1 NTF2L in complex with SARS-CoV-2 ITFG motif (PDB ID: 7SUO). Amino acid substitutions observed in different β-CoVs and their predicted effects on binding affinity are also reported.



	Ka (1/Ms)	Kd (1/s)	KD (M)	Rmax (RU)	Chi² (RU²)	U value
a	$1.005*10^5$	0.002295	$2.280*10^{-8}$	66.95	8.96	3
b	$5.856*10^8$	0.02973	$5.077*10^{-11}$	0.3774	0.200	95
c	$1*10^5$	0.001000	$1*10^{-8}$	1.281	166	95

Figure 3. G3BP1 binds to the N protein of SARS-CoV-2, but not to that of MERS-CoV and HCoV-OC43, *in vitro* and in infected cells. (a–c) SPR analysis of G3BP1 binding to N proteins from SARS-CoV-2 (a), HCoV-OC43 (b), and MERS-CoV (c). Recombinant G3BP1 was immobilized on a CM5 sensor chip surface, and recombinant N proteins from the three viruses, at increasing concentrations (15.625–500 nM), were injected over the chip surface for measurement. G3BP1 binding to the N protein of SARS-CoV-2 displays an equilibrium dissociation constant (K_d) of 2.280×10^{-8}, indicating strong interaction. Data are representative of three independent experiments. (d) LLC-MK2 cells were mock-infected or infected with SARS-CoV-2. Cells were harvested at 16 hpi, and cytoplasmic lysates were subjected to immunoprecipitation using antibodies against G3BP1 or N. Immunoprecipitates were analyzed by immunoblotting with anti-N or anti-G3BP1 antibodies. One representative western blot from three independent experiments is shown. (e) LLC-MK2 cells were mock-infected or infected with HCoV-OC43. Cells were harvested at 24 hpi, and cytoplasmic lysates were subjected to immunoprecipitation using antibodies against G3BP1 or N. Immunoprecipitates were analyzed by immunoblotting with anti-N or anti-G3BP1 antibodies. One representative western blot from three independent experiments is shown. (f) HeLa cells were transiently transfected with plasmids expressing SARS-CoV-2 DDK-tagged N protein or the following mutants: N-I15V, N-F17A, and N-G18A. At 24 h post-transfection, cells were harvested, and total protein extracts were subjected to immunoprecipitation with an anti-G3BP1 antibody. Immunoprecipitates were analyzed by immunoblotting with anti-DDK and anti-G3BP1 antibodies. A representative blot from three reproducible experiments is shown.

To determine whether the VTFG motif present in the N protein from the bat-infecting merbecovirus subclade could interact with G3BP1, we introduced an I15V mutation in the N protein of SARS-CoV-2 to replicate the VTGF motif. As controls, two Ala mutants—i.e. N-F17A and N-G18A—were created, as these mutations were computationally predicted by MutaBind2 to disrupt binding (data not shown). Plasmids expressing the wild-type (WT) N protein or the mutants were transfected into HeLa cells, and their ability to interact with G3BP1 was tested through co-IP. The results show that G3BP1 can immunoprecipitate comparable amounts of the N-I15V mutant and the WT N protein, whereas the two Ala controls display significantly weaker interactions with G3BP1 (Fig. 3f).

Overall, these findings, alongside the ancestral reconstruction analysis, indicate that a ΦxFG motif capable of interacting with G3BP1 independently emerged in sarbecoviruses and a subset of bat-infecting merbecoviruses, suggesting a convergent evolutionary mechanism by which β-CoVs use their N protein to prevent G3BP1 from forming functional SGs, effectively neutralizing their antiviral activity.

Retention of p4a RNA-binding activity in merbecoviruses suggests independent evolution of G3BP1 antagonism mechanisms

Previous studies have shown that infection with the highly pathogenic MERS-CoV fails to induce SG formation (Rabouw *et al.* 2016). This outcome appears to be mediated by the MERS-CoV accessory protein 4a (p4a), which antagonizes PKR-mediated stress response *via* its RNA-binding domain, an activity essential for its function (Rabouw *et al.* 2016). To determine whether the emergence of the G3BP1-binding motif in bat-infecting merbecoviruses might be related to p4a alterations, we analyzed

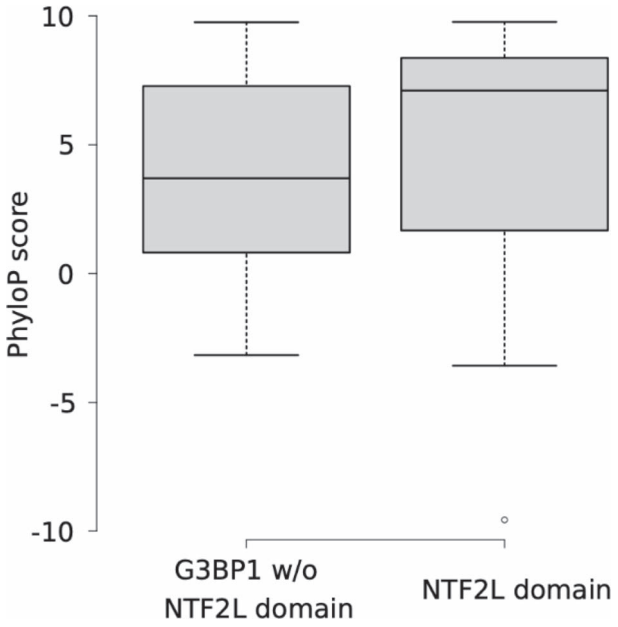


Figure 4. PhyloP conservation scores for G3BP1 in 100 vertebrate species. The Wilcoxon rank sum test *P*-value is reported.

the genomes of all merbecoviruses included in this study. Despite low sequence identity across the clade, all merbecoviruses were found to encode the 4a accessory protein with conserved charged residues critical for RNA binding (Fig. S1), implying that the RNA-binding function of p4a has been maintained across this lineage.

High evolutionary conservation of G3BP1 NTF2L domain highlights its functional importance in host–virus interactions

G3BP1, unlike other host proteins targeted by CoVs, has not undergone positive selection, indicating that it has remained relatively unchanged in bats or primates (Cariou *et al.* 2022). This stability is particularly noteworthy as previous studies have shown that fast-evolving viral motifs within IDRs often target highly conserved host proteins to enhance viral adaptability and exploit host cellular mechanisms (Shuler and Hagai 2022, Guan *et al.* 2023).

To investigate the evolution of G3BP1, we employed phyloP scores, which estimate evolutionary constraint at specific protein sites, derived from an alignment of 100 vertebrates. Positive phyloP scores indicate strong conservation, while negative scores reflect rapid evolution. By comparing the evolutionary rates of the G3BP1 NTF2L domain, which binds to the ITFG motif, with the rest of the protein, we observed significantly higher PhyloP scores in the NTF2L domain (Fig. 4 and Fig. S2), indicating the evolutionary conservation of the ITFG-binding pocket and its functional importance.

Discussion

Extensive research has established that many viruses, including polioviruses, porcine epidemic diarrhea virus, encephalomyocarditis virus, foot-and-mouth disease virus, alphaviruses, and noroviruses, disrupt SG assembly through direct targeting of G3BP1 (White *et al.* 2007, Panas *et al.* 2012, Ng *et al.* 2013, Panas *et al.* 2014, Hosmillo *et al.* 2019, Pandey *et al.* 2019, Visser *et al.* 2019, Jayabalan *et al.* 2021). For instance, Semliki Forest virus uses a motif similar to the ITFG motif in sarbecoviruses to bind to the NTF2L domain of G3BP1(Schulte *et al.* 2016, Yang *et al.* 2024). In CoVs, however, SG inhibition mechanisms are particularly complex, involving multiple viral proteins. For example, the N proteins of SARS-CoV-2 and HCoV-OC43 inhibit SG formation through distinct pathways (Dolliver *et al.* 2022). Moreover, the non-structural protein 1 (Nsp1), encoded by both viruses, also antagonizes SG assembly (Dolliver *et al.* 2022). In contrast, MERS-CoV relies exclusively on the p4a accessory protein to inhibit SG formation (Rabouw *et al.* 2016). Our data indicate that the β-CoV ancestor did not encode an N protein containing a motif capable of G3BP1 binding, and that the ability to interact with G3BP1 independently emerged twice during the evolution of β-CoVs. Although sequence-level convergence is incomplete, functional convergence is evident, as the N proteins encoded by both sarbecoviruses and a small clade of bat merbecoviruses are able to bind G3BP1. In the bat merbecovirus clade, this capability arose through a single amino acid substitution. Conversely, it remains difficult to speculate on the intermediate evolutionary steps that led to the emergence of the ITFG motif in sarbecoviruses. Our data indicate that the most critical substitution was A18G, since we demonstrate that the resulting VTFG motif is sufficient for G3BP1 binding. In any case, the convergent evolution of a functional G3BP1-binding motif strongly suggests a selective advantage conferred by this interaction.

Overall, these observations raise fundamental questions about the diversity of SG inhibition strategies: Why do SARS-CoV-2 and HCoV-OC43 employ distinct mechanisms to block SGs? What molecular targets are exploited by the HCoV-OC43 N protein? What evolutionary pressures led to the independent emergence of the ΦxFG motif in sarbecoviruses and a subset of bat-infecting merbecoviruses?

A plausible hypothesis is that G3BP1 targeting by the N protein bearing the ΦxFG motif also serves additional functions that are not limited to SG inhibition. For instance, while the SARS-CoV-2 N protein interacts with G3BP1, the N protein of HCoV-OC43, which does not bind to G3BP1, may instead engage with other SG-associated proteins to achieve similar or complementary effects. Given the distinct clinical outcomes of HCoV-OC43 and SARS-CoV-2 infections in humans, future studies comparing the molecular interactions of their N proteins—particularly with host factors such as SG components—may provide insight into virus-host adaptation mechanisms. However, we acknowledge that multiple viral and host factors contribute to virulence, and this remains a hypothesis *requiring further* experimental *investigation*.

Recent work by Burke *et al.* (2024) has shown that, in the absence of viral interference, G3BP1 uses its phase separation ability to sequester SARS-CoV-2 genomic RNA and other RNA viruses into viral aggregated RNA condensates (VARCs), thereby suppressing replication (Burke *et al.* 2024). Intriguingly, cells expressing a mutant G3BP1 variant (G3BP1-F124W), which supports SG formation but disrupts its interaction with the N protein, display reduced SARS-CoV-2 replication (Burke *et al.* 2024). Furthermore, G3BP-N complexes have been reported to promote highly localized translation of viral mRNAs near double membrane vesicles (DMVs), thus enhancing viral gene expression and replication (Long *et al.* 2024). These observations suggest that antiviral SGs and proviral granules involving G3BP1 may coexist in SARS-CoV-2-infected cells, collectively shaping viral replication and gene expression. Moreover, these findings emphasize the complex role of the interaction between β-CoV N proteins and G3BP1 in immune suppression, extending beyond SG disruption.

Whether MERS-CoV uses alternative mechanisms to prevent VARC formation and to promote viral mRNA translation remains an unresolved question. MERS-CoV has the highest pathogenic

potential among human-infecting β-CoVs. Previous studies have shown that a mutant virus lacking accessory protein 4a fails to inhibit SG formation and replicates less efficiently in human cells (Nakagawa *et al.* 2018). These effects are not mediated by the accumulation of viral mRNAs in SGs, but rather by the inefficient translation of viral proteins due to the sequestration of translation factors in SGs. These observations indicate that SGs have no proviral function during MERS-CoV infection and that β-CoVs have evolved highly different strategies to hijack the cellular machinery. In this context, identifying the interaction partner(s) of protein 4a may unveil novel therapeutic targets. Notably, our data indicate that a clade of bat-infecting viruses within the same genus as MERS-CoV has acquired the ability to interact with G3BP1. This adaptation does not appear to be a consequence of the loss of protein 4a, which remains relatively conserved among merbecoviruses. It will be extremely interesting to experimentally dissect the relative contribution of the G3BP1-binding motif and of protein 4a to SG disruption in these viruses. This observation suggests that, in bat-infecting merbecoviruses, the interaction between the N protein and G3BP1 may serve functions beyond stress granule inhibition. While SG interference appears to be a strategy used by several coronavirus genera to modulate host responses, the specific role of N–G3BP1 interactions in merbecovirus infection dynamics remains to be fully elucidated. Whether these viruses similarly exploit this host factor to enhance viral mRNA translation is an open and compelling question.

Finally, it is worth mentioning that we found the ΦxFG-binding pocket of G3BP1 to be highly conserved across vertebrates. This observation aligns with previous findings that viral motifs *often* hijack highly constrained host proteins, which are locked in their ability to evolve adaptive changes (Shuler and Hagai 2022). This strategy ensures that the viral mechanism for disrupting SG formation remains effective across a wide range of host species (Cariou *et al.* 2022). Therefore, our data provide insight into the evolutionary dynamics of host-pathogen interactions and the emergence of viral motifs within IDRs that can target conserved host proteins.

Acknowledgments

We thank Marcello Arsura for critically reviewing the manuscript. The authors thank Valeria Caneparo at CAAD, Center for Translational Research and Autoimmune and Allergic Disease for assistance in the SPR assay, and Serena Delbue for granting access to the biosafety level 3 (BSL3) facility at the University of Milan to obtain SARS-CoV-2-infected samples. The authors acknowledge support from the Department of Translational Medicine (DIMET) of the University del Piemonte Orientale through the APC initiative.

Author contributions

Conceptualization, C.B., M.G., M.S.; methodology, S.T., I.C., D.F., R.C., A.M.; formal analysis, S.T., I.C., C.B., D.F., M.S.; investigation, S.T., I.C., D.F., R.C., A.M.; funding acquisition, C.B., M.G., M.S., R.C.; writing, C.B., M.G., M.S.

Supplementary data

Supplementary data is available at *VEVOLU Journal* online.

Conflict of interest: The authors declare no competing interests.

Funding

This work was supported by Ministry for University and Research -MUR- (PRIN 2020: 2020KSY3KL_002 and PRIN 2022 PNRR: 2022FRE3RH to M.G.; PRIN 2022 PNRR: P20222HHXA to C.B.) and by the Italian Ministry of Health ('Ricerca Corrente' to RC).

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

References

Burke JM, Ratnayake OC, Watkins JM *et al.* G3BP1-dependent condensation of translationally inactive viral RNAs antagonizes infection. *Sci Adv* 2024;**10**:eadk8152. https://doi.org/10.1126/sciadv.adk8152

Cariou M, Picard L, Guéguen L *et al.* Distinct evolutionary trajectories of SARS-CoV-2-interacting proteins in bats and primates identify important host determinants of COVID-19. *Proc Natl Acad Sci USA* 2022;**119**:e2206610119. https://doi.org/10.1073/pnas.2206610119

Deng X, Baker SC. Coronaviruses: Molecular biology (Coronaviridae). *Encycl Virol* 2021;**2**:198–207. https://doi.org/10.1016/B978-0-12-814515-9.02550-9

Dolliver SM, Kleer M, Bui-Marinos MP *et al.* Nsp1 proteins of human coronaviruses HCoV-OC43 and SARS-CoV2 inhibit stress granule formation. *PLoS Pathog* 2022;**18**:e1011041.

Emenecker RJ, Griffith D, Holehouse AS. Metapredict: a fast, accurate, and easy-to-use predictor of consensus disorder and structure. *Biophys J* 2021;**120**:4312–9. https://doi.org/10.1016/j.bpj.2021.08.039

Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019;**20**:238. https://doi.org/10.1186/s13059-019-1832-y

Foley G, Mora A, Ross CM *et al.* Engineering indel and substitution variants of diverse and ancient enzymes using graphical representation of ancestral sequence predictions (GRASP). *PLoS Comput Biol* 2022;**18**:e1010633. https://doi.org/10.1371/journal.pcbi.1010633

Guan Y, Wang Y, Fu X *et al.* Multiple functions of stress granules in viral infection at a glance. *Front Microbiol* 2023;**14**:1138864. https://doi.org/10.3389/fmicb.2023.1138864

Hosmillo M, Lu J, McAllaster MR *et al.* Noroviruses subvert the core stress granule component G3BP1 to promote viral VPg-dependent translation. *eLife*. 2019;**8**:e46681. https://doi.org/10.7554/eLife.46681

Huang W, Ju X, Tian M *et al.* Molecular determinants for regulation of G3BP1/2 phase separation by the SARS-CoV-2 nucleocapsid protein. *Cell Discov* 2021;**7**:69. https://doi.org/10.1038/s41421-021-00306-w

Jayabalan AK, Adivarahan S, Koppula A *et al.* Stress granule formation, disassembly, and composition are regulated by alphavirus ADP-ribosylhydrolase activity. *Proc Natl Acad Sci USA* 2021;**118**:e2021719118. https://doi.org/10.1073/pnas.2021719118

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;**30**:772–80. https://doi.org/10.1093/molbev/mst010

Kosakovsky Pond SL, Posada D, Gravenor MB *et al.* Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol* 2006;**23**:1891–901. https://doi.org/10.1093/molbev/msl051

Kruse T, Benz C, Garvanska DH *et al.* Large scale discovery of coronavirus-host factor protein interaction motifs reveals SARS-CoV-2 specific mechanisms and vulnerabilities. *Nat Commun* 2021;**12**:6761. https://doi.org/10.1038/s41467-021-26498-z

Lo Cigno I, Calati F, Borgogna C *et al.* Human papillomavirus E7 oncoprotein subverts host innate immunity via SUV39H1-mediated epigenetic silencing of immune sensor genes. *J Virol* 2020;**94**:e01812–9. https://doi.org/10.1128/JVI.01812-19

Long S, Guzyk M, Perez Vidakovics L *et al.* SARS-CoV-2 N protein recruits G3BP to double membrane vesicles to promote translation of viral mRNAs. *Nat Commun* 2024;**15**:10607. https://doi.org/10.1038/s41467-024-54996-3

McCormick C, Khaperskyy DA. Translation inhibition and stress granules in the antiviral immune response. *Nat Rev Immunol* 2017;**17**:647–60. https://doi.org/10.1038/nri.2017.63

Minh BQ, Schmidt HA, Chernomor O *et al.* IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;**37**:1530–4. https://doi.org/10.1093/molbev/msaa015

Nakagawa K, Narayanan K, Wada M *et al.* Inhibition of stress granule formation by Middle East respiratory syndrome coronavirus 4a accessory protein facilitates viral translation, leading to efficient virus replication. *J Virol* 2018;**92**:e00902–18. https://doi.org/10.1128/JVI.00902-18

Ng CS, Jogi M, Yoo J-S *et al.* Encephalomyocarditis virus disrupts stress granules, the critical platform for triggering antiviral innate immune responses. *J Virol* 2013;**87**:9511–22. https://doi.org/10.1128/JVI.03248-12

Panas MD, Varjak M, Lulla A *et al.* Sequestration of G3BP coupled with efficient translation inhibits stress granules in Semliki Forest virus infection. *MBoC.* 2012;**23**:4701–12. https://doi.org/10.1091/mbc.e12-08-0619

Panas MD, Ahola T, McInerney GM. The C-terminal repeat domains of nsP3 from the old world alphaviruses bind directly to G3BP. *J Virol* 2014;**88**:5888–93. https://doi.org/10.1128/JVI.00439-14

Pandey K, Zhong S, Diel DG *et al.* GTPase-activating protein-binding protein 1 (G3BP1) plays an antiviral role against porcine epidemic diarrhea virus. *Vet Microbiol* 2019;**236**:108392. https://doi.org/10.1016/j.vetmic.2019.108392

Pond SLK, Frost SDW, Muse SV. HyPhy: hypothesis testing using phylogenies. *Bioinformatics.* 2005;**21**:676–9. https://doi.org/10.1093/bioinformatics/bti079

Rabouw HH, Langereis MA, Knaap RCM *et al.* Middle East respiratory coronavirus accessory protein 4a inhibits PKR-mediated antiviral stress responses. *PLoS Pathog* 2016;**12**:e1005982. https://doi.org/10.1371/journal.ppat.1005982

Salata C, Calistri A, Parolin C *et al.* Coronaviruses: a paradigm of new emerging zoonotic diseases. *Pathogens Dis* 2019;**77**:ftaa006. https://doi.org/10.1093/femspd/ftaa006

Schulte T, Liu L, Panas MD *et al.* Combined structural, biochemical and cellular evidence demonstrates that both FGDF motifs in alphavirus nsP3 are required for efficient replication. *Open Biol* 2016;**6**:160078. https://doi.org/10.1098/rsob.160078

Shuler G, Hagai T. Rapidly evolving viral motifs mostly target biophysically constrained binding pockets of host proteins. *Cell Rep* 2022;**40**:111212. https://doi.org/10.1016/j.celrep.2022.111212

Visser LJ, Medina GN, Rabouw HH *et al.* Foot-and-mouth disease virus leader protease cleaves G3BP1 and G3BP2 and inhibits stress granule formation. *J Virol* 2019;**93**:e00922–18. https://doi.org/10.1128/JVI.00922-18

White JP, Lloyd RE. Regulation of stress granules in virus systems. *Trends Microbiol* 2012;**20**:175–83. https://doi.org/10.1016/j.tim.2012.02.001

White JP, Cardenas AM, Marissen WE *et al.* Inhibition of cytoplasmic mRNA stress granule formation by a viral proteinase. *Cell Host Microbe* 2007;**2**:295–305. https://doi.org/10.1016/j.chom.2007.08.006

Yang Z, Johnson BA, Meliopoulos VA *et al.* Interaction between host G3BP and viral nucleocapsid protein regulates SARS-CoV-2 replication and pathogenicity. *Cell Rep* 2024;**43**:113965. https://doi.org/10.1016/j.celrep.2024.113965

Ye Z-W, Yuan S, Yuen K-S *et al.* Zoonotic origins of human coronaviruses. *Int J Biol Sci* 2020;**16**:1686–97. https://doi.org/10.7150/ijbs.45472

Zhang N, Chen Y, Lu H *et al.* MutaBind2: predicting the impacts of single and multiple mutations on protein-protein interactions. *iScience.* 2020;**23**:100939. https://doi.org/10.1016/j.isci.2020.100939

Virus Evolution, 2025, **11**, 1–8
https://doi.org/10.1093/ve/veaf063
Research Article