

PERSPECTIVE

Open Access

# New Consensus pattern in Spike CoV-2: potential implications in coagulation process and cell–cell fusion

Silvia Buonvino<sup>1</sup> and Sonia Melino<sup>1</sup>

## Abstract

Coagulopathy and syncytial formation are relevant effects of the SARS-CoV-2 infection, but the underlying molecular mechanisms triggering these processes are not fully elucidated. Here, we identified a potential consensus pattern in the Spike S glycoprotein present within the cytoplasmic domain; this consensus pattern was detected in only 79 out of 561,000 proteins (UniProt bank). Interestingly, the pattern was present in both human and bat the coronaviruses S proteins, in many proteins involved in coagulation process, cell–cell interaction, protein aggregation and regulation of cell fate, such as von Willebrand factor, coagulation factor X, fibronectin and Notch, characterized by the presence of the cysteine-rich EGF-like domain. This finding may suggest functional similarities between the matched proteins and the CoV-2 S protein, implying a new possible involvement of the S protein in the molecular mechanism that leads to the coagulopathy and cell fusion in COVID-19 disease.

The severe acute respiratory syndrome coronavirus (SARS-1) of 2002, the Middle East respiratory syndrome coronavirus (MERS-CoV) of 2012, and now the SARS-2 of 2019 (causing the COVID-19 disease) are all due to coronaviruses of the beta subgroup<sup>1</sup>. This positive-sense single-stranded RNA virus family possesses the structural proteins spike (S), membrane (M) and envelope (E) proteins, along with the nucleocapsid (N) protein<sup>2</sup>.

SARS spike glycoprotein is a trimeric protein that shows a large mass of 500 kDa for trimer and, striking, this appears like a three-bladed propeller with a radius of 90 Å<sup>3</sup> (see Fig. 1). The SARS spike protein is characterized by the presence of four structural domains (Fig. 1). While the two large ecto-domains S1 and S2 are responsible for receptor binding and membrane fusion respectively, the cytoplasmic domain (CD) has an important function in the assembly of several enveloped viruses, as described for other viral membrane proteins. In the case of

alphaviruses, for instance, the CD of the E2 glycoprotein plays a pivotal role in the interaction with the capsid protein during particle formation<sup>4,5</sup>. A critical role for the cytoplasmic tails in this process has been reported for several members including Simian virus 5 (SV5)<sup>6,7</sup>, Sendai virus<sup>8,9</sup>, and measles virus<sup>10</sup>. In Sendai virus the matrix protein was found to interact independently of the cytoplasmic tails of the HN and F glycoproteins<sup>11,12</sup>. In orthomyxovirus influenza A, the cytoplasmic tails of the two glycoproteins, HA and NA, influence budding efficiency as well as particle morphology. Their separate removal caused only limited effects while the lack of both tails resulted in severely impaired formation of deformed particles<sup>13–16</sup>. Interestingly, many viral ectodomain fragments of fusion protein without transmembrane (TM) and ENDO domains fold into a post-fusion states<sup>17,18</sup>, suggesting that membrane-anchoring parts help maintain functional metastable high energy conformations. In the case of HIV-1-gp41, for instance, Lu et al. report that antibodies (IgG) against LLP1–2 and LLP2 (lentivirus lytic peptide  $\alpha$  helix 1 and 2) regions inhibited HIV-1 Envelope-mediated cell fusion and bound to the interface

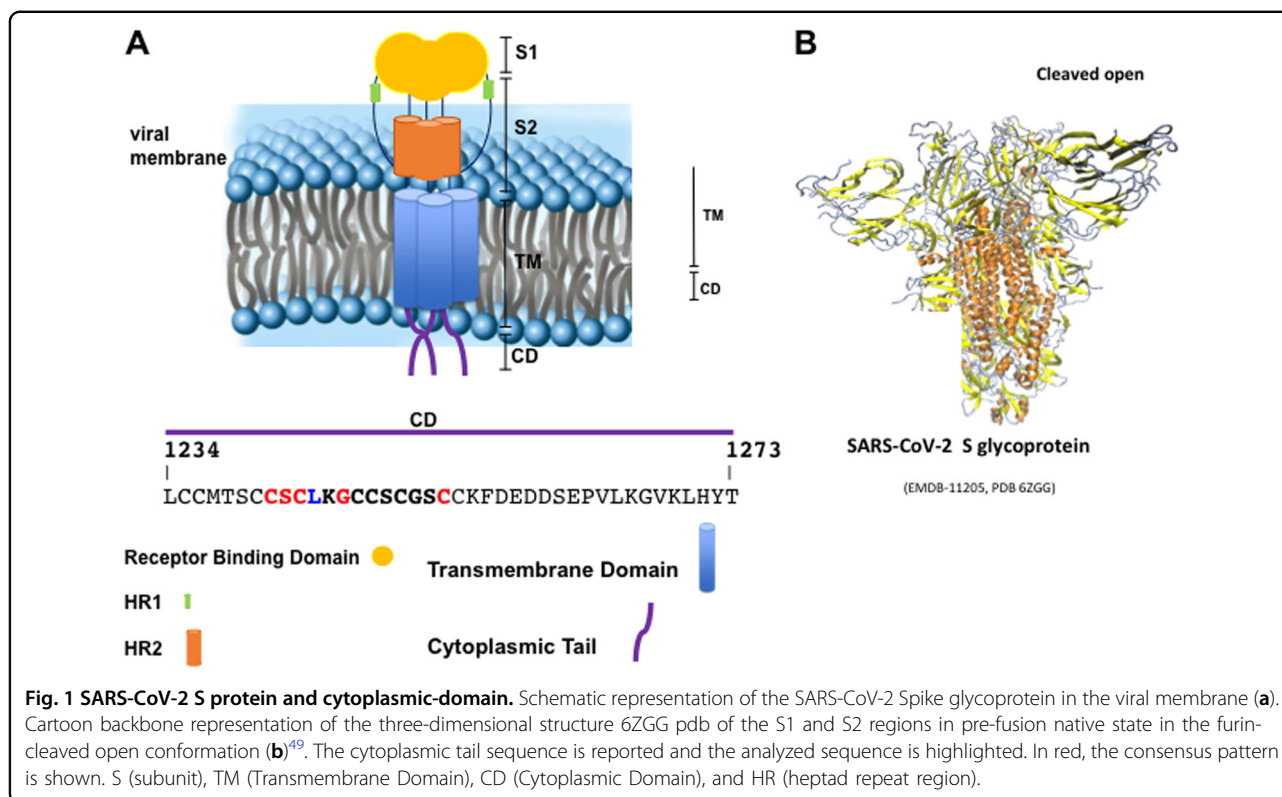
Correspondence: Sonia Melino ([melinos@uniroma2.it](mailto:melinos@uniroma2.it))

<sup>1</sup>Department of Chemical Science and Technologies, University of Rome “Tor Vergata”, via della Ricerca Scientifica 1, 00133 Rome, Italy  
Edited by R. A. Knight

© The Author(s) 2020



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.



between effector and target cells suggesting that LLP1–2, especially the LLP2 region located inside the viral membrane is transiently exposed on the membrane surface during the fusion process<sup>19</sup>.

For coronaviruses is not entirely clear how the intravirion parts of the fusion protein influence reactions that are carried out by the much larger exterior portion of the protein.

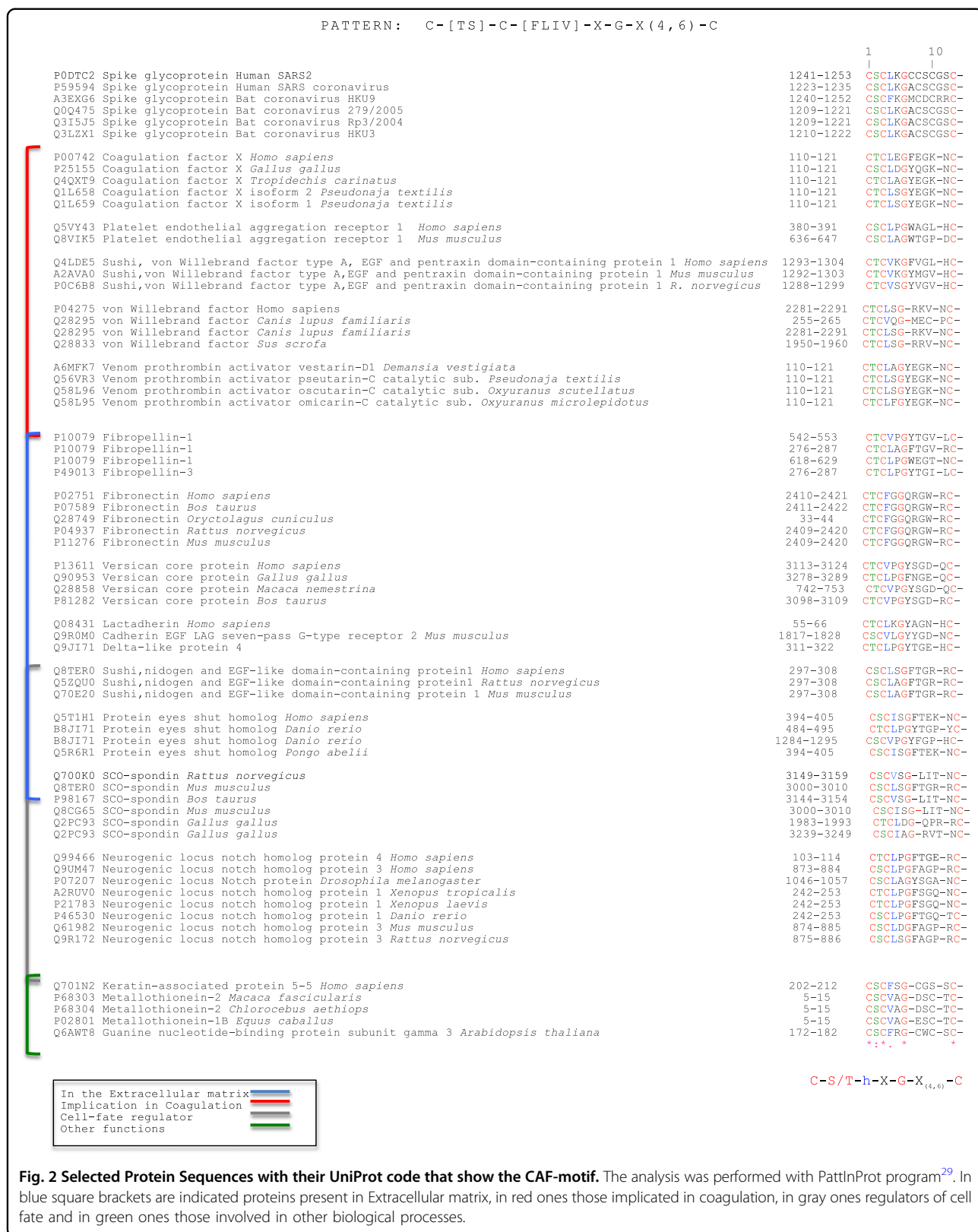
In the carboxy-terminal domain of the Coronaviral S protein there are two areas of conservation: one is at the transition of TM and ectodomain, i.e. where the S protein exits the viral membrane and it is characterized by a conspicuous, highly conserved 8-residues sequence (KWPWY/WVWL), probably important for membrane fusion but not for S protein incorporation into particles. The other area is located in the membrane-proximal part of the CD and it shows a conserved abundance of cysteines (Fig. 1). The carboxy-terminal truncations reveal that it is this specific domain that mediates particle assembly of the coronaviral spikes.

The importance of the cysteine-rich region for membrane fusion has already been established<sup>20,21</sup>. A spike mutant, with part of the cysteine-rich region deleted, was able to promote hemi-fusion, but was blocked in fusion pore formation. Whether this effect was due by preventing acylation is not completely clear, but it is possible that the membrane-inserted hydrophobic acyl chains are

implicated in fusion pore formation. A positive role of cysteine palmitoylation in cell fusion has been reported for influenza virus HA protein<sup>22,23</sup>, while a negative role was observed for Vesicular Stomatitis virus (VSV)<sup>24</sup>, influenza virus<sup>25,26</sup> and the Murine leukemia virus fusion protein<sup>27</sup>. In CoV S protein, the cysteines, and/or their palmitate adducts of the endo-domain, can change the rate-limiting step of the membrane fusion reaction<sup>28</sup>. Therefore, in this contest the role of the CD should be better investigated.

We analyzed the CD sequence of CoV-2 S protein in order to identify potential consensus sequence patterns (Fig. 1). A bioinformatic analysis using PattInProt v5.4up ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa\\_pattinprot.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_pattinprot.html))<sup>29</sup>, setting 100% of similarity, was performed identifying, in 79 proteins out of 561,000 proteins of the UniProt bank, a new potential amino acid pattern. The consensus pattern was C-[TS]-C-h-X-G-X(4,6)-C, herein called CAF-motif (Cysteine Aggregation Fusion), where h is a hydrophobic residue and X any other residue. The consensus pattern and the sequence alignments between the proteins and the pattern are shown in Fig. 2 and the PattInProt analysis is shown in Fig. 1S.

Interestingly, the only viral proteins that showed the CAF-motif were S proteins of human coronaviruses SARS and SARS -2 and the S proteins of bat coronavirus. Moreover, other proteins involved in coagulation,



**Fig. 2 Selected Protein Sequences with their UniProt code that show the CAF-motif.** The analysis was performed with PattlnProt program<sup>29</sup>. In blue square brackets are indicated proteins present in Extracellular matrix, in red ones those implicated in coagulation, in gray ones regulators of cell fate and in green ones those involved in other biological processes.

extracellular recognition and cell fate presented the same pattern. Intriguingly, the CAF-motif occurs in proteins such as coagulation factor X, von Willebrand factor, platelet endothelial aggregation receptor 1 and some prothrombin activators venom toxins that are involved in the coagulation process. The identification of a common pattern could suggest a new function of the S protein in the pathological effects of the SARS -2 infection.

Biopsy and autopsy studies on COVID-19 patients showed pulmonary pathology alveolar damage with diffuse thickening of the alveolar wall, formation of hyaline membranes and macrophages and mononuclear cells infiltration<sup>30–36</sup>. Moreover, according to recent reports, the most severely ill patients show relevant coagulopathy<sup>37,38</sup>. Clinical studies have revealed that 71.4% of non-survivors of COVID-19 matched the grade of overt disseminated intravascular coagulation ( $\geq 5$  points according to the International Society on Thrombosis and Haemostasis criteria) and showed abnormal coagulation results during later stages of the disease such as particularly increased levels of D-dimer and other fibrin degradation products that were significantly associated with poor prognosis<sup>39,40</sup>. The molecular mechanisms at the base of coagulopathy in COVID-19 disease are not yet identified so that the identification of this common pattern could suggest a similar molecular mechanism in the coagulation induction.

Other matched proteins are found in the extracellular matrix (ECM) involved in cellular adhesion to the matrix, cell–cell interaction, and cell signaling such as: fibropellin, fibronectin, versican, cadherins, and proteins responsible for cell fate regulation such as spondin, nidogen, and the cell-surface receptor Notch, which is also involved in fusion cell fate<sup>41</sup> and its ligand Dll4 (Delta-like 4).

Among the protein list are metallothioneins and Keratin-associated protein 5-5 (KRTAP 5-5); the first ones are a family of small, highly conserved, cysteine-rich metal-binding proteins important for zinc and copper homeostasis, buffering against toxic heavy metals and protection from oxidative stress, while KRTAP 5-5 belongs to a large protein family involved in cross-linking keratin intermediate filaments during hair formation process.

Intriguingly, this consensus pattern is located in the epidermal growth factor (EGF)-like domain present in the great part of the found proteins and characterized by six cysteines which form disulfide bonds within the domain (C1–C3, C2–C4, and C5–C6).

The EGF-like domain is involved in receptor–ligand interactions, extracellular matrix formation, cell adhesion and signal transduction, and chemotaxis<sup>42</sup> and in many proteins such as coagulation factors the EGF-like domains are known to bind calcium ions with D and E residues

allowing  $\text{Ca}^{2+}$  mediated protein–protein interactions<sup>43,44</sup>. In Notch, for example, the EGF motifs show multiple functions, such as the prevention of constitutive activation, reciprocal interaction with the ligands, and lateral interaction for homodimerization, playing a crucial role in Notch signaling system<sup>45</sup>. Moreover, Notch signaling, a major regulator of cardiovascular function and inflammation, is also implicated in several biological processes mediating viral infections such as SARS-CoV-2, playing an important role in developing of myocarditis, heart failure, and lung inflammation in COVID-19 patients<sup>46</sup>. In macrophages Dll1,4/Notch signaling promotes the inflammatory cytokines storm, interleukin-6 (IL-6) among those, which in turn increases the expression of notch ligands (Dll1,4), thus amplifying the signal establishing a feedback loop<sup>46</sup>.

In view of the above, we propose a hypothetic active role of the Coronavirus S protein cytoplasmic domain in protein–protein aggregation for clots formation and cell–cell fusion SARS-2-S protein-driven<sup>47</sup>. Therefore, our findings suggest a new potential molecular mechanism linked to the infection in which after virus-cell fusion, the infected cells expose on their surface the CAF-motif, leading to clots formation and cell–cell fusion by protein–protein aggregation processes. Moreover, it should not to be rolled out the ability of CD's S protein to coordinate  $\text{Ca}^{2+}$  ions, similarly to EGF-like domain, for mediating cell–cell fusion, which is responsible for instance of syncytia formation<sup>48</sup>.

The identification of this new consensus pattern provides a first evidence of a functional similarity between the CoV S protein and proteins involved in coagulation, in regulation of cell fate and cell–cell interaction and fusion that are at the base of coagulopathy and syncytia formation in COVID-19 disease.

#### Acknowledgements

We thank Edoardo Trotta and Gabriella Santoro for the critical discussion of the results.

#### Author contributions

S.M.: conceptualization, investigation, and manuscript writing; S.B.: investigation and manuscript writing.

#### Conflict of interest

The authors declare that they have no conflict of interest.

#### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

The online version of this article (<https://doi.org/10.1038/s41420-020-00372-1>) contains supplementary material, which is available to authorized users.

Received: 27 October 2020 Revised: 4 November 2020 Accepted: 5 November 2020

Published online: 27 November 2020



## References

- Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C. & Garry, R. F. The proximal origin of SARS-CoV-2. *Nat. Med.* **26**, 450–452 (2020).
- Li, F. Structure, function, and evolution of coronavirus spike proteins. *Annu. Rev. Virol.* **3**, 237–261 (2016).
- Beniac, D. R., Andonov, A., Grudeski, E. & Booth, T. F. Architecture of the SARS coronavirus prefusion spike. *Nat. Struct. Mol. Biol.* **13**, 751–752 (2006).
- Suomalainen, M., Liljestrom, P. & Garoff, H. Spike protein-nucleocapsid interactions drive the budding of alphaviruses. *J. Virol.* **66**, 4737–4747 (1992).
- Zhao, H., Lindqvist, B., Garoff, H., von Bonsdorff, C. H. & Liljestrom, P. A tyrosine-based motif in the cytoplasmic domain of the alphavirus envelope protein is essential for budding. *EMBO J.* **13**, 4204–4211 (1994).
- Schmitt, A. P., He, B. & Lamb, R. A. Involvement of the cytoplasmic domain of the hemagglutinin-neuraminidase protein in assembly of the paramyxovirus simian virus 5. *J. Virol.* **73**, 8703–8712 (1999).
- Waning, D. L., Schmitt, A. P., Leser, G. P. & Lamb, R. A. Roles for the cytoplasmic tails of the fusion and hemagglutinin-neuraminidase proteins in budding of the paramyxovirus simian virus 5. *J. Virol.* **76**, 9284–9297 (2002).
- Fouillot-Coriou, N. & Roux, L. Structure-function analysis of the Sendai virus F and HN cytoplasmic domain: different role for the two proteins in the production of virus particle. *Virology* **270**, 464–475 (2000).
- Takimoto, T., Bousse, T., Coronel, E. C., Scroggs, R. A. & Portner, A. Cytoplasmic domain of Sendai virus HN protein contains a specific sequence required for its incorporation into virions. *J. Virol.* **72**, 9747–9754 (1998).
- Cathomen, T., Naim, H. Y. & Cattaneo, R. Measles viruses with altered envelope protein cytoplasmic tails gain cell fusion competence. *J. Virol.* **72**, 1224–1234 (1998).
- Ali, A., Avalos, R. T., Ponimaskin, E. & Nayak, D. P. Influenza virus assembly: effect of influenza virus glycoproteins on the membrane association of M1 protein. *J. Virol.* **74**, 8709–8719 (2000).
- Sanderson, C. M., Wu, H. H. & Nayak, D. P. Sendai virus M protein binds independently to either the F or the HN glycoprotein in vivo. *J. Virol.* **68**, 69–76 (1994).
- Jin, H., Leser, G. P. & Lamb, R. A. The influenza virus hemagglutinin cytoplasmic tail is not essential for virus assembly or infectivity. *EMBO J.* **13**, 5504–5515 (1994).
- Jin, H., Leser, G. P., Zhang, J. & Lamb, R. A. Influenza virus hemagglutinin and neuraminidase cytoplasmic tails control particle shape. *EMBO J.* **16**, 1236–1247 (1997).
- Mitnaul, L. J., Castrucci, M. R., Murti, K. G. & Kawaoka, Y. The cytoplasmic tail of influenza A virus neuraminidase (NA) affects NA incorporation into virions, virion morphology, and virulence in mice but is not essential for virus replication. *J. Virol.* **70**, 873–879 (1996).
- Zhang, J., Pekosz, A. & Lamb, R. A. Influenza virus assembly and lipid raft microdomains: a role for the cytoplasmic tails of the spike glycoproteins. *J. Virol.* **74**, 4634–4644 (2000).
- Yin, H. S., Paterson, R. G., Wen, X., Lamb, R. A. & Jardetzky, T. S. Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein. *Proc. Natl Acad. Sci. USA* **102**, 9288–9293 (2005).
- Markosyan, R. M., Cohen, F. S. & Melikyan, G. B. HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation. *Mol. Biol. Cell* **14**, 926–938 (2003).
- Lu, L. et al. Surface exposure of the HIV-1 env cytoplasmic tail LLP2 domain during the membrane fusion process: interaction with gp41 fusion core. *J. Biol. Chem.* **283**, 16723–16731 (2008).
- Bos, E. C., Heijnen, L., Luytjes, W. & Spaan, W. J. Mutational analysis of the murine coronavirus spike protein: effect on cell-to-cell fusion. *Virology* **214**, 453–463 (1995).
- Chang, K. W., Sheng, Y. & Gombold, J. L. Coronavirus-induced membrane fusion requires the cysteine-rich domain in the spike protein. *Virology* **269**, 212–224 (2000).
- Naevé, C. W. & Williams, D. Fatty acids on the A/Japan/305/57 influenza virus hemagglutinin have a role in membrane fusion. *EMBO J.* **9**, 3857–3866 (1990).
- Sakai, T., Ohuchi, R. & Ohuchi, M. Fatty acids on the A/USSR/77 influenza virus hemagglutinin facilitate the transition from hemifusion to fusion pore formation. *J. Virol.* **76**, 4603–4611 (2002).
- Whitt, M. A. & Rose, J. K. Fatty acid acylation is not required for membrane fusion activity or glycoprotein assembly into VSV virions. *Virology* **185**, 875–878 (1991).
- Naim, H. Y., Amarnah, B., Ktistakis, N. T. & Roth, M. G. Effects of altering palmitoylation sites on biosynthesis and function of the influenza virus hemagglutinin. *J. Virol.* **66**, 7585–7588 (1992).
- Steinhauer, D. A., Wharton, S. A., Skehel, J. J., Wiley, D. C. & Hay, A. J. Amantadine selection of a mutant influenza virus containing an acid-stable hemagglutinin glycoprotein: evidence for virus-specific regulation of the pH of glycoprotein transport vesicles. *Proc. Natl Acad. Sci. USA* **88**, 11525–11529 (1991).
- Yang, C. & Compans, R. W. Analysis of the cell fusion activities of chimeric simian immunodeficiency virus-murine leukemia virus envelope proteins: inhibitory effects of the R peptide. *J. Virol.* **70**, 248–254 (1996).
- Shulla, A. & Gallagher, T. Role of spike protein endodomains in regulating coronavirus entry. *J. Biol. Chem.* **284**, 32725–32734 (2009).
- Combet, C., Blanchet, C., Geourjon, C. & Delage, G. NPS@: network protein sequence analysis. *Trends Biochem. Sci.* **25**, 147–150 (2000).
- Chen, J. et al. COVID-19 infection: the China and Italy perspectives. *Cell Death Dis.* **11**, 438 (2020).
- Venkatakrishnan, A. J. et al. Benchmarking evolutionary tinkering underlying human-viral molecular mimicry shows multiple host pulmonary-arterial peptides mimicked by SARS-CoV-2. *Cell Death Discov.* **6**, 96 (2020).
- Gebicki, J. & Wiczorkowska, M. COVID-19 infection: mitohormetic concept of immune response. *Cell Death Discov.* **6**, 60 (2020).
- Xu, J. et al. Digestive symptoms of COVID-19 and expression of ACE2 in digestive tract organs. *Cell Death Discov.* **6**, 76 (2020).
- Shi, Y. et al. COVID-19 infection: the perspectives on immune responses. *Cell Death Differ.* **27**, 1451–1454 (2020).
- Matsuyama, T., Kubli, S. P., Yoshinaga, S. K., Pfeffer, K. & Mak, T. W. An aberrant STAT pathway is central to COVID-19. *Cell Death Differ.* **oct 9**, 1–17 (2020).
- Shi, C. S., Nabar, N. R., Huang, N. N. & Kehrl, J. H. SARS-Coronavirus Open Reading Frame-8b triggers intracellular stress pathways and activates NLRP3 inflammasomes. *Cell Death Discov.* **5**, 101 (2019).
- China National Health Commission. Diagnosis and treatment of novel coronavirus pneumonia in China (trial version 7). [https://www.who.int/docs/default-source/wpro-documents/countries/china/covid-19-briefing-nhc-1-clinical-protocols-for-the-diagnosis-and-treatment-of-covid-19v7.pdf?sfvrsn=c6cbfa4\\_2](https://www.who.int/docs/default-source/wpro-documents/countries/china/covid-19-briefing-nhc-1-clinical-protocols-for-the-diagnosis-and-treatment-of-covid-19v7.pdf?sfvrsn=c6cbfa4_2). Accessed 14 April 2020.
- Yao, X. H. et al. A pathological report of three COVID-19 cases by minimal invasive autopsies. *Zhonghua Bing Li Xue Za Zhi* **49**, 411–417 (2020).
- Zhou, F. et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* **395**, 1054–1062 (2020).
- Iba, T., Levy, J. H., Levi, M., Connors, J. M. & Thachil, J. Coagulopathy of coronavirus disease 2019. *Crit. Care Med.* **48**, 1358–1364 (2020).
- Artero, R. D., Castanon, I. & Bayliss, M. K. The immunoglobulin-like protein Hibris functions as a dose-dependent regulator of myoblast fusion and is differentially controlled by Ras and Notch signaling. *Development* **128**, 4251–4264 (2001).
- Ma, J. et al. Role of a novel EGF-like domain-containing gene NGX6 in cell adhesion modulation in nasopharyngeal carcinoma cells. *Carcinogenesis* **26**, 281–291 (2005).
- Rao, Z. et al. Crystallization of a calcium-binding EGF-like domain. *Acta Crystallogr. D Biol. Crystallogr.* **51**, 402–403 (1995).
- Elies, J. et al. *Calcium Signaling. Advances in experimental medicine and biology* (ed. Islam M.) Vol. 1131. 183–213 (Springer, Cham, 2020).
- Sakamoto, K., Chao, W. S., Katsube, K. & Yamaguchi, A. Distinct roles of EGF repeats for the Notch signaling system. *Exp. Cell Res.* **302**, 281–291 (2005).
- Rizzo, P. et al. COVID-19 in the heart and the lungs: could we “Notch” the inflammatory storm? *Basic Res. Cardiol.* **115**, 31 (2020).
- Hoffmann, M., Kleine-Weber, H. & Pohlmann, S. A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol. Cell* **78**, 779–784 (2020).
- Chen, E. H., Grote, E., Mohler, W. & Vignery, A. Cell-cell fusion. *FEBS Lett.* **581**, 2181–2193 (2007).
- Wrobel, A. G. et al. SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat. Struct. Mol. Biol.* **27**, 763–767 (2020).