



Universiteit
Leiden
The Netherlands

Recognizing species as a new focus of virus research

Gorbalenya, A.E.; Siddell, S.G.

Citation

Gorbalenya, A. E., & Siddell, S. G. (2021). Recognizing species as a new focus of virus research. *Plos Pathogens*, 17(3). doi:10.1371/journal.ppat.1009318

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3214302>

Note: To cite this publication please use the final published version (if applicable).

OPINION

Recognizing species as a new focus of virus research

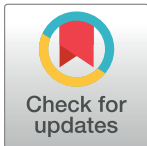
Alexander E. Gorbalenya^{1,2*}, Stuart G. Siddell³

1 Department of Medical Microbiology, Leiden University Medical Center, Leiden, the Netherlands, **2** Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia, **3** School of Cellular and Molecular Medicine, University of Bristol, Bristol, United Kingdom

* a.e.gorbalenya@lumc.nl

Abstract

Species taxa are the units of taxonomy most suited to measure virus diversity, and they account for more than 70% of all virus taxa. Yet, as evidenced by the content of GenBank entries and illustrated by the recent literature on SARS-CoV-2, they are the most neglected taxa of virus research. To correct this disparity, we propose to make species taxa a first choice for communicating virus taxonomy in publications concerning viruses. We see it as a key step toward promoting research on diverse viruses, including pathogens, at this fundamental level of biology.



OPEN ACCESS

Citation: Gorbalenya AE, Siddell SG (2021) Recognizing species as a new focus of virus research. *PLoS Pathog* 17(3): e1009318. <https://doi.org/10.1371/journal.ppat.1009318>

Editor: Thomas E. Morrison, University of Colorado Denver, UNITED STATES

Published: March 4, 2021

Copyright: © 2021 Gorbalenya, Siddell. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Work of Alexander E. Gorbalenya was partially supported by the Leiden University Fund. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

From virus pathogens to species

Viruses were discovered as miniscule agents causing infectious diseases [1–3], and this link to pathogenicity, which still pervades virology, is often reflected in the matching names for viruses and their associated diseases. Even in more recent times, where genomic characterization has come to play a central role, viruses continue to be viewed as phylogenetic clusters of entities that are closely related to the originally identified pathogens, sometimes with a division into subgroups such as genotypes or serotypes. The major virus pathogens of humans, including influenza virus, human immunodeficiency virus (HIV), and hepatitis C virus (HCV) were defined in this pathogen-centrist way [4–6]. However, with the advent of metagenomic (and transcriptomic) studies, which drive current virus discovery and account for the ever-increasing share of known viruses (see, e.g., [7,8]), a pathogen-based nucleation approach is not possible. Instead, these studies identify a virus as a genomic sequence from random sampling of the respective natural diversity, and typically, associated with no phenotype. The uncertain relationship between viruses that were defined as pathogens and those identified through metagenomics can best be resolved within the framework of a virus taxonomy that is based on a concept of species. This short opinion article calls for species to be recognized as the primary subjects of virology, to parallel a practice common elsewhere in biology.

Species in biology and virology

Species are the units of evolution, and surveys of the corresponding taxa are used as a census of the abundance of natural biodiversity [9]. Yet virology is the only discipline that has not yet

fully embraced this basic tenet of biology. We believe this is largely due to a fundamental and unresolved question: to what degree does virus taxonomy equate its species taxa with the biological entities, also named species, that are studied throughout the rest of biology? The issues of species definition, demarcation, and meaning in biology are legendary [10–12], but the same questions, which are just as relevant to virology and the classification of viruses [13–19], are not widely discussed among virologists. We maintain this is because the study of viruses in the context of their species has not entered the mainstream of virus research.

This can be illustrated by 2 examples. First, in GenBank, which is a popular entry point for everybody interested in biological information [20], the genome sequences of cellular organisms are specified using the respective species name. In contrast, the genome sequences of viruses are typically specified in the context of their virus names. Virus species are not commonly defined in the taxonomy hierarchy that GenBank uses to describe the position of viruses, although all other taxonomic ranks are listed. Indeed, the delineation of virus species at NCBI can most readily be located at pages exclusively dedicated to taxonomy, sending a message that this rank is of concern only for those interested in virus taxonomy, rather than to all virologists [14].

The second example we present is the literature relating to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the cause of the coronavirus disease 2019 (COVID-19) pandemic [21–23]. Although this virus was identified only a year ago, the rapidly accumulating literature already exceeds many tens of thousands of papers: a scale of attention afforded to only a few other viruses. The virus was named SARS-CoV-2 due to its assignment into the pre-existing species *Severe acute respiratory syndrome-related virus* by the *Coronaviridae* Study Group (CSG) [24]. The paper title and its Abstract include the species name, and the public release of the preprint on the bioRxiv server and the subsequent publication of the paper by *Nature Microbiology* were commented on by diverse media, including *Nature* and *Science*. Indeed, the name SARS-CoV-2 has been broadly accepted by the international community. Nevertheless, the readers of the SARS-CoV-2 literature will be hard-pressed to find a few mentions of the cognate species taxon of this virus (*Severe acute respiratory syndrome-related virus*) in these publications, regardless of whether or not the original CSG paper is acknowledged.

One might argue that everybody knows where SARS-CoV-2 belongs, making an explicit reference to its multirank hierarchical taxonomy excessive in papers produced at the time when practical matters of containing the pandemic are the absolute priority. Indeed, many researchers chose not to refer to the coronavirus taxonomy in their publications. However, if any mention of taxonomy is made, the reference is typically to the family *Coronaviridae*, the genus *Betacoronavirus* or the subgenus *Sarbecovirus*. This disparity is especially notable since all of the aforementioned taxa were delineated and named at very different times. For example, the increasingly popular subgenus *Sarbecovirus* was first recognized only in 2018 [25], more than a decade later than the species *Severe acute respiratory syndrome-related virus* [26]. Given that virus species account for more than 70% of all virus taxa (Fig 1), this example reveals a disconnect between what virologists consider to be the most informative ranks of virus taxonomy (namely, any rank other than species) and what taxonomists think of as the cornerstone of virus taxonomy (the species rank). Essentially, it means that the development of virology proceeds largely without the benefit of the species concept, which is at the foundation of modern biology.

Species in virology: The roots of neglect

So why do virologists not mention virus species in their work? There are probably many reasons, including (1) a reluctance to accept viruses as biological entities, despite extensive

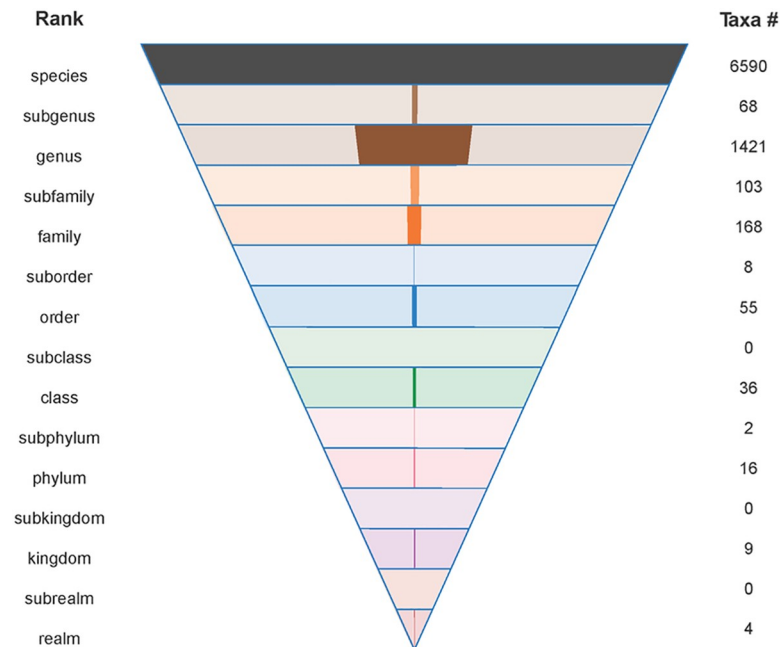


Fig 1. The 15 ranks of virus taxonomy are shown as an inverted pyramid, and the number of taxa currently assigned to each rank (December 2020) is shown both numerically and graphically.

<https://doi.org/10.1371/journal.ppat.1009318.g001>

evidence to the contrary; (2) a predominant focus of virus research and its applications on circulating (often pathogenic) viruses that are typically a small subset of the species taxon, although they may be sufficiently divergent to be mistaken for a “species,” tacitly or explicitly; (3) a tradition in virology for citing taxa at the family and genus ranks, mainly because these ranks were established first [27]; and (4) the persistent confusion of species taxa names with those of the viruses that populate these taxa. For instance, HIV-1 and HIV-2 may be easily mistaken as acronyms of the species *Human immunodeficiency virus 1* and *Human immunodeficiency virus 2*, respectively. This is incorrect. The acronyms refer to the viruses, human immunodeficiency virus type 1 and human immunodeficiency virus type 2, and not to the species [5]. A similar opportunity for confusion may be observed in many other instances, for example, influenza A-D viruses, and, until recently, HCV, whose species was renamed to *Hepacivirus C* to address this problem [28]. It is to be hoped that the recent introduction of a Linnaean-like 15 rank structure into virus taxonomy [29], and the proposed introduction of a binomial species nomenclature for viruses [30], may contribute to improving the discrimination of virus names and species taxa names in a more consistent and transparent manner. However, this change will only have a real impact in virology and beyond if researchers recognize the importance of virus species and start referring to species taxa.

Species in virology: Collaboration between virologists and the ICTV

Compared to the taxonomies of other biological entities, the development of virus taxonomy is uniquely supervised by a governing body, the International Committee on Taxonomy of Viruses (ICTV) [31]. This system provides a framework to develop a unified approach to virus classification and taxon nomenclature, which may be difficult to achieve otherwise, due to the unparalleled diversity of viruses and their ubiquitous infection of cellular organisms. However, this top-down arrangement, which has been applauded by some [32], might also create an

impression that the development of virus taxonomy is relevant only to the ICTV. Nothing could be further from the truth. Virus taxonomy is a basis for organizing knowledge about known viruses within a hierarchical evolutionary-based framework, which, in turn, facilitates knowledge access, exploration, and generalization by the entire research community, as well as others who contribute to or benefit from advancements in virology.

This connection of taxonomy to virus research may not always be apparent, but it is easily seen if one looks at how the taxonomy of distinct virus families is developed in practice. This responsibility rests mainly with experts on the viruses belonging to different families, who join together to form specialized Study Groups (SG), such as the CSG mentioned above. There are currently more than 100 SGs, and the number is likely to increase dramatically in the next few years. Taxonomic proposals may be initiated by any researcher, but it is generally the SGs that delineate taxa using the methods and demarcation criteria believed, at the time, to be best for partitioning the natural diversity of the respective virus groups. This modular structure also allows for flexibility as the SGs are largely independent of each other and develop their remit as they see fit. These SGs are the main conduits to connect the ICTV to all who are concerned with viruses.

Species in virology: Needs and benefits

We believe that to move the field forwards, the development of virus taxonomy has to be translated into research focusing on virus taxa, especially at the species level; a research area that is chronically underfunded, and how could it be otherwise? We are all aware of the chicken and egg mentality that bases research funding on prior research and creates a self-propagating circle of ignorance. What is needed is research that, for instance, includes the experimental verification of species by systematically probing the genetic compatibility of viruses at intra- and interspecies levels, as has been initiated for polioviruses and other viruses of the species *Enterovirus C* [33,34]. Another avenue to explore is the computational analysis of species demarcation, for example, the statistical evaluation of sequence discontinuity across different virus groups [35]. Only in this context will it become clear to virologists why, for example, HIV-1 and HIV-2 are assigned to 2 different species, while SARS-CoV and SARS-CoV-2 are 2 strains of a single species. Are these different assignments of a fundamental nature or the result of different approaches to species demarcation by the respective SGs? By resolving this question through empirical research on viruses at the species level, virologist will be able to understand better the relationship of taxonomy to biology and appreciate its implications, including the origins of pathogens. When studied in the context of their relatively long life span, the characterization of species can moderate the boom-and-bust cycles of interest and funding associated with virus pandemics, and advance virology in a conceptually coherent manner across the full spectrum of viruses.

Concluding remarks: Focusing research on virus species

We call on all virologists, whether or not they are interested in taxonomy, to reflect on the importance of species taxa in their research and take the lead in referring to them in their publications and other communications. We hope that this will encourage others including, for example, GenBank to follow. Moreover, there should be more awareness among the editors of journals and science communications, as well as administrators of regulatory bodies, regarding the proper and consistent communication of virus species. In this regard, the ICTV has an important role to play in providing accurate and up-to-date information, which can be found in the current Master Species List and Virus Metadata Resource at the ICTV website (<https://ictv.global/>).

In our opinion, it would be logical to make species taxa, the principal unit of taxonomy, a first choice for communicating virus taxonomy in publications and other documents concerning viruses. This would represent a first but significant step forward in the effort to define a position for the virus entity in the natural world. Using species abundance as a census of virus diversity in different hosts (notably humans) [36] and biomes [37–40] would be another step in this direction. Importantly, these developments would also facilitate research into the full extent of viruses at the species level, as is common elsewhere in biology [41]. In practical terms, recognizing virus species as the principal subjects of virology will also expand the scale of the spatiotemporal framework that connects the studies of natural virus variation, cross-host transmission, and pathogenicity, and thus contribute to the understanding and control of virus infections as part of biome characterization and exploration of the biosphere.

Acknowledgments

AEG and SGS have many years of experience in the study of viruses, using a variety of bioinformatic and laboratory-based approaches, and are former and current Vice-Presidents of the ICTV, respectively. This paper reflects their personal opinions.

References

1. Iwanowsky D. Über die Mosaikkrankheit der Tabakspflanze. Bulletin de l'Académie Impériale des Sciences de St-Petersbourg. 1892; 35:67–70.
2. Beijerinck MW. Über ein Contagium vivum fluidum als Ursache der Fleckenkrankheit der Tabaksblätter. Verhandlungen der Koninklijke akademie van Wetenschappen te Amsterdam. 1898; 65:1–22.
3. Loeffler F, Frosch P. Berichte der Kommission zur Erforschung der Maul- und Klauenseuche bei dem Institut für Infektionskrankheiten in Berlin. Abt. I. Centralblatt für Bakteriologie. Parasitenkunde und Infektionskrankheiten. 1898; 23:371–91.
4. A revised system of nomenclature for influenza viruses. Bull World Health Organ. 1971; 45 (1):119–24. PMID: [5316848](https://pubmed.ncbi.nlm.nih.gov/5316848/)
5. Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med. 2011 Sep; 1 (1):a006841. <https://doi.org/10.1101/cshperspect.a006841> PMID: [22229120](https://pubmed.ncbi.nlm.nih.gov/22229120/)
6. Simmonds P, Alberti A, Alter HJ, Bonino F, Bradley DW, Brechot C, et al. A proposed system for the nomenclature of hepatitis C viral genotypes. Hepatology. 1994 May; 19 (5):1321–4. PMID: [8175159](https://pubmed.ncbi.nlm.nih.gov/8175159/)
7. Zhang Y-Z, Chen Y-M, Wang W, Qin X-C, Holmes EC. Expanding the RNA Virosphere by Unbiased Metagenomics. Annu Rev Virol. 2019 Sep 29; 6 (1):119–39. <https://doi.org/10.1146/annurev-virology-092818-015851> PMID: [31100994](https://pubmed.ncbi.nlm.nih.gov/31100994/)
8. Wolf YI, Silas S, Wang Y, Wu S, Bocek M, Kazlauskas D, et al. Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. Nat Microbiol. 2020 Oct; 5 (10):1262–70. <https://doi.org/10.1038/s41564-020-0755-4> PMID: [32690954](https://pubmed.ncbi.nlm.nih.gov/32690954/)
9. Ereshefsky M (ed). The Units of Evolution: Essays on the Nature of Species. Cambridge, USA: MIT Press; 1992.
10. De Queiroz K. Species concepts and species delimitation. Syst Biol. 2007 Dec; 56 (6):879–86. <https://doi.org/10.1080/10635150701701083> PMID: [18027281](https://pubmed.ncbi.nlm.nih.gov/18027281/)
11. Campillo LC, Barley AJ, Thomson RC. Model-Based Species Delimitation: Are Coalescent Species Reproductively Isolated? Syst Biol. 2020 Jul 1; 69 (4):708–21. <https://doi.org/10.1093/sysbio/syz072> PMID: [31688934](https://pubmed.ncbi.nlm.nih.gov/31688934/)
12. Mallo D, Posada D. Multilocus inference of species trees and DNA barcoding. Philos Trans R Soc Lond Ser B Biol Sci. 2016 Sep; 5:371(1702). <https://doi.org/10.1098/rstb.2015.0335> PMID: [27481787](https://pubmed.ncbi.nlm.nih.gov/27481787/)
13. Lauber C, Gorbalenya AE. Toward genetics-based virus taxonomy: comparative analysis of a genetics-based classification and the taxonomy of picornaviruses. J Virol; 86 (7):3905–15. <https://doi.org/10.1128/JVI.07174-11> PMID: [22278238](https://pubmed.ncbi.nlm.nih.gov/22278238/)
14. Peterson AT. Defining viral species: making taxonomy useful. Virol J. 2014 Jul 23; 11:131. <https://doi.org/10.1186/1743-422X-11-131> PMID: [25055940](https://pubmed.ncbi.nlm.nih.gov/25055940/)
15. Bobay L-M, Ochman H. Biological species in the viral world. Proc Natl Acad Sci U S A. 2018 Jun 5; 115 (23):6040–5. <https://doi.org/10.1073/pnas.1717593115> PMID: [29784828](https://pubmed.ncbi.nlm.nih.gov/29784828/)

16. Alimpiev E. Rethinking the Virus Species Concept. http://stanford.edu/~alimpiev/thnk_ppr.pdf [Internet]. Stanford University; 2019. Available from: http://stanford.edu/~alimpiev/thnk_ppr.pdf.
17. Conix S. Measuring evolutionary independence: A pragmatic approach to species classification. *Biol Philos*. 2019 Oct 26; 34 (6):53.
18. Van Regenmortel MHV. The Species Problem in Virology. *Adv Virus Res*. 2018; 100:1–18. <https://doi.org/10.1016/bs.avir.2017.10.008> PMID: 29551132
19. Gibbs AJ, Gibbs MJ. A broader definition of “the virus species”. *Arch Virol*. 2006 Jul; 151 (7):1419–22. <https://doi.org/10.1007/s00705-006-0775-2> PMID: 16718358
20. Schoch CL, Ciuffo S, Domrachev M, Hottel CL, Kannan S, Khovanskaya R, et al. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database* [Internet]. 2020 [cited 2020 Jul 12];2020(baaa062). Available from: <https://doi.org/10.1093/database/baaa062>. <https://doi.org/10.1093/database/baaa062> PMID: 32761142
21. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020 Mar; 579 (7798):265–9. <https://doi.org/10.1038/s41586-020-2008-3> PMID: 32015508
22. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020 Mar; 579 (7798):270–3. <https://doi.org/10.1038/s41586-020-2012-7> PMID: 32015507
23. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med*. 2020; 382 (8):727–33.
24. Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol*. 2020 Apr 1; 5 (4):536–44. <https://doi.org/10.1038/s41564-020-0695-z> PMID: 32123347
25. Siddell SG, Walker PJ, Lefkowitz EJ, Mushegian AR, Adams MJ, Dutilh BE, et al. Additional changes to taxonomy ratified in a special vote by the International Committee on Taxonomy of Viruses (October 2018). *Arch Virol*. 2019 Mar; 164 (3):943–6. <https://doi.org/10.1007/s00705-018-04136-2> PMID: 30663020
26. Spaan WJM, Brian D, Cavanagh D, de Groot R, Enjuanes L, Gorbalenya A, et al. Coronaviridae. In: *Virus Taxonomy Eighth report of the International Committee on Taxonomy of Viruses* (eds Fauquet CM et al). Elsevier, Academic Press.; 2005. p. 947–64.
27. Wildy P. Classification and Nomenclature of Viruses. S. Karger; 1971. *Monographs in Virology Vol. 5* (ed. Melnick, J. L.).
28. Smith DB, Becher P, Bukh J, Gould EA, Meyers G, Monath T, et al. Proposed update to the taxonomy of the genera Hepacivirus and Pegivirus within the Flaviviridae family. *J Gen Virol*. 2016 Nov; 97 (11):2894–907. <https://doi.org/10.1099/jgv.0.000612> PMID: 27692039
29. Gorbalenya AE, Krupovic M, Mushegian A, Kropinski AM, Siddell SG, Varsani A, et al. The new scope of virus taxonomy: partitioning the virosphere into 15 hierarchical ranks. *Nat Microbiol*. 2020; 5 (5):668–74. <https://doi.org/10.1038/s41564-020-0709-x> PMID: 32341570
30. Siddell SG, Walker PJ, Lefkowitz EJ, Mushegian AR, Dutilh BE, Harrach B, et al. Binomial nomenclature for virus species: a consultation. *Arch Virol*. 2020 Feb; 165 (2):519–25. <https://doi.org/10.1007/s00705-019-04477-6> PMID: 31797129
31. Adams MJ, Lefkowitz EJ, King AMQ, Harrach B, Harrison RL, Knowles NJ, et al. 50 years of the International Committee on Taxonomy of Viruses: progress and prospects. *Arch Virol*. 2017; 162 (5):1441–6. <https://doi.org/10.1007/s00705-016-3215-y> PMID: 28078475
32. Conix S. In defence of taxonomic governance. *Organisms Diversity & Evolution*. 2019; 19 (2):87–97.
33. Jiang P, Faase JAJ, Toyoda H, Paul A, Wimmer E, Gorbalenya AE. Evidence for emergence of diverse polioviruses from C-cluster coxsackie A viruses and implications for global poliovirus eradication. *Proc Natl Acad Sci U S A*. 2007 May 29; 104 (22):9457–62. <https://doi.org/10.1073/pnas.0700451104> PMID: 17517601
34. Combelas N, Holmblat B, Joffret M-L, Colbère-Garapin F, Delpyroux F. Recombination between poliovirus and coxsackie A viruses of species C: a model of viral genetic plasticity and emergence. *Viruses*. 2011 Aug; 3 (8):1460–84.
35. Lauber C, Gorbalenya AE. Partitioning the genetic diversity of a virus family: approach and evaluation through a case study of picornaviruses. *J Virol*. 2012 Apr; 86 (7):3890–904. <https://doi.org/10.1128/JVI.07173-11> PMID: 22278230
36. Woolhouse MEJ, Howey R, Gaunt E, Reilly L, Chase-Topping M, Savill N. Temporal trends in the discovery of human viruses. *Proc Biol Sci*. 2008 Sep 22; 275 (1647):2111–5. <https://doi.org/10.1098/rspb.2008.0294> PMID: 18505720

37. Vlok M, Lang AS, Suttle CA. Application of a sequence-based taxonomic classification method to uncultured and unclassified marine single-stranded RNA viruses in the order Picornavirales. *Virus Evol.* 2019; 5 (2):vez056–6. <https://doi.org/10.1093/ve/vez056> PMID: 31908848
38. Virgin HW. The virome in mammalian physiology and disease. *Cell.* 2014 Mar 27; 157 (1):142–50. <https://doi.org/10.1016/j.cell.2014.02.032> PMID: 24679532
39. Gregory AC, Zayed AA, Conceição-Neto N, Temperton B, Bolduc B, Alberti A, et al. Marine DNA Viral Macro- and Microdiversity from Pole to Pole. *Cell.* 2019; 177 (5):1109, e14–23. <https://doi.org/10.1016/j.cell.2019.03.040> PMID: 31031001
40. Cobián Güemes AG, Youle M, Cantú VA, Felts B, Nulton J, Rohwer F. Viruses as Winners in the Game of Life. *Annu Rev Virol.* 2016; 3(1):197–214. <https://doi.org/10.1146/annurev-virology-100114-054952> PMID: 27741409
41. Garnett ST, Christidis L, Conix S, Costello MJ, Zachos FE, Bánki OS, et al. Principles for creating a single authoritative list of the world's species. *PLoS Biol.* 2020 Jul; 18 (7):e3000736. <https://doi.org/10.1371/journal.pbio.3000736> PMID: 32634138