



# Are There $10^{31}$ Virus Particles on Earth, or More, or Fewer?

 A. R. Mushegian<sup>a,b</sup>

<sup>a</sup>Molecular and Cellular Biology Division, National Science Foundation, Alexandria, Virginia, USA

<sup>b</sup>Clare Hall College, University of Cambridge, Cambridge, United Kingdom

**ABSTRACT** The number of virus particles on Earth is frequently reported in the scientific literature and in general-interest publications as being on the order of  $10^{31}$ , with some confusion about whether this is a high or low estimate. This number is often given without a source, although it should be attributed to a paper by Hendrix et al. published in 1999 (R. W. Hendrix, M. C. Smith, R. N. Burns, M. E. Ford, and G. F. Hatfull, *Proc Natl Acad Sci U S A* 96:2192–2197, 1999, <https://doi.org/10.1073/pnas.96.5.2192>). As with any oft-repeated statistic, it is informative to know how it has been derived and whether it should be revised in the light of new evidence. I review the history of the  $10^{31}$  estimate and use more recent assessments of the number of bacterial and viral particles in various habitats to conclude that the best estimate of the number of virus particles on Earth (“the Hendrix product”) remains close to  $10^{31}$  and is unlikely to be either much less or much more than that.

**KEYWORDS** bionumbers, global virome, virion numbers

For the first hundred years after the discovery of viruses at the end of the 19th century, virology was strongly motivated by the concerns of human health and disease, as well as the maladies of plants and animals of economic importance. Early on, it was also recognized that viruses, with their diversity of genome structures and expression strategies, as well as relative experimental tractability, were indispensable model systems in molecular studies, reliably illuminating basic molecular mechanisms, many of which proved relevant to the understanding of the workings of their host cells (2–4). Studies of many complete genome sequences of DNA and RNA viruses, at a time when the sequences of larger cellular genomes were far from completion, have played an essential, overlooked role in developing general methodologies for comparative genomics (5).

In this century, another side of viruses and virology has come to light: as metagenomic studies of microbes and viruses in various habitats have illuminated the enormous numbers of microorganisms on Earth, it has become clear that viruses play an important role in the biosphere, by actively participating, if not outright controlling, the rates of gene flow, biomass accretion, and elemental turnover in marine and terrestrial ecosystems (6, 7). With the background of the ongoing quest for quantitative approaches in biology and with the desire to improve numeracy in biology by obtaining the estimates of the key numbers that describe biological structure, function, evolution, and diversity (8), virologists and science journalists frequently present the number  $10^{31}$  as a key estimate of the abundance of virus particles on our planet. Confusingly,  $10^{31}$  is sometimes given as the maximal, and other times as the minimal, number of virus particles on Earth. For example, in two recent arbitrarily selected articles, we read (the emphasis in all quotations is mine), “It has been estimated by direct counts of extracellular (“free-floating”) virus-like particles (VLPs) that the global “virosphere” may contain up to  $10^{31}$  viral particles (2),” (9) and “At least  $10^{31}$  virus particles exist globally at any given time in most environments, including marine and

**Citation** Mushegian AR. 2020. Are there  $10^{31}$  virus particles on Earth, or more, or fewer? *J Bacteriol* 202:e00052-20. <https://doi.org/10.1128/JB.00052-20>.

**Editor** William Margolin, McGovern Medical School

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to mushegian2@gmail.com.

**Accepted manuscript posted online** 18 February 2020

**Published** 9 April 2020

freshwater habitats and metazoan gastrointestinal tracts, in which the number of detectable virus particles exceeds the number of cells by 10–100-fold (1–5)" (10).

In the first of these articles, their reference 10 is reference 7 here. In the second article, three of the five sources give no indication of the number of virions, instead dealing with the molecular diversity in metagenomic samples. The other two references are to a secondary source of information on virion numbers (11) and to a paper that does not provide the absolute virion counts but gives an important estimate of the virus-to-microbe ratio (VTM), which is more nuanced than the cited 10- to 100-fold value (12) (more on this below).

Thus, despite the wide acceptance of  $10^{31}$  virions as the key bionumber describing the virosphere, it is not clear whether this represents a lower or an upper bound. It is often given either without attribution or attributed to the secondary sources. However, the source and the derivation of the number are in fact clearly traceable, as I explain in this paper.

In a recent tribute to the life and work of Roger W. Hendrix (1943 to 2017), a preeminent molecular biologist who spent most of his career studying viruses of bacteria, Casjens and Hatfull wrote, "With his math prowess, Roger was quick to figure out that tailed phages are enormously abundant in nature, with an *estimated total abundance of  $10^{31}$* , making them the majority of all biological entities (38, 47)" (13). To the best of my knowledge, this attribution of the  $10^{31}$  estimate to Roger Hendrix (styled in the cited source as perhaps a mean) is correct. Examining the primary sources cited by S. R. Casjens and G. F. Hatfull, however, is illuminating. Reference 47 in the quotation above is to a 1999 paper coauthored by Roger Hendrix (1), where we read, "It has been estimated that there are  $4-6 \cdot 10^{30}$  prokaryotic cells in the biosphere (30), and direct counts on environmental samples typically show  $\approx 10$ -fold more tailed phage particles than cells (3)." The other reference (reference 38 in the Casjens and Hatfull quotation) cites Roger Hendrix' 2002 review (14), which says

There are typically  $10^7$  tailed phage particles per milliliter in coastal seawater and comparably high numbers in other sources examined (6); the global population is estimated to be *in excess of  $10^{30}$* . Because phage particles are typically present in *five- to 10-fold excess over bacterial cells*, it is plausible that tailed phages are not only the most abundant organisms on Earth but, in fact, constitute an absolute majority of organisms.

Thus, the earliest appearance of the number  $10^{31}$  (in effect  $>10^{31}$ , although perhaps not as much as  $10^{32}$ ) is indeed seen in two papers by Roger Hendrix, where it is implicit but can be obtained as the product of two key parameters, i.e., the number of prokaryotic cells in biosphere estimated by Whitman et al. (15) (reference 30 in the 1999 paper by Hendrix et al. [1]) and the virus-to-microbe ratio estimated by Bergh et al. (16) (reference 13 in the 1999 paper by Hendrix et al.). Both of these factors remain essential for determining the number of viruses on Earth, because bacterial and archaeal cells are more tractable and have been more extensively sampled than smaller-size virus particles, while the latter have been counted directly in fewer studies and at a smaller scale. Thus, a direct extrapolation of virus counts to the entire Earth is expected to be inferior to using the more numerous and more accurate data on bacterial and archaeal cell counts and multiplying them by the VTM value.

The 10-fold excess of phages over bacterial cells is given by Bergh et al. (16) without derivation; when looking at the data on VTM for each sample in their Table 1, one can see that the median measurement for all sites (omitting Raunefjorden in February, where no viruses were detected) is 9.9 and their geometric average is 12.6, probably explaining the estimated VTM value of 10. On the other hand, the estimation of the number of prokaryotic cells in Whitman et al. (15) was done explicitly by integrating many data points from the literature and recalculating them for the total estimated volume of each of the following habitats: aquatic (including marine shelf, open ocean, saline lakes, and freshwater lakes and rivers), soil, ocean subsurface, terrestrial sub-

face (mostly limited to the coastal plains and with a large uncertainty regarding the numbers of bacteria and archaea in the deep waters and aquifers), microbiomes of select animals and plants, and air. Only the first four of these numbers were large enough to influence the whole by an order of magnitude, and the counts for aquatic habitats were dominated by the open-ocean microbes. Taken together, the low bound of the number of prokaryotes on Earth was estimated as  $4 \times 10^{30}$ , as shown in the table. The Hendrix product of  $>10^{31}$  then is obtained easily.

A close reading of the mentions of the abundance of virus particles on Earth that were published in the last 2 decades suggests that all of them are variations on the  $10^{31}$  theme and that most of them either quote directly from Hendrix and coauthors (sometimes without attribution) or employ the same logic to multiply the same two numbers. In a relatively rare example of an independent assessment, C. A. Suttle has extrapolated the direct counts of viruses from various samples to the entire world ocean, arriving at an estimate of  $4 \times 10^{30}$  virus particles in oceanic waters alone, i.e., nearly four times as many as in the oceanic portion of the Hendrix product, and hypothesizing that the majority of the currently available direct virus counts are underestimates (7).

In the last 20 years, increasingly detailed sampling of the same habitats for bacteria, archaea, and viruses has taken place, and different kinds of mathematical models connecting microbe counts to VTM values have been fit to the data. Many of these reestimates have led to only modest correction of the values reported in the 1990s. At least two studies, however, reported the need for reevaluations of at least an order of magnitude. First, the number of bacterial and archaeal cells in oceanic seafloor sediment was reported to be only about  $0.3 \times 10^{30}$ , revising down the estimate from Whitman et al. by 10-fold (15); this is attributed to a combination of more extensive sampling, which, for the first time, included representative samples from microbe-poor as well as microbe-rich areas and more accurate accounting of the actual thicknesses of the sediment layers (17). Second, a survey of many sites where both virus and microbe abundances were estimated directly, again with the inclusion of microbe-rich as well as microbe-poor samples, suggested that even though a VTM of 10 has been observed in the surface waters of the ocean, microbial abundance is not a strong predictor of virus abundance in several other habitats, and that the general dependence of virus counts on prokaryotic cell counts is sublinear, i.e., once the concentration of microbes achieves certain levels, further increases may not result in equivalent increases in virus particle counts (12).

What do all these developments mean for the estimation of the Hendrix product? The most recent answer to this question is contained in the results of the census of the entire biomass of all living forms on our planet (18). In that triumph of estimation, the authors provided not only the biomass but also the numbers of individuals for many groups of living forms, including viruses. The authors have transparently provided links to large numbers of old and new sources of the data and to detailed calculations, including uncertainty estimates, for all numbers in the study. The main focus of the work, however, was on plants and animals, as well as on human contributions to the dynamics of planetary carbon. The update on virus counts and on its constituent parts is distributed across the main text and in supplementary material to that paper. Therefore, it is worthwhile to discuss the outcome of that recent thorough estimate and its effect on the Hendrix product.

Among several new elements in the work performed by Bar-On et al. (18), the following are particularly notable. First, several lines of evidence and "sanity checks" support the low estimate of the ocean subsurface microbes (17). Second, an explicit account of the microbes in the ocean crust is given for the first time. Third, a major reassessment of microbe abundance in terrestrial deep subsurface has been performed, with a 10-fold revision upwards. These developments have resulted in a new estimate of bacterial and archaeal abundance on Earth, which is about  $3 \times 10^{30}$ , i.e., 25% lower than that in reference 15 and the table therein. Fourth, the vast excess of the number of prokaryotic over eukaryotic cells on Earth is confirmed, putting on firmer ground the

**TABLE 1** Estimation of the number of prokaryotic cells and of the number of virus-like particles in various habitats on Earth

Sample source	No. of prokaryotes, based on <sup>a</sup> :		No. of viruses, based on Bar-On et al. <sup>a</sup>	VTM, based on Bar-On et al.	Habitat vol <sup>b</sup>	Concn of microbes <sup>c</sup>
	Whitman et al.	Bar-On et al.				
Ocean	0.12	0.12	2	17	10 <sup>3</sup>	1
Ocean subsurface	3.56	0.4	0.4	1	1	4 × 10 <sup>3</sup>
Ocean crust	Not assessed	0.2	6 <sup>d</sup> (2–20)	30 <sup>d</sup> (3–100)	10 <sup>3</sup>	~1
Soil	0.26	0.3	0.6	2	1	3.5 × 10 <sup>3</sup>
Terrestrial subsurface	0.25	2	2	1	15	1.5 × 10 <sup>3</sup>
Microbiomes	Insignificant	0.01 <sup>e</sup>	0.01	1		
Land water	Insignificant	Insignificant	Insignificant			
Air	Insignificant	Insignificant	Insignificant			
Total	4.19	3	11			

<sup>a</sup>The numbers are for cell or particle counts, divided by 10<sup>30</sup>. Whitman et al. is reference 15; Bar-On et al. is reference 18.

<sup>b</sup>Volumes are in cubic meters and divided by 10<sup>15</sup>. The volumes were estimated by taking the area of the ocean and area of land on Earth (3.6 × 10<sup>14</sup> and 1.5 × 10<sup>14</sup>, respectively) and multiplying by the effective thickness of each habitat from Bar-On et al. (18).

<sup>c</sup>Concentrations are cells or virions per cubic meter, divided by 10<sup>11</sup>. The values are obtained by dividing the virus counts by the habitat volume.

<sup>d</sup>This measure has a large uncertainty (18).

<sup>e</sup>The corrected number of bacteria plus archaea in the adult human body is 3.8 × 10<sup>13</sup> (22), which, for the entire Earth population of 7.5 × 10<sup>9</sup> humans (assuming adult sizes for all instead of the actual proportion of adults [70%], to obtain an upper bound), is 3 × 10<sup>23</sup> prokaryotes. Animal biomass is 300 times larger than human biomass; assuming that the animal microbiome size scales linearly with biomass (which is realistic given that the majority of the animal microbiome inhabits the digestive systems), there are at most 10<sup>26</sup> bacteria in the entire animal microbiome. Plant biomass (mostly trees) is 250 times greater than animal biomass; if plant microbiomes live mostly on plant surfaces, not inside largely metabolically inactive xylem, the plant microbiome size would scale as the power of 2/3 the mass. All told, just under 10<sup>28</sup> microbes may live in and on all macroorganisms on Earth. Considering the high density of microbes in these habitats, a low VTM applies, giving ~10<sup>28</sup> virus particles in the organismal virome on Earth.

previous idea that the Earth virome should be dominated by viruses of bacteria and archaea. Finally and importantly, instead of applying an average VTM value across the board, the authors have used a more involved model of virus-to-microbe ratio, along the lines proposed in reference 12. The result of the calculation is that the best estimate of the number of bacteria and archaea on Earth is about 3 × 10<sup>31</sup>, and the best estimate of the virus number is barely more than 10<sup>31</sup>. This argues against the earlier suggestion (7) that the number of virus particles on Earth has been systematically underestimated. The counts of prokaryotes in the air and those constituting the microbiomes of other living organisms were not explicitly considered by Bar-On et al. (18), but those are not expected to contribute much to the total; it is of some interest, however, that the counts of viruses living in and on all other living creatures may be only about 3 orders of magnitude lower than that in all other habitats combined (Table 1).

When all is said and done, what is the value of the Hendrix product? Are there 10<sup>31</sup> virus particles on Earth after all, or more, or less? The calculation by Bar-On et al. suggests that the answer is “more than 10<sup>31</sup>,” also taking into account the fact that the uncertainty of the estimate may be up to an order of magnitude (in part because of the large uncertainty on the bound-to-unbound particle ratio in the subterranean waters, which, when refined, may revise the estimate upwards). Unless a vast new reservoir of bacteria and archaea is discovered to add a large number of virus hosts or another major revision of the VTM values is provided, it is not obvious what could be a source of another order-of-magnitude revision.

It is worth remembering, however, that the direct measurements of virus abundances and VTM values report the number of virus-like particles; even if shown to contain DNA by staining with specific reagents, these particles are likely to include not only virions capable of productive infection but also defective viruses, capsids packaging nonvirus DNA, gene transfer agents, etc. These entities, important as they are in mediating gene flux, the virus-host arms race, and other functions in prokaryotes (19–21), are not viruses in the sense that is relevant to most virologists discussing the Hendrix product. Thus, if the real question of interest was how many conservatively defined, bona fide viruses exist, and depending on the prevalence of the nonproductive viruses in the general population, the number may have to be revised downwards, bringing it below, but barely below, 10<sup>31</sup> once again. The heterogeneity of the virus-like particle population, in which different subsets are likely to have vastly different stabilities, life cycles, and effects on various bacteria and archaea, is worth keeping in

mind when discussing the possible planetary effects of the world virus pangenome and virion biomass.

To summarize this very brief review of the history of a notion that, in my opinion, deserves to be called the Hendrix product or the Hendrix number, the current best estimate of the total number of virus-like particles on Earth remains close to  $10^{31}$ . When included in the scientific texts, it would be appropriate to credit this estimate to Hendrix et al. (1) or perhaps Hendrix et al. and Bar-On et al. (1, 18) to reflect its persistence in the light of the new data.

### ACKNOWLEDGMENTS

A.R.M. is a program director at the National Science Foundation (NSF), an agency of the U.S. Government; his work on this project and stay at Clare Hall College, University of Cambridge, is supported by the NSF Long-Term Professional Development Program, but the statements and opinions expressed here are made in a personal capacity and do not constitute an endorsement by the NSF or the government of the United States.

### REFERENCES

- Hendrix RW, Smith MC, Burns RN, Ford ME, Hatfull GF. 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc Natl Acad Sci U S A* 96:2192–2197. <https://doi.org/10.1073/pnas.96.5.2192>.
- Kornberg A. 2000. Ten commandments: lessons from the enzymology of DNA replication. *J Bacteriol* 182:3613–3618. <https://doi.org/10.1128/jb.182.13.3613-3618.2000>.
- Oldstone MB, Levine AJ. 2000. Virology in the next millennium. *Cell* 100:139–142. [https://doi.org/10.1016/s0092-8674\(00\)81690-6](https://doi.org/10.1016/s0092-8674(00)81690-6).
- Enquist LW, Editors of the Journal of Virology. 2009. Virology in the 21st century. *J Virol* 83:5296–5308. <https://doi.org/10.1128/JVI.00151-09>.
- Mushegian AR. 2007. Foundations of comparative genomics. Academic Press, Boston, MA.
- Wommack KE, Colwell RR. 2000. Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev* 64:69–114. <https://doi.org/10.1128/mnbr.64.1.69-114.2000>.
- Suttle CA. 2005. Viruses in the sea. *Nature* 437:356–361. <https://doi.org/10.1038/nature04160>.
- Milo R, Jorgensen P, Moran U, Weber G, Springer M. 2010. BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Res* 38:D750–D753. <https://doi.org/10.1093/nar/gkp889>.
- Zablocki O, Adriaenssens EM, Cowan D. 2016. Diversity and ecology of viruses in hyperarid desert soils. *Appl Environ Microbiol* 82:770–777. <https://doi.org/10.1128/AEM.02651-15>.
- Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, Carstens EB, Davison AJ, Delwart E, Gorbalenya AE, Harrach B, Hull R, King AMQ, Koonin EV, Krupovic M, Kuhn JH, Lefkowitz EJ, Nibert ML, Orton R, Roossinck MJ, Sabanadzovic S, Sullivan MB, Suttle CA, Tesh RB, van der Vlugt RA, Varsani A, Zerbini FM. 2017. Virus taxonomy in the age of metagenomics. *Nat Rev Microbiol* 15:161–168. <https://doi.org/10.1038/nrmicro.2016.177>.
- Mokili JL, Rohwer F, Dutilh BE. 2012. Metagenomics and future perspectives in virus discovery. *Curr Opin Virol* 2:63–77. <https://doi.org/10.1016/j.coviro.2011.12.004>.
- Wightington CH, Sonderegger D, Brussaard CPD, Buchan A, Finke JF, Fuhrman JA, Lennon JT, Middelboe M, Suttle CA, Stock C. 2016. Re-examination of the relationship between marine virus and microbial cell abundances. *Nat Microbiol* 1:15024. <https://doi.org/10.1038/nmicrobiol.2015.24>.
- Casjens SR, Hatfull GF. 2018. Roger Hendrix: gentle provocateur. *J Bacteriol* 200:e00058-18. <https://doi.org/10.1128/JB.00058-18>.
- Hendrix RW. 2002. Bacteriophages: evolution of the majority. *Theor Popul Biol* 61:471–480. <https://doi.org/10.1006/tpbi.2002.1590>.
- Whitman WB, Coleman DC, Wiebe WJ. 1998. Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A* 95:6578–6583. <https://doi.org/10.1073/pnas.95.12.6578>.
- Bergh O, Børsheim KY, Bratbak G, Haldal M. 1989. High abundance of viruses found in aquatic environments. *Nature* 340:467–468. <https://doi.org/10.1038/340467a0>.
- Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S. 2012. Global distribution of microbial abundance and biomass in subseafloor sediment. *Proc Natl Acad Sci U S A* 109:16213–16216. <https://doi.org/10.1073/pnas.1203849109>.
- Bar-On YM, Phillips R, Milo R. 2018. The biomass distribution on Earth. *Proc Natl Acad Sci U S A* 115:6506–6511. <https://doi.org/10.1073/pnas.1711842115>.
- Hynes AP, Villion M, Moineau S. 2014. Adaptation in bacterial CRISPR-Cas immunity can be driven by defective phages. *Nat Commun* 5:4399. <https://doi.org/10.1038/ncomms5399>.
- Grüll MP, Mulligan ME, Lang AS. 2018. Small extracellular particles with big potential for horizontal gene transfer: membrane vesicles and gene transfer agents. *FEMS Microbiol Lett* 365:fny192. <https://doi.org/10.1093/femsle/fny192>.
- Redfield RJ, Soucy SM. 2018. Evolution of bacterial gene transfer agents. *Front Microbiol* 9:2527. <https://doi.org/10.3389/fmicb.2018.02527>.
- Sender R, Fuchs S, Milo R. 2016. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol* 14:e1002533. <https://doi.org/10.1371/journal.pbio.1002533>.