Annual Review of Virology

The Case for Studying New Viruses of New Hosts

Mark D. Stenglein

Center for Vector-Borne Infectious Diseases, Department of Microbiology, Immunology, and Pathology, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, Colorado, USA; email: mark.stenglein@colostate.edu

Annu. Rev. Virol. 2022. 9:157–72

First published as a Review in Advance on June 7, 2022

The *Annual Review of Virology* is online at virology.annualreviews.org

https://doi.org/10.1146/annurev-virology-100220-112915

Keywords

viruses, virology, basic research, metagenomics, pandemic preparedness, model organisms

Abstract

Virology has largely focused on viruses that are pathogenic to humans or to the other species that we care most about. There is no doubt that this has been a worthwhile investment. But many transformative advances have been made through the in-depth study of relatively obscure viruses that do not appear on lists of prioritized pathogens. In this review, I highlight the benefits that can accrue from the study of viruses and hosts off the beaten track. I take stock of viral sequence diversity across host taxa as an estimate of the bias that exists in our understanding of host-virus interactions. I describe the gains that have been made through the metagenomic discovery of thousands of new viruses in previously unsampled hosts as well as the limitations of metagenomic surveys. I conclude by suggesting that the study of viruses that naturally infect existing and emerging model organisms represents an opportunity to push virology forward in useful and hard to predict ways.

INTRODUCTION

It feels somewhat absurd to be making the case for studying obscure viruses during a pandemic. But the point of this review is not that virology should not be spending time and money on pathogens that are clearly important to public health or economic security (1). Rather, it addresses the question of overall resource allocation and the return on investment from a diversified virology research portfolio (2, 3).

The idea that resources in virology might not be optimally allocated struck me when I published my first paper as a graduate student. The lab I had joined worked on proteins that blocked retroviruses, and we had found that they also blocked retrotransposons (4). Three other papers with more or less the same conclusion were published in quick succession before and after my own (5–7). There is value in confirmation, and competition can spur research, but wouldn't it have been better if our labs hadn't been quadruplicating efforts? Wouldn't the overall quality of science be better if virologists were working less redundantly on a more diverse array of questions, without rushing because of the looming fear of being scooped? In this review, I highlight the long-term payoffs from basic virology and take stock of the fertile fields of inquiry opened by unprecedented metagenomic virus discovery.

WHAT BENEFITS ACCRUE FROM BROAD AND IN-DEPTH STUDY OF VIRUSES?

The rationale for working on viruses that are pathogenic to humans or to species that we care about is clear. But why spend time and money on viruses that are not obviously pathogenic or that do not infect a species that we view as important? The rationale may not seem as obvious, but the fact is that enormous benefits have accrued from basic research on viruses and hosts from across the taxonomy, often when the value of the research was not immediately apparent.

Basic virology has produced significant advances in our understanding of fundamental biological processes. There is perhaps no better single example than that of Rous sarcoma virus, a tumor-causing chicken retrovirus named after Peyton Rous, who began working on this virus in the early twentieth century (8). Rous showed that a sarcoma could be transmitted to other chickens via injection of a filtered homogenate from the original tumor (9). Decades of work on Rous sarcoma virus and related viruses led to transformative advances in cancer biology and three Nobel prizes: for demonstrating that viruses can cause cancer, for the discovery of reverse transcriptase (10, 11), and for the existence of oncogenes (12).

Virology has also generated significant insights into cellular function and molecular biology. These include advances resulting from the study of bacteriophages during the early years of molecular biology: the determination that DNA is the genetic material (13) and the discovery that 3 bases make a codon (14), for instance (15). Similarly, studies of an array of animal viruses were instrumental to determining the capped, spliced, and polyadenylated structure of eukaryotic messenger RNA (mRNA) (16–19).

The catalogs of biotechnology companies are full of virus-derived enzymes that are the workhorses of modern molecular biology. The mRNA vaccines that have proven so effective at preventing severe coronavirus disease (COVID) are produced using viral enzymes: Bacteriophage RNA polymerases synthesize spike-encoding mRNAs, which are capped by vaccinia virus guanylyltransferase (20–23). Baculovirus-vectored protein production in insect cells is one of the best ways to generate large quantities of recombinant protein with appropriate eukaryotic cell modifications (24).

Other useful tools have resulted from wide-ranging studies of host immune responses. Single-domain antibodies (nanobodies), derived from camels and related species, are smaller than

traditional antibodies and have a number of resulting advantages as reagents and therapeutics (25, 26). Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas tools, components of bacterial immune systems, have enabled radical advances in the laboratory and the clinic (27, 28).

Our understanding of the evolution and function of the immune system has been improved through studies of a diverse range of viruses and hosts. B cells were shown to be responsible for antibody production through experiments in chickens (29, 30). (B cells are actually named after the bursa of Fabricius, an organ in birds that is functionally analogous to mammalian bone marrow.) The defensive role of Toll(-like) receptors in innate immunity was first identified in fruit flies (31, 32). Studies of the immune systems of phylogenetically far-flung organisms such as lampreys, placazoans, and choanoflagellates have proved critical to the understanding of the evolution of the adaptive and innate immune systems (33–35).

The extent to which we have been prepared for pandemics often reflects past work on relatively unknown viruses. When the human immunodeficiency virus (HIV) pandemic was first recognized in the 1980s, decades of research on other animal retroviruses such as Rous sarcoma virus provided a foundation of knowledge and tools (36). Prior work on other coronaviruses provided essential capabilities and understanding when the original severe acute respiratory syndrome coronavirus (SARS-CoV) emerged in 2003 (37). For example, infectious clone systems that had been developed for coronaviruses pathogenic to mice and swine provided a ready platform to study the new human pathogen (38–40). Similarly, prior work stabilizing the spike protein of common cold coronavirus HKU1 allowed the rapid structural determination of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike protein and enabled faster development of spike-based vaccines (41, 42).

Finally, there is what is sometimes called the One Health rationale. Despite our belief that we are special, humans are animals, and our self-regard does not prevent zoonotic virus infection. A frightening array of human pathogens have origins in other animals, and today's animal virus is tomorrow's human virus (43–46). Furthermore, all species are part of an interconnected web of life, and there are straightforward veterinary and conservation rationales for studying viruses of noneconomically important species.

The above examples highlight the benefits that accrue from basic virology research, and the advances I have described share a number of properties. First, this research was proactive. Investigators were in most cases not responding to "the viral concerns of today" (2, p. vii), but were engaged in basic research on questions they thought were interesting and important. Second, these advances resulted from in-depth focus on a problem over years or decades. Third, in many cases, the long-term benefits of the research were not always immediately obvious. A recent recounting of the decades of work that led to the extraordinary success of the SARS-CoV-2 mRNA vaccines noted the "happenstance" involved along the "unpredictable, zigzagging path" of basic research (47). Finally, a property these advances do not share is focus on a single type of virus or host. Instead, the diversity of systems is the conspicuous feature.

HOW (IM)BALANCED IS VIROLOGY RESEARCH?

So, if it is useful to study a broad diversity of viruses and hosts, to what extent is this being done? How would we know if we were devoting sufficient resources to basic virology research? Many authors have noted that virology tends to focus on a relatively low number of viruses from relatively few hosts (e.g., 2, 48–54). This seems correct, but it can be difficult to actually quantify the extent to which this is true. One way to try to do this would be to count virus sequences in public databases and to assess their distribution across viral and host taxa.

There are a number of obvious limitations to this strategy. It depends on sequence annotation, which can be incorrect or missing: In the National Center for Biotechnology Information (NCBI)

Nucleotide database, for instance, only 6.2×10^6 of the 6.8×10^6 viral sequences (91%) have an annotated host (55). It also depends on classification of viruses and hosts and the representation of this taxonomy in the NCBI Taxonomy database (56, 57). More importantly, it assesses virus sequences and not viruses themselves or knowledge about viruses (58). Indeed, virology existed for decades before the advent of sequencing (59). And most virology research neither requires nor produces new sequences.

Nevertheless, scientists have been sequencing viruses for more than 50 years (60), and generating virus sequences requires time and money. So, the number of sequences from a particular virus or host is directly proportional to resources expended by virologists. Sequences in public databases therefore provide a reasonable estimate of virology effort and interest.

I obtained metadata for the nearly 7 million virus sequences from the NCBI Virus portal (61) and tabulated the number of sequences per viral taxa and the number of virus sequences associated with individual hosts (**Tables 1–4**, **Figure 1**). An extended and reproducible version of this analysis is available (62).

Table 1 The 20 most-sequenced viruses[a]

Virus species	Representative virus or abbreviation	Thousand virus sequences	Percentage of all virus sequences	Cumulative percentage of virus sequences
Severe acute respiratory syndrome–related coronavirus	SARS-CoV-2	3,036.1	49.2	49.2
Human immunodeficiency virus 1	HIV-1	985.5	16.0	65.2
Influenza A virus	Influenza A virus	725.9	11.8	77.0
Hepacivirus C	Hepatitis C virus	155.0	2.5	79.5
Influenza B virus	Influenza B virus	112.0	1.8	81.3
Hepatitis B virus	Hepatitis B virus	89.0	1.4	82.7
Rotavirus A	Rotavirus A	85.0	1.4	84.1
Simian immunodeficiency virus	SIV	49.4	0.8	84.9
Norwalk virus	Norovirus	37.3	0.6	85.5
Enterovirus A	Human enterovirus A	32.0	0.5	86.0
Dengue virus	Dengue virus 1	31.6	0.5	86.5
Porcine reproductive and respiratory syndrome virus	PRRSV	27.6	0.4	87.0
Human orthopneumovirus	Respiratory syncytial virus	25.6	0.4	87.4
Rabies lyssavirus	Rabies virus	22.4	0.4	87.8
Pestivirus A	Bovine viral diarrhea virus 1	18.5	0.3	88.1
Enterovirus B	Coxsackievirus B3	18.4	0.3	88.4
Measles morbillivirus	Measles virus	16.9	0.3	88.6
Orthohepevirus A	Hepatitis E virus	16.4	0.3	88.9
Simian-human immunodeficiency virus	SHIV	16.3	0.3	89.2
Human immunodeficiency virus[b]	HIV	15.4	0.2	89.4

[a]As represented in the National Center for Biotechnology Information (NCBI) Nucleotide database. Values based on metadata downloaded from the NCBI Virus portal on January 4, 2022.

[b]These sequences have been classified as "Human immunodeficiency virus." The NCBI Taxonomy database notes that "[a]ll of them are probably 'Human immunodeficiency virus type 1' [sequences]."

Table 2 The hosts with the most (virus sequences)

Host taxon	Host common name	Thousand virus sequences	Percentage of all virus sequences	Cumulative percentage of virus sequences
Homo sapiens	Human	5,224.4	84.7	84.7
Sus scrofa	Pig	178.9	2.9	87.6
Gallus gallus	Chicken	93.3	1.5	89.1
Anas platyrhynchos	Mallard	50.2	0.8	89.9
Bos taurus	Cattle	45.9	0.7	90.7
Macaca mulatta	Rhesus monkey	34.9	0.6	91.2
Anatidae	Waterfowl	31.7	0.5	91.7
Simiiformes	Simians	20.9	0.3	92.1
Canis lupus	Dog and wolf	20.2	0.3	92.4
Equus caballus	Horse	17.7	0.3	92.7
Arenaria interpres	Ruddy turnstone	13.6	0.2	92.9
Ovis aries	Sheep	10.1	0.2	93.1
Aves	Birds	9.3	0.2	93.2
Vitis vinifera	Wine grape	8.9	0.1	93.4
Solanum lycopersicum	Tomato	8.6	0.1	93.5
Spatula discors	Blue-winged teal	8.4	0.1	93.6
Felis catus	Domestic cat	8.0	0.1	93.8
Anas acuta	Northern pintail	6.5	0.1	93.9
Meleagris gallopavo	Turkey	6.0	0.1	94.0
Macaca nemestrina	Pig-tailed macaque	5.5	0.1	94.1

As might be expected, virus sequences are strongly biased toward certain types of viruses (**Tables 1–4**, **Figure 1**). Pandemic viruses are by far the most sequenced. The SARS-CoV-2 pandemic has produced an unprecedented flood of genome sequencing, and the greater than 3×10^6 SARS-CoV-2 sequences in the NCBI Nucleotide database account for nearly half of all virus sequences (49.7%). By the time this review is published, there is no doubt SARS-CoV-2 sequences will be the majority of virus sequences in NCBI. In fact, there are already more SARS-CoV-2 sequences ($>7 \times 10^6$) in the Global Initiative on Sharing Avian Influenza Data (GISAID) database than there are virus sequences in NCBI (63). Human and simian immunodeficiency viruses and influenza viruses are the next most-sequenced types of viruses. Together, sequences from sarbecoviruses (SARS-CoV-2 and related viruses), lentiviruses (HIV and related viruses), and influenza viruses account for 80.4% of all virus sequences (**Figure 1**).

The effect of this can be seen in a plot of the cumulative distribution of virus sequences plotted against individual virus species ordered by their number of virus sequences (**Figure 1*a***). This curve goes almost straight up, with nearly all virus sequences accounted for by the most sequenced viruses. The 20 most-sequenced viruses account for 89.4% of sequences (**Table 1**).

Virus sequences are even more biased toward particular hosts, with the top 20 hosts accounting for 94.1% of all virus sequences (**Table 2**). The 5.2×10^6 sequences from humans account for 84.7% of all virus sequences. Other hosts with the most virus sequences include agriculturally important animals and plants (e.g., pigs, chickens, grapes, and tomatoes), nonhuman hosts of lentiviruses and influenza viruses (simians, pigs, and waterfowl), and companion animals (dogs, cats, and horses). Viruses from pigs are highly sequenced because pigs are both agriculturally important and host to influenza viruses.

It could be argued that the over-representation of pandemic virus sequences confounds this analysis. But even if sarbecoviruses, lentiviruses, and influenza viruses are removed from

Table 3 The most-sequenced viruses, excluding sarbecoviruses, influenza viruses, and lentiviruses

Virus species	Representative virus or abbreviation	Thousand remaining virus sequences	Percentage of remaining virus sequences	Cumulative percentage of remaining virus sequences
Hepacivirus C	Hepatitis C virus	155.0	12.8	12.8
Hepatitis B virus	Hepatitis B virus	89.0	7.4	20.2
Rotavirus A	Rotavirus A	85.0	7.0	27.3
Norwalk virus	Norovirus	37.3	3.1	30.4
Enterovirus A	Human enterovirus A	32.0	2.7	33.0
Dengue virus	Dengue virus 1	31.6	2.6	35.6
Porcine reproductive and respiratory syndrome virus	PRRSV	27.6	2.3	37.9
Human orthopneumovirus	Respiratory syncytial virus	25.6	2.1	40.0
Rabies lyssavirus	Rabies virus	22.4	1.9	41.9
Pestivirus A	Bovine viral diarrhea virus 1	18.5	1.5	43.4
Enterovirus B	Coxsackievirus B3	18.4	1.5	44.9
Measles morbillivirus	Measles virus	16.9	1.4	46.3
Orthohepevirus A	Hepatitis E virus	16.4	1.4	47.7
Alphapapillomavirus 9	Human papillomavirus 16	15.3	1.3	49.0
Mumps orthorubulavirus	Mumps virus	10.8	0.9	49.9
Foot-and-mouth disease virus	Foot-and-mouth disease virus	10.2	0.8	50.7
Avian coronavirus	Infectious bronchitis virus	9.7	0.8	51.5
Avian orthoavulavirus 1	Newcastle disease virus	8.0	0.7	52.2
Enterovirus C	Poliovirus	7.4	0.6	52.8
Porcine circovirus 2	Porcine circovirus 2	7.0	0.6	53.4

analysis, the list of most-sequenced viruses remains dominated by human and livestock pathogens (**Table 3**). Even without the contribution of the highly sequenced pandemic viruses, humans remain the host with the most associated virus sequences, accounting for 56.8% of remaining virus sequences (**Table 4**).

This paints a picture of a focused field. But there is another message from this analysis. Although it is true that the cumulative distribution curves have steep initial slopes, accounted for by the most-sequenced viruses, they also have extremely long flat tails (**Figure 1*a*,*c***). In other words, at the same time that a few viruses account for most sequences, there are sequences from thousands of different viruses and hosts (35,781 viral taxa and 8,322 hosts).

In fact, most viral taxa and many hosts are represented by only a single virus sequence (**Figure 1*b*,*d***). Although the mean number of viruses per virus species is 172 and the mean number of viruses per host is 741, the median values are only 1 sequence per virus and 4 per host. The majority of viral taxa (20,712, 58%) are represented by a single sequence. Randomly selected examples of such viruses include Wenling bighead beaked sandfish astrovirus (50), moosepox virus GoldyGopher14 (64), and *Streptococcus* phage Javan520 (65). Likewise, 2,068 hosts (24.8%) are associated with only 1 virus sequence. Randomly selected hosts from among this list include *Falsistrellus mackenziei*, an Australian bat (66), and *Gymnosporia buxifolia*, an African plant known as a spike-thorn (67).

Table 4 The hosts with the most virus sequences, excluding sarbecoviruses, influenza viruses, and lentiviruses

Host taxon	Host common name	Thousand remaining virus sequences	Percentage of remaining virus sequences	Cumulative percentage of remaining virus sequences
Homo sapiens	Human	685.9	56.8	56.8
Sus scrofa	Pig	77.3	6.4	63.2
Bos taurus	Cattle	44.7	3.7	66.9
Gallus gallus	Chicken	26.5	2.2	69.1
Canis lupus	Dog and wolf	16.7	1.4	70.5
Vitis vinifera	Wine grape	8.9	0.7	71.3
Solanum lycopersicum	Tomato	8.6	0.7	72.0
Equus caballus	Horse	7.0	0.6	72.6
Ovis aries	Sheep	6.3	0.5	73.1
Felis catus	Domestic cat	6.2	0.5	73.6
Apis mellifera	Honey bee	4.9	0.4	74.0
Solanum tuberosum	Potato	4.6	0.4	74.4
Sus	Wild and domestic pigs	4.6	0.4	74.8
Salmo salar	Atlantic salmon	4.1	0.3	75.1
Prunus persica	Peach	4.1	0.3	75.5
Capra hircus	Goat	3.5	0.3	75.7
Macaca mulatta	Rhesus monkey	3.4	0.3	76.0
Manihot esculenta	Cassava	3.1	0.3	76.3
Oryctolagus cuniculus	European rabbit	3.1	0.3	76.5
Capsicum annuum	Pepper plant	2.9	0.2	77.0

Clearly the description of viral diversity is nowhere close to being exhausted (52). Not only are most virus species and most hosts represented by one or a few virus sequences, but also the vast majority of species have no associated virus sequences. There are records for 626,899 species of cellular life in the NCBI Taxonomy database. This captures only a small fraction—perhaps 10%—of actual species diversity, meaning that virus sequences have been described for only something like 0.1% of hosts (8,322/626,899/10) (56).

So, yes, there has been a strong bias toward a few particularly important viruses. At the same time, there has been an enormous amount of study of many viruses beyond the usual suspects (as reflected in the basic virology advances described in the previous section) and a recent massive expansion in the number of sequenced viruses. In the next section, I describe the impact of metagenomics on our understanding of viral diversity and evolution, and discuss the limits of metagenomics.

METAGENOMICS IS TRANSFORMING OUR UNDERSTANDING OF THE VIROSPHERE, BUT IS IT ENOUGH?

Metagenomic virus discovery—the identification of virus sequences in shotgun sequencing data sets—is the main reason that there is such a long tail in the cumulative distribution plots of sequences per virus or host (**Figure 1**). Prior to metagenomics, virology focused for the most part on the subset of viruses for which infection produced an obvious phenotype. This was by necessity. There was no way to know that a viral infection was present if it did not produce a measurable phenotype. Often the phenotype was disease or cytopathic effect in cultured cells, but not always. For instance, sigma viruses of *Drosophila* were noticed and studied because they caused flies to

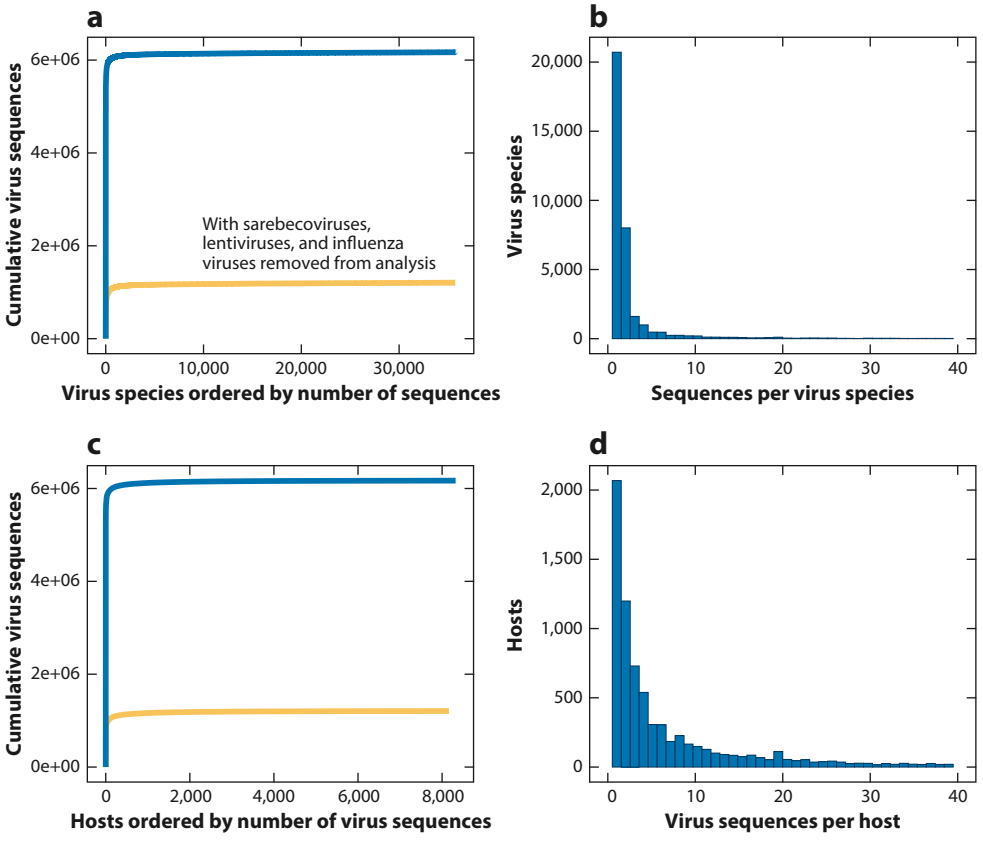

Figure 1

A very small number of viruses and hosts account for most virus sequences. At the same time, most viruses and hosts are represented by only a single virus sequence. (*a*) The cumulative number of virus sequences associated with individual virus species, which are ordered by their number of sequences. The blue line is all virus sequences in the National Center for Biotechnology Information Nucleotide database. The yellow line is with sequences from the taxa *Sarbecovirus*, *Lentivirus*, *Alphainfluenzavirus*, *Betainfluenzavirus*, *Gammainfluenzavirus*, and *Deltainfluenzavirus* removed. (*b*) A histogram of the number of sequences per virus species. The *x* axis is cut off at 40 to emphasize the point that most viral taxa and most hosts are represented by only one or a few virus sequences. Data in panels *c* and *d* are as in panels *a* and *b* but depicting virus sequences per host.

fail to recover from carbon dioxide anesthetization, which L'Heritier and Teissier (68) noted as a "physiological anomaly." Cryptic viruses of plants and fungi provide another example. As their name suggested (former genus *Cryptovirus*; current classification *Partitiviridae*), these viruses did not produce apparent phenotypes (69, 70). Instead, they were discovered because they produced abundant virus particles evident in electron micrographs of samples from apparently healthy plants (71).

In contrast, metagenomic virus discovery does not require that you be able to culture a virus or see the effect of infection. Instead, metagenomics uses random sequencing of nucleic acid to identify virus sequences among the sequences from hosts and other organisms in a sample. This methodological advance has unleashed a golden age of virus discovery and revealed a previously unimaginable diversity and abundance of viruses. Many excellent reviews have been written on

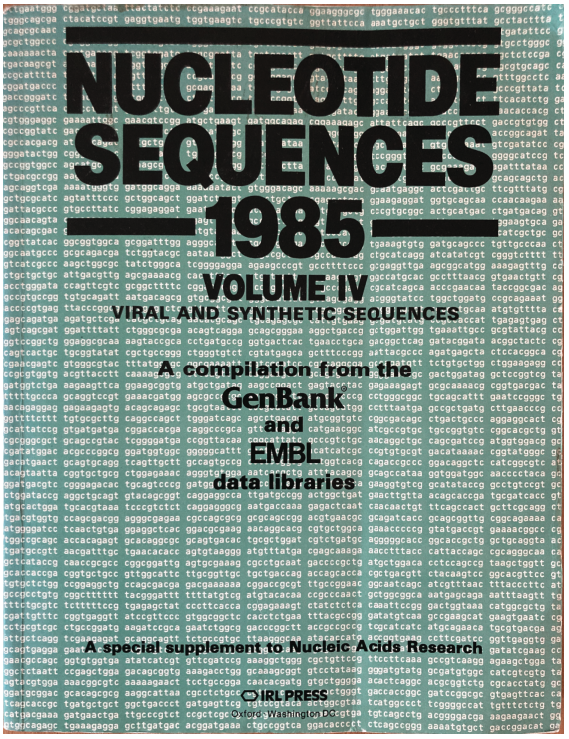

Figure 2

All the virus sequences used to fit in a book. The author's copy of the 1985 *Nucleic Acids Research* special supplement contained all of the then-existing virus sequences in GenBank (89). Now, the 9.9×10^{10} bases of virus sequence would need a book 33 million pages long (assuming 3,000 bases per page, an estimate from the pictured volume).

metagenomic viral discovery, and this review does not go in depth into the history or technical details of this field (e.g., 52, 54, 72–75). Early metagenomic surveys used cloning and Sanger sequencing or microarrays to identify new viruses (76–78). Now, metagenomic virus discovery uses next-generation sequencing to identify virus sequences. In the early days of sequencing, determination of even part of a viral genome sequence could be sufficient for a high-profile paper (79–81) (**Figure 2**). Nowadays, leading papers describe hundreds or thousands of new viral genome sequences in one go (e.g., 48–50, 82–85). Sequencing data sets generated for other purposes (e.g., gene expression analysis) often contain virus sequences, and existing public data sets remain a rich and largely untapped resource for virus discovery (84, 86–88).

All of these new virus sequences have produced significant advances in the understanding of viral diversity and evolution (52, 54, 72, 90). New viruses have led to the establishment of entirely new families and orders of viruses and indeed a wholesale reorganization of virus taxonomy (57, 75). The known host range of most groups of viruses has been substantially expanded. It is now clear that the genome structure of viruses is much more plastic than previously thought and that horizontal gene transfer between viruses and between viruses and hosts is common, at least on evolutionary timescales (49).

Metagenomic sequencing has also produced candidate etiological agents for infectious diseases that had proven intractable to traditional diagnostic methods. This has been a boon for infectious diseases of hosts such as reptiles that lack the economic importance of, say, livestock (91–100). In

the case of a new high-priority pathogen such as SARS-CoV-2, metagenomic sequencing can be used to determine a genome sequence within hours or days of sample collection (101).

Despite its extraordinary power, metagenomics has limitations (58). The fundamental issue is that sequencing only produces sequences, and there is a limit to how much you can infer about the biology of a virus from its sequence. For instance, it is often not clear from metagenomic data what the host of a virus (sequence) is. Even when a sample derives from a single host, virus sequences might actually derive from the putative host's diet or its microbiota, or even from reagent contamination (102). As one example, Twyford virus, originally identified in association with *Drosophila melanogaster*, proved to have a fungal host (a possibility the discoverers were careful to note in the original paper) (53, 103). There is even more uncertainty about the host when metagenomics is based on environmental samples, complex samples such as feces, or pools of samples from different organisms, which have been the basis for many of the most fruitful studies (49, 104).

Even when it is possible to link a virus sequence to a host (105), metagenomics reveals little about the effect of infection on the host. The sequence of a virus does not tell you how pathogenic a virus is or even if it is pathogenic at all (51). Sequences alone do not reveal anything about tissue tropism, mechanisms of transmission, or interactions with the host's immune system [sequencing of small RNAs in organisms with antiviral RNA interference provides one exception (53)]. Candidate etiologic agents from metagenomics must be validated by other means such as a strong association with disease or experimental infection studies (106).

This has led to something of a paradox. At the same time that so many new viruses are being discovered, so little is known about most of them beyond their phylogenetic placement and primary sequence. As an example, large-scale sequencing of pools of invertebrates revealed viruses so divergent that they ended up founding a completely new order of RNA viruses, the *Jingchuvirales*, a cousin taxon to mononegaviruses, orthomyxoviruses, and bunyaviruses (48, 57, 107). Since their initial discovery, dozens of papers using metagenomics have revealed that jingchuviruses are widespread in a diverse array of hosts (e.g., 50, 108–112). But still vanishingly little is known about their basic biology (113, 114). This is like discovering an entire new continent and having explored only a tiny fraction of its interior, and leads to questions about the proper balance between discovery of new viruses and experiments to actually study them.

THE WAY FORWARD?

So where to go from here? It is safe to assume that there is value in studying some of these new viruses, but which ones? Metagenomics has created an embarrassment of riches: far more possible targets than there are time and money to study. One possible way forward would be to recognize that scientific progress is unpredictable and to not even try to pick winners. Another perhaps more practical approach would be to focus on natural viruses of existing and emerging model organisms. This strategy would narrow down the long list of candidates to viruses that infect hosts that at least have existing genomic resources and experimental capabilities.

Traditional model organisms such as mice and fruit flies have been a mainstay of virus research for decades (115–117). More recently, there has been a renewed interest in studying viruses that naturally infect these organisms. This is in part because of work highlighting the importance of natural microbiota to immune system development (118). Metagenomic surveys of wild-caught individuals from these species have also contributed (53, 119–122). Natural viruses of other long-studied model species, such as *Caenorhabditis elegans* (nematodes) and *Danio rerio* (zebrafish), have also only recently been identified (123–127). The powerful existing resources associated with these organisms mean that mechanistic studies of host-virus interactions are relatively straightforward to design and implement (127–131).

Established models are great, but they represent only a small fraction of overall host diversity. Recent advances in genomic and gene-editing technologies have increased the feasibility of working with nontraditional model organisms (132, 133). In response, biologists have begun working increasingly with phylogenetically disparate species to answer questions about a range of biological processes such as cellular regeneration, the origins of multicellularity, and survival in extreme conditions (134–136). Viruses of these emerging models represent another promising avenue of research. Collaboration between virologists and scientists who have been using these organisms to answer other questions (and, importantly, who have been learning how to work with them) have great potential (35). In tandem, development of even more new models and targeted metagenomic surveys will continue to expand the list of candidates.

It is an exciting time to be a virologist. Advances in metagenomics are revolutionizing our understanding of virus evolution and diversity. At the same time, we have just scratched the surface of viral diversity, and new discoveries provide fertile ground for exploration. New techniques offer previously unimaginable opportunities to dissect host-virus interaction, even in species that were previously considered too hard to work with. The terrible cost of the SARS-CoV-2 pandemic has reinforced the importance of virology in the public eye. Many virologists pivoted to contribute to the pandemic response, the effectiveness of which was thanks in large part to decades of prior basic research. When it comes time to pivot back, virologists and funders would do well to remember the value of basic virology and invest in the promise of discovery.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Thanks to Alex Greninger and Medora Huseby for helpful discussions. Research in the Stenglein lab is supported in part by National Science Foundation IOS 2048214.

LITERATURE CITED

1. Imperiale MJ, Casadevall A. 2015. The importance of virology at a time of great need and great jeopardy. *mBio* 6(2):e00236-15
2. Goodin MM, Hatfull GF, Malik HS. 2016. A diversified portfolio. *Annu. Rev. Virol.* 3:vi–viii
3. Thompson JR. 2021. Fundamental virology: same objectives, changing tools. *Front. Virol.* 1:6
4. Stenglein MD, Harris RS. 2006. APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. *J. Biol. Chem.* 281(25):16837–41
5. Chen H, Lilley CE, Yu Q, Lee DV, Chou J, et al. 2006. APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr. Biol.* 16(5):480–85
6. Bogerd HP, Wiegand HL, Hulme AE, Garcia-Perez JL, O'Shea KS, et al. 2006. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *PNAS* 103(23):8780–85
7. Muckenfuss H, Hamdorf M, Held U, Perković M, Löwer J, et al. 2006. APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.* 281(31):22161–72
8. Weiss RA, Vogt PK. 2011. 100 years of Rous sarcoma virus. *J. Exp. Med.* 208(12):2351–55
9. Rous P. 1911. A sarcoma of the fowl transmissible by an agent separable from the tumor cells. *J. Exp. Med.* 13(4):397–411
10. Baltimore D. 1970. Viral RNA-dependent DNA polymerase: RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226(5252):1209–11
11. Temin HM, Mizutani S. 1970. Viral RNA-dependent DNA polymerase: RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226(5252):1211–13

12. Stehelin D, Varmus HE, Bishop JM, Vogt PK. 1976. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* 260(5547):170–73
13. Hershey AD, Chase M. 1952. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.* 36(1):39–56
14. Crick FH, Barnett L, Brenner S, Watts-Tobin RJ. 1961. General nature of the genetic code for proteins. *Nature* 192:1227–32
15. Ofir G, Sorek R. 2018. Contemporary phage biology: from classic models to new insights. *Cell* 172(6):1260–70
16. Furuichi Y, Morgan M, Muthukrishnan S, Shatkin AJ. 1975. Reovirus messenger RNA contains a methylated, blocked 5′-terminal structure: $m^7G(5')ppp(5')G^mpCp$-. *PNAS* 72(1):362–66
17. Nevins JR, Wilson MC. 1981. Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing. *Nature* 290(5802):113–18
18. Fitzgerald M, Shenk T. 1981. The sequence 5′-AAUAAA-3′ forms parts of the recognition site for polyadenylation of late SV40 mRNAs. *Cell* 24(1):251–60
19. Berget SM, Moore C, Sharp PA. 1977. Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. *PNAS* 74(8):3171–75
20. Pardi N, Hogan MJ, Porter FW, Weissman D. 2018. mRNA vaccines—a new era in vaccinology. *Nat. Rev. Drug Discov.* 17(4):261–79
21. Martin SA, Paoletti E, Moss B. 1975. Purification of mRNA guanylyltransferase and mRNA (guanine-7-)methyltransferase from vaccinia virions. *J. Biol. Chem.* 250(24):9322–29
22. Davanloo P, Rosenberg AH, Dunn JJ, Studier FW. 1984. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *PNAS* 81(7):2035–39
23. Melton DA, Krieg PA, Rebagliati MR, Maniatis T, Zinn K, Green MR. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 12(18):7035–56
24. Kost TA, Condreay JP, Jarvis DL. 2005. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat. Biotechnol.* 23(5):567–75
25. Muyldermans S. 2013. Nanobodies: natural single-domain antibodies. *Annu. Rev. Biochem.* 82:775–97
26. Muyldermans S. 2021. Applications of nanobodies. *Annu. Rev. Anim. Biosci.* 9:401–21
27. Doudna JA, Charpentier E. 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346(6213):1258096
28. Wright AV, Nuñez JK, Doudna JA. 2016. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* 164(1–2):29–44
29. Glick B, Chang TS, Jaap RG. 1956. The bursa of Fabricius and antibody production. *Poult. Sci.* 35(1):224–25
30. Cooper MD, Peterson RDA, Good RA. 1965. Delineation of the thymic and bursal lymphoid systems in the chicken. *Nature* 205(4967):143–46
31. Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. 1996. The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in Drosophila adults. *Cell* 86(6):973–83
32. Zambon RA, Nandakumar M, Vakharia VN, Wu LP. 2005. The Toll pathway is important for an antiviral response in *Drosophila*. *PNAS* 102(20):7257–62
33. Hirano M, Das S, Guo P, Cooper MD. 2011. The evolution of adaptive immunity in vertebrates. *Adv. Immunol.* 109:125–57
34. Kamm K, Schierwater B, DeSalle R. 2019. Innate immunity in the simplest animals—placozoans. *BMC Genom.* 20:5
35. Woznica A, Kumar A, Sturge CR, Xing C, King N, Pfeiffer JK. 2021. STING mediates immune responses in the closest living relatives of animals. *eLife* 10:e70436
36. Fauci AS, Lane HC. 2020. Four decades of HIV/AIDS—much accomplished, much to do. *N. Engl. J. Med.* 383(1):1–4
37. Weiss SR. 2020. Forty years with coronaviruses. *J. Exp. Med.* 217(5):e20200537
38. Almazán F, González JM, Pénzes Z, Izeta A, Calvo E, et al. 2000. Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *PNAS* 97(10):5516–21

39. Yount B, Curtis KM, Baric RS. 2000. Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. *J. Virol.* 74(22):10600–11
40. Yount B, Curtis KM, Fritz EA, Hensley LE, Jahrling PB, et al. 2003. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *PNAS* 100(22):12995–3000
41. Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, et al. 2016. Pre-fusion structure of a human coronavirus spike protein. *Nature* 531(7592):118–21
42. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, et al. 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 367(6483):1260–63
43. Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, et al. 2017. Pathways to zoonotic spillover. *Nat. Rev. Microbiol.* 15(8):502–10
44. Grange ZL, Goldstein T, Johnson CK, Anthony S, Gilardi K, et al. 2021. Ranking the risk of animal-to-human spillover for newly discovered viruses. *PNAS* 118(15):e2002324118
45. Woolhouse M, Gaunt E. 2007. Ecological origins of novel human pathogens. *Crit. Rev. Microbiol.* 33(4):231–42
46. Lednicky JA, Tagliamonte MS, White SK, Elbadry MA, Alam MM, et al. 2021. Independent infections of porcine deltacoronavirus among Haitian children. *Nature* 600(7887):133–37
47. Kolata G, Mueller B. 2022. Halting progress and happy accidents: how mRNA vaccines were made. *New York Times*, Jan. 15, p. A1
48. Li C-X, Shi M, Tian J-H, Lin X-D, Kang Y-J, et al. 2015. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife* 4:e05378
49. Shi M, Lin X-D, Tian J-H, Chen L-J, Chen X, et al. 2016. Redefining the invertebrate RNA virosphere. *Nature* 540(7634):539–43
50. Shi M, Lin X-D, Chen X, Tian J-H, Chen L-J, et al. 2018. The evolutionary history of vertebrate RNA viruses. *Nature* 556(7700):197–202
51. Roossinck MJ. 2015. Move over, bacteria! Viruses make their mark as mutualistic microbial symbionts. *J. Virol.* 89(13):6532–35
52. Greninger AL. 2018. A decade of RNA virus metagenomics is (not) enough. *Virus Res.* 244:218–29
53. Webster CL, Waldron FM, Robertson S, Crowson D, Ferrari G, et al. 2015. The discovery, distribution, and evolution of viruses associated with *Drosophila melanogaster*. *PLOS Biol.* 13(7):e1002210
54. Koonin EV, Dolja VV. 2018. Metaviromics: a tectonic shift in understanding virus evolution. *Virus Res.* 246:A1–3
55. Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, et al. 2021. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 49(D1):D10–17
56. Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, et al. 2020. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database* 2020:baaa062
57. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, et al. 2020. Global organization and proposed megataxonomy of the virus world. *Microbiol. Mol. Biol. Rev.* 84(2):e00061-19
58. Racaniello V. 2016. Moving beyond metagenomics to find the next pandemic virus. *PNAS* 113(11):2812–14
59. Oldstone MBA. 2014. History of virology. In *Encyclopedia Microbiology*, ed. TM Schmidt, pp. 608–12. Cambridge, MA: Academic
60. De Wachter R, Verhassel JP, Fiers W. 1968. The 5′-terminal tetranucleotide sequence of bacteriophage MS2 ribonucleic acid. *Arch. Int. Physiol. Biochim.* 76(3):580–81
61. Hatcher EL, Zhdanov SA, Bao Y, Blinkova O, Nawrocki EP, et al. 2017. Virus Variation Resource—improved response to emergent viral outbreaks. *Nucleic Acids Res.* 45(D1):D482–90
62. Stenglein MD. 2022. 2022_Annual_Review_of_Virology. *Code Repository*. **https://github.com/stenglein-lab/2022_Annual_Review_of_Virology**
63. Bogner P, Capua I, Lipman DJ, Cox NJ. 2006. A global initiative on sharing avian flu data. *Nature* 442(7106):981
64. Armién AG, Wolf TM, Mor SK, Ng TFF, Bracht AJ, et al. 2020. Molecular and biological characterization of a Cervidpoxvirus isolated from moose with necrotizing dermatitis. *Vet. Pathol.* 57(2):296–310

65. Rezaei Javan R, Ramos-Sevillano E, Akter A, Brown J, Brueggemann AB. 2019. Prophages and satellite prophages are widespread in *Streptococcus* and may play a role in pneumococcal pathogenesis. *Nat. Commun.* 10(1):4852
66. Prada D, Boyd V, Baker ML, O'Dea M, Jackson B. 2019. Viral diversity of microbats within the South West Botanical Province of Western Australia. *Viruses* 11(12):1157
67. Strydom E, Pietersen G. 2018. Diversity of partial RNA-dependent RNA polymerase gene sequences of soybean blotchy mosaic virus isolates from different host-, geographical- and temporal origins. *Arch. Virol.* 163(5):1299–305
68. L'Heritier PH, Teissier G. 1937. Une anomalie physiologique héréditaire chez la Drosophile. *C. R. Acad. Sci. Paris* 231:192–94
69. Roossinck MJ. 2010. Lifestyles of plant viruses. *Philos. Trans. R. Soc. B* 365(1548):1899–905
70. Ghabrial SA, Castón JR, Jiang D, Nibert ML, Suzuki N. 2015. 50-plus years of fungal viruses. *Virology* 479–480:356–68
71. Boccardo G, Lisa V, Luisoni E, Milne RG. 1987. Cryptic plant viruses. *Adv. Virus Res.* 32:171–214
72. Obbard DJ. 2018. Expansion of the metazoan virosphere: progress, pitfalls, and prospects. *Curr. Opin. Virol.* 31:17–23
73. Wang D. 2015. Fruits of virus discovery: new pathogens and new experimental models. *J. Virol.* 89(3):1486–88
74. Shi M, Zhang Y-Z, Holmes EC. 2018. Meta-transcriptomics and the evolutionary biology of RNA viruses. *Virus Res.* 243:83–90
75. Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, et al. 2017. Consensus statement: virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol.* 15(3):161–68
76. Allander T, Emerson SU, Engle RE, Purcell RH, Bukh J. 2001. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *PNAS* 98(20):11609–14
77. Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall AM, et al. 2002. Genomic analysis of uncultured marine viral communities. *PNAS* 99(22):14250–55
78. Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, et al. 2002. Microarray-based detection and genotyping of viral pathogens. *PNAS* 99(24):15687–92
79. Min Jou W, Haegeman G, Ysebaert M, Fiers W. 1972. Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. *Nature* 237(5350):82–88
80. Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, et al. 1977. Nucleotide sequence of bacteriophage φX174 DNA. *Nature* 265(5596):687–95
81. Racaniello VR, Baltimore D. 1981. Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the viral genome. *PNAS* 78(8):4887–91
82. Angly FE, Felts B, Breitbart M, Salamon P, Edwards RA, et al. 2006. The marine viromes of four oceanic regions. *PLOS Biol.* 4(11):e368
83. Tisza MJ, Pastrana DV, Welch NL, Stewart B, Peretti A, et al. 2020. Discovery of several thousand highly diverse circular DNA viruses. *eLife* 9:e51971
84. de la Higuera I, Kasun GW, Torrance EL, Pratt AA, Maluenda A, et al. 2020. Unveiling crucivirus diversity by mining metagenomic data. *mBio* 11(5):e01410-20
85. Gregory AC, Zayed AA, Conceição-Neto N, Temperton B, Bolduc B, et al. 2019. Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177(5):1109–23.e14
86. Roux S, Hallam SJ, Woyke T, Sullivan MB. 2015. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *eLife* 4:e08490
87. Gilbert KB, Holcomb EE, Allscheid RL, Carrington JC. 2019. Hiding in plain sight: new virus genomes discovered via a systematic analysis of fungal public transcriptomes. *PLOS ONE* 14(7):e0219207
88. Nayfach S, Páez-Espino D, Call L, Low SJ, Sberro H, et al. 2021. Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. *Nat. Microbiol.* 6(7):960–70
89. Armstrong J. 1985. *Nucleotide Sequences 1985: A Compilation from the GenBank and EMBL Data Libraries: A Special Supplement to* Nucleic Acids Research. Washington, DC: IRL Press
90. Roossinck MJ. 2015. Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles. *Front. Microbiol.* 5:767

91. Stenglein MD, Sanders C, Kistler AL, Ruby JG, Franco JY, et al. 2012. Identification, characterization, and *in vitro* culture of highly divergent arenaviruses from boa constrictors and annulated tree boas: candidate etiological agents for snake inclusion body disease. *mBio* 3(4):e00180-12
92. Bodewes R, Kik MJL, Raj VS, Schapendonk CME, Haagmans BL, et al. 2013. Detection of novel divergent arenaviruses in boid snakes with inclusion body disease in the Netherlands. *J. Gen. Virol.* 94(Part 6):1206–10
93. Hetzel U, Sironen T, Laurinmäki P, Liljeroos L, Patjas A, et al. 2013. Isolation, identification, and characterization of novel arenaviruses, the etiological agents of boid inclusion body disease. *J. Virol.* 87(20):10918–35
94. Hyndman TH, Marschang RE, Wellehan JFX, Nicholls PK. 2012. Isolation and molecular identification of Sunshine virus, a novel paramyxovirus found in Australian snakes. *Infect. Genet. Evol.* 12(7):1436–46
95. Hyndman TH, Shilton CM, Stenglein MD, Wellehan JFX. 2018. Divergent bornaviruses from Australian carpet pythons with neurological disease date the origin of extant *Bornaviridae* prior to the end-Cretaceous extinction. *PLOS Pathog.* 14(2):e1006881
96. Stenglein MD, Jacobson ER, Wozniak EJ, Wellehan JFX, Kincaid A, et al. 2014. Ball python nidovirus: a candidate etiologic agent for severe respiratory disease in *Python regius*. *mBio* 5(5):e01484-14
97. Bodewes R, Lempp C, Schürch AC, Habierski A, Hahn K, et al. 2014. Novel divergent nidovirus in a python with pneumonia. *J. Gen. Virol.* 95(Part 11):2480–85
98. Zhang J, Finlaison DS, Frost MJ, Gestier S, Gu X, et al. 2018. Identification of a novel nidovirus as a potential cause of large scale mortalities in the endangered Bellinger River snapping turtle (*Myuchelys georgesi*). *PLOS ONE* 13(10):e0205209
99. O'Dea MA, Jackson B, Jackson C, Xavier P, Warren K. 2016. Discovery and partial genomic characterisation of a novel nidovirus associated with respiratory disease in wild shingleback lizards (*Tiliqua rugosa*). *PLOS ONE* 11(11):e0165209
100. Sarker S, Hannon C, Athukorala A, Bielefeldt-Ohmann H. 2021. Emergence of a novel pathogenic poxvirus infection in the endangered green sea turtle (*Chelonia mydas*) highlights a key threatening process. *Viruses* 13(2):219
101. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, et al. 2020. A new coronavirus associated with human respiratory disease in China. *Nature* 579(7798):265–69
102. Naccache SN, Greninger AL, Lee D, Coffey LL, Phan T, et al. 2013. The perils of pathogen discovery: origin of a novel parvovirus-like hybrid genome traced to nucleic acid extraction spin columns. *J. Virol.* 87(22):11966–77
103. Coyle MC, Elya CN, Bronski M, Eisen MB. 2018. Entomophthovirus: an insect-derived iflavirus that infects a behavior manipulating fungal pathogen of dipterans. bioRxiv 371526. **https://doi.org/10.1101/371526**
104. Li L, Victoria JG, Wang C, Jones M, Fellers GM, et al. 2010. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. *J. Virol.* 84(14):6955–65
105. Roux S, Hawley AK, Torres Beltran M, Scofield M, Schwientek P, et al. 2014. Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell- and meta-genomics. *eLife* 3:e03125
106. Fredericks DN, Relman DA. 1996. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. *Clin. Microbiol. Rev.* 9(1):18–33
107. Walker PJ, Siddell SG, Lefkowitz EJ, Mushegian AR, Adriaenssens EM, et al. 2021. Changes to virus taxonomy and to the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses 2021. *Arch. Virol.* 166(9):2633–48
108. Batson J, Dudas G, Haas-Stapleton E, Kistler AL, Li LM, et al. Single mosquito metatranscriptomics identifies vectors, emerging pathogens and reservoirs in one assay. *eLife* 10:e68353
109. Hermida Lorenzo RJ, Cadar D, Koundouno FR, Juste J, Bialonski A, et al. 2021. Metagenomic snapshots of viral components in Guinean bats. *Microorganisms* 9(3):599
110. Geoghegan JL, Di Giallonardo F, Wille M, Ortiz-Baez AS, Costa VA, et al. 2021. Virome composition in marine fish revealed by meta-transcriptomics. *Virus Evol.* 7(1):veab005
111. Tokarz R, Lipkin WI. 2020. Discovery and surveillance of tick-borne pathogens. *J. Med. Entomol.* 58(4):1525–35

112. Modha S, Hughes J, Bianco G, Ferguson HM, Helm B, et al. 2019. Metaviromics reveals unknown viral diversity in the biting midge *Culicoides impunctatus*. *Viruses* 11(9):865
113. Dezordi FZ, dos Santos Vasconcelos CR, Rezende AM, Wallau GL. 2020. In and outs of *Chuviridae* endogenous viral elements: origin of a potentially new retrovirus and signature of ancient and ongoing arms race in mosquito genomes. *Front. Genet.* 11:1291
114. Hahn MA, Rosario K, Lucas P, Dheilly NM. 2020. Characterization of viruses in a tapeworm: phylogenetic position, vertical transmission, and transmission to the parasitized host. *ISME J.* 14(7):1755–67
115. Bryant JL. 2008. Animal models in virology. In *Sourcebook of Models for Biomedical Research*, ed. PM Conn, pp. 557–63. Totowa, NJ: Humana
116. Xu J, Cherry S. 2014. Viruses and antiviral immunity in *Drosophila*. *Dev. Comp. Immunol.* 42(1):67–84
117. L'heritier P. 1958. The hereditary virus of Drosophila. *Adv. Virus Res.* 5:195–245
118. Rosshart SP, Herz J, Vassallo BG, Hunter A, Wall MK, et al. 2019. Laboratory mice born to wild mice have natural microbiota and model human immune responses. *Science* 365(6452):eaaw4361
119. Phan TG, Kapusinszky B, Wang C, Rose RK, Lipton HL, Delwart EL. 2011. The fecal viral flora of wild rodents. *PLOS Pathog.* 7(9):e1002218
120. Rosshart SP, Vassallo BG, Angeletti D, Hutchinson DS, Morgan AP, et al. 2017. Wild mouse gut microbiota promotes host fitness and improves disease resistance. *Cell* 171(5):1015–28.e13
121. Wallace MA, Coffman KA, Gilbert C, Ravindran S, Albery GF, et al. 2021. The discovery, distribution, and diversity of DNA viruses associated with *Drosophila melanogaster* in Europe. *Virus Evol.* 7(1):veab031
122. Williams SH, Che X, Garcia JA, Klena JD, Lee B, et al. 2018. Viral diversity of house mice in New York City. *mBio* 9(2):e01354-17
123. Félix M-A, Ashe A, Piffaretti J, Wu G, Nuez I, et al. 2011. Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to nodaviruses. *PLOS Biol.* 9(1):e1000586
124. Félix M-A, Wang D. 2019. Natural viruses of *Caenorhabditis* nematodes. *Annu. Rev. Genet.* 53:313–26
125. Binesh CP. 2013. Mortality due to viral nervous necrosis in zebrafish *Danio rerio* and goldfish *Carassius auratus*. *Dis. Aquat. Organ.* 104(3):257–60
126. Altan E, Kubiski SV, Boros Á, Reuter G, Sadeghi M, et al. 2019. A highly divergent picornavirus infecting the gut epithelia of zebrafish (*Danio rerio*) in research institutions worldwide. *Zebrafish* 16(3):291–99
127. Balla KM, Rice MC, Gagnon JA, Elde NC. 2020. Linking virus discovery to immune responses visualized during zebrafish infections. *Curr. Biol.* 30(11):2092–103.e5
128. Cross ST, Maertens BL, Dunham TJ, Rodgers CP, Brehm AL, et al. 2020. Partitiviruses infecting *Drosophila melanogaster* and *Aedes aegypti* exhibit efficient biparental vertical transmission. *J. Virol.* 94(20):e01070-20
129. Franz CJ, Renshaw H, Frezal L, Jiang Y, Félix M-A, Wang D. 2014. Orsay, Santeuil and Le Blanc viruses primarily infect intestinal cells in *Caenorhabditis* nematodes. *Virology* 448:255–64
130. Palmer WH, Medd NC, Beard PM, Obbard DJ. 2018. Isolation of a natural DNA virus of *Drosophila melanogaster*, and characterisation of host resistance and immune responses. *PLOS Pathog.* 14(6):e1007050
131. Palmer WH, Joosten J, Overheul GJ, Jansen PW, Vermeulen M, et al. 2019. Induction and suppression of NF-κB signalling by a DNA virus of *Drosophila*. *J. Virol.* 93(3):e01443-18
132. Goldstein B, King N. 2016. The future of cell biology: emerging model organisms. *Trends Cell Biol.* 26(11):818–24
133. Russell JJ, Theriot JA, Sood P, Marshall WF, Landweber LF, et al. 2017. Non-model model organisms. *BMC Biol.* 15(1):55
134. Tang SKY, Marshall WF. 2017. Self-repairing cells: how single cells heal membrane ruptures and restore lost structures. *Science* 356(6342):1022–25
135. Hibshman JD, Clegg JS, Goldstein B. 2020. Mechanisms of desiccation tolerance: themes and variations in brine shrimp, roundworms, and tardigrades. *Front. Physiol.* 11:592016
136. Woznica A, King N. 2018. Lessons from simple marine models on the bacterial regulation of eukaryotic development. *Curr. Opin. Microbiol.* 43:108–16

Annual Review
of Virology

Volume 9, 2022

Contents

History

Opportunities, Technology, and the Joy of Discovery
Patricia G. Spear 1

The Ups and Downs of an Out-of-the-Box Scientist with a Curious Mind
Marilyn J. Roossinck 19

Tobacco Mosaic Virus and the History of Molecular Biology
Angela N.H. Creager 39

Ecology and Evolution

Understanding the Impacts of Bacteriophage Viruses: From Laboratory Evolution to Natural Ecosystems
Britt Koskella, Catherine A. Hernandez, and Rachel M. Wheatley 57

Amoebae: Hiding in Plain Sight: Unappreciated Hosts for the Very Large Viruses
Victória Fulgêncio Queiroz, Rodrigo Araújo Lima Rodrigues, Paulo Victor de Miranda Boratto, Bernard La Scola, Julien Andreani, and Jônatas Santos Abrahão 79

Viruses in Subsurface Environments
Jennifer Wirth and Mark Young 99

Natural Selection, Intracellular Bottlenecks of Virus Populations, and Viral Superinfection Exclusion
Camila Perdoncini Carvalho, Ruifan Ren, Junping Han, and Feng Qu 121

A Trait-Based Approach to Predicting Viral Host-Range Evolvability
Hannah M. Strobel, Elizabeth C. Stuart, and Justin R. Meyer 139

The Case for Studying New Viruses of New Hosts
Mark D. Stenglein 157

The Ecology of Viral Emergence
Edward C. Holmes 173

Genome Replication, Regulation of Gene Expression, and Biosynthesis

Crowning Touches in Positive-Strand RNA Virus Genome Replication Complex Structure and Function
Masaki Nishikiori, Johan A. den Boon, Nuruddin Unchwaniwala, and Paul Ahlquist 193

The Role of Viral RNA Degrading Factors in Shutoff of Host Gene Expression
Léa Gaucherand and Marta Maria Gaglia 213

The Life Cycle of the Vaccinia Virus Genome
Matthew D. Greseth and Paula Traktman 239

Nuclear Capsid Uncoating and Reverse Transcription of HIV-1
Thorsten G. Müller, Vojtech Zila, Barbara Müller, and Hans-Georg Kräusslich 261

Virus Cell Biology

Liquid Phase Partitioning in Virus Replication: Observations and Opportunities
Chao Wu, Alex S. Holehouse, Daisy W. Leung, Gaya K. Amarasinghe, and Rebecca Ellis Dutch 285

Replication Compartments of Eukaryotic and Bacterial DNA Viruses: Common Themes Between Different Domains of Host Cells
David M. Knipe, Amy Prichard, Surendra Sharma, and Joe Pogliano 307

Viral G Protein–Coupled Receptors Encoded by β- and γ-Herpesviruses
Mette M. Rosenkilde, Naotaka Tsutsumi, Julius M. Knerr, Dagmar F. Kildedal, and K. Christopher Garcia 329

The Role of Viruses in Identifying and Analyzing RNA Silencing
David C. Baulcombe 353

APOBEC3: Friend or Foe in Human Papillomavirus Infection and Oncogenesis?
Cody J. Warren, Mario L. Santiago, and Dohun Pyeon 375

Systems Biology of Virus-Host Protein Interactions: From Hypothesis Generation to Mechanisms of Replication and Pathogenesis
Priya S. Shah, Nitin S. Beesabathuni, Adam T. Fishburn, Matthew W. Kenaston, Shiaki A. Minami, Oanh H. Pham, and Inglis Tucker 397

Pathogenesis

Citrus Tristeza Virus: From Pathogen to Panacea
Svetlana Y. Folimonova and Yong-Duo Sun 417

Advances in Understanding Neuropathogenesis of Rift Valley Fever Virus
Kaleigh A. Connors and Amy L. Hartman 437

Immunity

Cyclic Nucleotide Signaling in Phage Defense and Counter-Defense
Januka S. Athukoralage and Malcolm F. White 451

Shifting the Immune Memory Paradigm: Trained Immunity in Viral Infections
Esther J.M. Taks, Simone J.C.F.M. Moorlag, Mihai G. Netea, and Jos W.M. van der Meer 469

Vaccines

Lessons from Acquired Natural Immunity and Clinical Trials to Inform Next-Generation Human Cytomegalovirus Vaccine Development
Xintao Hu, Hsuan-Yuan Wang, Claire E. Otero, Jennifer A. Jenks, and Sallie R. Permar 491

RNA-Based Vaccination of Plants for Control of Viruses
Andreas E. Voloudakis, Athanasios Kaldis, and Basavaprabhu L. Patil 521

Errata

An online log of corrections to *Annual Review of Virology* articles may be found at http://www.annualreviews.org/errata/virology