

Preplanned Studies

A Landscape Analysis on Virus: based on NCBI Database

Mingchen Zhao¹; Jingyuan Chen²; Qiang Wang³; Zuhong Lu⁴; Zhongwei Jia^{1,5,6,#}**Summary****What is already known about this topic?**

Studies indicate that viruses could spread across species, but it is difficult to know when and where such small probability events occur because it is almost impossible to design an observational study on the whole landscape.

What is added by this report?

We did a comprehensive analysis on the National Center for Biotechnology Information database and tried to find the time, place, and host that the viruses stayed in their long evolutionary history.

What are the implications for public health practice?

Public databases are helpful to understand the risk of virus infection in humans and also a cost-effective method for monitoring public health and safety events.

According to the International Committee on Taxonomy of Viruses Master Species List 2020, more than 9,000 virus species have been identified on earth (1), of which the World Health Organization (WHO) announced that more than 200 species were known as zoonotic viruses (2). Previous studies have also shown that zoonoses (hantavirus, Ebola virus, highly pathogenic avian influenza, West Nile virus, Rift Valley fever virus, norovirus, severe acute respiratory syndrome coronavirus 1, Marburg virus, influenza A virus) infected more than 2.5 billion people every year, among which 2.7 million died (3). Zoonotic viruses have aroused broad concerns in recent years so that people have been encouraged to avoid eating wild animals, and a series of animal protection laws and regulations were enacted, such as the Convention on International Trade in Endangered Species of Wild Fauna and Flora (4). Researchers from different fields and countries have tried to collaborate to explore virus associations between animals and humans worldwide (5). However, many studies only focused on investigating a specific virus when it received enough attention as in a localized or global epidemic, such as Ebola virus, H1N1, Zika virus, and severe acute

respiratory syndrome coronavirus 2 (SARS-CoV-2) (6). Even so, it's difficult for people to know exactly when, where and how the virus jumped from animal or environment to humans because it is impossible to design an observational study on the whole landscape around a virus. To understand the overall evolution law of viruses, this study aimed to find the spatial and temporal distribution of viruses and related hosts through the public database.

The National Center for Biotechnology Information (NCBI) virus database is a database of gene sequences maintained by the National Institutes of Health that aggregates and annotates all publicly available nucleotide and protein sequences, and it has been used for studies exploring quantitative snapshots of viral genomic trends and overviewing virus real-time quantitative polymerase chain reaction (RT-qPCR) method performance (7–8). In this study, we performed a spatiotemporal analysis on the viruses in the NCBI database and tried to disclose the potential time, place, and hosts that the virus appeared or stayed in their long evolutionary history. Our study indicated the top five widely distributed viruses and found that no virus had been reported in all countries/regions. In addition, the reported areas of viruses did not completely overlap with areas where their suspected hosts live. We also found that the 249 viruses isolated from humans were also isolated from 705 other mammals and 938 non-mammals. We attempted to map the distribution and evolution of viruses and identify suspected hosts globally based on the NCBI database, which was helpful for understanding the risk of virus infection in humans and was also a cost-effective method for monitoring and predicting public health and safety events globally.

The data were downloaded from the NCBI Virus database and covered available data up to September 2021 (www.ncbi.nlm.nih.gov/labs/virus). The data included virus genomic sequence submission information (accession number, submitter, and release date), virus types (species, genus, and family), and biological sample description (location, organisms of sample, isolation, and collection date). Data that were

duplicated, missing, or had unidentified collected dates, locations, or organisms of the biological samples were excluded. Because the NCBI only reported that the virus was isolated from an organism, not whether the organism was the host of the virus, we referred to the organisms as suspected hosts. Based on the standard of NCBI taxonomy, suspected hosts were classified into animals, plants, and microbes.

The numbers were calculated by virus species, suspected hosts, suspected host locations, and collection time. The spatiotemporal distribution was mapped in country/region units and the path was analyzed by linking the reported location of the viruses or suspected hosts. Potential zoonotic diseases were explored by observing the suspected hosts from which the viruses were isolated.

All statistical analyses were performed in R statistical software (version 4.1.0, The R Foundation for Statistical Computing, Vienna, Austria). Animal silhouettes representing hosts were downloaded from PhyloPic (<http://www.phylopic.org>). All maps were made by using ArcGIS (version 10.7, Esri Inc, Redlands, CA, USA).

A total of 605,504 records covering 24,234 viruses from 240 countries/regions with sample collection date from 1865 to September 22, 2021 were involved in the final analysis. We observed that 12,243 viruses were isolated from 4,187 animals, 2,856 from 2,074 plants, and 9,176 from 965 microbes (Figure 1). However, more than 90% viruses (21,845 of 24,234) were reported in a single country/region, and the remaining 10% (2,389 of 24,234) of viruses were from 236 countries/regions and were isolated from 4,742 suspected hosts in 238 countries/regions.

The top five widely distributed viruses were influenza A virus, human immunodeficiency virus 1, hepatitis B virus, rabies lyssavirus, dengue virus, and measles morbillivirus (dengue virus and measles morbillivirus tied for fifth place), which were reported in 155, 151, 145, 128, 122, and 122 countries/regions, respectively (Supplementary Table S1, available in <http://weekly.chinacdc.cn/>). Influenza A virus was reported over approximately 118 years, ranging from the first report in Italy in 1902 to the last report in Tanzania in 2020, and was reported in 155 countries/regions from 422 suspected hosts living in 231 countries/regions. Human immunodeficiency virus 1 was reported over approximately 120 years, ranging from Russia in 1899 to Iraq in 2019, and was reported in 151 countries/regions from 1 suspected

host (*Homo sapiens*) in 225 countries/regions; Hepatitis B virus was reported over approximately 120 years, ranging from Indonesia in 1900 to Ireland in 2020, and was reported in 145 countries/regions from 35 suspected hosts in 229 countries/regions; Rabies lyssavirus was reported over approximately 138 years, ranging from France in 1882 to Equatorial Guinea in 2020, and was reported in 128 countries/regions from 192 suspected hosts in 233 countries/regions; Dengue virus was reported over approximately 75 years, ranging from Colombia in 1944 to Sudan in 2019, and was reported in 122 countries/regions from 25 suspected hosts in 227 countries/regions; Measles morbillivirus was reported over approximately 101 years, ranging from New Zealand in 1919 to United Arab Emirates 2020, and was reported in 122 countries/regions from 3 suspected hosts (*Chlorocebus aethiops*, *Homo sapiens*, and *Macaca fascicularis*) in 225 countries/regions (Figure 2, Supplementary Figure S1A, available in <http://weekly.chinacdc.cn/>).

The top five widely distributed suspected hosts were *Homo sapiens*, *Bos taurus*, *Canis lupus familiaris*, *Gallus gallus*, and *Sus scrofa*, which were reported in 225, 143, 123, 122 and 110 countries/regions, respectively (Supplementary Table S1). *Homo sapiens* were reported in 225 countries/regions in the past 154 years (from the United States in 1866 to the Faroe Islands in 2020), and approximately 1,075 viruses were isolated from this host. *Bos taurus* were reported in 143 countries/regions during the past 138 years (from France in 1882 to Luxembourg in 2020) and from which 273 viruses were isolated; *Canis lupus familiaris* were reported in 123 countries/regions during the past 89 years (from China in 1931 to Ghana in 2020) and from which 106 viruses were isolated; *Gallus gallus* were reported in 122 countries/regions during the past 118 years (from Italy in 1902 to Timor-Leste in 2020) and from which 287 viruses were isolated; *Sus scrofa* were reported in 110 countries/regions during the past 99 years (from China in 1921 to Paraguay in 2020) and from which 276 viruses were isolated (Supplementary Figure S1B, available in <http://weekly.chinacdc.cn/>).

There were 27 viruses in the database reported to be isolated from animals as well as plants (animals-plants). Similarly, 9 and 5 viruses were isolated from animals-microbes and plants-microbes, respectively (Figure 3A). Specifically, 249 viruses from humans were also isolated from the other 1,643 animals, among which 705 were mammals. Most of the 249

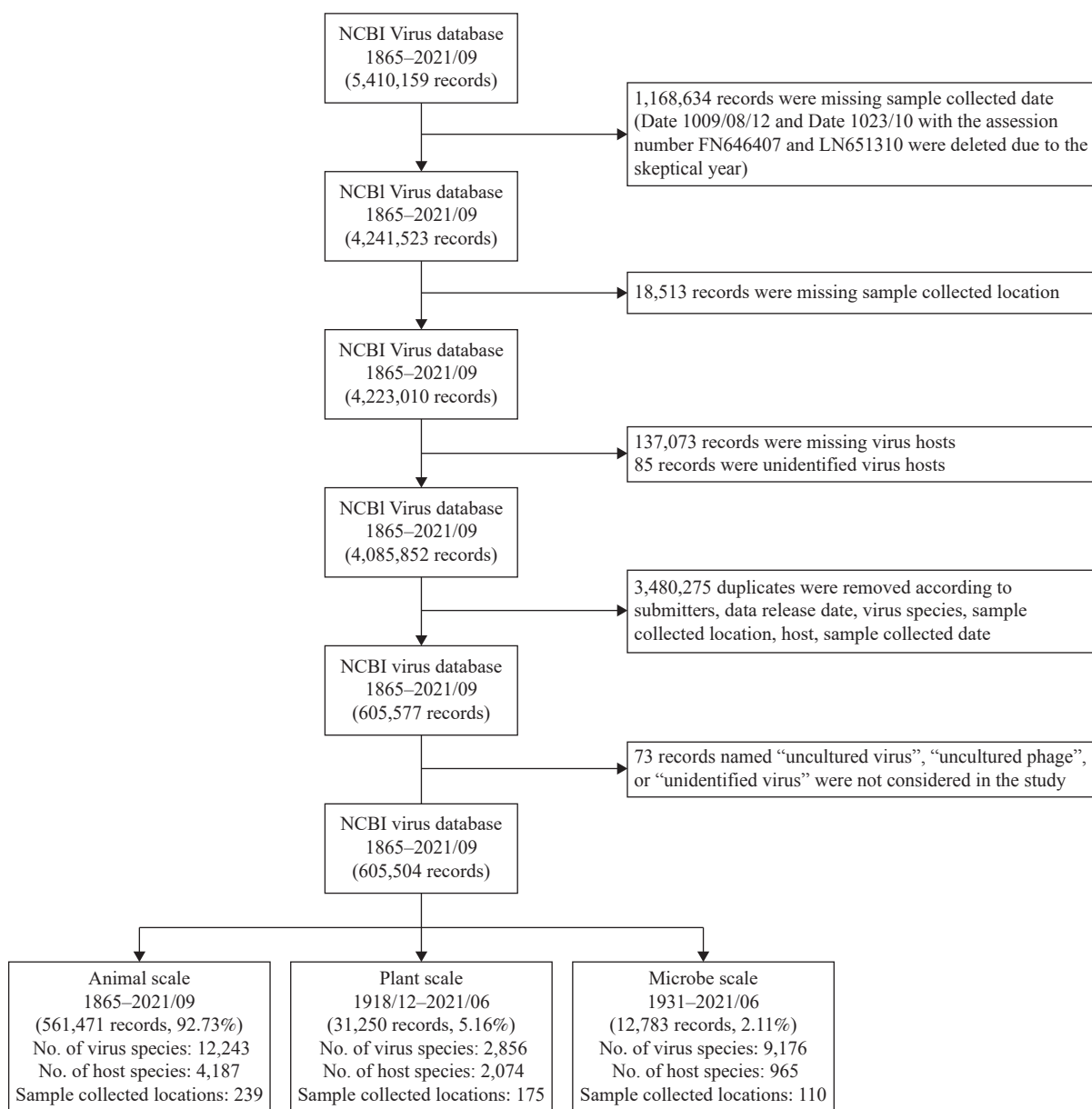


FIGURE 1. Flow diagram depicting the process of identification and inclusion of selected records. Abbreviation: NCBI=National Center for Biotechnology Information.

multi-suspected host viruses were reported in the United States (125 of 249), China (141 of 249), and Brazil (107 of 249), with 402, 344, and 187 suspected hosts, respectively (Figure 3B). The first reported human-animal multi-suspected host virus was the Vaccinia virus, isolated from *Homo sapiens* in the United States in 1866 (Supplementary Table S2, available in <http://weekly.chinacdc.cn/>).

In the study period, 181 countries/regions reported at least 1 new virus. The top 5 countries reporting the highest number of new virus species were the United States, China, Brazil, Australia, and Italy, with 6,628,

4,518, 1,032, 994, and 730 new virus species, respectively. Newly reported viruses showed a rapid increase since 2003, peaking in 2013, and then showed a jump decline until the bottom of 2019 (Supplementary Figure S2, available in <http://weekly.chinacdc.cn/>). In 2021, 14 new viruses were reported, of which Belarus, Canada, South Africa, the Republic of Korea, Spain, and Switzerland reported 1 new virus, Vietnam reported 2 new viruses, and the United States and China both reported 3 new virus species (Supplementary Table S3, available in <http://weekly.chinacdc.cn/>).

DISCUSSION

To our knowledge, this is the first systematic, global landscape analysis of viruses utilizing the NCBI virus database. Although the database may not cover all known viruses, the NCBI includes almost all viruses that have had significant impact on humans. The study indicates that approximately 24,234 viruses have been investigated as of September 22, 2021, covering 240 countries/regions and involving suspected hosts of 4,187 animals, 2,074 plants and 965 microbes (Figure 1). The top five most widely distributed viruses were influenza A virus, human immunodeficiency virus 1, Hepatitis B virus, rabies lyssavirus, dengue virus and measles morbillivirus, and the top five most widely distributed suspected hosts were *Homo sapiens*, *Bos taurus*, *Canis lupus familiaris*, *Gallus gallus*, and *Sus scrofa*, which all covered more than 100 countries/regions (Supplementary Table S1, available in <http://weekly.chinacdc.cn/>). These outcomes pose several concerning issues.

First, the top five most widely distributed viruses have been investigated and reported in most countries/regions (Supplementary Table S1, available in <http://weekly.chinacdc.cn/>), but no virus has been reported in all countries/regions. The surprising finding is the distribution of influenza A virus — a respiratory virus closely related to humans. We presumed it should be distributed everywhere and reported by all 240 countries/regions. Our study indicates that influenza A virus is the most reported location but 85 countries/regions have not yet reported it. It is possible that the NCBI database has not collected these viruses in these countries/regions or that these areas have not covered these viruses. In either case, these countries/regions deserve attention from the perspective of virus monitoring.

Second, the reported areas of viruses do not completely overlap with areas where their suspected hosts live. For example, influenza A virus has been reported in 155 countries/regions, but its 422 hosts live in 231 countries/regions. The results have two explanations: first, the 76 countries/regions may have influenza A virus but do not investigate and report it; second, the 76 countries/regions could have no influenza A virus, which implies that these areas are susceptible to influenza A virus. However, both viruses and suspected hosts all take approximately one hundred years to survive in a region (Supplementary Figure S1, available in <http://weekly.chinacdc.cn/>).

This long period of history leaves many opportunities for people to find and be involved in their evolution.

Third, we found that 249 viruses from humans were also isolated from other 705 mammals and 938 non-mammals. It is not surprising that most mammals' zoonoses are from *Sus scrofa*, *Gallus gallus*, *Anas platyrhynchos*, *Canis lupus familiaris*, and *Bos taurus*, because these domestic mammals have a close relationship with humans. Bats (152 kinds of species in the database) and *Paguma larvata* covered 44 zoonoses and 3 zoonoses with humans in the reported data. We also found 27 viruses from animals and plants that are involved in insects, whether it is related with humans has not been reported (Figure 2).

This study was subject to some limitations. First, NCBI is a public database, and the data quality may be uneven. However, NCBI has a form for submitted data, which guarantees that the basic information of submitted data is consistent at some level, and these data can meet the requirements of our spatiotemporal analysis. Second, we classified viruses and suspected hosts according to their submitted names. It will be regarded as two different viruses or suspected hosts if their names are spelled differently. For example, *Enterovirus A*, *Enterovirus B*, *Enterovirus C*, and *Enterovirus sp.* are regarded as different viruses in this study. This might be a slight overestimation of the virus species and suspected hosts submitted in the NCBI database but cannot change their whole spatiotemporal distribution, especially in the long history of their evolution. Further, we identified 249 zoonotic viruses in NCBI, which was consistent with the figure published by the WHO (2). Third, data reported by countries may not be complete, which could lead to bias in the analysis. Also, the advent of next generation sequencing (NGS) would cause bias on the temporal discovery of viruses because it was hard to perform large-scale genome sequencing of viruses before NGS became relatively affordable. Fourth, we did not analyze the path of each virus and their suspected hosts. Analysis of the overlap of viruses and their hosts is important for predicting the risk of a new virus outbreak in local areas, but beyond the scope of this study.

In conclusion, we attempted to map the distribution and evolution of viruses and suspected hosts globally based on the NCBI database, which is helpful for understanding the risk of virus infection in humans and is also a cost-effective method for monitoring and predicting public health and safety events globally.

Conflicts of interest: No conflicts of interest.

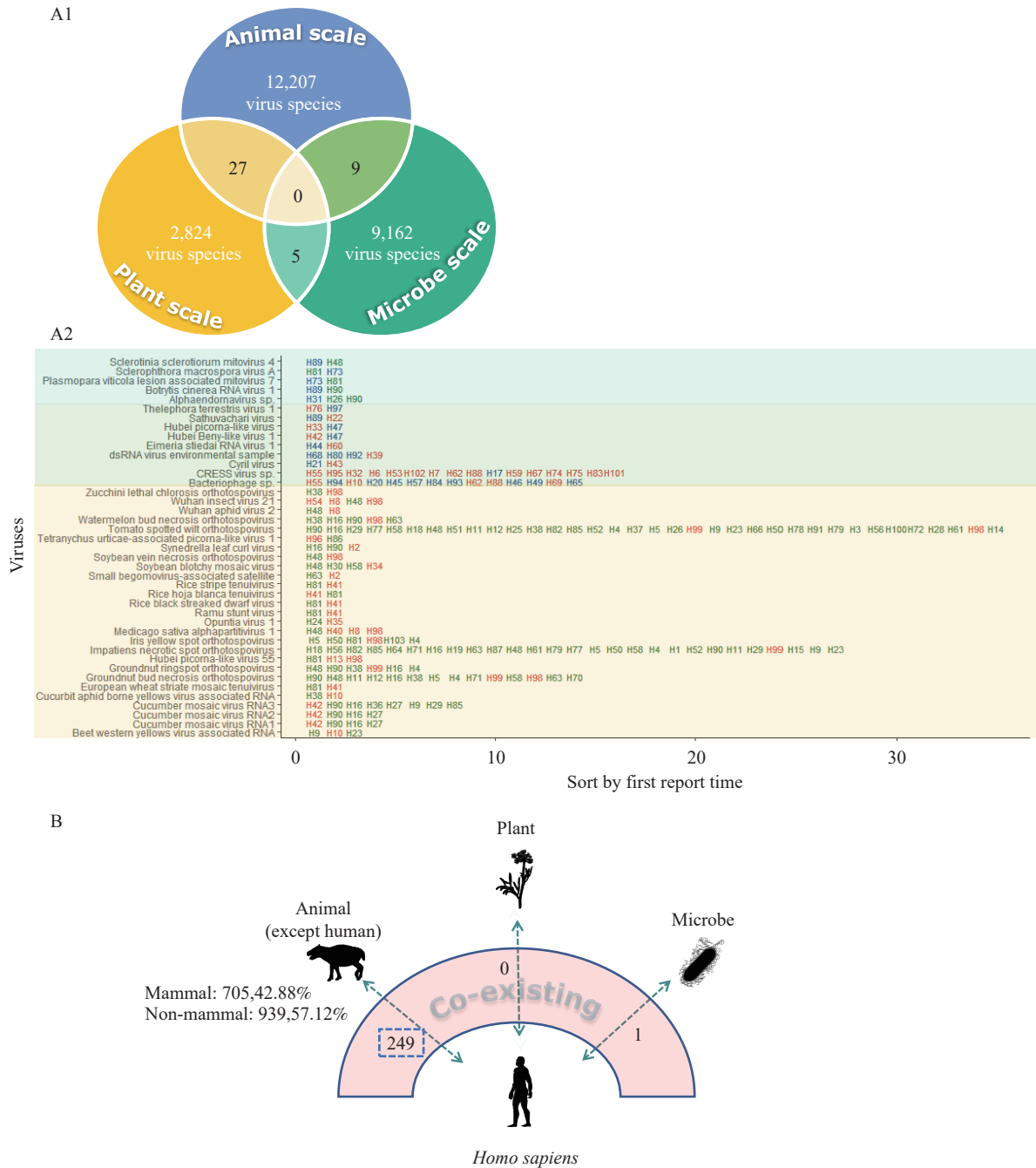


FIGURE 2. Coexisting virus species among animals, plants, and microbes. (A1) The number and intersection of virus species reported from NCBI Virus Database between 1865 and 2021/9. (A2) First reported host family sorted by first report time among coexisting virus species; (B) Coexisting virus species between *Homo sapiens* and animals, plants, or microbes. Notes: The host families represented by the tags were shown in Supplementary Table S4, available in <http://weekly.chinacdc.cn/>.

Funding: This work was supported by the National Key Research and Development Program of China [No. 2020YFC0849200, 2021YFC0863400] and NSFC projects [No. 72174004, 91546203, 91846302].

doi: 10.46234/ccdcw2022.019

Corresponding author: Zhongwei Jia, urchinjj@163.com.

¹ School of Public Health, Peking University, Beijing, China; ² West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, Sichuan, China; ³ Goodwill Hessian Health Technology Co., Ltd., Beijing, China; ⁴ State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing, Jiangsu, China; ⁵ Center for Intelligent Public Health, Institute for Artificial Intelligence, Peking University,

Beijing, China; ⁶ Center for Drug Abuse Control and Prevention, National Institute of Health Data Science, Peking University, Beijing, China.

Submitted: December 13, 2021; Accepted: January 12, 2022

REFERENCES

1. Walker PJ, Siddell SG, Lefkowitz EJ, Mushegian AR, Adriaenssens EM, Alfenas-Zerbini P, et al. 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* 2021;166(9):2633 – 48. <http://dx.doi.org/10.1007/s00705-021-05156-1>.
2. World Health Organization. Zoonoses. <https://www.who.int/news-room/fact-sheets/detail/zoonoses>. [2021-12-01].
3. Gebreyes WA, Dupouy-Camet J, Newport MJ, Oliveira CJB, Schlesinger LS, Saif YM, et al. The global one health paradigm: challenges and opportunities for tackling infectious diseases at the human, animal, and environment interface in low-resource settings. *PLoS Negl Trop Dis* 2014;8(11):e3257. <http://dx.doi.org/10.1371/journal.pntd.0003257>.
4. The Convention on International Trade in Endangered Species of Wild Fauna and Flora. 2021. <https://cites.org/eng/disc/text.php>. [2021-12-01].
5. World Health Organization. Regional Office for the Western Pacific. Zoonotic diseases: a guide to establishing collaboration between animal and human health sectors at the country level. Geneva: WHO Regional Office for the Western Pacific. 2008. <https://iris.wpro.who.int/handle/10665.1/10415>.
6. Wilder-Smith A, Osman S. Public health emergencies of international concern: a historic overview. *J Travel Med* 2020;27(8):taaa227. <http://dx.doi.org/10.1093/jtm/taaa227>.
7. Mahmoudabadi G, Phillips R. A comprehensive and quantitative exploration of thousands of viral genomes. *eLife* 2018;7:e31955. <http://dx.doi.org/10.7554/eLife.31955>.
8. Vanneste K, Garland L, Broeders S, Van Gucht S, Roosens NH. Application of whole genome data for in silico evaluation of primers and probes routinely employed for the detection of viral species by RT-qPCR using dengue virus as a case study. *BMC Bioinform* 2018;19(1):312. <http://dx.doi.org/10.1186/s12859-018-2313-0>.

SUPPLEMENTARY TABLE S1. The top five widely distributed viruses and hosts.

Virus species	Distributed countries/regions
Influenza A virus	Afghanistan; Albania; Algeria; Angola; Antarctica; Argentina; Australia; Austria; Azerbaijan; Bahrain; Bangladesh; Belarus; Belgium; Benin; Bhutan; Bolivia; Bosnia and Herzegovina; Brazil; Bulgaria; Burkina Faso; Cambodia; Cameroon; Canada; Central African Republic; Chad; Chile; China; Colombia; Cook Islands; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Cyprus; Czech Republic; Czechoslovakia; Democratic Republic of the Congo; Denmark; Djibouti; Dominican Republic; Ecuador; Egypt; El Salvador; Estonia; Ethiopia; Fiji; Finland; France; French Guiana; French Polynesia; Georgia; Germany; Ghana; Gibraltar; Greece; Greenland; Guatemala; Haiti; Honduras; Hong Kong, China; Hungary; Iceland; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Kosovo; Kuwait; Kyrgyzstan; Laos; Latvia; Lebanon; Libya; Lithuania; Luxembourg; Macao, China; Madagascar; Malaysia; Mali; Mauritius; Mexico; Moldova; Mongolia; Montenegro; Montserrat; Morocco; Myanmar; Namibia; Nepal; Netherlands; New Caledonia; New Zealand; Nicaragua; Niger; Nigeria; Democratic People's Republic of Korea; Norway; Oman; Pakistan; Palestine; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Qatar; Republic of Serbia; Reunion; Romania; Russia; Saudi Arabia; Senegal; Singapore; Slovakia; Slovenia; Solomon Islands; South Africa; Republic of Korea; Spain; Sri Lanka; Sudan; Suriname; Sweden; Switzerland; Taiwan, China; Tajikistan; Thailand; Togo; Tonga; Tunisia; Turkey; Turkmenistan; Uganda; Ukraine; United Arab Emirates; United Kingdom; United Republic of Tanzania; United States of America; Uruguay; Uzbekistan; Venezuela; Vietnam; West Bank; Zambia; Zimbabwe (155)
Human immunodeficiency virus 1	Afghanistan; Algeria; Angola; Argentina; Armenia; Australia; Austria; Azerbaijan; Bangladesh; Barbados; Belarus; Belgium; Belize; Benin; Bolivia; Bosnia and Herzegovina; Botswana; Brazil; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Cape Verde; Central African Republic; Chad; Chile; China; Colombia; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Cyprus; Czech Republic; Democratic Republic of the Congo; Denmark; Djibouti; Dominican Republic; Ecuador; Egypt; El Salvador; Equatorial Guinea; Eritrea; Estonia; Ethiopia; Fiji; Finland; France; French Guiana; Gabon; Gambia; Georgia; Germany; Ghana; Greece; Greenland; Guadeloupe; Guatemala; Guinea; Guinea-Bissau; Haiti; Honduras; Hong Kong, China; Hungary; Iceland; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Kazakhstan; Kenya; Kuwait; Kyrgyzstan; Laos; Latvia; Liberia; Libya; Lithuania; Luxembourg; Macao, China; Madagascar; Malawi; Malaysia; Mali; Martinique; Mauritania; Mauritius; Mexico; Mongolia; Morocco; Mozambique; Myanmar; Nepal; Netherlands; Niger; Nigeria; Norway; Pakistan; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Republic of Serbia; Romania; Russia; Rwanda; Saudi Arabia; Senegal; Seychelles; Sierra Leone; Singapore; Slovakia; Slovenia; Somalia; South Africa; Republic of Korea; Spain; Sri Lanka; Sudan; Suriname; Swaziland; Sweden; Switzerland; Taiwan, China; Tajikistan; Thailand; Togo; Trinidad and Tobago; Tunisia; Turkey; Uganda; Ukraine; United Kingdom; United Republic of Tanzania; United States of America; Uruguay; Uzbekistan; Venezuela; Vietnam; Zambia; Zimbabwe (151)
Hepatitis B virus	Afghanistan; Albania; Algeria; Angola; Argentina; Armenia; Australia; Azerbaijan; Bangladesh; Belarus; Belgium; Bhutan; Bolivia; Bosnia and Herzegovina; Botswana; Brazil; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Cape Verde; Central African Republic; Chad; Chile; China; Colombia; Comoros; Cote d'Ivoire; Croatia; Cuba; Czech Republic; Democratic Republic of the Congo; Dominican Republic; Ecuador; Egypt; Eritrea; Estonia; Ethiopia; Fiji; Finland; France; French Guiana; French Polynesia; Gabon; Gambia; Germany; Ghana; Greece; Greenland; Guadeloupe; Guinea; Guinea-Bissau; Guyana; Haiti; Hong Kong, China; Hungary; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Japan; Jordan; Kenya; Kiribati; Kuwait; Kyrgyzstan; Laos; Latvia; Lebanon; Liberia; Libya; Lithuania; Macedonia; Madagascar; Malaysia; Maldives; Mali; Mauritania; Mauritius; Mayotte; Mexico; Moldova; Mongolia; Morocco; Mozambique; Myanmar; Nepal; Netherlands; New Caledonia; New Zealand; Niger; Nigeria; Norway; Pakistan; Palestine; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Republic of Serbia; Reunion; Romania; Russia; Rwanda; Samoa; Saudi Arabia; Senegal; Sierra Leone; Singapore; Slovakia; Slovenia; Somalia; South Africa; Republic of Korea; Spain; Sri Lanka; Sudan; Suriname; Sweden; Switzerland; Syria; Taiwan, China; Thailand; Tonga; Tunisia; Turkey; Uganda; Ukraine; United Kingdom; United Republic of Tanzania; United States of America; Uruguay; Uzbekistan; Vanuatu; Venezuela; Vietnam; Zimbabwe (145)
Rabies lyssavirus	Afghanistan; Algeria; Argentina; Austria; Azerbaijan; Bangladesh; Belgium; Benin; Bhutan; Bosnia and Herzegovina; Botswana; Brazil; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Central African Republic; Chad; Chile; China; Colombia; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Czech Republic; Democratic Republic of the Congo; Dominican Republic; Ecuador; Egypt; El Salvador; Equatorial Guinea; Estonia; Ethiopia; Finland; France; French Guiana; Gabon; Gambia; Georgia; Germany; Ghana; Greece; Greenland; Grenada; Guatemala; Guinea; Guyana; Haiti; Honduras; Hungary; India; Indonesia; Iran; Iraq; Israel; Italy; Japan; Jersey; Jordan; Kazakhstan; Kenya; Laos; Latvia; Lebanon; Lesotho; Liberia; Lithuania; Macedonia; Madagascar; Malaysia; Mali; Mauritania; Mexico; Mongolia; Montenegro; Morocco; Mozambique; Myanmar; Namibia; Nepal; Netherlands; Niger; Nigeria; Norway; Oman; Pakistan; Paraguay; Peru; Philippines; Poland; Qatar; Republic of Serbia; Romania; Russia; Rwanda; Saudi Arabia; Senegal; Sierra Leone; Slovenia; Somalia; South Africa; Republic of Korea; Spain; Sri Lanka; Sudan; Svalbard; Swaziland; Switzerland; Syria; Taiwan, China; Tajikistan; Thailand; Tunisia; Turkey; Uganda; Ukraine; United Arab Emirates; United Kingdom; United Republic of Tanzania; United States of America; Uruguay; Vietnam; Zaire; Zambia; Zimbabwe (128)

TABLE S1. (Continued)

Virus species	Distributed countries/regions
Dengue virus	Angola; Anguilla; Antigua and Barbuda; Arctic Ocean; Argentina; Aruba; Australia; Bangladesh; Barbados; Belize; Benin; Bhutan; Bolivia; Borneo; Brazil; British Virgin Islands; Brunei; Burkina Faso; Cambodia; Cameroon; Cape Verde; Chile; China; Colombia; Comoros; Cook Islands; Costa Rica; Cote d'Ivoire; Cuba; Democratic Republic of the Congo; Djibouti; Dominica; Dominican Republic; East Timor; Ecuador; El Salvador; Eritrea; Federated States of Micronesia; Fiji; France; French Guiana; French Polynesia; Gabon; Germany; Ghana; Grenada; Guadeloupe; Guatemala; Guinea; Guyana; Haiti; Honduras; India; Indonesia; Iran; Ireland; Italy; Jamaica; Japan; Kenya; Kiribati; Laos; Madagascar; Malaysia; Maldives; Mali; Marshall Islands; Martinique; Mauritania; Mexico; Mozambique; Myanmar; Nauru; Nepal; New Caledonia; Nicaragua; Nigeria; Niue; Pacific Ocean; Pakistan; Palau; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Portugal; Reunion; Russia; Saint Barthelemy; Saint Kitts and Nevis; Saint Lucia; Saint Vincent and the Grenadines; Samoa; Saudi Arabia; Senegal; Seychelles; Singapore; Solomon Islands; Somalia; Republic of Korea; Spain; Sri Lanka; Sudan; Suriname; Sweden; Taiwan, China; Thailand; The Bahamas; Timor-Leste; Togo; Tonga; Trinidad and Tobago; Tuvalu; United Republic of Tanzania; United States of America; Uruguay; Vanuatu; Venezuela; Vietnam; Wallis and Futuna; Yemen (122)
Measles morbillivirus	Afghanistan; Algeria; Angola; Argentina; Australia; Austria; Bahrain; Bangladesh; Belarus; Belgium; Benin; Bhutan; Bosnia and Herzegovina; Botswana; Brazil; Brunei; Bulgaria; Cameroon; Canada; Central African Republic; Chile; China; Colombia; Cote d'Ivoire; Croatia; Cuba; Cyprus; Czech Republic; Democratic Republic of the Congo; Ecuador; Estonia; Ethiopia; Fiji; Finland; France; Gabon; Germany; Ghana; Gibraltar; Greece; Guatemala; Hong Kong, China; Hungary; Iceland; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Japan; Kazakhstan; Kenya; Kyrgyzstan; Latvia; Lebanon; Lesotho; Liberia; Lithuania; Luxembourg; Macedonia; Malawi; Malaysia; Mali; Mauritania; Mexico; Moldova; Mongolia; Montenegro; Morocco; Mozambique; Myanmar; Namibia; Nepal; Netherlands; New Caledonia; New Zealand; Niger; Nigeria; Norway; Oman; Pakistan; Panama; Papua New Guinea; Philippines; Poland; Portugal; Qatar; Republic of Serbia; Romania; Russia; Rwanda; Senegal; Sierra Leone; Slovakia; Slovenia; Solomon Islands; Somalia; South Africa; Republic of Korea; Spain; Sri Lanka; Sudan; Swaziland; Sweden; Switzerland; Taiwan, China; Thailand; Togo; Tunisia; Turkey; Uganda; Ukraine; United Arab Emirates; United Kingdom; United States of America; Uzbekistan; Venezuela; Vietnam; Zambia; Zimbabwe (122)
Host	Distributed countries
<i>Homo sapiens</i>	Afghanistan; Albania; Algeria; Andorra; Angola; Anguilla; Antigua and Barbuda; Arctic Ocean; Argentina; Armenia; Aruba; Australia; Austria; Azerbaijan; Bahrain; Bangladesh; Barbados; Belarus; Belgium; Belize; Benin; Bermuda; Bhutan; Bolivia; Borneo; Bosnia and Herzegovina; Botswana; Brazil; British Virgin Islands; Brunei; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Cape Verde; Cayman Islands; Central African Republic; Chad; Chile; China; Colombia; Comoros; Cook Islands; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Cyprus; Czech Republic; Democratic Republic of the Congo; Denmark; Djibouti; Dominica; Dominican Republic; East Timor; Ecuador; Egypt; El Salvador; Equatorial Guinea; Eritrea; Estonia; Ethiopia; Faroe Islands; Federated States of Micronesia; Fiji; Finland; France; French Guiana; French Polynesia; Gabon; Gambia; Georgia; Germany; Ghana; Gibraltar; Greece; Greenland; Grenada; Guadeloupe; Guatemala; Guinea; Guinea-Bissau; Guyana; Haiti; Honduras; Hong Kong, China; Hungary; Iceland; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Kiribati; Kosovo; Kuwait; Kyrgyzstan; Laos; Latvia; Lebanon; Lesotho; Liberia; Libya; Liechtenstein; Lithuania; Luxembourg; Macao, China; Macedonia; Madagascar; Malawi; Malaysia; Maldives; Mali; Malta; Marshall Islands; Martinique; Mauritania; Mauritius; Mayotte; Mexico; Moldova; Mongolia; Montenegro; Montserrat; Morocco; Mozambique; Myanmar; Namibia; Nauru; Nepal; Netherlands; New Caledonia; New Zealand; Nicaragua; Nigeria; Niue; Norway; Oman; Pacific Ocean; Pakistan; Palau; Palestine; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Qatar; Republic of Serbia; Reunion; Romania; Russia; Rwanda; Saint Barthelemy; Saint Kitts and Nevis; Saint Lucia; Saint Martin; Saint Vincent and the Grenadines; Samoa; Sao Tome and Principe; Saudi Arabia; Senegal; Seychelles; Sierra Leone; Singapore; Slovakia; Slovenia; Solomon Islands; Somalia; South Africa; Republic of Korea; South Sudan; Spain; Sri Lanka; Sudan; Suriname; Swaziland; Sweden; Switzerland; Syria; Taiwan, China; Tajikistan; Thailand; The Bahamas; Timor-Leste; Togo; Tokelau; Tonga; Trinidad and Tobago; Tunisia; Turkey; Turkmenistan; Turks and Caicos Islands; Tuvalu; Uganda; Ukraine; United Arab Emirates; United Kingdom; United Republic of Tanzania; United States of America; Uruguay; Uzbekistan; Vanuatu; Venezuela; Vietnam; Wallis and Futuna; West Bank; Western Sahara; Yemen; Yugoslavia; Zaire; Zambia; Zimbabwe (225)
<i>Bos taurus</i>	Afghanistan; Albania; Algeria; Angola; Argentina; Armenia; Australia; Austria; Azerbaijan; Bahrain; Bangladesh; Belarus; Belgium; Benin; Bhutan; Bolivia; Bosnia and Herzegovina; Botswana; Brazil; Bulgaria; Burkina Faso; Cambodia; Cameroon; Canada; Chad; Chile; China; Colombia; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Cyprus; Denmark; Ecuador; Egypt; El Salvador; Eritrea; Estonia; Ethiopia; France; French Guiana; Gaza Strip; Georgia; Germany; Ghana; Greece; Guatemala; Guinea; Honduras; Hong Kong, China; Hungary; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Japan; Jordan; Kazakhstan; Kenya; Kosovo; Kuwait; Kyrgyzstan; Laos; Latvia; Lebanon; Lesotho; Libya; Liechtenstein; Lithuania; Luxembourg; Madagascar; Malawi; Malaysia; Mali; Martinique; Mauritania; Mauritius; Mayotte; Mexico; Moldova; Mongolia; Montenegro; Morocco; Mozambique; Myanmar; Namibia; Nepal; Netherlands; New Zealand; Niger; Nigeria; Democratic People's Republic of Korea; Norway; Oman; Pakistan; Palestine; Panama; Papua New Guinea; Paraguay; Peru; Poland; Portugal; Republic of Serbia; Romania; Russia; Rwanda; San Marino; Saudi Arabia; Senegal; Slovakia; Slovenia; Somalia; South Africa; Republic of Korea; Spain; Sri Lanka; Sudan; Sweden; Switzerland; Syria; Taiwan, China; Tajikistan; Thailand; Togo; Trinidad and Tobago; Tunisia; Turkey; Uganda; Ukraine; United Arab Emirates; United Kingdom; United Republic of Tanzania; United States of America; Uruguay; Venezuela; Vietnam; Yemen; Zambia; Zimbabwe (143)

TABLE S1. (Continued)

Virus species	Distributed countries/regions
<i>Canis lupus familiaris</i>	Afghanistan; Albania; Algeria; Argentina; Australia; Austria; Azerbaijan; Bangladesh; Belgium; Benin; Bhutan; Bosnia and Herzegovina; Botswana; Brazil; Bulgaria; Burkina Faso; Cambodia; Cameroon; Canada; Central African Republic; Chad; Chile; China; Colombia; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Czech Republic; Democratic Republic of the Congo; Dominican Republic; Ecuador; Egypt; Estonia; Ethiopia; Finland; France; French Guiana; Gabon; Gambia; Georgia; Germany; Ghana; Greece; Greenland; Grenada; Guatemala; Guinea; Guyana; Haiti; Honduras; Hong Kong, China; Hungary; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Japan; Kazakhstan; Kenya; Laos; Latvia; Lesotho; Liberia; Lithuania; Luxembourg; Madagascar; Malaysia; Mali; Mauritania; Mexico; Mongolia; Morocco; Mozambique; Myanmar; Namibia; Nepal; Netherlands; New Zealand; Nicaragua; Niger; Nigeria; Oman; Pakistan; Paraguay; Peru; Philippines; Poland; Portugal; Republic of Serbia; Romania; Russia; Rwanda; Saint Kitts and Nevis; Senegal; Slovenia; Somalia; South Africa; Republic of Korea; Spain; Sri Lanka; Sudan; Swaziland; Switzerland; Taiwan, China; Tajikistan; Thailand; Tunisia; Turkey; Uganda; Ukraine; United Arab Emirates; United Kingdom; United Republic of Tanzania; United States of America; Uruguay; Vietnam; Zaire; Zambia; Zimbabwe (123)
<i>Gallus gallus</i>	Afghanistan; Algeria; Argentina; Australia; Austria; Azerbaijan; Bangladesh; Belgium; Belize; Benin; Bhutan; Botswana; Brazil; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Chile; China; Colombia; Costa Rica; Cote d'Ivoire; Croatia; Czech Republic; Democratic Republic of the Congo; Denmark; Dominican Republic; Ecuador; Egypt; El Salvador; Ethiopia; Fiji; Finland; France; Georgia; Germany; Ghana; Greece; Grenada; Guatemala; Honduras; Hong Kong, China; Hungary; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Japan; Jordan; Kazakhstan; Kenya; Kuwait; Kyrgyzstan; Laos; Latvia; Lebanon; Libya; Luxembourg; Macedonia; Madagascar; Malaysia; Mali; Mexico; Morocco; Mozambique; Myanmar; Namibia; Nepal; Netherlands; New Zealand; Nicaragua; Niger; Nigeria; Democratic People's Republic of Korea; Oman; Pakistan; Peru; Philippines; Poland; Portugal; Republic of Serbia; Romania; Russia; Saudi Arabia; Senegal; Singapore; Slovakia; Slovenia; South Africa; Republic of Korea; Spain; Sri Lanka; Sudan; Sweden; Switzerland; Syria; Taiwan, China; Tajikistan; Thailand; Timor-Leste; Togo; Trinidad and Tobago; Tunisia; Turkey; Uganda; Ukraine; United Arab Emirates; United Kingdom; United Republic of Tanzania; United States of America; Uruguay; Venezuela; Vietnam; West Bank; Yemen; Zambia; Zimbabwe (122)
<i>Sus scrofa</i>	Afghanistan; Albania; Angola; Argentina; Armenia; Australia; Austria; Bangladesh; Belarus; Belgium; Benin; Bhutan; Bolivia; Brazil; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Chile; China; Colombia; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Czech Republic; Democratic Republic of the Congo; Denmark; Dominican Republic; Egypt; Estonia; Ethiopia; Finland; France; Gambia; Germany; Ghana; Greece; Guatemala; Haiti; Hong Kong, China; Hungary; India; Indonesia; Ireland; Israel; Italy; Japan; Kazakhstan; Kenya; Laos; Latvia; Lithuania; Luxembourg; Madagascar; Malawi; Malaysia; Mali; Malta; Mauritius; Mexico; Mongolia; Mozambique; Myanmar; Namibia; Nepal; Netherlands; New Caledonia; New Zealand; Nicaragua; Nigeria; Democratic People's Republic of Korea; Norway; Paraguay; Peru; Philippines; Poland; Portugal; Republic of Serbia; Romania; Russia; Saint Kitts and Nevis; Sao Tome and Principe; Senegal; Singapore; Slovakia; Slovenia; South Africa; Republic of Korea; Spain; Sri Lanka; Sweden; Switzerland; Taiwan, China; Thailand; Togo; Turkey; Uganda; Ukraine; United Kingdom; United Republic of Tanzania; United States of America; Uruguay; Venezuela; Vietnam; Zaire; Zambia; Zimbabwe (110)

SUPPLEMENTARY TABLE S2. Characteristics of first reported viruses.

Virus species	Type	Why report	First reported time	First reported location	First reported host
Avipoxvirus	Animal virus	First animal virus reported in database	1865	Peru	Leptotila rufaxilla
Cotton leaf curl Multan virus	Plant virus	First plant virus reported in database	1918	China	Hibiscus rosa-sinensis
Saccharomyces paradoxus virus	Microbe virus	First microbe virus reported in database	1931	Russia	Saccharomyces paradoxus
Influenza A virus	Animal virus	The most widely distributed virus in database	1902	Italy	Gallus gallus and unspecified Aves
SARSr-Cov	Animal virus	The most submitted virus in database	2003	China	Homo sapiens
Aeromonas phage PZL-Ah8	Microbe virus	Latest new virus reported in database	2021	China	Aeromonas hydrophila
Vaccinia virus	Animal virus	The earliest reported human-animal co-virus	1866	the United States	Homo sapiens

SUPPLEMENTARY TABLE S3. New viruses reported times in each country or region.

Country/region	No. of report times
United States of America	6628
China	4518
Brazil	1032
Australia	994
Italy	730
Germany	590
Spain	552
India	529
United Kingdom	424
Japan	379
France	304
Canada	300
Republic of Korea	297
New Zealand	291
Denmark	264
Mexico	248
South Africa	230
Hungary	220
Russia	212
Bangladesh	207
Kenya	198
Sweden	195
Thailand	167
Zambia	162
Cameroon	158
Antarctica	157
Netherlands	151
Finland	143
Ghana	137
Taiwan, China	133
Argentina	122
Iran	119
Pakistan	119
Portugal	108
Czech Republic	107
Nigeria	97
Austria	91
Switzerland	86
Colombia	85
Vietnam	83
Turkey	79
Uganda	77

Continued

Country/region	No. of report times
Poland	75
Cote d'Ivoire	74
Belgium	73
Tunisia	71
Malaysia	70
Indonesia	69
Hong Kong, China	65
Philippines	63
Gabon	62
Bulgaria	57
Bermuda	56
Democratic Republic of the Congo	55
Ireland	54
Madagascar	54
Peru	53
Tonga	51
Norway	49
United Republic of Tanzania	49
Israel	48
Panama	47
Egypt	45
Venezuela	41
Chile	39
Greece	37
Costa Rica	36
Central African Republic	34
Cuba	34
Slovenia	34
French Guiana	29
Senegal	29
Ecuador	27
Kuwait	27
Liberia	27
Rwanda	27
Nepal	25
South Georgia and the Islands	25
Trinidad and Tobago	25
United Arab Emirates	24
Croatia	22
Malawi	22
Saudi Arabia	22
Republic of Serbia	21
Georgia	20

Continued

Country/region	No. of report times
Papua New Guinea	20
Bolivia	19
Cambodia	19
Guadeloupe	19
Burkina Faso	18
Lithuania	18
Reunion	18
Myanmar	17
Ethiopia	16
Mauritius	16
Sudan	16
Pacific Ocean	14
Uruguay	14
Mozambique	12
Singapore	12
Ukraine	12
Azerbaijan	11
Iraq	11
Morocco	11
Vanuatu	11
Guatemala	10
Saint Kitts and Nevis	10
Gambia	9
Iceland	9
Oman	9
Sri Lanka	9
Benin	8
Guinea	8
Slovakia	8
Laos	7
Namibia	7
Saint Vincent and the Grenadines	7
Afghanistan	6
Algeria	6
Arctic Ocean	6
Dominican Republic	6
Nicaragua	6
The Bahamas	6
Belarus	5
Honduras	5
Jamaica	5
Jordan	5

Continued

Country/region	No. of report times
Mongolia	5
Romania	5
Yemen	5
Albania	4
Estonia	4
Kazakhstan	4
Kyrgyzstan	4
Mali	4
Mayotte	4
Solomon Islands	4
Cayman Islands	3
Latvia	3
Paraguay	3
Samoa	3
South Sudan	3
Syria	3
Togo	3
Turkmenistan	3
Zimbabwe	3
Angola	2
Armenia	2
Bhutan	2
Bosnia and Herzegovina	2
Christmas Island	2
Comoros	2
Cyprus	2
Czechoslovakia	2
Indian Ocean Territories	2
Lebanon	2
Libya	2
Moldova	2
Niger	2
Saint Barthelemy	2
Seychelles	2
Sierra Leone	2
Barbados	1
Belize	1
Botswana	1
Chad	1
East Timor	1
Faroe Islands	1
Fiji	1
Grenada	1

Continued

Country/region	No. of report times
Haiti	1
Kiribati	1
Mediterranean Sea	1
Midway Islands	1
Montenegro	1
New Caledonia	1
Somalia	1
Suriname	1
Uzbekistan	1
Yugoslavia	1

SUPPLEMENTARY TABLE S4. Co-existed viruses host family and tags.

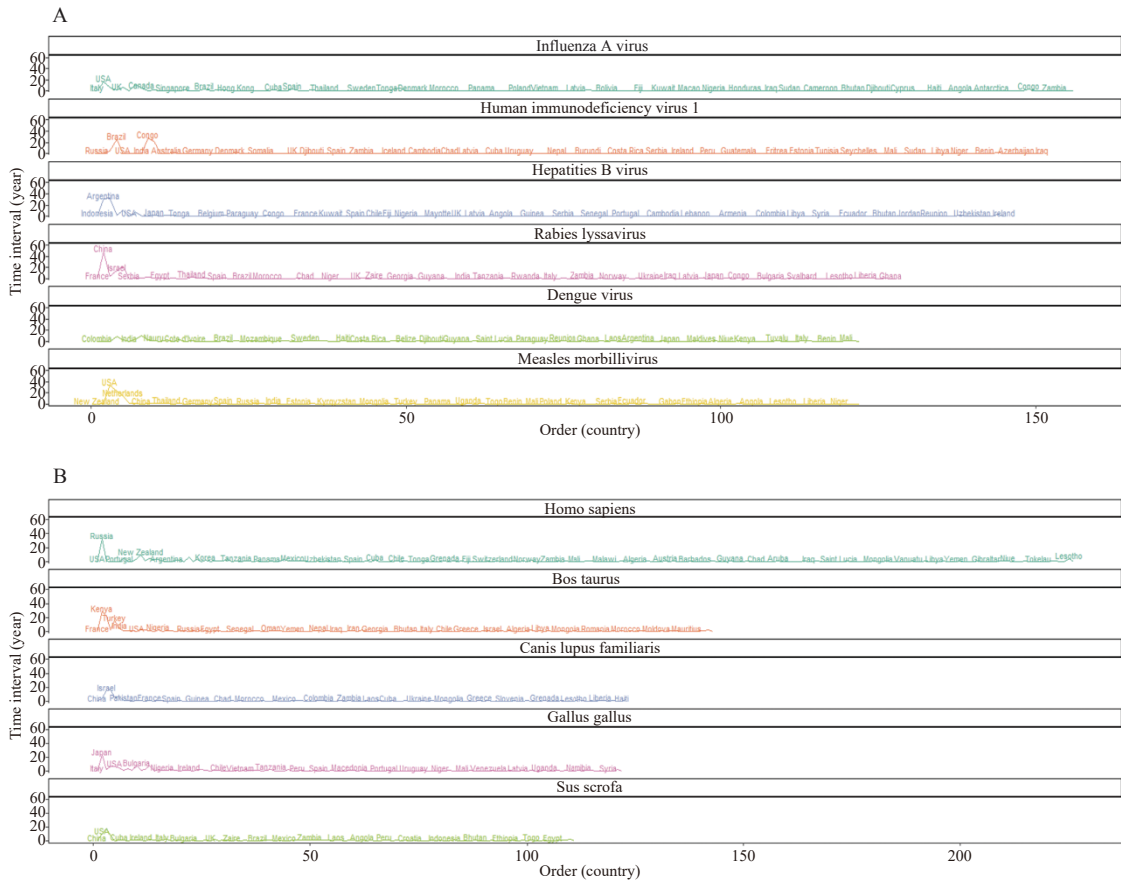
Host family	Tag
Aleyrodidae	H2
Alstroemeriaceae	H3
Amaranthaceae	H4
Amaryllidaceae	H5
Anatidae	H6
Animalia	H7
Aphididae	H8
Apiaceae	H9
Apidae	H10
Apocynaceae	H11
Araceae	H12
Araneae	H13
Arecaceae	H14
Asparagaceae	H15
Asteraceae	H16
Bacteria	H17
Balsaminaceae	H18
Begoniaceae	H19
Bifidobacteriaceae	H20
Botryosphaeriaceae	H21
Brassicaceae	H22
Bovidae	H23
Cactaceae	H24
Calceolariaceae	H25
Campanulaceae	H26
Cannaceae	H27
Caprifoliaceae	H28
Caryophyllaceae	H29
Celastraceae	H30
Ceratobasidiaceae	H31

Continued

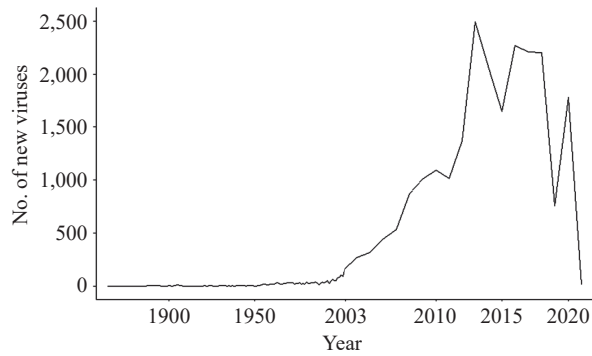
Host family	Tag
Cervidae	H32
Chiroptera	H33
Cicadellidae	H34
Coccoidea	H35
Commelinaceae	H36
Convolvulaceae	H37
Cucurbitaceae	H38
Culicidae	H39
Curculionidae	H40
Delphacidae	H41
Diptera	H42
Drosophilidae	H43
Eimeriidae	H44
Enterobacteriaceae	H45
Enterococcaceae	H46
Entomophthoraceae	H47
Fabaceae	H48
Firmicutes	H49
Gentianaceae	H50
Geraniaceae	H51
Gesneriaceae	H52
Gruidae	H53
Hexapoda	H54
Homo sapiens	H55
Iridaceae	H56
Lactobacillaceae	H57
Lamiaceae	H58
Laniidae	H59
Leporidae	H60
Liliaceae	H61
Lutjanidae	H62
Malvaceae	H63
Marantaceae	H64
Microcystaceae	H65
Moringaceae	H66
Muscicapidae	H67
Mycosphaerellaceae	H68
Nymphalidae	H69
Oleaceae	H70
Orchidaceae	H71
Pedaliaceae	H72
Peronosporaceae	H73
Phasianidae	H74

Continued

Host family	Tag
Phoenicopteridae	H75
Phthiracaridae	H76
Piperaceae	H77
Pittosporaceae	H78
Plantaginaceae	H79
Pleosporaceae	H80
Poaceae	H81
Primulaceae	H82
Prunellidae	H83
Pseudomonadaceae	H84
Ranunculaceae	H85
Rosaceae	H86
Rubiaceae	H87
Salmonidae	H88
Sclerotiniaceae	H89
Solanaceae	H90
Spermatophyta	H91
Sporormiaceae	H92
Staphylococcaceae	H93
Streptococcaceae	H94
Suidae	H95
Tetranychidae	H96
Thelephoraceae	H97
Thripidae	H98
Thysanoptera	H99
Tropaeolaceae	H100
Turdidae	H101
Viverridae	H102
Zygophyllaceae	H103



SUPPLEMENTARY FIGURE S1. Top 5 widely distributed (A) viruses and (B) hosts sorted by first report country or region.



SUPPLEMENTARY FIGURE S2. The trend of new reported viruses sorted by year.