

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Journal of Genetics and Genomics 44 (2017) 259-264

Contents lists available at ScienceDirect

Journal of Genetics and Genomics

Journal homepage: www.journals.elsevier.com/journal-of-geneticsand-genomics/

Original research

DRodVir: A resource for exploring the virome diversity in rodents

Lihong Chen, Bo Liu, Zhiqiang Wu, Qi Jin^{*}, Jian Yang^{*}

MOH Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100176, China

ARTICLE INFO

Article history: Received 24 November 2016 Received in revised form 27 April 2017 Accepted 30 April 2017 Available online 3 May 2017

Keywords: Rodents Viruses Database Zoonoses Infectious diseases

ABSTRACT

Emerging zoonotic diseases have received tremendous interests in recent years, as they pose a significant threat to human health, animal welfare, and economic stability. A high proportion of zoonoses originate from wildlife reservoirs. Rodents are the most numerous, widespread, and diverse group of mammals on the earth and are reservoirs for many zoonotic viruses responsible for significant morbidity and mortality. A better understanding of virome diversity in rodents would be of importance for researchers and professionals in the field. Therefore, we developed the DRodVir database (http://www.mgc.ac.cn/ DRodVir/), a comprehensive, up-to-date, and well-curated repository of rodent-associated animal viruses. The database currently covers 7690 sequences from 5491 rodent-associated mammal viruses of 26 viral families detected from 194 rodent species in 93 countries worldwide. In addition to virus sequences, the database provides detailed information on related samples and host rodents, as well as a set of online analytical tools for text query, BLAST search and phylogenetic reconstruction. The DRodVir database will help virologists better understand the virome diversity of rodents. Moreover, it will be a valuable tool for epidemiologists and zoologists for easy monitoring and tracking of the current and future zoonotic diseases. As a data application example, we further compared the current status of rodent-associated viruses with bat-associated viruses to highlight the necessity for including additional host species and geographic regions in future investigations, which will help us achieve a better understanding of the virome diversities in the two major reservoirs of emerging zoonotic infectious diseases.

Copyright © 2017, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. All rights reserved.

1. Introduction

Zoonotic diseases comprise a significant and increasing proportion of all emerging human infectious diseases, and most of them originate from wildlife (Kruse et al., 2004; Jones et al., 2008). Two mammalian orders, *Rodentia* (rodents) and *Chiroptera* (bats), represent the most relevant potential sources of new zoonoses (Calisher et al., 2006; Meerburg et al., 2009; Smith and Wang, 2013). The *Rodentia* is the single largest order of mammals. There are more than 2000 living species of rodents, which comprise ~40% of all mammalian species, including mice, rats, hamsters, guinea pigs, voles, chinchillas, chipmunks, gophers, muskrats, gerbils, woodchucks, and many others (Huchon et al., 2002). Despite their great species diversity, all rodents share a common feature – a single pair of continuously growing incisors, which are used to gnaw food, excavate burrows, and defend themselves. Rodents are

* Corresponding authors. E-mail addresses: zdsys@vip.sina.com (Q. Jin), yangj@ipbcams.ac.cn (J. Yang). native on all continents except Antarctica, and they show a wide range of lifestyles, ranging from terrestrial, subterranean, arboreal, to aquatic habitats (Wolff and Sherman, 2007).

Rodents are common hosts for pathogens that transmit diseases to humans and domestic animals because of their properties of close association with humans, large social group size, intense social interaction, high population density and widespread geographic distribution. Since the Middle Ages, rodents have been known to contribute to human diseases, as black rats were associated with the spread of the plague (Perry and Fetherston, 1997). However, rodents are also a threat to public health in modern times. They have been implicated as reservoir hosts of zoonotic pathogens, such as viral hemorrhagic fever viruses, including arenaviruses (Junin, Machupo, and Lassa) and hantaviruses (Hantaan, Dobrava, and Sin Nombre) (Enria and Pinheiro, 2000; Goeijenbier et al., 2013). Also, global climate change and continued urbanization have led to increased problems with rodent-associated zoonoses (Mills et al., 2010). Recent metagenomic studies have identified a wide diversity of novel viruses in rodents from families/







http://dx.doi.org/10.1016/j.jgg.2017.04.004

^{1673-8527/}Copyright © 2017, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. All rights reserved.

genera that contain important human pathogens, including new species/variants of picornaviruses, hepaciviruses and sapoviruses (Firth et al., 2014). Therefore, it is necessary to enhance our knowledge of the virome diversity in rodents to facilitate future prevention and control of emerging zoonotic diseases.

Bats are the second most diverse mammalian orders on the earth, with more than 1200 living species identified worldwide (Altringham, 2011). The importance of bats as natural hosts for several important viral agents, including Ebola virus, Marburg virus, Hendra virus, Nipah virus, Severe Acute Respiratory Syndrome (SARS) coronavirus, and Middle East Respiratory Syndrome (MERS) coronavirus, has been established (Calisher et al., 2006; Wong et al., 2007; Ge et al., 2013; Smith and Wang, 2013; Munster et al., 2016). Due to the wide-scale microbial surveillance programs conducted and novel high-throughput detection methods developed in recent years, there are a plethora of newly identified viruses in both rodents and bats. Our group already constructed the first database of bat-associated viruses (DBatVir) in 2014 (Chen et al., 2014). In this study, we present a newly established DRodVir database for a thorough and modern understanding of rodent-associated viral zoonoses (http://www.mgc.ac.cn/DRodVir/). Moreover, we used the information from two sister databases to perform a comparative analysis on the virome diversities of the two major reservoirs of emerging zoonotic infectious diseases.

2. Database construction and description

2.1. Database construction

DRodVir is a sequence-centric database, since molecular methods are now commonly used in virus detection and functional analysis. To retrieve all available sequences of rodent-associated viruses from the public domain, we first performed exhaustive searches in both the PubMed and Nucleotide databases of NCBI using a group of keywords pertaining to rodents. Then, the retrieved GenBank records were downloaded to a local system and parsed by an in-house BioPerl script to generate human-readable tables of the metadata, including the sampling time, location, rodent species, specimen type (e.g., feces, blood, or tissues), and viral detection method (e.g., PCR or metagenomics), for further review. For published sequences, additional meta information was extracted from the related literature using manual curation. As the DRodVir database focuses on natural rodent-associated mammal viruses that are most relevant to emerging zoonotic infectious diseases, the following filter criteria are employed to exclude unrelated records: i) samples derived from animals other than rodents; ii) sequences of phages; iii) insect or plant viruses; iv) laboratory-cultured viruses in model rodents. Taxonomic information of all viruses and rodents was derived from the Taxonomy database of NCBI. In addition, we collected information concerning the viruses (e.g., genome architecture, average size and virus graphics) and the rodents (e.g., common names and known geographic ranges) from the ViralZone databases (Masson et al., 2013) and the ICUN Red List (www.iucnredlist.org), respectively. Moreover, the established phylogenetic relationships between different rodent families from a previous study were carefully incorporated into the database (Meredith et al., 2011).

The DRodVir database reused the background data schema and the majority of the foreground Perl CGI and JavaScript codes of the DBatVir database because they are efficient to provide versatile and user-friendly interface (Chen et al., 2014). Given that many users of DRodVir and DBatVir databases may overlap, the similar architecture makes the new database instantly familiar with previous users. In addition, the two sister databases with similar web design and analytical functions will make it very easy to perform any comparative analysis on bat- and rodent-associated viruses for all users. The NCBI BLAST and MUSCLE programs were integrated into the database to allow users to conduct sequence similarity searches and multiple sequence alignment on the web page (Altschul et al., 1997; Edgar, 2004). The FastTree program and jsPhyloSVG library were used for online phylogenetic tree construction and visualization, respectively (Price et al., 2010; Smits and Ouverney, 2010).

2.2. Database description

The main page of DRodVir database provides a highly responsive and intuitive user interface, with a look similar to a desktop application rather than a traditional website. On the left side is a multifunctional menu panel for easy navigation, and on the right side is a tabbed content panel for presenting tables, figures, trees, BLAST outputs, and other results. To maximize the visible section of the content panel for users with limited screen size, the menu panel can be collapsed into a clickable vertical bar automatically (depends on visitors' screen resolution setting) or manually (by single click the icon on the top right corner) (Fig. 1A). Features that were previously available only in standalone applications, including collapsible menus, expandable trees, sortable grids, tabbed panels and live statistical pie charts, provide high performance and an improved user experience.

DRodVir inherits the major features from the DBatVir database (Chen et al., 2014). Users can either browse or search the database contents. In browse mode, the menu panel offers three submenus in accordion style for users to browse the database by categories of viruses/rodents or geographic regions. Users who are familiar with certain viruses can use the 'browse by virus' submenu to explore the current host range in rodents and their geographic distributions (Fig. 1B), whereas those who are interested in given types of rodents may try the 'browse by rodent' submenu to investigate the known virome diversity in different rodentia hosts (Fig. 1C). In addition, the 'browse by region' submenu allows the users to easily summarize the current research efforts conducted on rodentassociated viruses for any countries or continents (Fig. 1D). All three submenus are organized as hierarchical and expandable trees and can be easily switched on/off by single click on the title bar. Each branch of the trees in the submenus provides a direct link to individual tab in the main content panel, which presents a uniform sortable grid that includes brief information of the viruses, specimens, associated rodents, determined sequences and related literature (Fig. 1A). Each line of the table can be further expanded by double click to show additional information, such as virus detection method, GenBank accession number(s), sequence submission date, submitters and their affiliations. A clickable linear map is available for each full-length viral sequence to highlight the genome architecture of the virus (Fig. 1A). For brevity, the table provides up to 100 records by default. Handy toolbar is available for users to page up/down if more records exist, or alternatively, turn off paging to force to display all records within a single table. In addition, the main table offers various manipulating functions, i.e., sorting/ filtering the table by column contents, reordering the columns, adjusting column width or hiding some columns.

In search mode, two parallel query forms are provided to the users for text information search and BLAST sequence similarity search, respectively (Fig. 2A). The text search engine allows the extraction of virus/rodent/specimen/sequence information from the database using either a simple query for a quick start or combined search patterns for advanced usages. The search results are organized into an individual high-performance grid in the content panel as aforementioned. To facilitate online data analysis, two visualization tools are integrated into the result table: i) a statistical pie chart is available with a single click on the column title of virus

| | 🛚 Viral families 🌸 🐁 Browse rodents 🌸 🚱 Global map 🌸 | Occument | 🔍 Database search 🛎 🛃 Statistics | × Rodents phylogeny | × 🔄 Paramyxovirio | dae virus 🛎 | | | |
|---|---|---|--|---|---|--|--|---|--|
| | Viruses 🔺 | Viral family | From rodent | Rodent family | Sample type | Collection year | Sampling country | Sequence coding | References |
| ŧ | Beilong virus | Paramyxoviridae | Rattus norvegicus | Muridae | Culture | ND | ND | Full length | Virology 2006, 346(1):219-28 |
| Ð | Beilong virus ERN081008_1S | Paramyxoviridae | Rattus norvegicus | Muridae | Tissue | 2008 | China | Full length | Infect Genet Evol 2016, 45:311-319 |
| ŧ | Beilong virus HKIARN060309_1K | Paramyxoviridae | Rattus norvegicus | Muridae | Tissue | 2009 | China | Full length | Infect Genet Evol 2016, 45:311-319 |
| | Beilong virus YLRR120908_1K | Paramyxoviridae | Rattus rattus | Muridae | Tissue | 2008 | China | Full length | Infect Genet Evol 2016, 45:311-319 |
| | Beilong virus YLRR120908_3S | Paramyxoviridae | Rattus rattus | Muridae | Tissue | 2008 | China | Full length | Infect Genet Evol 2016, 45:311-319 |
| | Guinea pig parainfluenza virus TS-9 CavPIV-3 | Paramyxoviridae | Cavia porcellus | Caviidae | Culture | 1989 | United Kingdom | ND | Comp Med 2002, 52(6):548-54 |
| | J-virus JPV | Paramyxoviridae | Mus musculus | Muridae | Culture | 1972 | Australia | Full length | J Virol 2005, 79(16): 10690-700 |
| 9 | Mojiang virus Tongguan1 | Paramyxoviridae | Rattus rattus | Muridae | Swabs | 2012 | China | Full length | Emerg Intect Dis 2014, 20(6): 1064-6 |
| | Detection method: Metagenomics/HTS (high-throughput Sample collection date: 2012-12-02 | (bequerienig) | | | | | | | |
| | | , Sun,L., Zhu,Y. and | | cademy of Medical Science | es & Pelving Union Me | dical College, No. 6, Ron | gjing East Street, Yizhuang, Be | ijing 100176, China | |
| | Sample collection date: 2012-12-02 Sampling location: China Determined sequence: Submit date: 2010-2014 Submitted from: MOH Key Laboratory of Systems Biol Accession: #7225832 (RHSset; UC 205322) Length: 18406 tp (consider sequence) | , Sun,L., Zhu,Y. and igy of Pathogens, In | G 10000 | L | | | gjing East Street, Yuzhuang, Be | iying 100176. China | |
| | Sample collection date: 2012-12-02 Sampling location: China Determined sequence: Submited services (Sample Characterian Sampling) Submited pr. Wu Z., Yang, F., Ren X., Dong, J., Submited pr. Wu Z., Yang, J., Hen X., Dong, J., Submitted pr. Wu Z., Yang, K., Karakara, Sample Characterian Accession: (<u>ECTRISS</u>) (RefSeq: <u>ICC.05352</u>) Length: 10406 bp (<i>Complete Sample</i>) V P M KF278639 5000 Related publications: | , Sun,L., Zhu,Y. and igy of Pathogens, In | G 10000 | L | | | gjing East Street, Yizhuang, Be | ijing 100176, China | Vroboy 2003, 317 (2) 330-44 |
| 9 | Sample collection date: 2012-12-02 Sampling location: China Determined sequence: Submit dise: 2013-62-20 Submit dise: 2013-62-20 Submit dise: 2013-62-20 Submit dise: 2013-62-20 Length: 18406 bp (consider sequence) VP VP M KF2278639 5000 Related publications: Wu Z, Yang L, Yang F, Ren X, Jang J, Dong J, Sun L, Zh Mossman virus MoV_Ratus | , Sun,L., Zhu,Y. and gy of Pathogens, In F U Y, Zhou H, Jin Q, Paramyaowindae | G 10000 Novel Henipa-like virus, Mojiang Param Rattus novegicus | L 15000 yxovirus, in rats, China, 20 <i>Muridae</i> | 112. Emerg Intect Dis : Culture | 2014, 20(6):1064-6: 🚭 ND | Australia | <u>Full length</u> | Wrolsgy 2003, 317(2):330-44 |
| 9 | Sample collection date: 2012-12-02 Sampling location: China Determined sequence: Submit dise: 2013-62-20 Submit dise: 2013-62-20 Submit dise: 2013-62-20 Submit dise: 2013-62-20 Length: 18406 bp (consider sequence) VP VP M KF2278639 5000 Related publications: Wu Z, Yang L, Yang F, Ren X, Jang J, Dong J, Sun L, Zh Mossman virus MoV_Ratus | Sun,L., Zhu,Y. and gy of Pathogens, In F u Y, Zhou H, Jin Q, I Paramyoovindee Paramyoovindee | G G 10000 Novel Heripa-like virus, Mojiang Param Ratius novegicus Mus musculus | L 15000 yxovirus, in rats, China, 20 | 112. Emerg Infect Dis | 2014, 20(6): 1064-6, 🍄 ND 1980 | Australia United States of America | Eull length NS2, N, P, M, SH, G | Unpublished |
| 9 | Sample collection date: 2012-12-02 Sampling location: China Determined sequence: Submitted by: Wu Z, Vang, F, Ren X, Dong, J, Submitted by: Wu Z, Vang, F, Ren X, Dong, J, Submitted from (MCH Key, Laboratory of Systems Biolo Accession: <u>100710000</u> (Reflect: <u>10070300</u>) Length: 18405 by (Contactor accession) C V V V V V V V V V V | , Sun JL., Zhu Y. and gy of Pathogens, In F V Zhou H, Jin Q, Paramyowiridae Paramyowiridae | G G 10000 Novel Henipa-like virus, Majiang Param Rattus norvegicus Mus musculus Ziygodontomys brevicauda | L 15000 Wurklae Murklae | 112. Emerg Intect Dis 1 Culture Tissue | 2014, 20(6):1064-6: 🚭 ND | Australia United States of America Trinidad and Tobago | Full Length NS2, N, P, M, SH, G Full Length | Unpublished Arch Virol 2009, 154(2):199-207 |
| 9 | Sample collection date: 2012-12-02 Samples collection date: 2012-12-02 Sampling location: China Determined sequence: Submit date: 2013-02-00 Submit date: 2013-02-00 Submit date: 2013-02-00 Submit date: 2013-02-00 Submit date: 2013-02-00 Submit date: 2013-02-00 Submit date: 2013-02-00 PVP Market Sequence: VPVP Market Sequence: VP | , Sun,L., Zhu,Y. and igy of Pathogens, In F V, Zhou H, Jin Q, I Paramysoviridae Paramysoviridae Paramysoviridae | G C C C C C C C C C C C C C C C C C C C | L 15000 Muridae Muridae Cricetidae | 112. Emerg Intect Dis Culture Tissue Culture | 2014, 20(6):1064-6. 🍄 ND 1980 1962 | Australia United States of America Trinidad and Tobago United Kingdom | Eull length NS2, N, P, M, SH, G | Unpublished Arch Virol 2009, 154 (2):199-207 Vet Microbiol 2014, 170 (3-4):391-7 |
| 999999999999999999999999999999999999999 | Sample collection date: 2012-12-02 Samples collection date: 2012-12-02 Submit dec 2013-02-02 Submit dec 2013-02-02 Submit dec 2013-02-02 Submit de Dy: WU-2, Yang F, RenX, Dong J, Submit de Dy: WU-2, Yang F, RenX, Dong J, Length: (8405 hp (concelete sequence) V P V P M V C 2013 (Concelete sequence) V V Z, Yang L, Yang F, RenX, Jang J, Dong J, Sun L, Zh Mastran vrus MoV, Rattus Murine pneumonia vrus Y Narta vrus NAFY Paramysovirus A69-1 PenPv2 Paramysovirus A69-1 PenPv3 | , Sun JL., Zhu Y. and gy of Pathogens, In F V Zhou H, Jin Q, Paramyowiridae Paramyowiridae | G G 10000 Novel Henipa-like virus, Majiang Param Rattus norvegicus Mus musculus Ziygodontomys brevicauda | J5000 yuovirus, in rats, China, 20 Muridae Critedidae Sciuridae | D12. Emerg Intect Dis : Culture Tissue Culture Culture Culture | 2014, 20(6): 1064-6. 🍄 ND 1980 1962 2009 | Australia United States of America Trinidad and Tobago | EuilLength NS2, IV, F, M, SH, G EuilLength L | Unpublished Arch Virol 2009, 154(2):199-207 |

B

| Main menu | « | Main menu | * | | Main menu | |
|--|----------|---|--------------|---|--------------------------------|--------------|
| Browse by virus | - | C 🎂 Browse by virus | + | D | Browse by virus | • |
| Viruses 🔺 | Number | 1 Browse by rodent | - | | & Browse by rodent | |
| 🖌 🞝 dsDNA viruses, no RNA stage | 401 | Rodents + | # of viruses | | Browse by region | |
| Adenoviridae | 37 | - Chystricognathi | 16 | | Territory + | # of viruses |
| Description >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> | 181 | Caviidae | 5 | | ▷ ➡ AFRICA | 529 |
| Papillomaviridae | 19 | ▷ Dasyproctidae | 1 | | | 1853 |
| Polyomaviridae | 30 | ▷ Echimvidae | 4 | | Afghanistan | 1055 |
| Poxviridae | 134 | ▷ Erethizontidae | 1 | | China | 1136 |
| dsRNA viruses | 26 | | 5 | | Indonesia | |
| Picobirnaviridae | 6 | ▲ U Sciurognathi | 5462 | | indonesia | 132 |
| Reoviridae | 20 | Castoridae | 3 | | 🐑 📥 India | 3 |
| Retro-transcribing viruses | 18 | ▲ ↓ Cricetidae | 2262 | | Japan | 1 |
| Hepadnaviridae | 14 | Arvicolinae | 1411 | | | 119 |
| ▶ ⇒Retroviridae | 4 | ▷ Arvicolniae | 1411 | | Cambodia Cambodia Sorea, North | 20 |
| SsDNA viruses | 152 | ▷ Eothenomys | 57 | | | 17 |
| > Anelloviridae | 33 | ▷ ➡ Losiopodomys | 1 | | 🐚 🍽 Korea, South | 160 |
| > Circoviridae | 30 | ▷ Lennus | 2 | | Laos | 12 |
| Parvoviridae | 89 | ▷ ≦ ∠errindus | 402 | | Myanmar | 1 |
| SsRNA negative-strand viruses | 3922 | ▷ → Myodes | 923 | | | 21 |
| > Arenaviridae | 730 | Dependences | 13 | | Malaysia | 11 |
| ⊳ ⊜ Bornaviridae | 12 | Volemys | 2 | | | 4 |
| b Bunyaviridae | 2968 | ▷ Cricetinae | 100 | | Thailand | 38 |
| Drthomyxoviridae | 2 | Neotominae | 297 | | 🐚 🛄 Turkmenistan | 1 |
| Paramyxoviridae | 190 | ▷ ⇒ Sigmodontinae | 454 | | 🐑 🚾 Turkey | 32 |
| ▶ 🖨 Rhabdoviridae | 20 | ▷ □ Signobonunae ▷ □ Dipodidae | 454 | | Taiwan, Province of China | 3 |
| SsRNA positive-strand viruses, no DNA stage | 1192 | ▷ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ | 3 | | 🧐 💶 Vietnam | 136 |
| Arteriviridae | 11 | ▷ □→ Gindae | 6 | | | 1918 |
| ▷ ⇒ Astroviridae | 99 | ↓ ↓ Muridae | 3052 | | NORTH AMERICA | 788 |
| Caliciviridae | 121 | ► Deomvinae | 6 | | | 38 |
| Coronaviridae | 104 | Debriyinae | 7 | | ▲↓ SOUTH AMERICA | 496 |
| Flaviviridae | 212 | ▷ → Gerbininae ▷ → Murinae | 3039 | | 🔄 🔤 Argentina | 79 |
| ▶ ⇒ Hepeviridae | 212 | ▷ ▷ Nesomyidae | | | 🐑 🚾 Bolivia | 13 |
| ▷ Picornaviridae | 317 | Sciuridae | 16 | | 🐑 💽 Brazil | 161 |
| ▷ Dogaviridae | 116 | | 90 | | 🐑 🏪 Chile | 42 |
| andassified viruses | 18 | Spalacidae Junclassified Rodentia | 24 | | 🐑 💳 Colombia | 26 |
| - | | sta undassined Kodentia | 251 | | 🐑 🚔 Ecuador | 4 |
| Browse by rodent | + | | | | 🐑 陆 French Guiana | 9 |
| Browse by region | + | Browse by region | + | | 🐑 📭 Peru | 15 |

Fig. 1. The desktop-like interface of the DRodVir database and various data browse menus. **A**: An overview of the web interface of DRodVir with a collapsed menu panel (left vertical bar) and the multi-tabbed main content panel showing an information grid of viruses in the Paramyxoviridae family (right). One line of the table is expanded by double click to display additional information and an illustrative linear map of the virus genome. **B**: The first layer of the accordion style main menu for browsing the database by virus category. The hierarchical taxonomic tree shows virus families (bolded branches) by default for brevity. **C**: The second layer of the main menu for browsing the database by rodent category. Several branches are expended to show lower taxonomic levels (i.e., subfamily, genus and species). **D**: The third layer of the main menu for browsing the database by geographic regions. The geographic tree is organized by continents and countries.

family, rodent species/family, sample type and sampling country (Fig. 2B); ii) a global map with indicative markers is provided for the column of sampling country to better illustrate the geographic distribution of the rodent-associated viruses (http://www.mgc.ac. cn/cgi-bin/DRodVir/main.cgi?func=map). In addition, the result

table and the related sequences can be easily downloaded as Excel and FASTA files respectively for further offline analyses. To bridge the gap between virologists and zoologists, the information concerning the rodents including common names, known distribution and phylogenetic relationships of rodents is also available and L. Chen et al. / Journal of Genetics and Genomics 44 (2017) 259-264

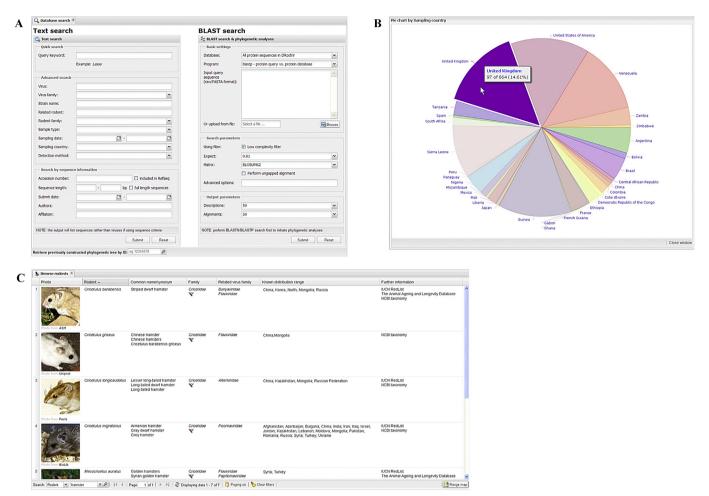


Fig. 2. Some additional utilities of the DRodVir database. **A**: The database search panel with two parallel query forms for text search (left) and BLAST search (right), respectively. **B**: A statistical pie chart of sampling countries for the 664 rodent-associated arenavirus identified so far. The largest proportion for United Kingdom is highlighted with a popup message by mouse over. **C**: An example of the rodent-related information table searched by the keyword 'hamster'. A simple search engine is provided in the bottom toolbar.

searchable in the DRodVir database (Fig. 2C). The BLAST sequence similarity search form enables the users to submit their own sequences for homology comparison among all known rodentassociated viruses. Moreover, an automatic pipeline for multiple sequence alignment and online phylogenetic tree construction is offered to facilitate follow-up sequence analysis. Detailed step-bystep instruction of all aforementioned usages of the DRodVir database is available in the online help page (http://www.mgc.ac. cn/DRodVir/howto.htm).

3. Database utility: an example of comparison

3.1. Virome diversities in rodents and bats

The genomic studies on rodent-associated viruses were initiated earlier than on bat-associated viruses. Both have been boosted by the genomic era during the past two decades. In addition, metagenomic approaches or high-throughput sequencing (HTS)-based methods have played a major role in the discovery of novel viruses in both rodents and bats in recent years (Fig. 3A). However, many GenBank records provide insufficient meta information about the detection methods applied, so the contributions of metagenomics or HTS-based approaches are systematically underestimated in Fig. 3A. Currently, ~25% sequences of rodent- or bat-associated viruses are deposited in public domains without related literature. To date (November 2016), 5491 viruses from 26 families and 6268 viruses from 24 families were identified from rodents and bats, respectively. However, only 21 viral families are shared by rodent- and bat-associated viruses. Viruses from families *Arena-viridae, Arteriviridae,* and *Picobirnaviridae,* carried by rodents, have not yet been detected in bats. Similarly, viruses from *Filoviridae,* identified in bats, have not been found in rodents so far (Fig. 3B). We avoided the use of lower taxonomic concepts such as genus or species because of the inconsistent criteria used for such definitions as described previously (Anthony et al., 2013) and because there are many newly identified viruses grouped into unclassified genera or species in public data.

Our current data show that the virome diversities in rodents and bats are similar at the family level, though a previous study suggests that bats harbor more zoonotic viruses per species than rodents (Luis et al., 2013). Remarkably, the aforementioned viruses were identified from 194 rodent species and 246 bat species. Given that there is nearly twice the number of living rodent species as bat species, the current investigation efforts cover only ~9% of the rodent species but ~21% of the bat species. Thus, the virome diversity in rodents has probably been underestimated compared with that of bats due to the limitation of the currently available data. Future studies to screen additional rodent species are required to achieve a better estimation of the virome diversity in rodents.

262

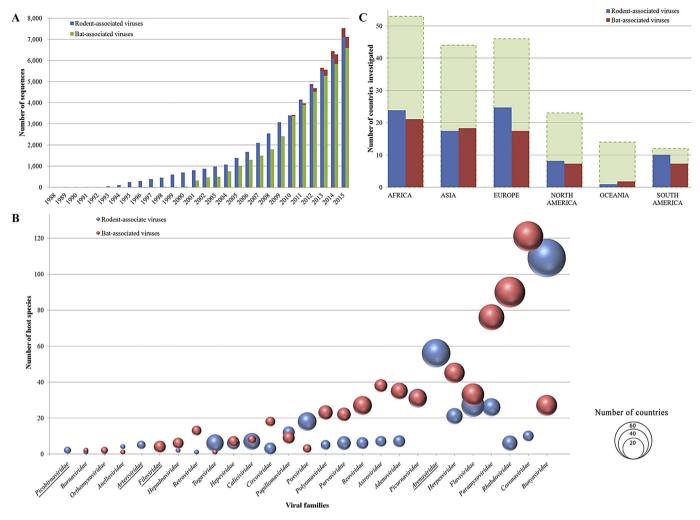


Fig. 3. Comparative statistics of rodent- and bat-associated viruses (as of November 2016). A: Increasing of public sequences available from GenBank by year. GenBank records without submission date are excluded. The proportions detected by metagenomic approaches or HTS-based methods are highlighted in red. B: Bubble chart of current identified viral families vs. related host species. The area of each bubble positively correlates with the number of sampling countries. Viral families detected solely in rodents or bats are underlined. C: Current sampling countries in each continent. Dashed shadows indicate the total number of countries in each continent.

3.2. Current research in rodent- and bat-associated viruses

The geographic distribution of the hosts being investigated is also an important factor to affect the estimation of virome diversity in both rodents and bats, since some viruses carried by the hosts are likely regional even for the same host species. For example, the Ebola virus and Marburg virus were exclusively identified from bats in Africa. Therefore, a better understanding of the virome diversity in either host requires the investigation efforts to cover as many countries as possible. Indeed, the current studies already include countries from all six continents. Thus far, the coverage of countries in most continents is far from comprehensive (Fig. 3C). Although the species diversity of rodents/bats in different countries is different and the current efforts in most countries could not cover all resident host species, further investigations dedicated to those unexplored regions of each continent are essential for our understanding of virome diversity in both rodents and bats.

The current data are disproportionally distributed in different viral families for both rodent- and bat-associated viruses. This could be a reflection of particular public concerns on some well-known zoonotic viruses carried by rodents or bats. For example, ~52% of the rodent-associated viruses currently identified were

from family Bunvaviridae (mostly are Hantaviruses), whereas ~41% of the bat-associated viruses reported so far were from family Rhabdoviridae (mostly are Lyssaviruses). Nevertheless, the biased data may also be a helpful measure for the potential distribution range of each viral family in the hosts. For instance, viruses from the family Bunyaviridae are the most widely distributed viruses among rodent-associated viruses, which have been detected in 109 rodent species from 64 countries worldwide (Fig. 3B). However, although the current data of viruses from the family Coronaviridae (i.e., coronavirus) in bats are less than half of those from family Rhabdoviridae, coronaviruses were surprisingly identified in 121 bat species from 39 countries, which makes them the most widely distributed bat-associated viruses thus far (Fig. 3B). Because the causative agents of both SARS and MERS pandemics are coronaviruses, their widespread occurrence in bats further highlights their potential threats to public health. Also, the current data show that viruses from some other families such as Rhabdoviridae, Paramyxoviridae, Flaviviridae, Herpesviridae, Arenaviridae, and Picornaviridae are broadly distributed in terms of host species (rodents and/or bats) and geography (Fig. 3B). Therefore, these viral families should be the hotspots for future studies dedicated to the prevention and control of zoonotic viral diseases.

4. Discussion

As of November 2016, the DRodVir database has collected information on 5491 rodent-associated animal viruses of 26 virus families detected from 194 rodent species in 93 countries worldwide. These data will give us an overview and snapshot of the current research regarding rodent-associated viruses and provide a substantial source for future attempts to assess and predict epidemic risks. Moreover, this work provides the scientific community with a framework and platform for further exploring the virome diversity of rodents.

Mammals are supposed to harbor over 320,000 undiscovered viruses (Anthony et al., 2013). Rodents and bats are important reservoirs for an increasing number of zoonotic infectious diseases with a significant impact on public health. The two most diverse and geographically widespread mammalian orders play vital roles in the maintenance, evolution, and spread of many emerging infectious disease agents. Our comparative analysis primarily explored the virome diversities in rodents and bats based on the summary of previous studies and highlighted the necessity to cover additional host species and geographic regions in future investigations. For easy comparative analysis on rodent- and bat-associated viruses in further studies, we plan to combine the two sister databases, DBatVir and DRodVir, to form an integrated resource for the two major reservoirs of emerging zoonoses in the future.

Acknowledgments

This work was supported by the National Major Science and Technology Project (2014ZX10004001) and the National Key Research and Development Program (2016YFC1202404) of China. J.Y. and Z.W. were supported by the Program for Changjiang Scholars and Innovative Research Team in University (IRT13007) and the CAMS Innovation Fund for Medical Sciences (2016-I2M-1-014), respectively.

References

- Altringham, J.D., 2011. Bats: from Evolution to Conservation, second ed. Oxford University Press, New York.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.
- Anthony, S.J., Epstein, J.H., Murray, K.A., Navarrete-Macias, I., Zambrana-Torrelio, C.M., Solovyov, A., Ojeda-Flores, R., Arrigo, N.C., Islam, A., Ali Khan, S., Hosseini, P., Bogich, T.L., Olival, K.J., Sanchez-Leon, M.D., Karesh, W.B., Goldstein, T., Luby, S.P., Morse, S.S., Mazet, J.A., Daszak, P., Lipkin, W.I., 2013. A strategy to estimate unknown viral diversity in mammals. MBio 4, e00598–00513.
- Calisher, C.H., Childs, J.E., Field, H.E., Holmes, K.V., Schountz, T., 2006. Bats: important reservoir hosts of emerging viruses. Clin. Microbiol. Rev. 19, 531–545.

- Chen, L., Liu, B., Yang, J., Jin, Q., 2014. DBatVir: the Database of Bat-associated viruses. Database (Oxford) 2014, bau021.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
- Enria, D.A., Pinheiro, F., 2000. Rodent-borne emerging viral zoonosis: hemorrhagic fevers and hantavirus infections in South America. Infect. Dis. Clin. North. Am. 14, 167–184.
- Firth, C., Bhat, M., Firth, M.A., Williams, S.H., Frye, M.J., Simmonds, P., Conte, J.M., Ng, J., Garcia, J., Bhuva, N.P., Lee, B., Che, X., Quan, P.L., Lipkin, W.I., 2014. Detection of zoonotic pathogens and characterization of novel viruses carried by commensal Rattus norvegicus in New York City. MBio 5, e01933–01914.
- Ge, X.Y., Li, J.L., Yang, X.L., Chmura, A.A., Zhu, G., Epstein, J.H., Mazet, J.K., Hu, B., Zhang, W., Peng, C., Zhang, Y.J., Luo, C.M., Tan, B., Wang, N., Zhu, Y., Crameri, G., Zhang, S.Y., Wang, L.F., Daszak, P., Shi, Z.L., 2013. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature 503, 535–538.
- Goeijenbier, M., Wagenaar, J., Goris, M., Martina, B., Henttonen, H., Vaheri, A., Reusken, C., Hartskeerl, R., Osterhaus, A., Van Gorp, E., 2013. Rodent-borne hemorrhagic fevers: under-recognized, widely spread and preventable – epidemiology, diagnostics and treatment. Crit. Rev. Microbiol. 39, 26–42.
- Huchon, D., Madsen, O., Sibbald, M.J., Ament, K., Stanhope, M.J., Catzeflis, F., de Jong, W.W., Douzery, E.J., 2002. Rodent phylogeny and a timescale for the evolution of Glires: evidence from an extensive taxon sampling using three nuclear genes. Mol. Biol. Evol. 19, 1053–1065.
- Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L., Daszak, P., 2008. Global trends in emerging infectious diseases. Nature 451, 990–993.
- Kruse, H., kirkemo, A.M., Handeland, K., 2004. Wildlife as source of zoonotic infections. Emerg. Infect. Dis. 10, 2067–2072.
- Luis, A.D., Hayman, D.T., O'Shea, T.J., Cryan, P.M., Gilbert, A.T., Pulliam, J.R., Mills, J.N., Timonin, M.E., Willis, C.K., Cunningham, A.A., Fooks, A.R., Rupprecht, C.E., Wood, J.L., Webb, C.T., 2013. A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? Proc. Biol. Sci. 280, 20122753.
- Masson, P., Hulo, C., De Castro, E., Bitter, H., Gruenbaum, L., Essioux, L., Bougueleret, L., Xenarios, I., Le Mercier, P., 2013. ViralZone: recent updates to the virus knowledge resource. Nucleic Acids Res. 41, D579–D583.
- Meerburg, B.G., Singleton, G.R., Kijlstra, A., 2009. Rodent-borne diseases and their risks for public health. Crit. Rev. Microbiol. 35, 221–270.
- Meredith, R.W., Janecka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C., Goodbla, A., Eizirik, E., Simao, T.L., Stadler, T., Rabosky, D.L., Honeycutt, R.L., Flynn, J.J., Ingram, C.M., Steiner, C., Williams, T.L., Robinson, T.J., Burk-Herrick, A., Westerman, M., Ayoub, N.A., Springer, M.S., Murphy, W.J., 2011. Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science 334, 521–524.
- Mills, J.N., Gage, K.L., Khan, A.S., 2010. Potential influence of climate change on vector-borne and zoonotic diseases: a review and proposed research plan. Environ. Health Perspect. 118, 1507–1514.
- Munster, V.J., Adney, D.R., van Doremalen, N., Brown, V.R., Miazgowicz, K.L., Milne-Price, S., Bushmaker, T., Rosenke, R., Scott, D., Hawkinson, A., de Wit, E., Schountz, T., Bowen, R.A., 2016. Replication and shedding of MERS-CoV in Jamaican fruit bats (*Artibeus jamaicensis*). Sci. Rep. 6, 21878.
- Perry, R.D., Fetherston, J.D., 1997. Yersinia pestis-etiologic agent of plague. Clin. Microbiol. Rev. 10, 35–66.
- Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2–approximately maximumlikelihood trees for large alignments. PLoS One 5, e9490.
- Smith, I., Wang, L.F., 2013. Bats and their virome: an important source of emerging viruses capable of infecting humans. Curr. Opin. Virol. 3, 84–91.
- Smits, S.A., Ouverney, C.C., 2010. jsPhyloSVG: a javascript library for visualizing interactive and vector-based phylogenetic trees on the web. PLoS One 5, e12267.
- Wolff, J.O., Sherman, P.W., 2007. Rodent Societies: an Ecological and Evolutionary Perspective. University of Chicago Press, Chicago, IL.
- Wong, S., Lau, S., Woo, P., Yuen, K.Y., 2007. Bats as a continuing source of emerging infections in humans. Rev. Med. Virol. 17, 67–91.