






# Gut Virome of the World's Highest-Elevation Lizard Species (*Phrynocephalus erythrurus* and *Phrynocephalus theobaldi*) Reveals Versatile Commensal Viruses

Juan Lu,<sup>a</sup> Shixing Yang,<sup>a</sup> Chunmei Wang,<sup>b</sup> Hao Wang,<sup>c</sup> Ga Gong,<sup>d</sup> Yuan Xi,<sup>a</sup> Jiamin Pan,<sup>a</sup> Xiaochun Wang,<sup>a</sup> Jian Zeng,<sup>a</sup> Ju Zhang,<sup>a</sup> Peng Li,<sup>e</sup>  Quan Shen,<sup>a</sup>  Tongling Shan,<sup>b</sup>  Wen Zhang<sup>a</sup>

<sup>a</sup>Department of Microbiology, School of Medicine, Jiangsu University, Zhenjiang, Jiangsu, China

<sup>b</sup>Shanghai Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Shanghai, China

<sup>c</sup>The Affiliated Huai'an Hospital, Xuzhou Medical University, Huai'an, Jiangsu, China

<sup>d</sup>Animal Science College, Tibet Agriculture and Animal Husbandry University, Nyingchi, Tibet, China

<sup>e</sup>Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing, Jiangsu, China

Juan Lu, Shixing Yang, Chunmei Wang, Hao Wang, and Ga Gong contributed to this article. Author order was determined in chronological order.

**ABSTRACT** The gut virome is a reservoir of diverse symbiotic and pathogenic viruses coevolving with their hosts, and yet limited research has explored the gut viromes of highland-dwelling rare species. Using viral metagenomic analysis, the viral communities of the *Phrynocephalus* lizards living in the Qinghai-Tibet Plateau were investigated. Phage-encoded functional genes and antibiotic resistance genes (ARGs) were analyzed. The viral communities of different lizard species were all predominated by bacteriophages, especially the *Caudovirales* order. The virome of *Phrynocephalus erythrurus* living around the Namtso Lake possessed a unique structure, with the greatest abundance of the *Parvoviridae* family and the highest number of exclusive viral species. Several vertebrate-infecting viruses were discovered, including caliciviruses, astroviruses, and parvoviruses. Phylogenetic analyses demonstrated that the virus hallmark genes of bacteriophages possessed high genetic diversity. After functional annotation, the majority of phage-associated functional genes were classified in the energy metabolism category. In addition, plenty of ARGs belonging to the multidrug category were discovered, and five ARGs were exclusive to the virome from *Phrynocephalus theobaldi*. This study provided the first insight into the structure and function of the virome in highland lizards, contributing to the protection of threatened lizard species. Also, our research is of exemplary significance for the gut virome research of lizard species and other cold-blooded and highland-dwelling animals, prompting a better understanding of the interspecific differences and transmission of commensal viruses.

**IMPORTANCE** The *Phrynocephalus* lizards inhabiting the Qinghai-Tibet Plateau (QTP) are considered to be the highest-altitude lizard species in the world, and they have been added to the IUCN list of threatened species. Living in the QTP with hypoxic, arid, natural conditions, the lizards presented a unique pattern of gut virome, which could provide both positive and negative effects, such as the enrichment of functional genes and the dissemination of antibiotic resistance genes (ARGs). This work provides the foundation for further research on the gut virome in these endangered lizard species and other cold-blooded and highland-dwelling animals, contributing to the maintenance of ecological balance on the plateau.

**KEYWORDS** viral metagenomics, bacteriophages, gut virome, lizards, Qinghai-Tibet Plateau

**Editor** Frederick S. B. Kibenge, University of Prince Edward Island

**Copyright** © 2022 Lu et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Wen Zhang, z0216wen@yahoo.com, Tongling Shan, shantongling@shvri.ac.cn, Quan Shen, shenquan@ujs.edu.cn, or Peng Li, lipeng@njnu.edu.cn.

The authors declare no conflict of interest.

**Received** 11 October 2021

**Accepted** 29 January 2022

**Published** 23 February 2022

The habitat elevation of *Phrynocephalus* lizards reaches the altitude of the Qinghai-Tibet Plateau (QTP), which is known as the “roof of the world” and has an average altitude of over 4,500 m (1). The *Phrynocephalus* genus contains over 40 species of toad-headed lizards, some of which are widely distributed in the QTP (2), including the species *Phrynocephalus erythrurus* and *Phrynocephalus theobaldi*, which are considered to be the highest-altitude lizard species in the world (3). Living restricted to the QTP with its harsh climate, these species need to have the adaptive ability to survive in the hypoxic, arid, natural conditions (4). As global warming continues, the QTP is being affected as a vulnerable ecological environment sensitive to natural and anthropogenic disturbances (5). Overwhelmed by the increasing environmental pressure, the two endangered lizard species have both been listed in the International Union for Conservation of Nature’s *IUCN Red List of Threatened Species* (6, 7). Conducting research on the endangered lizard species is of vital ecological significance in maintaining the ecological balance on the plateau.

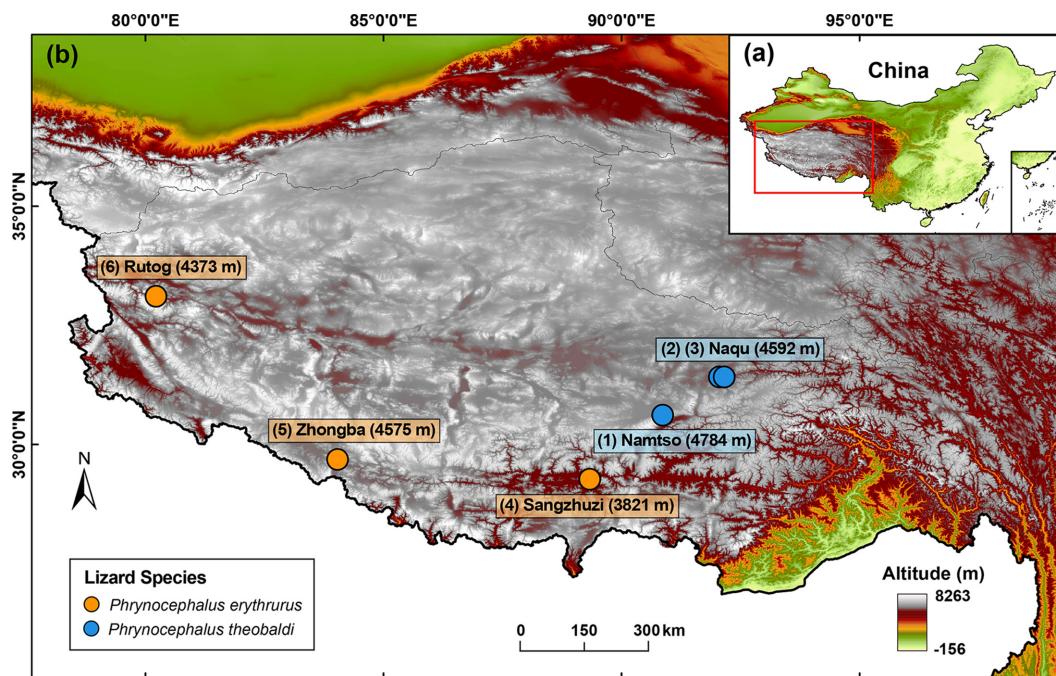
Previous studies revealed the genetic mechanisms for native animals to adapt to the extremely inhospitable highland, and additional functional genes related to energy metabolism were identified from animals living in the QTP, such as yaks (*Bos grunniens*) (8) and antelope (*Pantholops hodgsonii*) (9). Microbial genomes, known as the “second genome,” also play a vital role in the adaptation of highland-dwelling animals, and research on yak rumen and pika gut microbiomes highlights the adaptive contributions of microbial communities (10, 11). The gut microbiome of western fence lizards (*Sceloporus occidentalis*) also presents adaptive changes against different temperatures (12). Nevertheless, despite its being an indispensable part of the microbiome, consisting of eukaryotic and prokaryotic viruses from lizards, gut bacteria, and their food that have both pathogenic and beneficial effects, the gut virome in lizards is largely unexplored (13). On one hand, wild reptiles, including snakes, crocodiles, and lizards, could be potential reservoirs of pathogenic viruses that pose health risks to human and non-human animals, leading to the spread of zoonotic disease (14, 15). On the other hand, the transkingdom interactions between bacteriophages and their bacterial hosts in the intestine could contribute to host fitness by altering gene expression or mediating gene transduction (16). Given the complicated functions of the virome, the necessity to explore the composition of viral communities in highland lizards’ guts in order to protect these endangered species is great.

The biological functions of bacteriophages could extend to the dissemination of antibiotic resistance genes (ARGs) through horizontal gene transfer (HGT), which puts an increasing strain on public health (17). The abuse of antibiotics may bring about antibiotic-resistant bacteria and thus interfere with the treatment and control of infectious diseases (18), and phage transduction may accelerate the progress (19). Phage-associated ARGs have been discovered from various sources that are closely related to human activities, such as animals, including livestock and poultry (20), and aquatic environments, including wastewater and freshwater (21, 22). However, the ARGs of animals living in the desolate and underpopulated plateau have not yet been explored. Hence, there might be a unique pattern of phage-encoded ARGs in the two lizard species in the study.

In the present study, we investigated the viral communities of the *P. erythrurus* and *P. theobaldi* lizards living in the QTP and analyzed the genetic diversity of the main virus groups based on virus hallmark genes through phylogenetic analyses. Then, we explored the composition of phage-related functional genes and ARGs in the highland lizards’ gut viromes. The findings in this study could reveal the composition and function of the gut viromes in lizards inhabiting the plateau, contributing to the knowledge of the gut viromes in these endangered lizard species and other cold-blooded and highland-dwelling animals.

## RESULTS

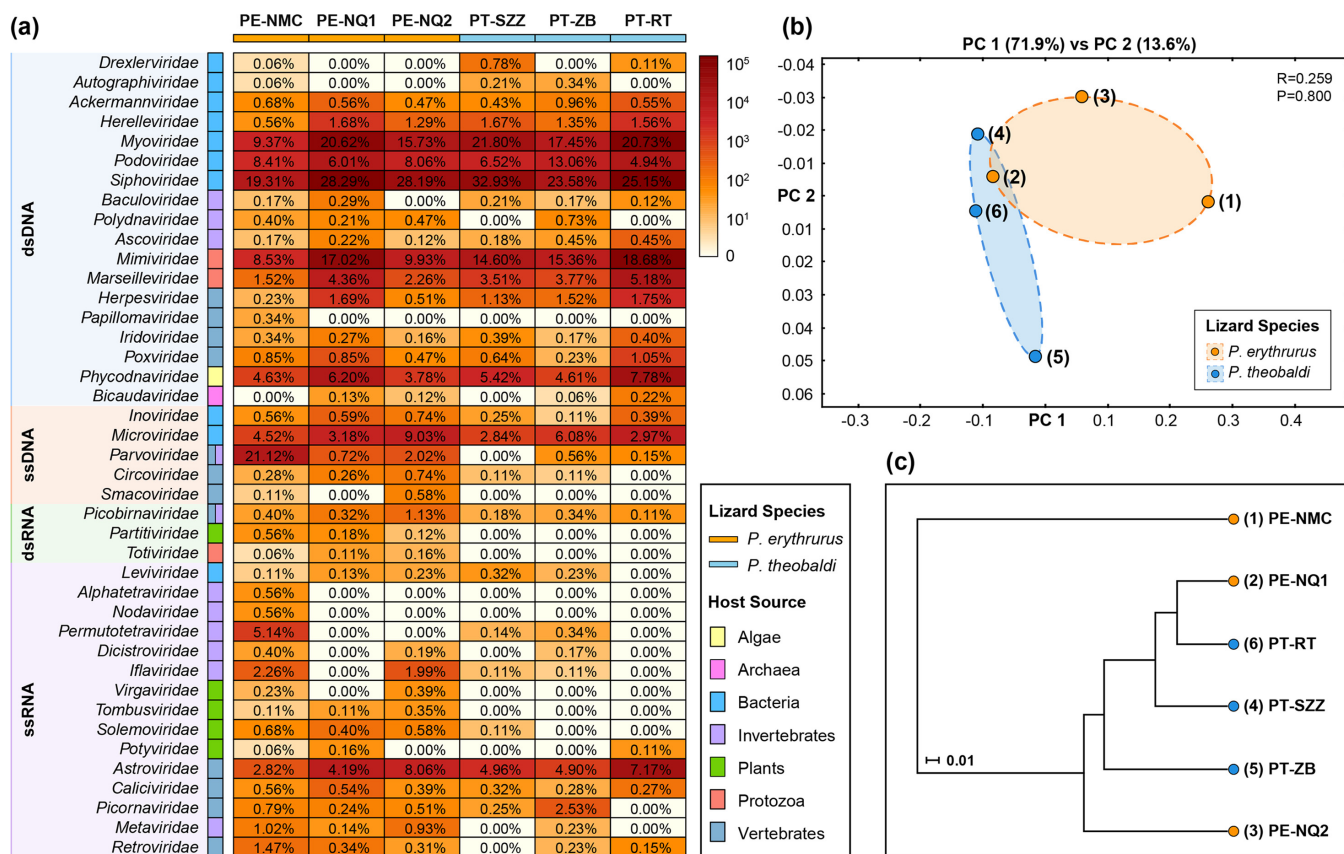
**Overview of sequencing outcomes.** To explore the gut virome of highland lizards, we carried out a complicated viral metagenomics investigation of both DNA and RNA viruses in 81 fecal samples from the lizards in the Qinghai-Tibetan Plateau (Fig. 1).



**FIG 1** Map of the Qinghai-Tibet Plateau with sampling sites. (a) Location within China. (b) Magnified view of the plateau in the red box of panel a. The sampling sites are indicated by dots and labeled with city names and altitudes. Lizard species are indicated by the color coding shown in the key. The map was created by the authors.

These fecal samples were pooled into six sample pools according to the two lizard species (*P. erythrurus* and *P. theobaldi*) and five sampling sites (Namtso, Naqu, Sangzhuzi, Zhongba, and Rutog). After library construction and next-generation sequencing on the Illumina NovaSeq platform, the six libraries generated a total of 86,640,498 raw sequence reads with an average length of 247 bp and an average GC content of 45.3%. The sequence reads were binned according to barcode and assembled into larger contigs. In total, 322,225 viral contigs (average length of 554.3 bp) were obtained through *de novo* assembly within the six bins and alignment against the viral protein database using BLASTx. The percentage of raw reads mapped to the viral contigs in each library ranged from 39.26% to 59.81% (Table S1 in the supplemental material). As for sequences in the blank control, no detectable DNA existed in the control library during quality inspection. Equal volumes of the seven libraries were included in the run of Illumina NovaSeq 6000 platform sequencing, where the control library generated a small number of reads ( $n = 14,162$ ). BLASTx searching based on the total reads in the control library revealed no viral sequences.

**Characterization of viral communities.** At the family level, viral contigs and singlet reads were classified into 41 viral families, including 18 double-stranded DNA (dsDNA) viral families, five single-stranded DNA (ssDNA) viral families, three double-stranded RNA (dsRNA) viral families, and 15 single-stranded RNA (ssRNA) viral families (Fig. 2a). The order *Caudovirales* accounted for the largest percentage of viral reads in the six libraries, mainly dominated by *Siphoviridae* ( $26.24\% \pm 4.26\%$  [mean  $\pm$  standard deviation]), *Myoviridae* ( $17.62\% \pm 4.24\%$ ), *Podoviridae* ( $7.83\% \pm 2.62\%$ ), and other bacteriophage families. Fewer reads were assigned to other dsDNA viral families, such as *Mimiviridae* ( $14.02\% \pm 3.64\%$ ) and *Marseilleviridae* ( $3.43\% \pm 1.237\%$ ), which infect amoebas, and *Phycodnaviridae* ( $5.40\% \pm 1.30\%$ ), which infect algae. The most abundant ssDNA viral family was *Microviridae* ( $4.77\% \pm 2.21\%$ ), whose members are bacteriophages, except in the virome from Namtso, which was predominated by *Parvoviridae* ( $21.12\%$ ). Relatively small numbers of reads were identified as being homologous to circular Rep-encoding single-stranded (CRESS) DNA viruses, including vertebrate-infecting *Circoviridae* ( $0.25\% \pm 0.24\%$ ) and *Smacoviridae* ( $0.12\% \pm 0.21\%$ ). The *Riboviria* realm was

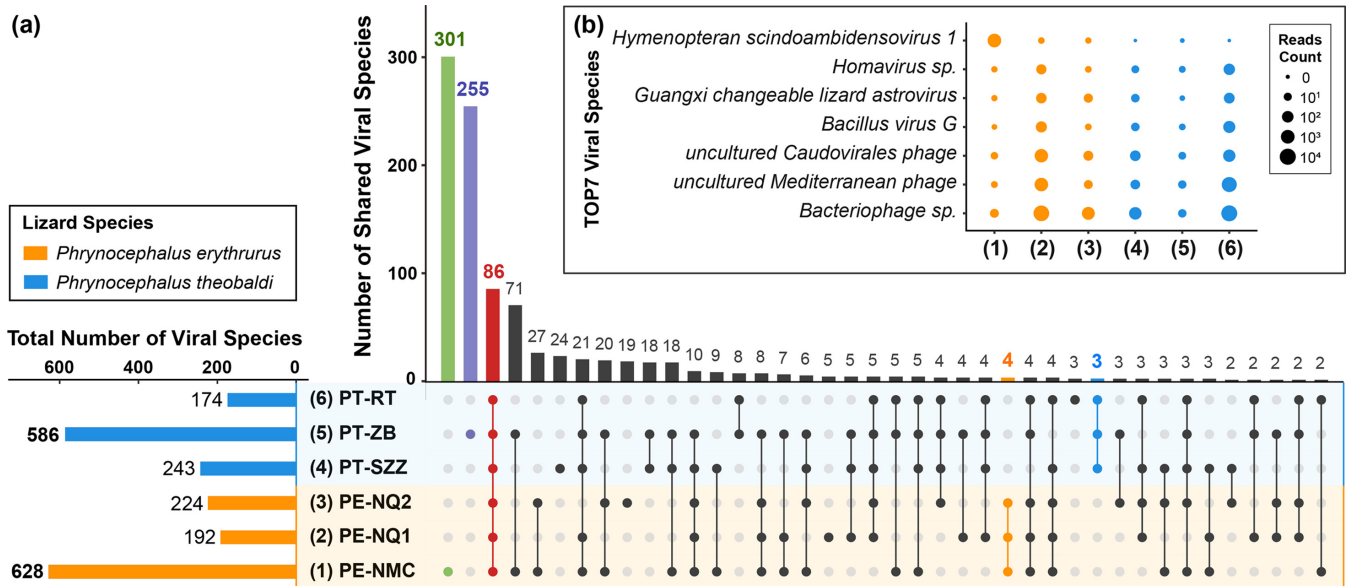


**FIG 2** Viral taxonomy analyses at the family level. (a) Heatmap representing the read number of each viral family in exponential form. Host sources and lizard species are indicated with the color coding shown in the key. Viral genome composition is represented by different-colored rectangles, with taxon names indicated on the left. The percentages of each viral family in the six viromes are shown. The sample names are defined in the form of lizard species-sampling site. PE, *Phrynocephalus erythrurus*; PT, *Phrynocephalus theobaldi*; NMC, Namtso, NQ, Naqu; SZZ, Sangzhu; ZB, Zhongba; RT, Rutog. (b, c) PCoA plot (b) and UPGMA taxonomic tree (c) showing the similarity of viral community structures based on the Bray-Curtis ecological distance matrix. The *P* value was calculated using ANOSIM.

mainly dominated by the family *Astroviridae* ( $5.35\% \pm 1.77\%$ ), whose members infect vertebrates. Other RNA viral families were distributed sporadically in the six viromes, such as invertebrate-infecting *Permutotetraviridae* in the virome from Namtso (5.14%) and vertebrate-infecting *Picornaviridae* in the virome from Zhongba (2.53%). The comparison of viral communities between *Phrynocephalus erythrurus* and *Phrynocephalus theobaldi* through principal-coordinate analysis (PCoA) suggested that there was no statistically significant difference between the two lizard groups at the family level ( $P > 0.05$ ) (Fig. 2b). Meanwhile, the unweighted pair group method with arithmetic mean (UPGMA) dendrogram presented a clear separation between the virome from Namtso and the other five viromes. The second and the sixth viromes from different lizard species were clustered together, and thus, they might share similar compositions and abundances (Fig. 2c).

At the species level, a total of 1,044 viral species were identified in the six viral communities. The virome from Namtso had the highest number of viral species, followed by the one from Zhongba. Meanwhile, the two viromes also contained considerable numbers of distinctive species, 301 and 255, respectively (Fig. 3a). The substantial numbers of exclusive viral species in the two viromes demonstrated their unique viral community signatures. However, a total of 86 viral species were shared in all viromes of the two lizard species, accounting for 13% to 49% of the total number of species identified in each of the six viromes. It is worth mentioning that not only did large proportions of viral species exist simultaneously in both lizard species, but also, few viral species were exclusive to either one of them, four and three species, respectively. These results revealed the slight influence of lizard species on viral community structure at the



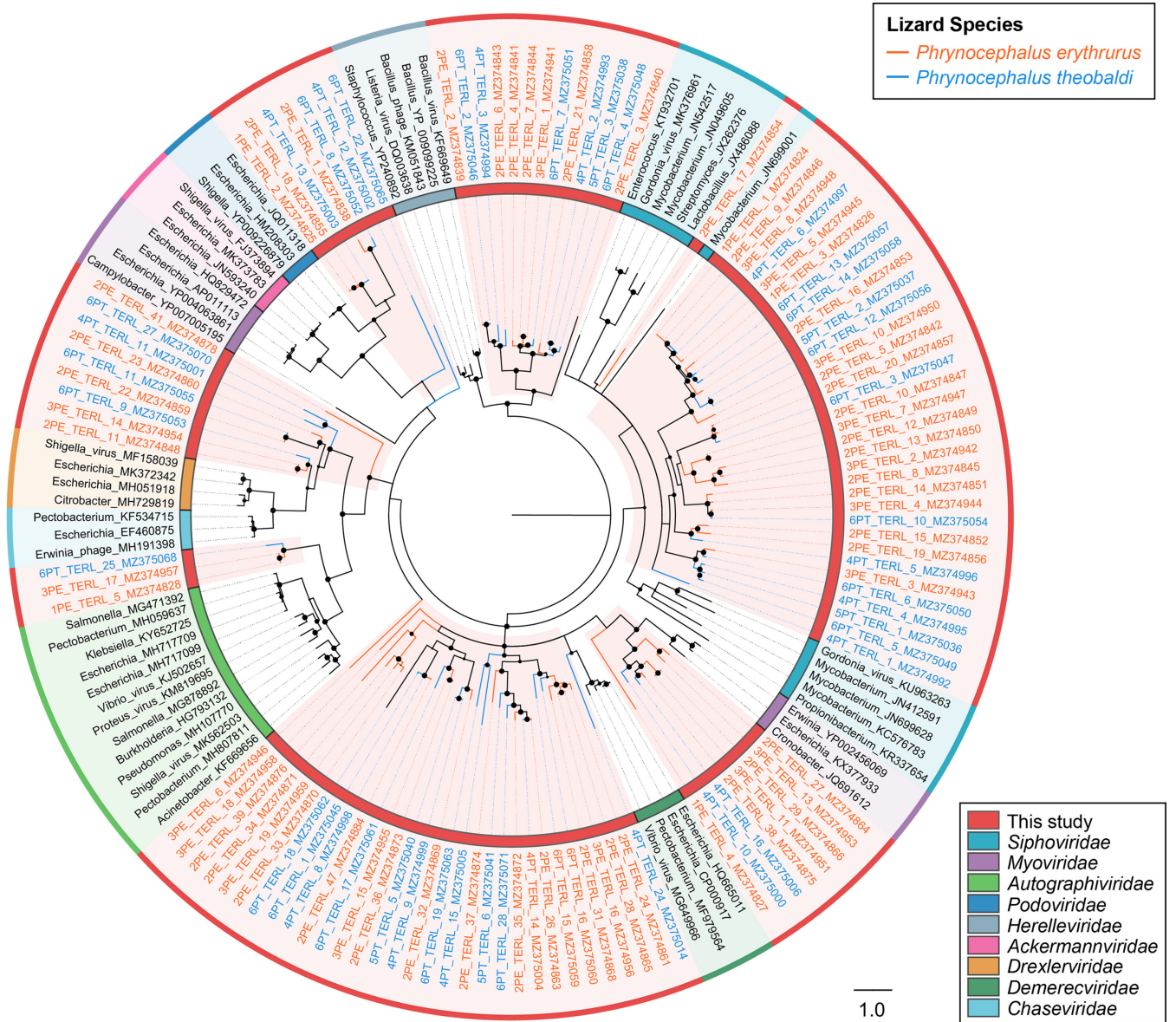


**FIG 3** Viral taxonomy analyses at the species level. (a) UpSet plot depicting the numbers of shared viral species among the six viromes. Filled dots with interconnecting vertical lines represent the intersections, and unfilled light gray dots represent sets that do not belong to the intersections. The bars above represent the numbers of species within the intersections, and the bars to the left depict the total number of viral species in each virome set. (b) Bubble chart showing the top 7 most abundant viral species in the six viromes. Bubble size indicates the abundance of reads assigned to each species. In both panels, the lizard species are indicated with the color coding shown in the key in panel a.

species level. For further analyses, the results for the top 7 most abundant viral species in the six viromes are depicted in Fig. 3b; these were dominated by viral species belonging to bacteriophages, such as uncultured *Caudovirales* phage and uncultured Mediterranean phage. In addition, species *Hymenopteran scindoambidensovirus 1* of the *Parvoviridae* family was particularly abundant in the virome from Namtso, indicating that the virome possibly possessed a unique viral community structure.

**Identification of virus hallmark gene sequences.** In this study, 465 viral sequences were generated after extension and annotation of contigs (Fig. S1; Table S2). These sequences belonged to the five abundant virus groups in the study and contained the virus hallmark genes (i.e., conserved domains) of each virus group (23, 24), including the genes encoding phage terminase large subunit (TerL) for *Caudovirales*, major capsid protein (MCP) for *Microviridae*, replication-associated protein (Rep) for CRESS DNA viruses, nonstructural protein 1 (NS1) for *Parvoviridae*, and RNA-dependent RNA polymerase (RdRp) domain for *Riboviria*. The BLASTx results showed that those sequences shared sequence identities with their best matches ranging from 23.04% to 100%, and their lengths ranged from 305 bp to 11,503 bp with an average of 2,294 bp. To phylogenetically analyze these viral sequences, 186 sequences among them with a complete coding sequence (CDS) of a virus hallmark gene were selected for further construction of phylogenetic trees.

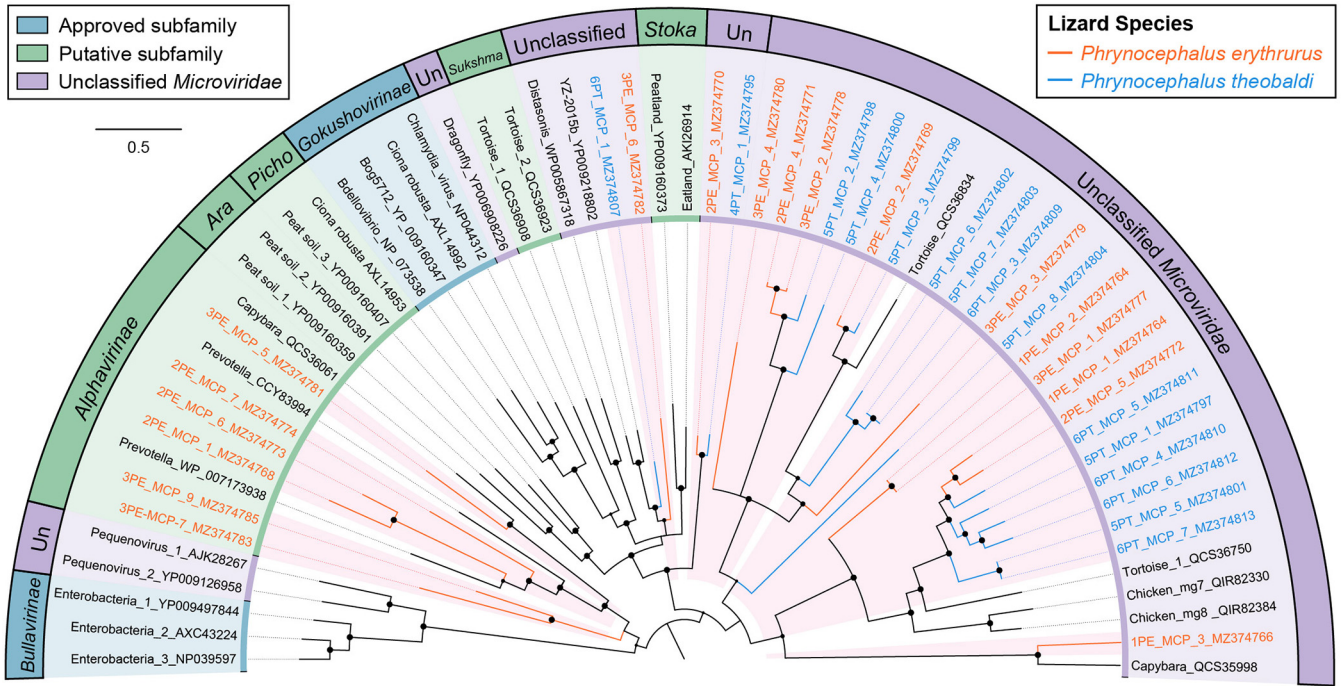
**Phylogenetic analyses of the main virus groups.** The phylogenies of the main virus groups in the study were constructed based on the amino acid sequences encoded by complete virus hallmark genes and the corresponding reference protein sequences in the GenBank database. In total, 186 sequences with complete marker regions were selected for phylogenetic analyses. The order *Caudovirales* and the family *Microviridae* represent the largest categories of dsDNA and ssDNA bacteriophages, respectively. The phage large terminase subunit (TerL) region is highly conserved in its evolution in the *Caudovirales* order, and thus, the phylogenetic analysis of dsDNA phages was performed based on the 105 TerL protein sequences that were identified in this study (Fig. 4). The sequences from *P. erythrurus* distributed alternating with those from *P. theobaldi*. The tree structure revealed that the majority of phage TerL sequences were too divergent to be classified into known families in the *Caudovirales* order and were



**FIG 4** The phylogeny of *Caudovirales* identified in feces of lizards. Bayesian inference tree was established based on amino acid sequences of TerL of *Caudovirales*. Representative strains of all families in *Caudovirales* are included and marked with the color coding in the key on the bottom right. The viruses found in this study are indicated by red-filled sectors and orange or blue lines according to lizard species. The sizes of the black dots on the nodes are positively correlated with the corresponding bootstrap scores. The scale bar indicates the amino acid substitutions.

grouped into several distinct clusters among the representative clades. The divergence was supported by high bootstrap values, indicating the hidden diversity of tailed bacteriophages in the study. Meanwhile, the phylogenetic tree based on 33 MCPs of *Microviridae* also revealed the substantial genetic variety of ssDNA bacteriophages (Fig. 5). Specifically, none of them showed a close relationship to the approved subfamilies, including *Bullavirinae* and *Gokushovirinae*. Moreover, six of these sequences from *P. erythrurus* were phylogenetically grouped into the putative subfamily *Alphavirinae*, and the rest of them were classified into separate clades genetically distinct from those classified into subfamilies. These results revealed the undiscovered diversity of bacteriophages from lizards living in the plateau.

However, phylogenetic analyses of other ssDNA viruses and RNA viruses presented a different picture, in which most of the sequences were closely clustered with known viral strains. As presented in the phylogenetic tree of replication protein (Rep)



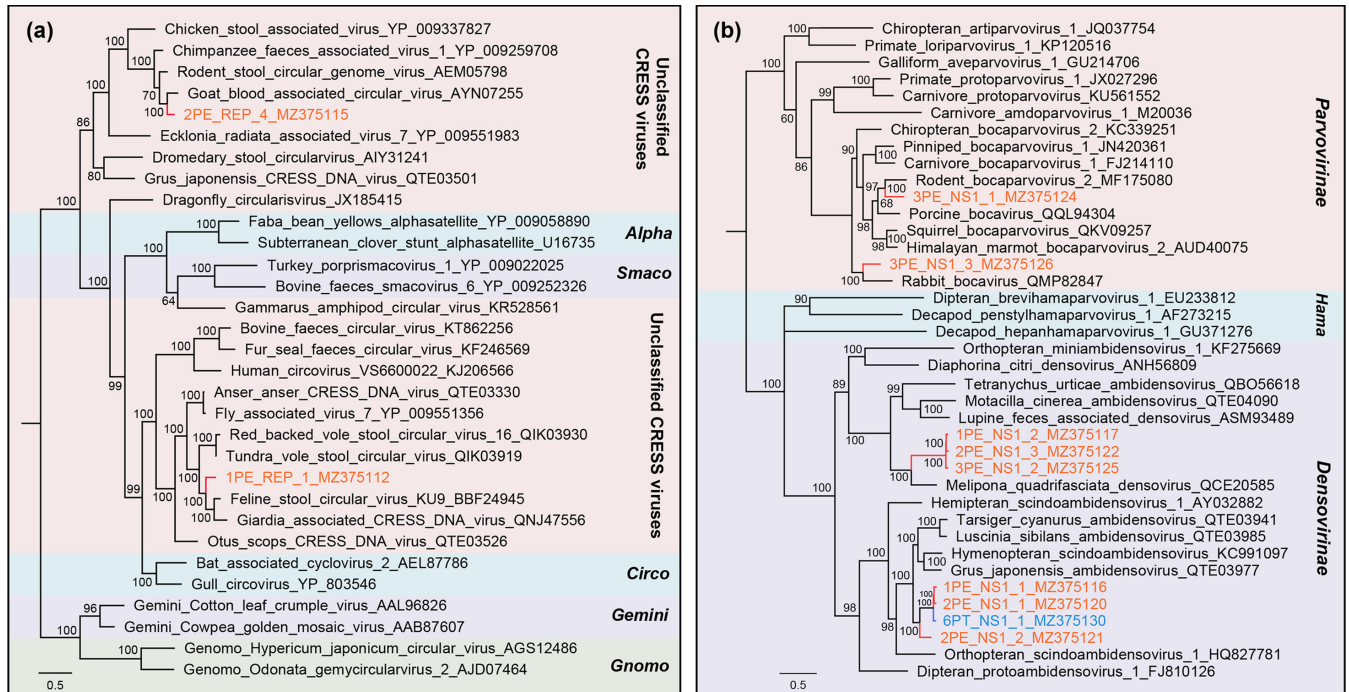
**FIG 5** The phylogeny of *Microviridae* identified in feces of lizards. Bayesian inference tree established based on amino acid sequences of MCP of *Microviridae*. The viruses found in this study are marked with colored lines and red-filled sectors. Different viral groups are marked with color coding shown in the key on the upper left. The sizes of the black dots on the nodes are positively correlated with the corresponding bootstrap scores. The scale bar indicates the amino acid substitutions. *Ara*, Aravirinae; *Picho*, Pichovirinae; *Sukshma*, Sukshma virinae; *Stoka*, Stokavirinae; Un, unclassified *Microviridae*.

sequences for CRESS DNA viruses, the two sequences from *P. erythrurus* were both phylogenetically clustered with known viral strains in the clade of unclassified CRESS DNA viruses, sharing >60% Rep protein sequence identity (Fig. 6a). Similarly, the tree based on NS1 for *Parvoviridae* also indicated that the nine sequences all showed a close relationship to previous viral species (Fig. 6b). Two of them were closely clustered with bocaparvoviruses that were identified from rodents and rabbits, and seven showed close relationships with ambidensoviruses from a wide range of invertebrates. Moreover, the dendrogram of RdRp proteins showed that the 37 sequences of RNA viruses were all clustered with known viral species, including caliciviruses and astroviruses, which infect vertebrates, and iflaviruses and alphapermutotetraviruses, which infect invertebrates (Fig. 7).

**Functional analysis of phage genes.** Considering that a large number of diverse bacteriophages were identified in the study and phages are capable of controlling and impacting bacterial hosts, we annotated the predicted open reading frames (ORFs) in phage contigs using the KEGG database to explore the latent functions of phages (Fig. 8; Table S3). The majority of ORFs were annotated with metabolism functions, such as amino acid, carbohydrate, and nucleotide metabolism. As for the genetic information processing category, functional genes related to replication and repair were the largest proportion. Moreover, it is worth mentioning that there were considerable numbers of genes directly correlated with the bacterial hosts, including those classified into prokaryote cellular community, antimicrobial drug resistance, and bacterial infectious disease categories, indicating the host-regulating function of bacteriophages in the lizard gut.

**Diversity and abundance of phage ARGs.** Bacteriophage-mediated transduction is an indispensable route for the dissemination of ARGs, and thus, phages might be a crucial reservoir of ARGs. To explore the functions of phages, the potential ARGs in phage contigs were identified using BLASTp against the Comprehensive Antibiotic Resistance Database (CARD). A total of 146 predicted ORFs were annotated as the corresponding 52 ARGs classified into 19 drug classes (Fig. 9a; Table S4). The predominant antibiotic drug class was multi-drug (62.70 to 11,406.68 ppm), followed by macrolide (12.90 to 2,248.56 ppm) and phenicol (10.31 to 816.55 ppm) (Fig. 9b), and 13 ARGs were widely detected in all six viromes. The





**FIG 6** The phylogenies of CRESS DNA viruses and *Parvoviridae* identified in feces of lizards. (a) Bayesian inference tree established based on amino acid sequences of Rep protein of CRESS DNA viruses. (b) Bayesian inference tree established based on amino acid sequences of NS1 protein of *Parvoviridae*. The viruses found in this study are marked with colored lines and letters. Different taxonomic clusters were represented by rectangles filled with different colors, and taxon names are indicated on the right. Scale bars indicate the amino acid substitutions per site.

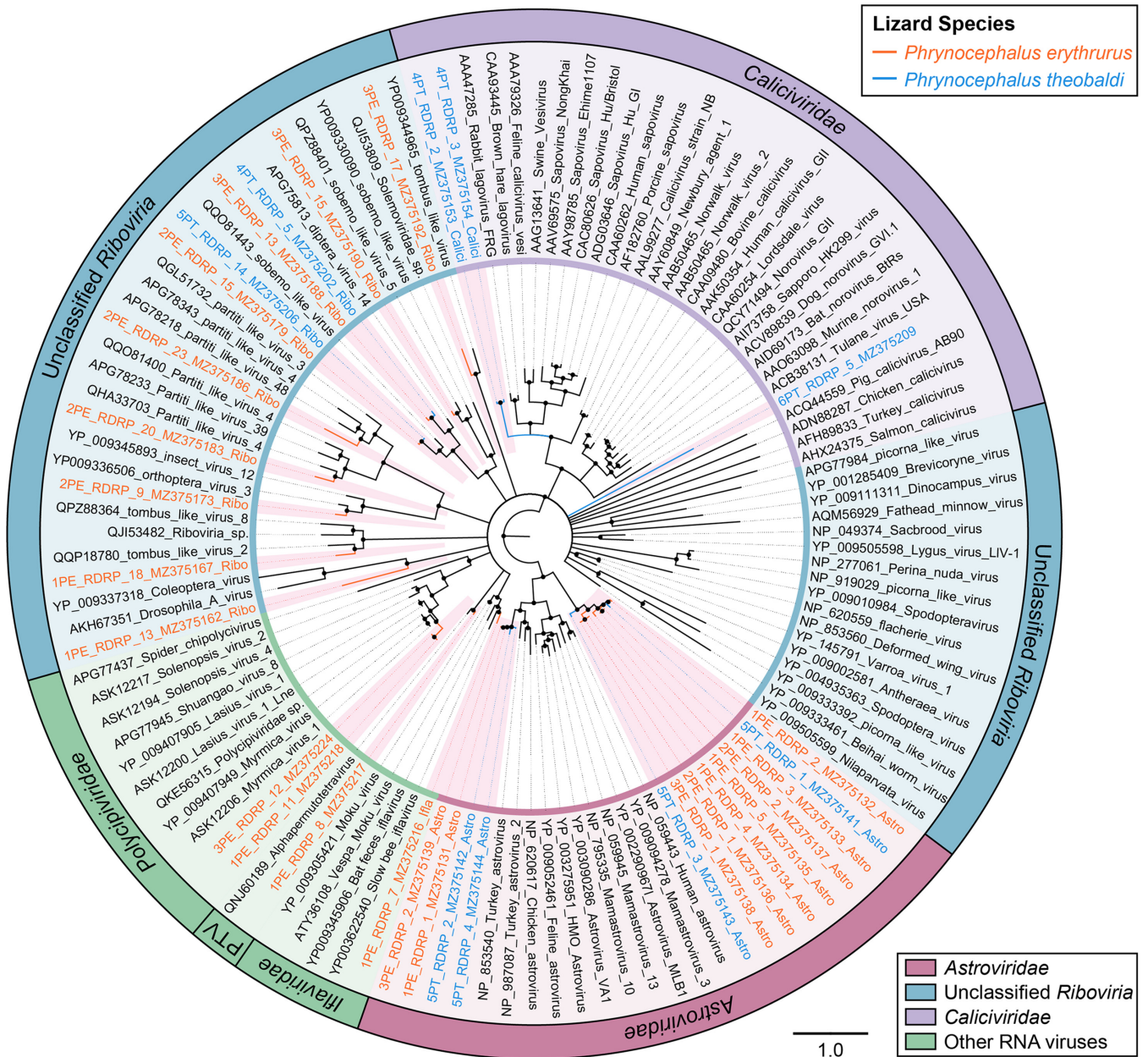
most abundant gene was *cmeB* (26.77 to 5,781.76 ppm), followed by *macB* (8.28 to 1,778.52 ppm), *oprM* (3.33 to 1,887.22 ppm), *cmeA* (10.87 to 1,368.48 ppm), etc. In addition, five ARGs were exclusive to the lizard species *Phrynocephalus theobaldi*, including *arnA* (11.95 to 116.15 ppm), *ugd* (31.87 to 498.60 ppm), *iles* (10.33 to 997.12 ppm), *tet36* (28.33 to 152.27 ppm), and *dfpE* (13.43 to 113.04 ppm), indicating a unique composition of ARGs in the lizard species (Fig. 9a). The relative abundances of resistance mechanisms in the six viromes are presented in Fig. 9c. Antibiotic efflux pumps formed the largest proportion of resistance mechanisms throughout the six viromes, ranging from 63.59% to 95.03%, followed by antibiotic target alteration (~3.70% to 23.32%), and the other three mechanisms were all less than 10%.

## DISCUSSION

The virome is an integral part of the microbiome that can provide both positive and negative effects during coevolution with prokaryotic and eukaryotic hosts (13). Considering the inhospitable environment in the QTP and the rarity of highland lizards, this study was carried out to explore the viral community structure in two endangered lizard species and provide clues to reveal the underlying function of phages in the gut virome, in order to provide a foundation for protecting threatened species.

The viral communities in the six virome were all predominated by bacteriophages, especially the *Caudovirales* order, mainly including *Myoviridae*, *Podoviridae*, and *Siphoviridae* families. In addition, another bacteriophage family, *Microviridae*, also accounted for a definite proportion of each of the six viromes. These results demonstrated that phages accounted for the overwhelming proportion of the lizard gut virome, consistent with the gut viromes in honeybees, birds, and mammals (25, 26). Furthermore, the conserved TerL and MCP genes of the *Caudovirales* order and *Microviridae* family all presented high genetic diversity that clustered phylogenetically between clades of classified species and even into several novel clades. The scenario is similar to that of the bacteriophages in the human gut virome, which possess a predominant virome core containing highly diverse viral sequences from *Caudovirales* and *Microviridae* (27). Therefore, although bacteriophages are characterized by

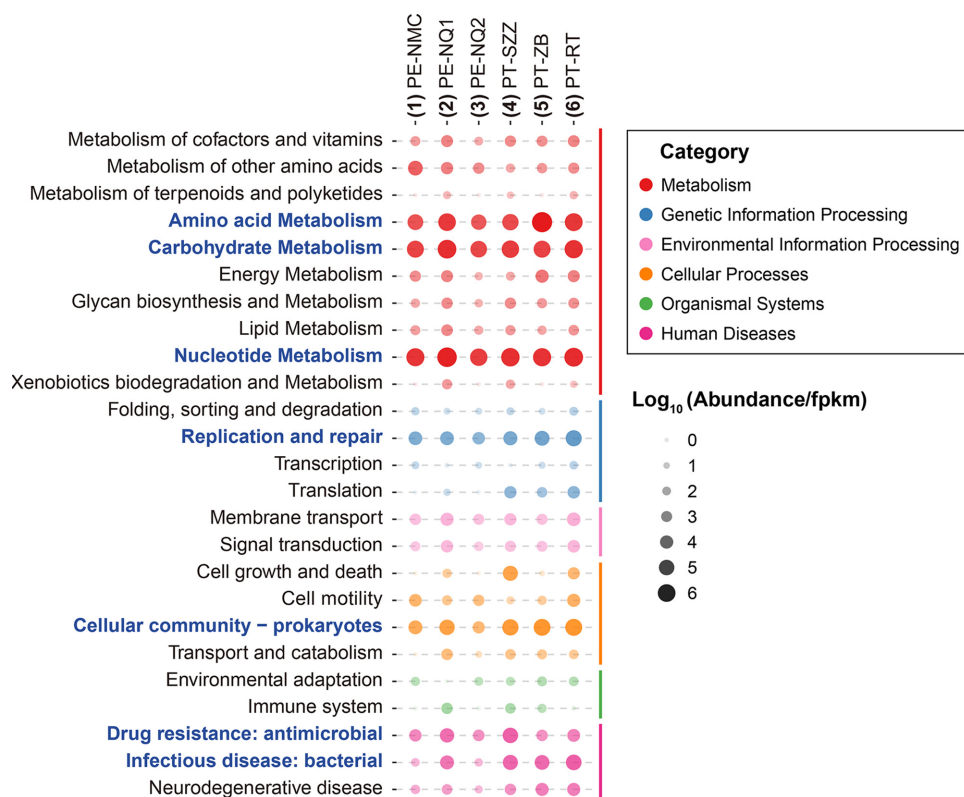




**FIG 7** The phylogeny of RNA viruses identified in feces of lizards. Bayesian inference tree established based on amino acid sequences of RdRp protein of RNA viruses. The viruses found in this study are marked with colored lines and red-filled sectors. Different viral groups are marked with color coding shown in the key on the bottom right. The sizes of the black dots on the nodes are positively correlated with the corresponding bootstrap scores. The scale bar indicates the amino acid substitutions. PTV, *Permutotetraviridae*.

the simplest structures and the smallest sizes, phages possess the highest abundance and the greatest diversity in gut viromes.

The viral community structures of the two lizard species presented no significant differences. Gut microbiomes are subject to multiple factors, such as living conditions (28) and eating patterns (29), and environments play a greater role in shaping gut microbiomes than host genetic backgrounds (30). Given the bleak and inhospitable characteristics of the plateau, it came as no surprise that the species factor alone would make no difference. Although the two sample groups collected from the Naqu area did not cluster closely together in PCoA, there was no statistically significant difference between the two groups, and the differences could be explained by individual variations and environmental factors. In addition, it should be taken into consideration that

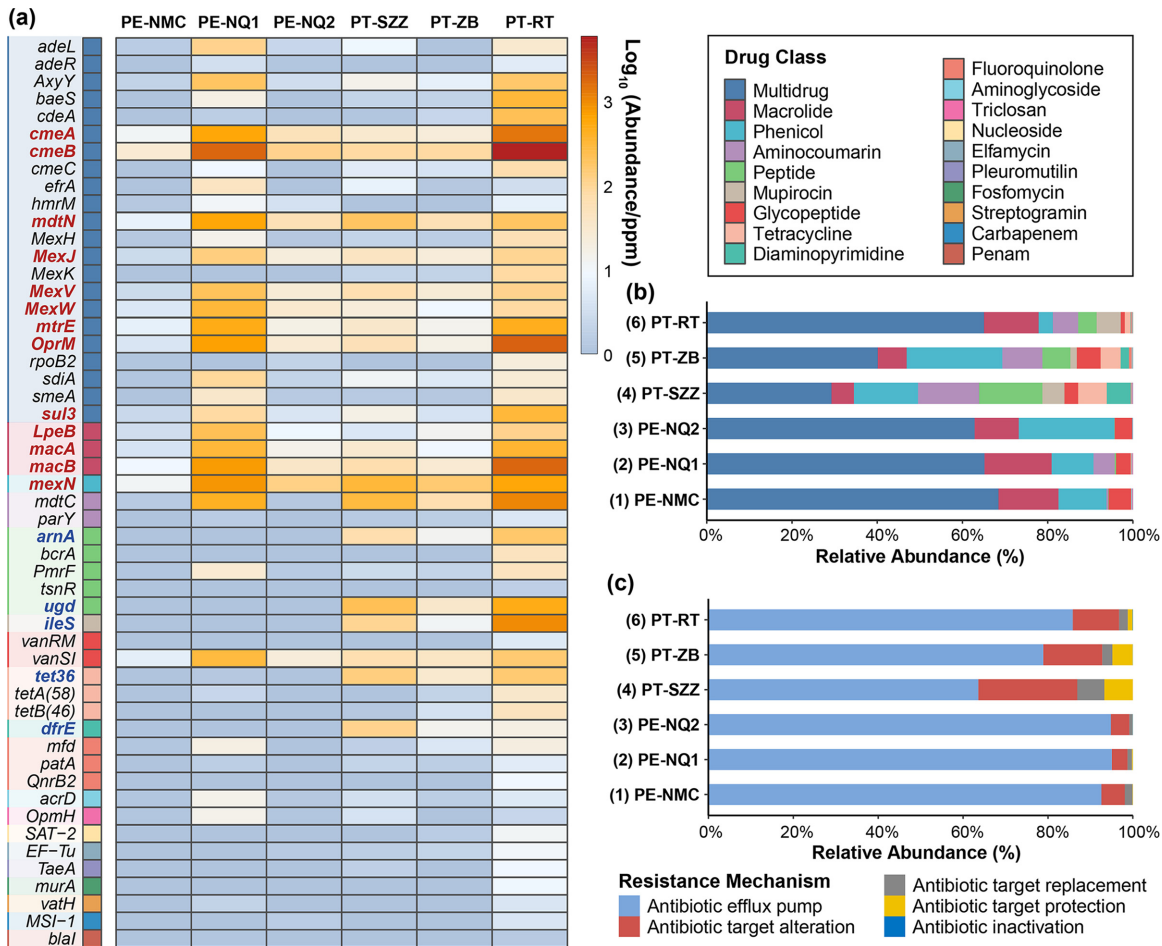


**FIG 8** Abundances of functional genes in phage contigs from the six viromes. Bubble chart presenting the functional annotation results for putative proteins in phage contigs based on the KEGG pathway database. The sizes of the bubbles are positively correlated with the abundances of the genes. Pathway subcategories mentioned in the text are highlighted in blue font.

the virome from *P. erythrus* lizards living around the Namtso Lake possessed a unique viral community structure. It had the highest number of viral species and contained more viruses belonging to the *Iflaviridae*, *Permutotetraviridae*, and *Parvoviridae* families, which infect a wide range of invertebrates, because the lake area was not only a natural habitat for many insects and other prey of lizards but also a tourist attraction with numerous visitors. And the UPGMA dendrogram obviously separated this virome from the others, further highlighting the role of environmental factors in the regulation of viral communities.

Several vertebrate-infecting viruses were discovered in the study, including caliciviruses and astroviruses, which are potentially pathogenic for lizards. Astroviruses and caliciviruses both are nonenveloped, positive-sense, single-stranded RNA viruses and infect a wide range of vertebrates, causing viral gastroenteritis accompanied by various symptoms, including diarrhea, listlessness, nervousness, loss of appetite, etc. However, astroviruses mainly infect avian and mammalian animal species, and there are no related reports about lizard-infecting astroviruses (31). Although caliciviruses have been found in other reptiles, such as snakes, the infected snakes presented no correlated symptoms of gastroenteritis (32). Given that the lizard subjects in the study were also seemingly asymptomatic, the pathogenicity and transmissibility of those viruses remain to be further studied.

Bacteriophages are capable of mediating the transduction of genetic material in bacterial communities, affecting and controlling the phenotypes of their bacterial hosts (16). Considerable functional gene reserves have been discovered in bacteriophages, contributing to the evolution of their hosts, including metabolic genes and virulence genes (33). In this study, the majority of phage-associated functional genes were classified into the energy metabolism and DNA repair subcategories. Native animals living in high-altitude regions are forced to generate adaptive evolution in order to survive in low-temperature and hypoxic



**FIG 9** Abundances of ARGs in phage contigs from the six viromes. (a) Heatmap depicting the abundances of dominant ARGs in the six viromes. Drug classes are indicated with the color coding shown in the key on the upper right. The ARGs exclusive to an individual lizard species are marked in blue font, and the ARGs existing in all six viromes are marked in red font. (b) Bar plot showing the relative abundances of different drug classes of ARGs. (c) Bar plot showing the proportions of different resistance mechanisms of ARGs in the six viromes.

environments. Functional genes related to nutrient metabolism and DNA repair are also enriched in highland-dwelling humans and other mammals, providing potential genetic mechanisms of highland adaptation (8, 9, 34). However, the lack of lowland-dwelling species of lizards in this study hindered comparative research based on different altitudes, and thus, the phenomena could not simply be attributed to the factor of elevation. Further research would be needed to identify the adaptive responses to extreme environments.

Bacteriophages can transfer ARGs through transduction, facilitating the evolution of drug-resistant bacteria (17). Previous research demonstrates that bacteriophages in both densely populated and sparsely populated areas store numerous ARGs, which are associated with the level of anthropogenic activities (35, 36). In the present study, most of the ARGs were classified into the multidrug category, and few of them encoded  $\beta$ -lactam resistance, whereas  $\beta$ -lactam resistance genes have been identified as a predominant resistance type in phage viromes from multiple animals and environments, such as livestock (37), poultry (20), agricultural soil (38), and wastewater (21). Thus, it is probably because the Plateau is a desolate area with a low urbanization level and would not be affected by contamination with common antibiotics that these highland-dwelling lizard species might possess a different composition of ARGs. In addition, some ARGs were exclusive to the virome from *P. theobaldi*, revealing that those genes might be correlated with the lizard species. Previous research identified the ARGs in different chicken cecum microbiomes and found that the lincosamide-streptogramin-macrolide (LSM) resistance genes were more abundant



in one chicken group (39). Thus, the uniqueness of ARGs in different species deserves further investigation.

In conclusion, this study first explored the structure and function of viral communities in two endangered lizard species, *P. erythrurus* and *P. theobaldi*, inhabiting the Qinghai-Tibet Plateau. The lizard gut viromes were predominated by a wide variety of bacteriophages, which included plenty of functional genes associated with energy metabolism and antibiotic resistance. Because of the rarity of the lizard species and the difficulty of sampling in the highland, sample sizes are inevitably limited in the study. Even so, our research contributed to a better understanding of viromes in highland lizard species, encouraging more research to focus on these vulnerable species and other cold-blooded and highland-dwelling animals in the future.

## MATERIALS AND METHODS

**Sample collection and preparation.** In July 2020, to investigate the viromes of two lizard species (*Phrynocephalus erythrurus* and *Phrynocephalus theobaldi*) living in the Qinghai-Tibetan Plateau, a total of 81 fecal samples from the two lizard species were collected from five sampling sites (Namtso, Naqu, Sangzhuzi, Zhongba, and Rutog) with an average altitude of over 4,000 m (Fig. 1; Table S1). All samples were collected using disposable materials and shipped on dry ice. Fecal samples were pooled in six sample pools according to lizard species and sampling sites. Samples were resuspended in 10 volumes of phosphate-buffered saline (PBS) and vigorously vortexed for 5 min and then freeze-thawed three times on dry ice. The supernatants were then collected after centrifugation (10 min at  $15,000 \times g$ ) and stored at  $-80^{\circ}\text{C}$  until use.

**Viral metagenomic library construction.** Five hundred microliters of each supernatant was filtered through a  $0.45\text{-}\mu\text{m}$  filter (Millipore) to remove eukaryotic- and bacterial-cell-sized particles. The filtrates enriched in viral particles were treated with DNase and RNase to digest unprotected nucleic acid at  $37^{\circ}\text{C}$  for 60 min (40–42). Then, the remaining total nucleic acid was isolated using a QIAamp viral RNA minikit (Qiagen) according to the manufacturer's protocol. For library construction, dsDNA was synthesized from viral RNA and DNA. For RNA viruses, a reverse transcription kit (SuperScript III reverse transcriptase) was used for reverse transcribing RNA into cDNA, and then the DNA polymerase I large fragment (Klenow fragment) was added to synthesize the second strand of cDNA (dsDNA). Specifically,  $12\ \mu\text{L}$  of nucleic acid extracts was added to the reaction mixture to synthesize dsDNA (total reaction mixture volume,  $20\ \mu\text{L}$ ). For ssDNA viruses, ssDNA was converted to dsDNA using the Klenow reaction and all the dsDNA products were used to construct libraries. Overall, six libraries, along with a control library, were then constructed using a Nextera XT DNA sample preparation kit (Illumina), and the quality was inspected using agarose gel electrophoresis and the Agilent bioanalyzer 2100. All libraries were sequenced on an Illumina NovaSeq 6000 platform with 250-bp paired-end reads with dual barcoding for each sample pool (43). The total read number of each library is presented in Table S1.

**Bioinformatics analysis.** Paired-end reads of 250 bp generated by NovaSeq were debarcoded using vendor software from Illumina. An in-house analysis pipeline running on a 32-node Linux cluster was utilized to process the data. Reads were considered duplicates if bases 5 to 55 were identical, and only one randomly chosen copy of any pair of duplicates was kept. Low-sequencing-quality tails were trimmed using a Phred quality score of 10 as the threshold. Adaptors were trimmed using the default parameters of VecScreen, which is based on NCBI BLASTn with specialized parameters designed for adapter removal. The cleaned reads were *de novo* assembled within each barcode, and the chimeras among them were detected. Then, the reads were filtered by length using the ENSEMBLE assembler with the default parameters (44). Contigs and singlet reads were then matched against a customized viral proteome database using BLASTx with an E-value cutoff of  $<10^{-5}$ . The virus BLASTx database was compiled using the NCBI virus reference proteome (<ftp://ftp.ncbi.nih.gov/refseq/release/viral/>) and viral protein sequences from NCBI nr FASTA files (based on annotation taxonomy in Virus Kingdom). Candidate viral hits were then compared to an in-house nonvirus nonredundant (NVNR) protein database with an E-value cutoff of  $<10^{-5}$  to remove false-positive viral hits. The NVNR database was compiled using nonviral protein sequences extracted from NCBI nr FASTA files (based on annotation taxonomy excluding Virus Kingdom). Contigs without significant BLASTx similarity to the viral proteome database were searched against viral protein families in the vFam database (45) using HMMER3 (46–48) to detect remote viral protein similarities.

**Viral community analysis.** Composition similarity analysis of the six viromes was performed using MEGAN software (MEtaGenome Analyzer, version 6.21.7) (49) under the compare option. The results were presented by constructing an unweighted pair group method with arithmetic mean (UPGMA) taxonomic tree and using canonical correspondence analysis (CCA) under the cluster analysis option and the Bray-Curtis ecological distance matrix with default parameters. Analysis of similarities (ANOSIM) was used to compare differences between groups in viral communities using the R version 4.0.4 package *vegan* (version 2.5-7; <https://CRAN.R-project.org/package=vegan>). The viral community structure and richness results were visualized in a heatmap and UpSet plot, which were generated using the R version 4.0.4 package *phemap* (version 1.0.12; <https://CRAN.R-project.org/package=phemap>) and UpSetR (version 1.4.0; <https://CRAN.R-project.org/package=UpSetR>), respectively.

**Viral sequence extension and annotation.** Viral contigs that might have been from the same genome but without overlaps were merged using the Low Sensitivity/Fastest parameter in Geneious software version 11.1.2 (50). The individual contigs were used as the reference for mapping to the raw reads of their original barcodes using the Low Sensitivity/Fastest parameter. Putative viral open reading frames (ORFs) were



predicted by using Geneious version 11.1.2 with built-in parameters (minimum size, 300; genetic code, standard; start codons, ATG) (50) and were further checked through comparison to related viruses by BLASTp in NCBI. The annotations of these ORFs were based on comparisons to the Conserved Domain Database using RPS-BLAST with an E-value cutoff of  $<10^{-5}$ . Those contigs annotated with virus hallmark genes of the main virus groups were selected, among which those identified as complete ORFs were included for further phylogenetic analyses (the virus hallmark genes used were MCP for *Microviridae*, NS1 for *Parvoviridae*, Rep for CRESS DNA viruses, TerL for *Caudovirales*, and RdRp for *Riboviria*). All sequences with virus hallmark genes are presented in scatterplots drawn using R package ggplot2 (version 3.2.1; <https://ggplot2.tidyverse.org>).

**Phylogenetic analysis.** Phylogenetic analyses were performed based on the predicted protein sequences of virus hallmark genes identified in this study and protein sequences of reference strains belonging to different groups of viruses that were downloaded from the NCBI GenBank database. Related protein sequences were aligned using MUSCLE in MEGA version 10.1.8 (51) with the default settings. Sites containing more than 50% gaps were temporarily removed from alignments. Bayesian inference trees were then constructed using MrBayes version 3.2.7 (52). The Markov chain was run for a maximum of 1 million generations, sampled every 50 generations, and the first 25% of Markov chain Monte Carlo (MCMC) samples were discarded as burn-in. Maximum-likelihood trees were also constructed to confirm all the Bayesian inference trees, using MEGA software version 10.1.8 (51).

**Phage ARG characterization and functional annotation.** Contigs and singlet reads classified as phage sequences in the previous taxonomy annotation based on BLASTx were included for further ARG characterization and functional annotation. The ORF predictions for the phage sequences were measured in VirSorter with default settings (53). Then, the ORFs were used as queries to search for ARGs and functional genes using BLASTp against the Comprehensive Antibiotic Resistance Database (CARD) (54) and Kyoto Encyclopedia of Genes and Genomes (KEGG) database with an E-value cutoff of  $<10^{-5}$ . To quantify gene abundance, raw sequence reads were aligned to the ORFs using Bowtie 2 with default parameters (55). The numbers of read hits were then normalized to the library size and the coding length of each gene, manifested as parts per million (ppm) or fragments per kilobase per million (FPKM) (56, 57). The annotation information and the abundances of gene hits in each library are shown in Tables S3 and S4.

**Quality control.** As a blank control, sterile double-distilled water (ddH<sub>2</sub>O; Sagon, Shanghai, China) was prepared simultaneously and further processed under the same conditions. Particular attention was given to minimizing the risk of cross contamination and nucleic acid degradation. Aerosol filter pipet tips were used to avoid possible cross contamination among samples. All experimental materials (including microcentrifuge tubes, pipet tips, etc.) that directly contacted nucleic acid samples were RNase and DNase free, and all nucleic acid samples were dissolved in diethyl pyrocarbonate (DEPC)-treated water with RNase inhibitors added. All experimental processes were performed in a biological safety cabinet.

**Data availability.** The raw sequence read data analyzed in this study are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive database under accession numbers [SRX11831358](https://ncbi.nlm.nih.gov/sra/SRX11831358), [SRX11831382](https://ncbi.nlm.nih.gov/sra/SRX11831382), [SRX11831397](https://ncbi.nlm.nih.gov/sra/SRX11831397), [SRX11831406](https://ncbi.nlm.nih.gov/sra/SRX11831406), [SRX11831413](https://ncbi.nlm.nih.gov/sra/SRX11831413), [SRX11831441](https://ncbi.nlm.nih.gov/sra/SRX11831441), and [SRX11870785](https://ncbi.nlm.nih.gov/sra/SRX11870785) (control library). All viral sequences with virus hallmark genes identified in this study were deposited in the GenBank database under the accession numbers [MZ374764](https://ncbi.nlm.nih.gov/nucl/MZ374764), [MZ375112](https://ncbi.nlm.nih.gov/nucl/MZ375112), [MZ375114](https://ncbi.nlm.nih.gov/nucl/MZ375114), and [MZ375229](https://ncbi.nlm.nih.gov/nucl/MZ375229).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, PDF file, 1.6 MB.

## ACKNOWLEDGMENTS

This work was supported by National Key Research and Development Programs of China for Virome in Important Wildlife under grant no. 2017YFC1200201, Jiangsu Provincial Key Research and Development Projects under grant no. BE2017693, Donghai People's Hospital—School of Medicine UJS Joint Laboratory Foundation (grant number 20210467), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (grant number KYCX20\_1241).

Wen Zhang, Tongling Shan, and Quan Shen designed the study and methods. Peng Li, Shixing Yang, and Hao Wang constructed the libraries. Juan Lu, Wen Zhang, Chunmei Wang, Ga Gong, Yuan Xi, Jiamin Pan, Jian Zeng, Ju Zhang, and Xiaochun Wang completed the data analysis. The first draft was prepared by Juan Lu, and the manuscript was substantially reviewed and revised by all authors.

We declare no competing interests.

## REFERENCES

1. Jin Y, Brandt DY, Li J, Wo Y, Tong H, Shchur V. 2021. Elevation as a selective force on mitochondrial respiratory chain complexes of the Phrynocephalus lizards in the Tibetan plateau. *Curr Zool* 67:191–199. <https://doi.org/10.1093/cz/zoaa056>.
2. Jin YT, Brown RP. 2013. Species history and divergence times of viviparous and oviparous Chinese toad-headed sand lizards (*Phrynocephalus*) on the Qinghai-Tibetan Plateau. *Mol Phylogenet Evol* 68:259–268. <https://doi.org/10.1016/j.ympev.2013.03.022>.

3. Jin Y, Wo Y, Tong H, Song S, Zhang L, Brown RP. 2018. Evolutionary analysis of mitochondrially encoded proteins of toad-headed lizards, *Phrynocephalus*, along an altitudinal gradient. *BMC Genomics* 19:185. <https://doi.org/10.1186/s12864-018-4569-1>.
4. Sun Y, Liu S, Liu Y, Dong Y, Li M, An Y, Shi F, Beazley R. 2021. Effects of the interaction among climate, terrain and human activities on biodiversity on the Qinghai-Tibet Plateau. *Sci Total Environ* 794:148497. <https://doi.org/10.1016/j.scitotenv.2021.148497>.
5. Pan T, Zou X, Liu Y, Wu S, He G. 2017. Contributions of climatic and non-climatic drivers to grassland variations on the Tibetan Plateau. *Ecol Eng* 108:307–317. <https://doi.org/10.1016/j.ecoleng.2017.07.039>.
6. Manthey U. 2010. *Phrynocephalus theobaldi*. IUCN Red List Threat Species 2010:e.T170415A6779694. <https://doi.org/10.2305/IUCN.UK.2010-4.RLTS.T170415A6779694.en>. Accessed May 8, 2018.
7. Wang Y, Shi L, Huang S, Guo X. 2019. *Phrynocephalus erythrurus*. IUCN Red List Threat Species 2019:e.T107660333A107660347. <https://doi.org/10.2305/IUCN.UK.2019-2.RLTS.T107660333A107660347.en>. Accessed May 8, 2018.
8. Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, Cao C, Hu Q, Kim J, Larkin DM, Auvil L, Capitanu B, Ma J, Lewin HA, Qian X, Lang Y, Zhou R, Wang L, Wang K, Xia J, Liao S, Pan S, Lu X, Hou H, Wang Y, Zang X, Yin Y, Ma H, Zhang J, Wang Z, Zhang Y, Zhang D, Yonezawa T, Hasegawa M, Zhong Y, Liu W, Zhang Y, Huang Z, Zhang S, Long R, Yang H, Wang J, Lenstra JA, Cooper DN, Wu Y, Wang J, Shi P, Wang J, Liu J. 2012. The yak genome and adaptation to life at high altitude. *Nat Genet* 44:946–949. <https://doi.org/10.1038/ng.2343>.
9. Ge RL, Cai Q, Shen YY, San A, Ma L, Zhang Y, Yi X, Chen Y, Yang L, Huang Y, He R, Hui Y, Hao M, Li Y, Wang B, Ou X, Xu J, Zhang Y, Wu K, Geng C, Zhou W, Zhou T, Irwin DM, Yang Y, Ying L, Bao H, Kim J, Larkin DM, Ma J, Lewin HA, Xing J, Platt RN, II, Ray DA, Auvil L, Capitanu B, Zhang X, Zhang G, Murphy RW, Wang J, Zhang YP, Wang J. 2013. Draft genome sequence of the Tibetan antelope. *Nat Commun* 4:1858. <https://doi.org/10.1038/ncomms2860>.
10. Zhang Z, Xu D, Wang L, Hao J, Wang J, Zhou X, Wang W, Qiu Q, Huang X, Zhou J, Long R, Zhao F, Shi P. 2016. Convergent evolution of rumen microbiomes in high-altitude mammals. *Curr Biol* 26:1873–1879. <https://doi.org/10.1016/j.cub.2016.05.012>.
11. Wang Y, Zhou R, Yu Q, Feng T, Li H. 2020. Gut microbiome adaptation to extreme cold winter in wild plateau pika (*Ochotona curzoniae*) on the Qinghai-Tibet Plateau. *FEMS Microbiol Lett* 367:fnaa134. <https://doi.org/10.1093/femsle/fnaa134>.
12. Moeller AH, Ivey K, Cornwall MB, Herr K, Rede J, Taylor EN, Gunderson AR. 2020. The lizard gut microbiome changes with temperature and is associated with heat tolerance. *Appl Environ Microbiol* 86:e01181-20. <https://doi.org/10.1128/AEM.01181-20>.
13. Cadwell K. 2015. The virome in host health and disease. *Immunity* 42:805–813. <https://doi.org/10.1016/j.immuni.2015.05.003>.
14. Marschang RE. 2011. Viruses infecting reptiles. *Viruses* 3:2087–2126. <https://doi.org/10.3390/v3112087>.
15. Hoon-Hanks LL, Stohr AC, Anderson AJ, Evans DE, Nevarez JG, Diaz RE, Rodgers CP, Cross ST, Steiner HR, Parker RR, Stenglein MD. 2020. Serpentinovirus (Nidovirus) and Orthoreovirus coinfection in captive veiled chameleons (*Chamaeleo calytratus*) with respiratory disease. *Viruses* 12:1329. <https://doi.org/10.3390/v12111329>.
16. Virgin HW. 2014. The virome in mammalian physiology and disease. *Cell* 157:142–150. <https://doi.org/10.1016/j.cell.2014.02.032>.
17. von Wintersdorff CJ, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, Savelkoul PH, Wolfs PF. 2016. Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Front Microbiol* 7:173. <https://doi.org/10.3389/fmicb.2016.00173>.
18. Gillings MR, Stokes HW. 2012. Are humans increasing bacterial evolvability? *Trends Ecol Evol* 27:346–352. <https://doi.org/10.1016/j.tree.2012.02.006>.
19. Modi SR, Lee HH, Spina CS, Collins JJ. 2013. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 499:219–222. <https://doi.org/10.1038/nature12212>.
20. Colomer-Lluich M, Imamovic L, Jofre J, Muniesa M. 2011. Bacteriophages carrying antibiotic resistance genes in fecal waste from cattle, pigs, and poultry. *Antimicrob Agents Chemother* 55:4908–4911. <https://doi.org/10.1128/AAC.00535-11>.
21. Lekunberri I, Villagrasa M, Balcazar JL, Borrego CM. 2017. Contribution of bacteriophage and plasmid DNA to the mobilization of antibiotic resistance genes in a river receiving treated wastewater discharges. *Sci Total Environ* 601–602:206–209. <https://doi.org/10.1016/j.scitotenv.2017.05.174>.
22. Moon K, Jeon JH, Kang I, Park KS, Lee K, Cha CJ, Lee SH, Cho JC. 2020. Freshwater viral metagenome reveals novel and functional phage-borne antibiotic resistance genes. *Microbiome* 8:75. <https://doi.org/10.1186/s40168-020-00863-4>.
23. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, Yutin N, Zerbini FM, Kuhn JH. 2020. Global organization and proposed megataxonomy of the virus world. *Microbiol Mol Biol Rev* 84:e00061-19. <https://doi.org/10.1128/MMBR.00061-19>.
24. Koonin EV, Senkevich TG, Dolja VV. 2006. The ancient virus world and evolution of cells. *Biol Direct* 1:29. <https://doi.org/10.1186/1745-6150-1-29>.
25. Bonilla-Rosso G, Steiner T, Wichmann F, Bexkens E, Engel P. 2020. Honey bees harbor a diverse gut virome engaging in nested strain-level interactions with the microbiota. *Proc Natl Acad Sci U S A* 117:7355–7362. <https://doi.org/10.1073/pnas.2000228117>.
26. Wang H, Ling Y, Shan T, Yang S, Xu H, Deng X, Delwart E, Zhang W. 2019. Gut virome of mammals and birds reveals high genetic diversity of the family Microviridae. *Virus Evol* 5:vez013. <https://doi.org/10.1093/ve/vez013>.
27. Shkoporov AN, Clooney AG, Sutton TDS, Ryan FJ, Daly KM, Nolan JA, McDonnell SA, Khokhlova EV, Draper LA, Forde A, Guerin E, Velayudhan V, Ross RP, Hill C. 2019. The human gut virome is highly diverse, stable, and individual specific. *Cell Host Microbe* 26:527–541.e5. <https://doi.org/10.1016/j.chom.2019.09.009>.
28. Amato KR, Yeoman CJ, Kent A, Righini N, Carbonero F, Estrada A, Gaskins HR, Stumpf RM, Yildirim S, Torralba M, Gillis M, Wilson BA, Nelson KE, White BA, Leigh SR. 2013. Habitat degradation impacts black howler monkey (*Alouatta pigra*) gastrointestinal microbiomes. *ISME J* 7:1344–1353. <https://doi.org/10.1038/ismej.2013.16>.
29. Nelson TM, Rogers TL, Carlini AR, Brown MV. 2013. Diet and phylogeny shape the gut microbiota of Antarctic seals: a comparison of wild and captive animals. *Environ Microbiol* 15:1132–1145. <https://doi.org/10.1111/1462-2920.12022>.
30. Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, Costea PI, Godneva A, Kalka IN, Bar N, Shilo S, Lador D, Vila AV, Zmora N, Pevsner-Fischer M, Israeli D, Kosower N, Malka G, Wolf BC, Avnit-Sagi T, Lotan-Pompan M, Weinberger A, Halpern Z, Carmi S, Fu J, Wijemanga C, Zhenakova A, Elinav E, Segal E. 2018. Environment dominates over host genetics in shaping human gut microbiota. *Nature* 555:210–215. <https://doi.org/10.1038/nature25973>.
31. Wohlgemuth N, Honce R, Schultz-Cherry S. 2019. Astrovirus evolution and emergence. *Infect Genet Evol* 69:30–37. <https://doi.org/10.1016/j.meegid.2019.01.009>.
32. Smith AW, Anderson MP, Skilling DE, Barlough JE, Ensley PK. 1986. First isolation of calicivirus from reptiles and amphibians. *Am J Vet Res* 47:1718–1721.
33. Boyd EF. 2012. Bacteriophage-encoded bacterial virulence factors and phage-pathogenicity island interactions. *Adv Virus Res* 82:91–118. <https://doi.org/10.1016/B978-0-12-394621-8.00014-5>.
34. Simonson TS, Yang Y, Huff CD, Yun H, Qin G, Witherspoon DJ, Bai Z, Lorenzo FR, Xing J, Jorde LB, Prchal JT, Ge R. 2010. Genetic evidence for high-altitude adaptation in Tibet. *Science* 329:72–75. <https://doi.org/10.1126/science.1189406>.
35. Calero-Caceres W, Ye M, Balcazar JL. 2019. Bacteriophages as environmental reservoirs of antibiotic resistance. *Trends Microbiol* 27:570–577. <https://doi.org/10.1016/j.tim.2019.02.008>.
36. Lekunberri I, Subirats J, Borrego CM, Balcazar JL. 2017. Exploring the contribution of bacteriophages to antibiotic resistance. *Environ Pollut* 220:981–984. <https://doi.org/10.1016/j.envpol.2016.11.059>.
37. Wang M, Liu P, Zhou Q, Tao W, Sun Y, Zeng Z. 2018. Estimating the contribution of bacteriophage to the dissemination of antibiotic resistance genes in pig feces. *Environ Pollut* 238:291–298. <https://doi.org/10.1016/j.envpol.2018.03.024>.
38. Larranaga O, Brown-Jaque M, Quiros P, Gomez-Gomez C, Blanch AR, Rodriguez-Rubio L, Muniesa M. 2018. Phage particles harboring antibiotic resistance genes in fresh-cut vegetables and agricultural soil. *Environ Int* 115:133–141. <https://doi.org/10.1016/j.envint.2018.03.019>.
39. Kumar H, Park W, Lim D, Srikanth K, Kim JM, Jia XZ, Han JL, Hanotte O, Park JE, Oyola SO. 2020. Whole metagenome sequencing of cecum microbiomes in Ethiopian indigenous chickens from two different altitudes reveals antibiotic resistance genes. *Genomics* 112:1988–1999. <https://doi.org/10.1016/j.ygeno.2019.11.011>.
40. Zhang W, Li L, Deng X, Blümel J, Nübling CM, Hunfeld A, Baylis SA, Delwart E. 2016. Viral nucleic acids in human plasma pools. *Transfusion* 56:2248–2255. <https://doi.org/10.1111/trf.13692>.
41. Zhang W, Li L, Deng X, Kapusinszky B, Pesavento PA, Delwart E. 2014. Faecal virome of cats in an animal shelter. *J Gen Virol* 95:2553–2564. <https://doi.org/10.1099/vir.0.069674-0>.
42. Zhang W, Yang S, Shan T, Hou R, Liu Z, Li W, Guo L, Wang Y, Chen P, Wang X, Feng F, Wang H, Chen C, Shen Q, Zhou C, Hua X, Cui L, Deng X, Zhang Z, Qi

- D, Delwart E. 2017. Virome comparisons in wild-diseased and healthy captive giant pandas. *Microbiome* 5:90. <https://doi.org/10.1186/s40168-017-0308-0>.
43. Liu Z, Yang S, Wang Y, Shen Q, Yang Y, Deng X, Zhang W, Delwart E. 2016. Identification of a novel human papillomavirus by metagenomic analysis of vaginal swab samples from pregnant women. *Virology* 13:1–7. <https://doi.org/10.1186/s12985-016-0583-6>.
  44. Deng X, Naccache SN, Ng T, Federman S, Li L, Chiu CY, Delwart EL. 2015. An ensemble strategy that significantly improves de novo assembly of microbial genomes from metagenomic next-generation sequencing data. *Nucleic Acids Res* 43:e46. <https://doi.org/10.1093/nar/gkv002>.
  45. Skewes-Cox P, Sharpton TJ, Pollard KS, DeRisi JL. 2014. Profile hidden Markov models for the detection of viruses within metagenomic sequence data. *PLoS One* 9:e105067. <https://doi.org/10.1371/journal.pone.0105067>.
  46. Eddy SR. 2009. A new generation of homology search tools based on probabilistic inference. *Genome Inform* 23:205–211.
  47. Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39:W29–W37. <https://doi.org/10.1093/nar/gkr367>.
  48. Johnson LS, Eddy SR, Portugaly E. 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* 11: 431. <https://doi.org/10.1186/1471-2105-11-431>.
  49. Huson DH, Auch AF, Qi J, Schuster SC. 2007. MEGAN analysis of metagenomic data. *Genome Res* 17:377–386. <https://doi.org/10.1101/gr.5969107>.
  50. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.
  51. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>.
  52. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542. <https://doi.org/10.1093/sysbio/sys029>.
  53. Roux S, Enault F, Hurwitz BL, Sullivan MB. 2015. VirSorter: mining viral signal from microbial genomic data. *PeerJ* 3:e985. <https://doi.org/10.7717/peerj.985>.
  54. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, Kalan L, King AM, Koteva K, Morar M, Mulvey MR, O'Brien JS, Pawlowski AC, Piddock LJ, Spanogiannopoulos P, Sutherland AD, Tang I, Taylor PL, Thaker M, Wang W, Yan M, Yu T, Wright GD. 2013. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* 57:3348–3357. <https://doi.org/10.1128/AAC.00419-13>.
  55. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
  56. Chen B, Yang Y, Liang X, Yu K, Zhang T, Li X. 2013. Metagenomic profiles of antibiotic resistance genes (ARGs) between human impacted estuary and deep ocean sediments. *Environ Sci Technol* 47:12753–12760. <https://doi.org/10.1021/es403818e>.
  57. Bayard Q, Meunier L, Peneau C, Renault V, Shinde J, Nault JC, Mami I, Couchy G, Amaddeo G, Tubacher E, Bacq D, Meyer V, La Bella T, Debaillon-Vesque A, Bioulac-Sage P, Seror O, Blanc JF, Calderaro J, Deleuze JF, Imbeaud S, Zucman-Rossi J, Letouze E. 2018. Cyclin A2/E1 activation defines a hepatocellular carcinoma subclass with a rearrangement signature of replication stress. *Nat Commun* 9:5235. <https://doi.org/10.1038/s41467-018-07552-9>.