



**HAL**  
open science

## The influence of habitat on viral diversity in neotropical rodent hosts

Sourakhata Tirera, Benoit de Thoisy, Damien Donato, Christiane Bouchier,  
Vincent Lacoste, Alain Franc, Anne Lavergne

► **To cite this version:**

Sourakhata Tirera, Benoit de Thoisy, Damien Donato, Christiane Bouchier, Vincent Lacoste, et al..  
The influence of habitat on viral diversity in neotropical rodent hosts. *Viruses*, 2021, 13 (9), pp.1-29.  
10.3390/v13091690 . hal-03370479

**HAL Id: hal-03370479**

**<https://hal.inrae.fr/hal-03370479>**

Submitted on 8 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

## Article

# The Influence of Habitat on Viral Diversity in Neotropical Rodent Hosts

Sourakhata Tirera <sup>1</sup>, Benoit de Thoisy <sup>1</sup>, Damien Donato <sup>1</sup>, Christiane Bouchier <sup>2</sup>, Vincent Lacoste <sup>1,3,4</sup>, Alain Franc <sup>5,6</sup> and Anne Lavergne <sup>1,\*</sup>

<sup>1</sup> Laboratoire des Interactions Virus-Hôtes, Institut Pasteur de la Guyane, BP 6010, 97306 Cayenne, France; stirera@pasteur-cayenne.fr (S.T.); bdehoisy@pasteur-cayenne.fr (B.d.T.); ddonato@pasteur-cayenne.fr (D.D.); vincent.lacoste@pasteur.fr (V.L.)

<sup>2</sup> Institut Pasteur, 25-28 Rue du Docteur Roux, CEDEX 15, 75724 Paris, France; bouchier@pasteur.fr

<sup>3</sup> Département de Virologie, Institut Pasteur, 75015 Paris, France

<sup>4</sup> Arbovirus & Emerging Viral Diseases Laboratory, Institut Pasteur du Laos, Vientiane 3560, Laos

<sup>5</sup> UMR BIOGECO, INRAE, University Bordeaux, 33612 Cestas, France; alain.franc@inra.fr

<sup>6</sup> Pleiade, EPC INRIA-INRAE-CNRS, University Bordeaux, 33405 Talence, France

\* Correspondence: alavergne@pasteur-cayenne.fr

**Abstract:** Rodents are important reservoirs of numerous viruses, some of which have significant impacts on public health. Ecosystem disturbances and decreased host species richness have been associated with the emergence of zoonotic diseases. In this study, we aimed at (a) characterizing the viral diversity in seven neotropical rodent species living in four types of habitats and (b) exploring how the extent of environmental disturbance influences this diversity. Through a metagenomic approach, we identified 77,767 viral sequences from spleen, kidney, and serum samples. These viral sequences were attributed to 27 viral families known to infect vertebrates, invertebrates, plants, and amoeba. Viral diversities were greater in pristine habitats compared with disturbed ones, and lowest in peri-urban areas. High viral richness was observed in savannah areas. Differences in these diversities were explained by rare viruses that were generally more frequent in pristine forest and savannah habitats. Moreover, changes in the ecology and behavior of rodent hosts, in a given habitat, such as modifications to the diet in disturbed vs. pristine forests, are major determinants of viral composition. Lastly, the phylogenetic relationships of four vertebrate-related viral families (*Polyomaviridae*, *Flaviviridae*, *Togaviridae*, and *Phenuiviridae*) highlighted the wide diversity of these viral families, and in some cases, a potential risk of transmission to humans. All these findings provide significant insights into the diversity of rodent viruses in Amazonia, and emphasize that habitats and the host's dietary ecology may drive viral diversity. Linking viral richness and abundance to the ecology of their hosts and their responses to habitat disturbance could be the starting point for a better understanding of viral emergence and for future management of ecosystems.

**Keywords:** virome; alpha diversity; rodents; Amazonia; viral ecology; viral phylogenies



**Citation:** Tirera, S.; de Thoisy, B.; Donato, D.; Bouchier, C.; Lacoste, V.; Franc, A.; Lavergne, A. The Influence of Habitat on Viral Diversity in Neotropical Rodent Hosts. *Viruses* **2021**, *13*, 1690. <https://doi.org/10.3390/v13091690>

Academic Editor: Ayato Takada

Received: 2 August 2021

Accepted: 15 August 2021

Published: 26 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Viruses have conquered all living systems, infecting other microbes (bacteria, fungi, and parasites) and more complex organisms, such as plants, invertebrates, and vertebrates. Most viruses have small genomes with high mutation rates [1], giving them the ability to evolve and adapt quickly to new environments and potentially the ability to infect new hosts.

The development of metagenomic approaches applied to viruses (viromics) [2,3] has improved our knowledge of the extent of viral diversity and of the host spectra of several viral families. This is, for instance, the case for hepaciviruses, with both the descriptions of novel viral species in mammals and the evidence of infection in non-mammal species [4,5]. Viromic studies have also led to the discovery of new viral genotypes, helping us understand their evolutionary history [6] and providing new insights into the roles of viruses in

ecosystems [7–9]. Despite these advances, a large number of viral species remain unknown and many biomes continue to be under- or unexplored [10]. In the context of natural habitat disturbances, the disruptions of population dynamics and ecology that favor contacts between species, cross-species transmissions, spill-over, amplification, spread of viruses, and increased contact with wild fauna may lead, under certain circumstances, to the emergence of infectious diseases (EIDs) in human populations [11–13]. It is now well established that more than 70% of EIDs originated from animals and mainly from wildlife [14]. Recent examples of virus spill-overs, such as the 2002 severe acute respiratory syndrome (SARS) in China and the 2012 Middle East respiratory syndrome (MERS) in Saudi Arabia, or even the 2014 Ebola epidemic in West Africa, have had severe public health and economic consequences [15,16]. The impact of the COVID-19 pandemic, though yet to be assessed, appears to be far more damaging [17]. Hence, both the identification of potentially zoonotic viruses, and the understanding of their transmission mechanisms and their ecological contexts, have gained attention in the last few decades [18], as illustrated by the PREDICT global surveillance program [19] and the Global Virome Project [20,21]. Despite limited knowledge of the processes of zoonotic viral emergence, several groups of species, such as primates, birds, bats, and rodents, are described as major reservoirs due to their ability to host wide ranges of viruses [22,23].

*Rodentia* is composed of approximately 2277 species [24,25] and is the most speciose mammalian order. They occupy most of the earth's ecosystems, from highly anthropized to pristine natural habitats. The richness of rodent species, along with certain traits such as their fast-paced lives, population dynamics, opportunism, and synanthropism makes them efficient amplifiers, spreaders, and transmitters of viruses [23,26–28]. This is illustrated by the 31 mammalian viral families reported from the DRodVir online database, as of 21 July 2021 [29,30]. Rodents also play a key role in viral emergence phenomena, with many species being reservoirs and carriers of viral pathogens, for example, hantaviruses or mammarenaviruses, some of which are associated with human diseases [31–35].

Amazonia is known for its high biodiversity of mammals, plants, invertebrates, and microbes [36–38], comprising a large set of potential hosts for viruses. This implies that high viral diversity is expected from the region and thus a high number of potential zoonotic viruses [39,40]. Despite this context, Amazonia remains one of the least explored areas in terms of viral diversity, and to date, the few studies investigating the viral diversity of the main reservoir species have only been conducted with bats [8,41,42]. In French Guiana, 36 rodent species are present, and live in different habitats (various types of forest, savannah, agricultural areas, peri-urban zones, and urban zones). Among them, several species have been described to host viruses relevant to human health [33]. As an example, hantaviruses, belonging to the *Bunyavirales* order, are widespread worldwide and in South America cause a cardiopulmonary syndrome with a high mortality rate. In French Guiana, six human cases of cardiopulmonary syndrome have been identified to date, four of which were fatal [43–45]. Screening rodents captured in more likely areas of contamination made it possible to identify two species (*Zygodontomys brevicauda* and *Oligoryzomys delicatus*) as potential reservoirs of this virus [32]. To date, no human cases of mammarenavirus infection have been registered in French Guiana, although two members of this viral family (*Arenaviridae*), which in severe cases can cause hemorrhagic fevers and meningitis, were recently identified in rodents [46,47].

We first explored the viral diversity of seven Neotropical rodent species from two families: Echimyidae (*Proechimys cuvieri* and *P. guyannensis*) and Cricetidae (*Zygodontomys brevicauda*, *Oecomys bicolor*, *O. auyantepui*, *Hylaeamys yunganus*, and *Hylaeamys megacephalus*) to determine how viral diversity is distributed across species and habitats. Second, we focused on the phylogenetic relationships of viral sequences related to four viral families known to infect vertebrates including arthropod-borne viruses (*Polyomaviridae*, *Flaviviridae*, *Togaviridae*, and *Phenuiviridae*). These results allowed us to identify novel vertebrate-related viral sequences and shed new light on the role of habitats in shaping viral diversity in Amazonian rodent species.

## 2. Materials and Methods

### 2.1. Ethical and Legal Statements

All animals were captured, handled, and sampled following ASM guidelines under the supervision of researchers granted the French animal experimentation level 1 diploma [48]. The use of genetic resources was approved by the French Ministry of the Environment under reference number ABSCH-IRCC-FR-252439-1, 9 June 2020, in compliance with the Access and Benefit Sharing procedure implemented by the *Loi pour la Reconquête de la Biodiversité*.

### 2.2. Capturing

Within the framework of a long-term project dedicated to assessing the role of wildlife in the emergence of different viruses in French Guiana, rodents were sampled across the region from a total of 12 sites during the 2001–2014 period. Four environments were considered: disturbed and pristine forests, savannahs, and peri-urban habitats, according to the 2015 land use classification [49]. This classification was based on the Europe CORINE (coordinated information on the environment) Land Cover methodology [50] but was adapted to the country, since forms of urbanization and natural habitat categories differ greatly between continental Europe and French Guiana.

According to the sites, 20–30 trap-lines were organized in successive stations spaced 20 m apart with four traps at each station: two Sherman traps (Sherman Trap Co., Tallahassee, FL, USA) and two BTTm traps (BTTm, Besançon Trap Service mécanique, Besançon, France). The traps were baited with apples and checked every morning. The mean sampling effort per site was  $1016 \pm 420$  traps/night, for a total effort of 28,470 traps/night and more than 500 rodent captures (belonging to 19 species) [33]. Rodents were caught alive, brought back to the laboratory facilities, anesthetized, and sampled for blood and/or euthanized to preserve their organs (kidney and spleen). When necessary, sacrifice was done chemically with barbituric acid after anesthesia, one of the recommended methods for wild-caught rodents, according to the AVMA Guidelines for the Euthanasia of Animals, 2020 Edition. The animals were identified using external morphological features and confirmed molecularly using the cytochrome oxidase I sequence [51,52]. Seven species were selected for the virome analysis according to their presence in distinct habitats: *P. cuvieri*, *P. guyannensis*, *Z. brevicauda*, *O. bicolor*, *O. auyantepui*, *H. yunganus*, and *H. megacephalus* (Table 1).

### 2.3. Sample Processing

Prior to processing, samples from the same species, the same organs, and the same environment were pooled (e.g., the sample k\_Pguy\_PF is a pool of kidneys from *P. guyannensis* collected in pristine forest), resulting in 36 different pools, i.e., 36 different viromes. Pools, at the organ level, included 2–31 individuals according to sample availability. Overall, 442 organs and sera from 187 individuals were included in this study (Table 1).

For serum samples, 50  $\mu$ L from each collecting tube was used to constitute the pools. For kidney and spleen samples, 100 mg of organs was crushed in 400  $\mu$ L of DMEM, and 200  $\mu$ L of suspension from each collecting tube was used to constitute the pools. All pools were processed as previously described [41]: Pools were cleared of debris by low-speed centrifugation (5 min,  $10,000 \times g$ , 4 °C). Eukaryotic and prokaryotic cell-sized particles were removed from supernatants through three successive filtrations (0.8, 0.45, and 0.22  $\mu$ m), using cellulose acetate membrane filters (Nalgene). The filtrates were cleared of persistent high-density particles with low-speed centrifugation (15 min,  $10,000 \times g$ , 4 °C); then, viral particles were pelleted with a 1-h ultracentrifugation step ( $100,000 \times g$ , 4 °C). All viral pellets were resuspended in 40  $\mu$ L of nuclease-free water.

**Table 1.** Sampling data of individuals by organ, species, and habitat. The composition of each pool, including organ, species, habitat, and number of individuals, is given. The last column corresponds to the total number of different individuals used to constitute each pool of each species in a given habitat. The last row corresponds to the totals of organs and individual samples for this study.

Pool Number	Sample ID	Organ	Species	Habitat	Nb. Individuals	Total of Different Individuals
1	k_Pguy_PF	Kidney	<i>Proechimys guyannensis</i>	Pristine Forest	5	12
2	r_Pguy_PF	Spleen	<i>Proechimys guyannensis</i>	Pristine Forest	8	
3	s_Pguy_PF	Sera	<i>Proechimys guyannensis</i>	Pristine Forest	12	
4	k_Pguy_DF	Kidney	<i>Proechimys guyannensis</i>	Disturbed Forest	24	29
5	r_Pguy_DF	Spleen	<i>Proechimys guyannensis</i>	Disturbed Forest	27	
6	s_Pguy_DF	Sera	<i>Proechimys guyannensis</i>	Disturbed Forest	27	
7	k_Pcuv_PF	Kidney	<i>Proechimys cuvieri</i>	Pristine Forest	9	35
8	r_Pcuv_PF	Spleen	<i>Proechimys cuvieri</i>	Pristine Forest	24	
9	s_Pcuv_PF	Sera	<i>Proechimys cuvieri</i>	Pristine Forest	31	
10	k_Pcuv_DF	Kidney	<i>Proechimys cuvieri</i>	Disturbed Forest	18	18
11	r_Pcuv_DF	Spleen	<i>Proechimys cuvieri</i>	Disturbed Forest	18	
12	s_Pcuv_DF	Sera	<i>Proechimys cuvieri</i>	Disturbed Forest	15	
13	k_Oaui_DF	Kidney	<i>Oecomys auyantepui</i>	Disturbed Forest	5	14
14	r_Oaui_DF	Spleen	<i>Oecomys auyantepui</i>	Disturbed Forest	11	
15	s_Oaui_DF	Sera	<i>Oecomys auyantepui</i>	Disturbed Forest	13	
16	k_Obic_DF	Kidney	<i>Oecomys bicolor</i>	Disturbed Forest	13	15
17	r_Obic_DF	Spleen	<i>Oecomys bicolor</i>	Disturbed Forest	14	
18	s_Obic_DF	Sera	<i>Oecomys bicolor</i>	Disturbed Forest	12	
19	k_Hyun_DF	Kidney	<i>Hylaeamys yunganus</i>	Disturbed Forest	11	11
20	r_Hyun_DF	Spleen	<i>Hylaeamys yunganus</i>	Disturbed Forest	11	
21	s_Hyun_DF	Sera	<i>Hylaeamys yunganus</i>	Disturbed Forest	10	
22	k_Hmeg_PF	Kidney	<i>Hylaeamys megacephalus</i>	Pristine Forest	2	5
23	r_Hmeg_PF	Spleen	<i>Hylaeamys megacephalus</i>	Pristine Forest	4	
24	s_Hmeg_PF	Sera	<i>Hylaeamys megacephalus</i>	Pristine Forest	3	
25	k_Hmeg_DF	Kidney	<i>Hylaeamys megacephalus</i>	Disturbed Forest	4	5
26	r_Hmeg_DF	Spleen	<i>Hylaeamys megacephalus</i>	Disturbed Forest	4	
27	s_Hmeg_DF	Sera	<i>Hylaeamys megacephalus</i>	Disturbed Forest	4	

Table 1. Cont.

Pool Number	Sample ID	Organ	Species	Habitat	Nb. Individuals	Total of Different Individuals
28	k_Zbre_SV	Kidney	<i>Zygodontomys brevicauda</i>	Savannah	6	12
29	r_Zbre_SV	Spleen	<i>Zygodontomys brevicauda</i>	Savannah	6	
30	s_Zbre_SV	Sera	<i>Zygodontomys brevicauda</i>	Savannah	8	
31	k_Zbre_DF	Kidney	<i>Zygodontomys brevicauda</i>	Disturbed Forest	7	10
32	r_Zbre_DF	Spleen	<i>Zygodontomys brevicauda</i>	Disturbed Forest	7	
33	s_Zbre_DF	Sera	<i>Zygodontomys brevicauda</i>	Disturbed Forest	10	
34	k_Zbre_PU	Kidney	<i>Zygodontomys brevicauda</i>	Peri-urban	20	21
35	r_Zbre_PU	Spleen	<i>Zygodontomys brevicauda</i>	Peri-urban	20	
36	s_Zbre_PU	Sera	<i>Zygodontomys brevicauda</i>	Peri-urban	19	
Totals		-	-	-	442	187

### 2.3.1. Nuclease Treatment and Viral Nucleic Acid Extraction

All resuspended viral pellets were treated with a mixture of DNases (Turbo DNase from Ambion and Benzonase from Novagen) and RNase One (Promega) to digest non-enveloped nucleic acids (i.e., those not in viral capsids) [53]. All viral nucleic acids were then extracted using the NucliSENS easyMAG<sup>®</sup> bio-robot (bioMérieux).

### 2.3.2. Reverse Transcription and Amplification

For each pool, the RNA virus-only and DNA virus-only libraries were constructed using a whole transcriptome (WTA) or a whole genome (WGA) amplification method, respectively, as previously described [41].

### 2.4. Next-Generation Sequencing

For each pool, 1 µg from each library was pooled together, whenever possible, to construct RNA plus DNA viral libraries. High-throughput sequencing was carried out at the genomics center of the Institut Pasteur, Paris. Shotgun libraries were prepared by standard Illumina protocols using 1 µg of total genomic DNA. Each sample (sera, kidney, and spleen) was tagged according to its provenance (species, organs, and habitats) using Illumina adaptor-specific primers. High-throughput shotgun metagenomic sequencing was carried out in two different sessions. The first round was completed using an Illumina MiSeq platform with 300-bp paired-end reads (eight samples, see Table 2). The second round was performed with an Illumina HiSeq 2500 platform with 250-bp paired-end reads (28 samples, see Table 2).

### 2.5. Bioinformatic Analyses

Globally, after a first cleaning step, reads were assembled de novo (Step 1, Figure 1). Then, clean reads were mapped back to contigs in order to obtain the number of reads aligned to each contig (Step 2, Figure 1). The taxonomic assignment of contigs was achieved through BLASTn and BLASTx (Step 3, Figure 1). Finally, a matrix corresponding to the number of viral reads at the genus/subfamily level for each species–habitat was built for statistical analysis (Step 4, Figure 1).

All sequences were submitted to FaQCs [54] with an automated search for PhiX sequences, and quality filtering, after removal of adapters and poly-A tails (“-phiX yes -adapter yes -polyA yes”). The resulting clean sequence files were submitted to de novo assembly with MEGAHIT [55,56] using default parameters, which provide a set of k-mers from 21 to 141 with a step of 12 used in the assembly process, and the minimum contig length was set at 200 nucleotides (Step 1, Figure 1). Then cleaned reads were mapped back

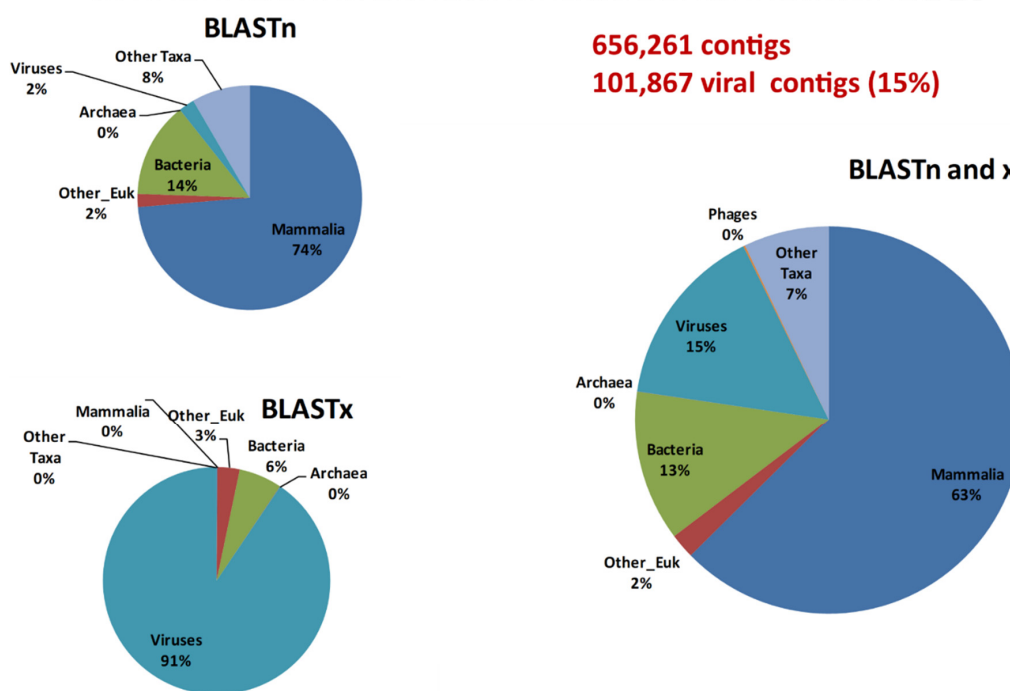
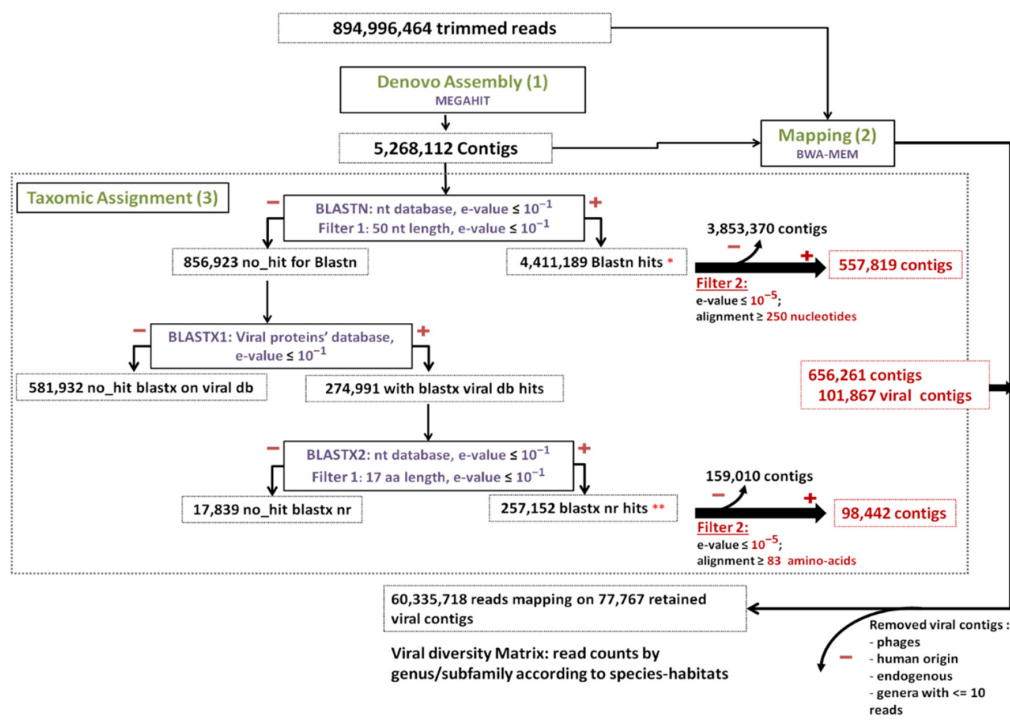
to the contigs using the BWA-MEM mapper [57] and Samtools [58] in order to obtain the number of reads for each contig (Step 2, Figure 1).

**Table 2.** Sequencing and bioinformatic processing data by organ, species, and habitat. It shows the number of raw reads, the corresponding Illumina Platform, the percentage of reads kept after trimming and cleansing, the number of assembled contigs, the percentage of cleaned reads mapped back to contigs, and the number of viral contigs.

Sample ID	No. of Raw Reads	Sequencing Type	% Cleaned Reads	No. of Contigs	% Mapped Reads	No. of Viral Contigs
k_Pguy_PF	39,505,174	Hiseq2500	98.12	323,113	97.81	101
k_Pguy_DF	24,142,444	Hiseq2500	98.99	143,955	97.88	161
k_Pcuv_PF	39,878,026	Hiseq2500	98.15	267,539	98.33	398
k_Pcuv_DF	32,782,446	Hiseq2500	99.56	88,325	99.30	4783
k_Oaui_DF	43,211,478	Hiseq2500	98.01	225,793	98.36	86
k_Obic_DF	3,999,474	MiSeq	97.55	31,337	98.92	2208
k_Hmeg_PF	35,419,066	Hiseq2500	98.50	285,067	97.96	71
k_Hmeg_DF	6,259,058	MiSeq	96.62	53,664	98.99	2998
k_Hyun_DF	31,133,790	Hiseq2500	98.48	159,779	98.77	142
k_Zbre_PU	34,912,952	Hiseq2500	98.48	181,251	98.79	150
k_Zbre_DF	6,754,874	MiSeq	98.35	56,113	98.75	1698
k_Zbre_SV	4,395,890	MiSeq	98.16	25,215	99.25	1656
r_Pguy_PF	40,349,120	Hiseq2500	98.01	330,777	97.87	79
r_Pguy_DF	29,988,950	Hiseq2500	98.22	231,108	97.69	125
r_Pcuv_PF	16,840,948	Hiseq2500	98.77	124,874	97.58	384
r_Pcuv_DF	27,294,898	Hiseq2500	98.28	267,976	97.05	909
r_Oaui_DF	33,188,412	Hiseq2500	98.57	184,913	98.19	109
r_Obic_DF	5,964,414	MiSeq	98.74	67,976	98.40	3506
r_Hmeg_PF	35,586,626	Hiseq2500	98.22	282,402	98.10	121
r_Hmeg_DF	7,633,778	MiSeq	97.32	53,117	99.38	1109
r_Hyun_DF	33,617,438	Hiseq2500	98.54	159,882	98.40	317
r_Zbre_PU	44,909,868	Hiseq2500	99.02	176,663	99.18	13,343
r_Zbre_DF	5,029,582	MiSeq	97.65	59,922	98.53	2591
r_Zbre_SV	4,296,504	MiSeq	95.59	79,251	97.40	617
s_Pguy_PF	20,326,680	Hiseq2500	98.86	56,461	98.83	2115
s_Pguy_DF	25,433,042	Hiseq2500	99.49	54,835	99.01	3473
s_Pcuv_PF	30,224,896	Hiseq2500	98.99	86,555	99.30	12,162
s_Pcuv_DF	25,361,476	Hiseq2500	98.15	347,957	96.80	3341
s_Oaui_DF	18,055,176	Hiseq2500	98.00	71,707	98.84	255
s_Obic_DF	26,895,348	Hiseq2500	99.23	103,205	98.55	2114
s_Hmeg_PF	32,497,436	Hiseq2500	99.48	75,533	99.19	248
s_Hmeg_DF	22,859,508	Hiseq2500	98.62	70,539	99.04	1862
s_Hyun_DF	24,873,840	Hiseq2500	98.78	258,141	96.09	6109
s_Zbre_PU	26,110,716	Hiseq2500	99.35	87,791	98.60	10,330
s_Zbre_DF	37,050,082	Hiseq2500	99.09	103,974	99.07	13,622
s_Zbre_SV	30,948,566	Hiseq2500	98.94	91,402	98.89	8574
Total	907,731,976	-	-	5,268,112	-	101,867

We used BLAST [59,60] for the taxonomic assignment of contigs (Step 3, Figure 1). All contigs were submitted to DISCONTIGOUS MEGABLAST BLASTn (e-value  $\leq 10^{-1}$ ). Those without results in BLASTn were submitted to BLASTx. The BLASTx process comprised two steps: (1) using an in-house viral protein database, which was created by clustering (CD-HIT, 100% homology) the NCBI-nr database viral protein sequences (August 2018); (2) the set of positive contigs was subsequently submitted to BLASTx (e-value  $\leq 10^{-1}$ ) against the whole NCBI-nr database. Both BLASTn and BLASTx results were filtered using in-house python scripts, which selected the best scoring match for each contig (max e-value of  $10^{-1}$  and coverage of, respectively, >50 nucleotides and 17 amino acids in length for BLASTn and BLASTx) (Filter 1). Taxonomies were deduced from the BLAST results and fragments were assigned to selected viruses matching Taxids against the full name

lineages file from NCBI (f) [61]. These two data sets were filtered again (Filter 2) according to the e-values and coverage according to both BLASTn and BLASTx (e-value =  $10^{-5}$ ; coverage  $\geq 250$  nt or 83 amino acids in length) so as to consolidate the results. The remaining contigs were used for counting in the results. To provide a more accurate definition of the virus’s taxonomic status, each viral taxonomic identification was associated with host types, such as bacteria, vertebrates, invertebrates, amoeba, and plants, by consulting the ICTV [62], ViralZone, and virus host database websites [63,64].



**Figure 1.** Main bioinformatic processing steps, data streams, and resulting taxonomic categories. “Other\_Euk”: Other (non-mammalian eukaryotes).



As the presence of bacteriophages is unexpected in compartments such as the kidney, spleen, and blood, and it is difficult to attribute such bacteriophages to bacterial infections, contigs assigned to bacteriophages were discarded from the data set. In addition, a manual inspection was carried out and viruses known to be amplified in the laboratory (*Herpesviridae* and *Papillomaviridae*) and endogenous virus (*filovirus*) were discarded from the data set. Finally, based on mapping data, viral taxa (at the genus/subfamily level) identified with  $\leq 10$  reads were discarded from the data set in order to avoid the spurious presence of contigs due to potential contamination.

From these data, a data set associating the number of contigs with viral families was created. The number of viral families (categorized according to host type) associated with each species and its habitats was plotted and a heatmap was developed (Rstudio, “pheatmap” library). The heatmap represents the number of contigs associated with each viral family by species–habitats. The viral genomes’ completeness of assigned contigs was tested using CHECKV (version v0.7.0) and its associated database [65].

Lastly, a quantitative data set associating each contig with its number of reads was constructed with the number of reads associated with each taxonomic category at the subfamily or genus level.

The main steps are shown in solid-line boxes in green font characters, with the number in brackets representing the corresponding step in order of execution (de novo assembly, read mapping to contigs, and taxonomic assignment).

BLASTn and BLASTx, as sub-steps of taxonomic assignment, are shown in solid-line boxes with purple font characters. For each output, (+) and (–) stand for positive/kept and negative/discarded results, respectively, where they appear; taxonomic information acquisition and the filter-1 step are represented by red stars.

Subsequent filtering of contigs based on host (phages/human), endogenous status, and number of reads was carried out to obtain the final data set of both contigs and the corresponding reads matrix.

The taxonomic categories (Eukaryotes, Mammalia, Bacteria, Archaea, Viruses) to which contigs were assigned are shown as pie charts separately for BLASTn and BLASTx, and they were merged (BLASTn and x).

## 2.6. Statistical Analysis

To assess whether or not viral diversity is related to habitat type, diversity analyses were run for the four species present in at least two different habitats (pristine and disturbed forest for *P. guyannensis*, *P. cuvieri*, and *H. megacephalus*; disturbed forest, savannah, and peri-urban areas for *Z. brevicauda*) using the number of mapped reads on contigs assigned to each viral genus/subfamily (Supplemental data, Table S1).

As a prerequisite, before assessing local alpha diversities for the nine species–habitat combinations, we tested that (i) richness was not related to sequencing type using the Welch two-sample *t*-test; (ii) the number of viruses detected (genus/subfamily level) was not impacted by the number of individuals in each sequencing pool using a Pearson correlation test (R standard library) (Supplemental data, Table S2); (iii) each pool had been sufficiently sequenced to represent its diversity using rarefaction curves with the *rarefy* and *rarecurve* functions (Vegan R package) [66] with sampling at 231,725 reads (minimum number of viral reads across samples) and a step of 2000 reads (Supplemental data, Figure S1).

We then computed diversity indices for each of the nine species and habitat combinations and compared viral diversities, species by species, in their respective habitats. A synthesis between the most commonly accepted indices has been proposed [67] as a family of indices inspired in statistical physics [68], extending the link between diversity and entropy [69]. Based on  $\alpha$ , related to Rényi’s entropy,  $H_\alpha$  by  $N_\alpha = \exp(H_\alpha)$ , three measures of diversity can be recovered: The total number of species (richness,  $\alpha = 0$ ), Shannon’s entropy ( $\alpha = 1$ ), and the inverse of Simpson’s dominance index ( $\alpha = 2$ ) [67,70]. The lower the  $\alpha$  value is, the higher the weight given to rare species. The Hill  $\alpha$ -diversity index values were generated with  $\alpha$  in [0, 0.25, 0.5, 0.75, 1, 2], using the Vegan R package. A community

A can be considered as more diverse than a community B if all Rényi's entropy values for A are higher than for B when  $\alpha$  runs over a given range (here  $0 < \alpha < 2$ ) [71]. Therefore, to enhance comparisons between samples, we plotted Rényi's entropy instead of Hill's diversity for a given value of  $\alpha$  followed by comparisons according to the parameters above [71].

Furthermore, to highlight the importance of the abundance of viral taxa (i.e., the number of reads per viral genus/subfamily) and to explore how rare species have contributed to diversity according to the habitat, we recomputed richness by progressive deletion of taxa with the lowest number of reads (i.e., setting their counts to 0 if they were under a defined threshold T). Thereafter, we quantified the impacts of these deletions on richness by calculating, for each case, the difference between the value of richness without deletion ( $R_0$ ) and the richness value after suppression ( $R_T$ , richness value after deletion under T threshold): The higher the difference ( $R_0 - R_T$ ), the more numerous the rare species. The thresholds for deletion of a taxon were selected as fractions of the sample of the smallest size, i.e., fractions of 200,000 reads. We deleted all taxa with numbers of reads smaller than 0.01, 0.1, 1, 2, and 5% of this value, successively (i.e., 20, 200, 2000, 4000, 10,000) in all nine species-habitat combinations and calculated richness at each step.

### 2.7. Phylogenetic Analyses

Contigs assigned to four viral families (*Polyomaviridae*, *Flaviviridae*, *Togaviridae*, and *Phenuiviridae*) were selected from assembly files. Contigs were checked manually using the NCBI BLASTx web tool for the presence of stop codons. Then sequences presenting stop codons were deleted from further analysis. In order to infer phylogenetic relationships, we selected the closest sequences provided by BLAST during the taxonomic assignment process and some representatives of the family to be analyzed. In addition, for the *Polyomaviridae* and *Flaviviridae* analyses, sequences detected in rodents worldwide were also included. For the *Alphavirus* and *Phlebovirus* analysis, the data set including the closest sequences provided by BLAST was completed by sequences representative of the known antigenic complexes. All reference sequences were downloaded from the NCBI-nt database.

Accession numbers of viral sequences used to infer the phylogenetic trees are given in the respective phylogenetic reconstructions. After selecting the best-suited region for phylogenetic analyses, the Muscle algorithm [72] was used for multiple sequence alignments with default parameters. Pairwise sequence identity (at the nucleotide and amino acid levels) for each selected region were calculated using uncorrected p-distances. The best-fitted model of nucleotide or amino acid substitution for each analysis was selected using jModelTest 2 [73] and ProtTest 3 [74], respectively, under corrected Akaike information criteria (AICc). Bayesian phylogenetic analyses were performed using MrBayes 3.2 [75,76]. The Markov chain Monte Carlo (MCMC) algorithm was run with four chains with 2 million generations each, with trees sampled every 500 generations and a 25% burn-in. Validation of the inference was assessed based on the standard deviation of split frequencies, less than the expected threshold value of 0.01 in MrBayes and by inspecting the effective sampling size ( $ESS > 500$ ) criterion in Tracer version 1.6 [77].

### Nucleotide Sequence Accession Numbers

All virus sequences reported in this study were deposited in the GenBank nucleotide database under accession numbers MT732099 to MT732117. The data from Illumina sequencing were deposited in the GenBank Sequence Reads Archive under accession numbers SAMN15496919 to SAMN15496954.

## 3. Results

### 3.1. Illumina Sequencing and Bioinformatics Analyses

Overall, 453,865,988 paired-end raw reads (907,731,976 individual reads) were obtained that were 250–300 bp in length. After trimming by FaQcs, 96.62–99.56% of the reads were kept, totaling 894,996,464 reads.

These cleaned reads were assembled de novo for a total of 5,268,112 contigs using MEGAHIT (Figure 1 and Supplemental data, Table S3 for assembly statistics). The number of contigs ranged from 25,215 to 347,957. The mean number was 146,336 contigs/sample (Table 2). All these contigs were submitted to taxonomic assignment.

After megaBLASTn, 4,411,189 (83.73%) contigs were assigned to an organism, among which 50,431 were attributed to viruses. The remaining 856,923 (16.27%) unassigned contigs were first submitted to a BLASTx search against the in-house viral protein database (BLASTx1) (Figure 1). Only 274,991 contigs matched viral proteins, and they were all put through a second BLASTx step against the entire nr protein database (BLASTx2) (Figure 1). Overall, after this second BLASTx step (BLASTx2), 257,152 contigs were re-assigned to viruses, 3776 were assigned to other types of organisms, and 17,839 were not assigned at all (Figure 1).

To further avoid artifacts and false-positive results, the virus-assigned contigs were filtered at a coverage of  $\geq 250$  bp for BLASTn or  $\geq 83$  amino acids for BLASTx results and an e-value of  $\leq e^{-5}$  (Filter 2), resulting in 101,867 viral contigs, accounting for 1.93% of the total initial number of contigs and 15% of the assigned contigs (Figure 1).

Among the 101,867 viral contigs, 945 assigned to bacteriophages were discarded from further analyses. After a manual inspection, 20 sequences assigned to the *Herpesviridae* (human herpesviruses;  $>97\%$  nucleotide identity), *Papillomaviridae* (human papillomaviruses;  $>90\%$  nucleotide identity), and *Filoviridae* (endogenous) families were also removed (Figure 1). After removal of contigs with less than 11 reads mapped, 77,767 viral contigs were kept for further analyses. Among them, 118 contigs submitted to CHECKV were high quality, with completeness above 90%, and 72, 23, and 23 contigs were assigned to *Anelloviridae*, *Circoviridae*, and *Genomoviridae*, respectively.

### 3.2. Viral Diversity Detected through Species and Environments

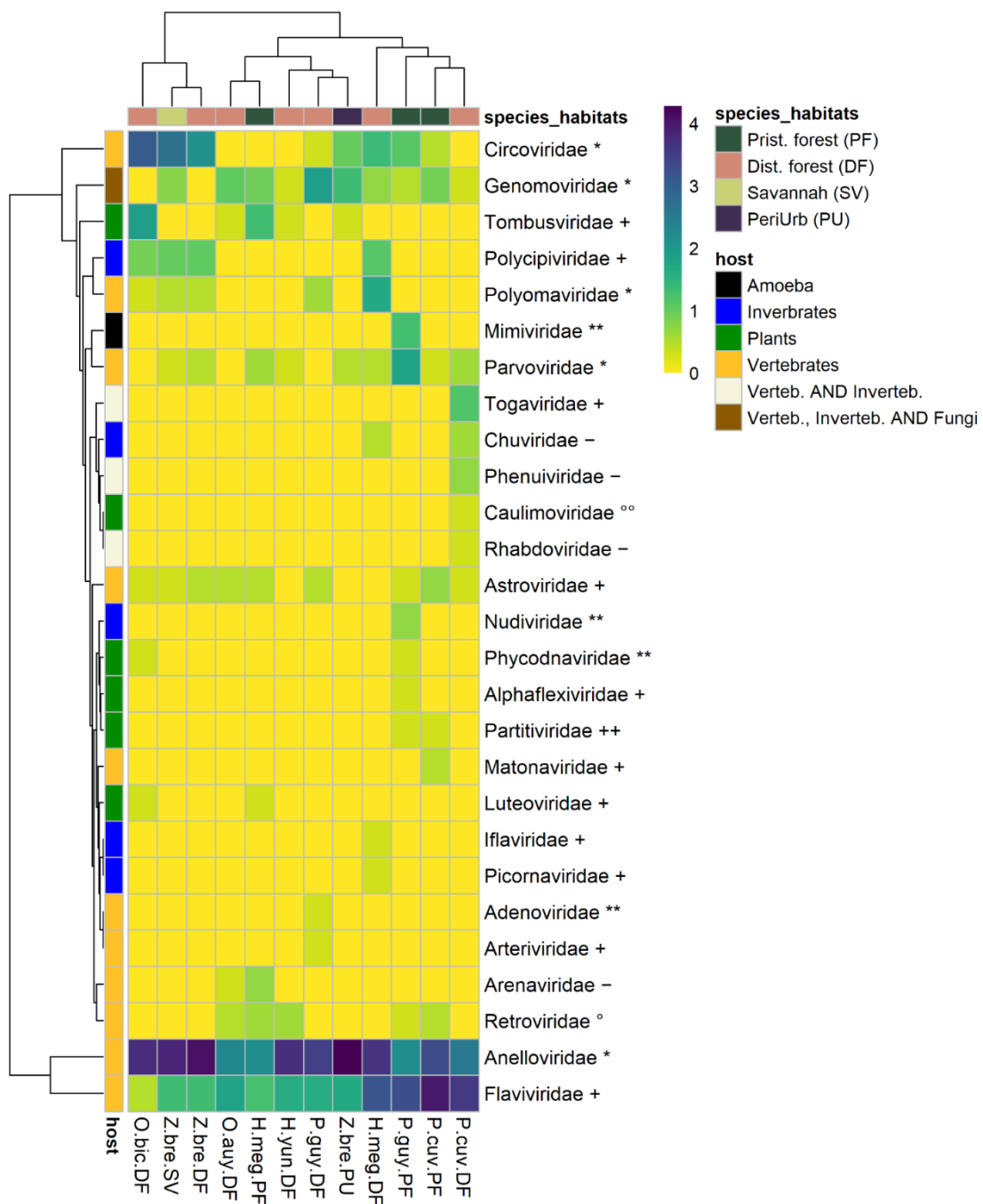
For the description of viral diversity, the results of viromes of different organs from a given species and in a given habitat were pooled together. The 77,767 viral contigs obtained were assigned to 27 families known to infect vertebrates, invertebrates, plants, and amoeba (Figure 2 and Table S4). Viral family presence varied largely across species–habitat categories. Indeed, we observed a pattern of ubiquity for some vertebrate viruses, along with those of the *Genomoviridae* family (which are commonly found in vertebrates, invertebrates, and fungi). On the other hand, some viruses, especially plant and invertebrate viruses, seemed more specific to the few species–habitat categories that they were found in. The patterns of ubiquity/specificity across species–habitat categories seemed to follow the patterns of the hosts (vertebrates, invertebrates, plants).

#### 3.2.1. Plant Viruses

Six plant-infecting viral families were detected from eight species–habitats, accounting for 49 contigs. Disturbed forest-originating samples from *P. guyannensis*, *H. megacephalus*, and *Z. brevicauda*, along with those from *Z. brevicauda* from savannah, contained no plant-infecting viruses. The most common plant-infecting viral family was the *Tombusviridae* family found in five species–habitats, followed by the *Luteoviridae*, *Partitiviridae*, and *Phycodnaviridae* families present in two species–habitats each. Lastly, *Alphaflexiviridae* and *Caulimoviridae* were all found in a single pool (Figure 2 and Table S2).

#### 3.2.2. Invertebrate Viruses

Five viral families of insect and invertebrate tropism were detected in six different species–habitats, totaling 106 contigs. The most common families were *Polycipiviridae* and *Chuviridae* present in four and two species–habitats, respectively. The remaining families (*Iflaviridae*, *Nudiviridae*, *Picornaviridae*) were detected in a single species–habitat (Figure 2 and Table S4). *P. cuvieri* from pristine forest; *P. guyannensis*, *O. auyantepui*, *H. yunganus*, and *H. megacephalus*, all from disturbed forest; and *Z. brevicauda* from peri-urban habitat contained no invertebrate viruses.



**Figure 2.** Heatmap of viral families’ numbers of contigs by species–habitat. Each cell representing a viral family in a species-habitat contains  $\log(1 + N)$ , (where N is the number of contigs assigned to a viral family in a species–habitat). The left row represents host-type (vertebrate, invertebrate, plants). Viral family names are marked with genome type, as follows: \*\* = dsDNA, \* = ssDNA, ++ = dsRNA, + = ssRNA (+), – = ssRNA (–), °° = DNA-retrotranscribing, ° = RNA-retrotranscribing. P.guy.PF: *Proechimys guyannensis* from pristine forest, P.guy.DF: *Proechimys guyannensis* from disturbed forest, P.cuv.PF: *Proechimys cuvieri* from pristine forest, P.cuv.DF: *Proechimys cuvieri* from disturbed forest, H.meg.PF: *Hylaeamys megacephalus* from pristine forest, H.meg.DF: *Hylaeamys megacephalus* from pristine forest, Z.bre.SV: *Zygodontomys brevicauda* from savannah, Z.bre.DF: *Zygodontomys brevicauda* from disturbed forest, Z.bre.PU: *Zygodontomys brevicauda* from peri-urban areas.

### 3.2.3. Vertebrate Viruses

A total of 11 viral families strictly associated with vertebrates were detected, accounting for 77,438 contigs. ssDNA virus families such as *Anelloviridae*, *Circoviridae*, *Parvoviridae*, and *Polyomaviridae* were detected in 12, nine, eight, and five species–habitats, respectively. *Adenoviridae* (dsDNA virus)-assigned contigs were found only in *P. cuvieri* from disturbed forest. RNA viruses (*Riboviria*) accounted for six families. Several positive-sense RNA viral families were also detected: *Astroviridae*, *Arteriviridae*, *Flaviviridae* (*Hepacivirus*), and *Matonaviridae*. *Matonaviridae*-attributed sequences were found in only one species–habitat (*P. cuvieri* in pristine forest), just as *Arteriviridae* sequences were detected only in *P. guyannensis* from disturbed forest. By contrast, *Flaviviridae* (*Hepacivirus*) and *Astroviridae* had a greater presence across species–habitats, respectively, 12 of 12 and nine of 12 species–habitats. An *Arenaviridae*, an assigned sequence close to “Patawa virus” [46], was detected in *O. auyantepui* from disturbed forest (Figure 2 and Table S4).

### 3.2.4. Potential Vector-Borne Viruses

Viral genera such as *Alphavirus* (*Togaviridae*) and *Phlebovirus* (*Phenuiviridae*), and the family *Rhabdoviridae*, are recognized as potential vector-borne viruses, since they infect both vertebrate and invertebrate hosts, were detected in one species–habitat. All contigs (19) attributed to vector-borne viruses were detected in *P. cuvieri* from disturbed forest (Figure 2 and Table S4).

## 3.3. Statistical and Ecological Analyses of the Viromes According to Hosts and Environments

### 3.3.1. Sampling Effort on Viral Diversity

The Pearson correlation between the number of individuals per sample in a pool and the number of viral genera detected showed no significant association ( $R^2 = -0.07$ ,  $p = 0.694$ ). The sequencing type had no impact on the viral diversity detected (two-sided unpaired Student’s *t*-test,  $t = -1.76$ ,  $p = 0.105$ ).

The rarefaction curves on the number of reads per sample, established for the nine species–habitat combinations, reached their asymptotes or started to plateau, suggesting that saturation was almost achieved if not in viral sequencing (Figure S1, Supplementary data). Hence, the sampling effort for our data set, both regarding the number of rodent individuals and the number of reads per sample, was adequate for diversity comparisons.

### 3.3.2. Diversity across Species–Habitats

Viral richness oscillated between 10 for *Zygodontomys* in peri-urban habitats and 21 for *H. megacephalus* in disturbed forests (Table 3). On the other hand, Rényi’s entropy tended to converge between species and habitats for  $\alpha = 2$  (Figure 3). Most of the differentiation was between  $0.5 < \alpha < 1.0$ . Regarding the habitats of a given species, we observed that for *P. guyannensis* the diversity in pristine forest was higher than in disturbed forest (Figure 3). For *Z. brevicauda*, the diversity was higher in savannahs, followed by disturbed forests, and was lowest in peri-urban habitats. For *H. megacephalus* and *P. cuvieri*, the diversity was higher in disturbed forests for  $\alpha < 0.25$  and  $\alpha < 0.75$ , respectively. For these species, diversity was higher in pristine forests for  $\alpha > 0.25$  and  $\alpha > 0.75$  (Figure 3). Thus, for two of the four rodent species trapped in two or more habitats, the diversity trend associated with different habitats was preserved over all  $\alpha$  values between 0 (genera richness) and 2 (Simpson index).

The richness loss, calculated for the nine species–habitat combinations with different removal thresholds of rare genera, is shown in Figure 4. *P. guyannensis* and *P. cuvieri* showed, at all threshold values, greater richness loss in pristine forest compared with disturbed forest, revealing that viromes from pristine habitats possessed higher numbers of rare viral entities (at the genus and subfamily levels according to their taxonomic classification) and fewer dominant ones than disturbed environments (Figure 4). *Z. brevicauda* showed higher richness loss in savannah compared with disturbed forest and peri-urban areas, showing that rare viral entities are more frequent in savannah than in the two other types

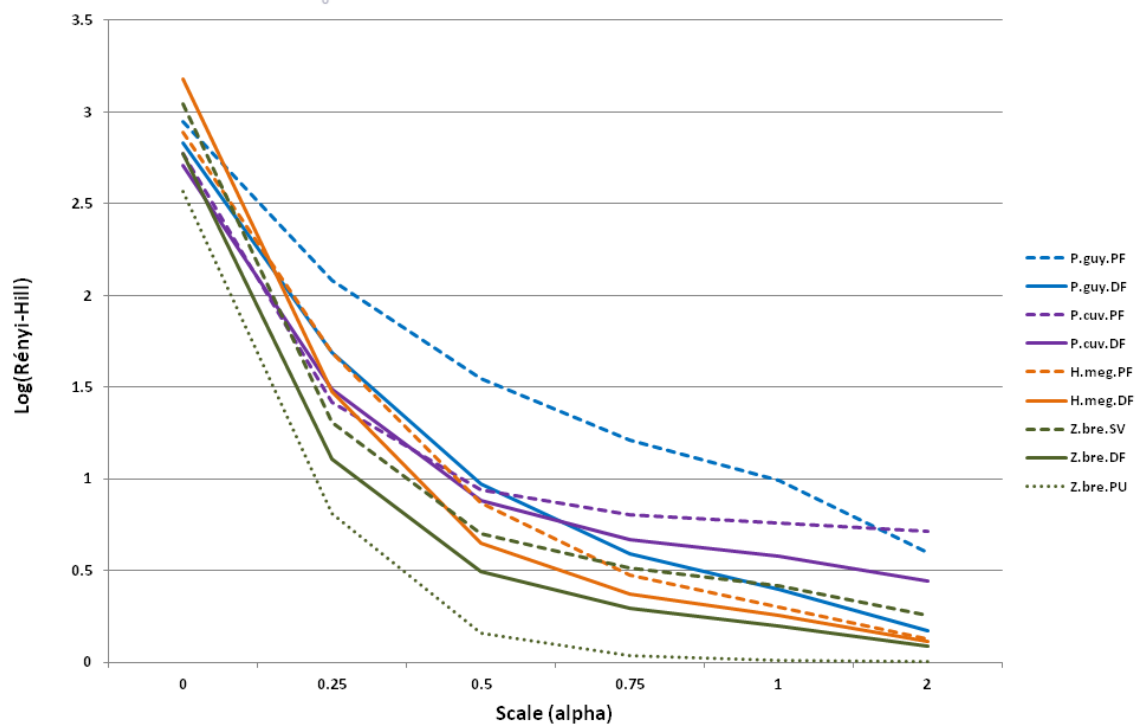
of habitats. On the other hand, greater richness loss was observed for *H. megacephalus* in disturbed forest compared with pristine forest, highlighting the importance of rare entities in disturbed environments for this species (Figure 4). We observed that the hierarchy of richness loss between habitats was similar over all threshold values for each rodent species. Some of the differences in viral diversity between habitats for a given species are predominantly due to rare viral entities, rather than fractions of abundant ones.

### 3.4. Phylogenetic Relationships of Selected Viruses

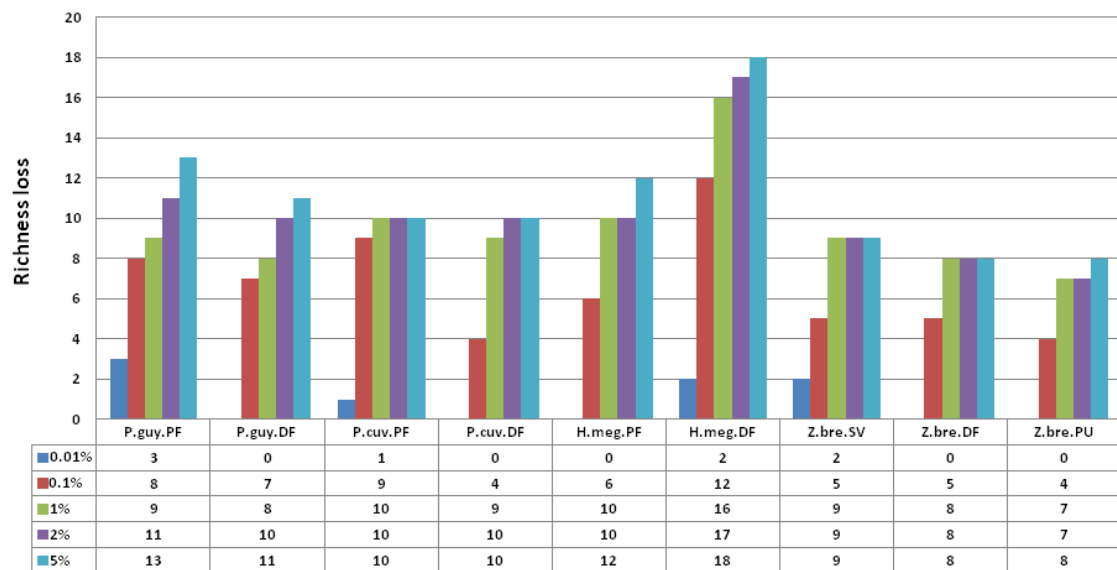
For phylogenetic analyses, we chose viral families for their frequent presence in the samples (*Polyomaviridae* and *Flaviviridae-Hepacivirus*) and their interest as potential EID agents because they are arthropod-borne (*Phlebovirus* and *Alphavirus*).

**Table 3.** Richness and Rényi’s entropy values for different values of  $\alpha$  between 0 and 2 for the nine species–habitat combinations.

	P.guy.PF	P.guy.DF	P.cuv.PF	P.cuv.DF	H.meg.PF	H.meg.DF	Z.bre.SV	Z.bre.DF	Z.bre.PU
Richness	17	16	14	12	13	21	12	11	10
Renyi ( $\alpha = 0$ )	2.94	2.83	2.77	2.71	2.89	3.18	3.04	2.77	2.56
Renyi ( $\alpha = 0.25$ )	2.08	1.69	1.42	1.49	1.69	1.47	1.31	1.11	0.81
Renyi ( $\alpha = 0.5$ )	1.55	0.97	0.94	0.89	0.87	0.65	0.70	0.49	0.16
Renyi ( $\alpha = 0.75$ )	1.21	0.59	0.81	0.67	0.48	0.37	0.51	0.29	0.04
Renyi ( $\alpha = 1$ )	0.99	0.40	0.76	0.58	0.30	0.25	0.42	0.20	0.01
Renyi ( $\alpha = 2$ )	0.60	0.17	0.72	0.44	0.13	0.12	0.25	0.09	0.00



**Figure 3.** Ordination plot of Rényi’s entropy values by scale numbers ( $\alpha = 0, 0.25, 0.5, 0.75, 1, \text{ and } 2$ ) for the four species present in different types of habitats (nine species–habitat combinations). Pristine forest and savannah are represented by dashed lines, disturbed forest by solid lines, and peri-urban areas by the dotted line. P.guy.PF: *Proechimys guyannensis* from pristine forest, P.guy.DF: *Proechimys guyannensis* from disturbed forest, P.cuv.PF: *Proechimys cuvieri* from pristine forest, P.cuv.DF: *Proechimys cuvieri* from disturbed forest, H.meg.PF: *Hylaeamys megacephalus* from pristine forest, H.meg.DF: *Hylaeamys megacephalus* from disturbed forest, Z.bre.SV: *Zygodontomys brevicauda* from savannah, Z.bre.DF: *Zygodontomys brevicauda* from disturbed forest, Z.bre.PU: *Zygodontomys brevicauda* from peri-urban areas.



**Figure 4.** Bar plot of richness loss between original data set and different thresholds (minimum percent abundance) of suppressions by species–habitat combinations. Percent thresholds were 0.01, 0.1, 1, 2 and 5%. P.guy.PF: *Proechimys guyannensis* from pristine forest, P.guy.DF: *Proechimys guyannensis* from disturbed forest, P.cuv.PF: *Proechimys cuvieri* from pristine forest, P.cuv.DF: *Proechimys cuvieri* from disturbed forest, H.meg.PF: *Hylaeamys megacephalus* from pristine forest, H.meg.DF: *Hylaeamys megacephalus* from disturbed forest, Z.bre.SV: *Zygodontomys brevicauda* from savannah, Z.bre.DF: *Zygodontomys brevicauda* from disturbed forest, Z.bre.PU: *Zygodontomys brevicauda* from peri-urban areas.

### 3.4.1. Rodent Polyomaviruses

The *Polyomaviridae* family is composed of four genera: *Alphapolyomavirus*, *Betapolyomavirus*, *Gammapolyomavirus*, and *Deltapolyomavirus* [78]. Each genome is composed of a circular dsDNA of approximately 5 kb. Polyomaviruses (PyVs) may be transmitted through either direct contact or by aerial or fecal–oral routes.

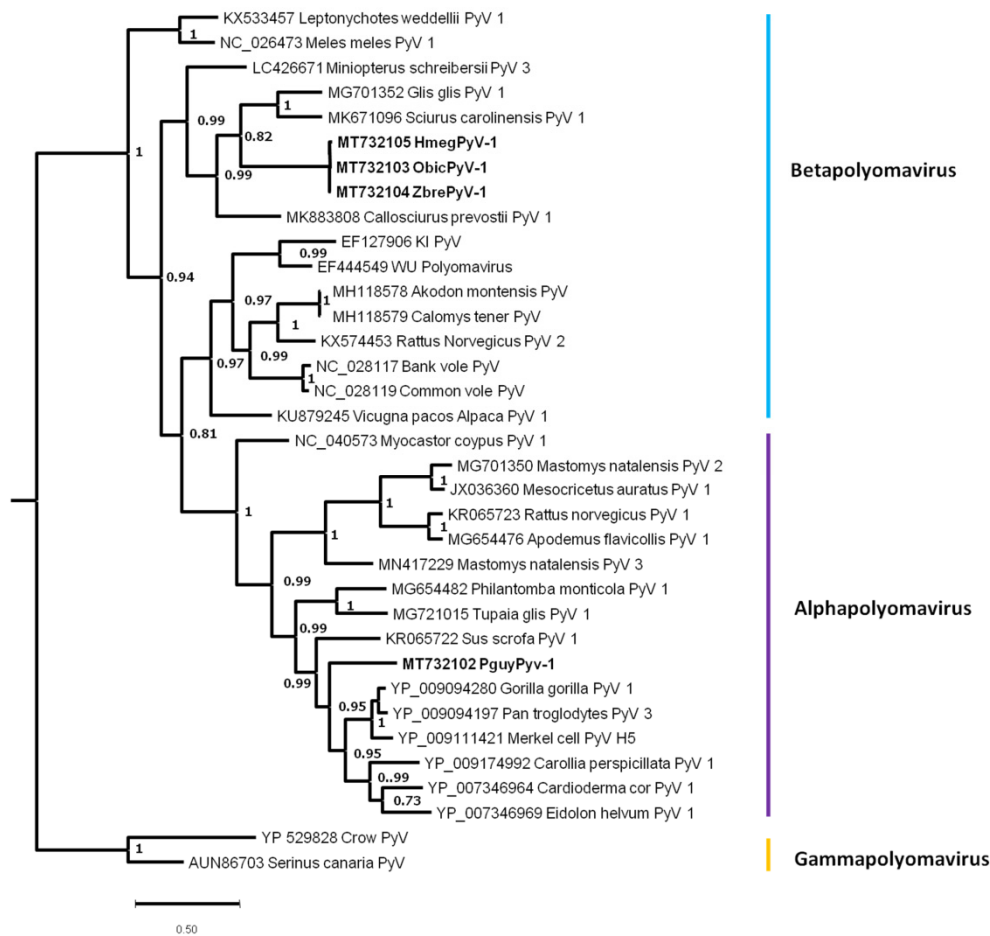
Overall, 54 contigs were assigned to the *Polyomaviridae* family and detected in four species: *O. bicolor*, *Z. brevicauda*, *H. megacephalus*, and *P. guyannensis* (Table S4). In kidney samples, one contig was detected in *O. bicolor* (disturbed forest); six were detected in *Z. brevicauda* (three from savannah and three from disturbed forest); 139 were detected in *H. megacephalus* (disturbed forest); and five were detected in *P. guyannensis* (disturbed forest). In spleen samples, two contigs were identified in *O. bicolor* (disturbed forest), two in *Z. brevicauda* (savannah and disturbed forest), and one in *H. megacephalus* (disturbed forest).

After alignment, the longest common sequences identified from the four species covered a fragment of 498 nucleotides of the Large T antigen (*LTag*). Sequence fragments were named *O. bicolor polyomavirus 1* (*ObicPyV-1*, GenBank accession number MT732103), *Z. brevicauda polyomavirus 1* (*ZbrePyV-1*, GenBank accession number MT732104), *P. guyannensis polyomavirus 1* (*PguyPyV-1*, GenBank accession number MT732102), and *H. megacephalus polyomavirus 1* (*HmegPyV-1*, GenBank accession number MT732105).

The *ObicPyV-1* and *ZbrePyV-1* sequences shared 100% nucleotide and amino acid identity. *HmegPyV-1* displayed 99.22% nucleotide identity and 98.83% amino acid identity with the two previous sequences (Supplemental data, Table S5). *PguyPyV-1* displayed substantial divergence from those two sequences—50.80% nucleotide identity and 39.16% amino acid identity with the *ObicPyV-1* and *ZbrePyV-1* sequences—and showed 50% and 37.95% nucleotide identity and amino acid identity, respectively, with *HmegPyV-1*. Compared with other polyomavirus sequences, *ObicPyV-1* and *ZbrePyV-1* showed the highest percentages of identity (61.45% and 57.23% in nucleotide identity and amino acid identity, respectively) with *Sciurus carolinensis polyomavirus 1* (GenBank accession number MK671096);

and *HmegPyV-1* showed 60.04% nucleotide identity and 54.02% amino acid identity with it (Supplemental data, Table S5). *PguyPyV-1* showed the highest level of nucleotide identity (67.78%) with *Pan troglodytes PyV3* (GenBank accession number YP\_009094197) and 63.25% amino acid identity with *Merkel cell Polyomavirus* (GenBank accession number YP\_009111421). *PguyPyV-1* also showed 59.44% nucleotide identity and 52.99% amino acid identity with the Alphapolyomavirus *Myocastor coypus polyomavirus 1* (GenBank accession number NC\_040573) from a rodent species originating from South America.

The phylogenetic analysis carried out on 166 amino acid-long sequences of the *LTA*g identified the two monophyletic clades corresponding to the *Alphapolyomavirus* (posterior probability = 1) and *Gammmapolyomavirus* (posterior probability = 1) genera, whereas *Betapolyomavirus* genus was not supported. *ObicPyV-1*, *ZbrePyV-1*, and *HmegPyV-1* polyomavirus sequences were clustered with polyomavirus sequences derived from Sciuridae (*Sciurus carolinensis*) and Gliridae (*Glis glis* and *Callosciurus prevostii*), a group of sequences that belongs to the *Betapolyomavirus* genus. These sequences are also associated with *Miniapterus schreibersii polyomavirus 3* (posterior probability = 0.99). Among alphapolyomaviruses, *PguyPyV-1* was not related to any other PyVs from rodents. It possessed a basal position of a clade consisting of viruses originating from primates and bats (posterior probability = 0.99) and to a lesser extent to PyVs from Artiodactyla and Scandentia (Figure 5).



**Figure 5.** Phylogenetic analysis of partial sequences of the Large T antigen (*LTA*g) (alignment of 166 amino acid positions) of *Polyomaviridae* representatives (Alpha, Beta, and Gammmapolyomaviruses with Gammmapolyomaviruses as the outgroup). The tree was inferred from amino acid sequences using the Bayesian method with the WAG model. Sequence identifiers include the NCBI accession number and the isolate name. Posterior probabilities of the Bayesian analysis (>70%) are shown next to the nodes. The scale bar indicates amino acid substitutions per site. Sequences identified in this study are shown in bold.



### 3.4.2. Rodent Hepaciviruses

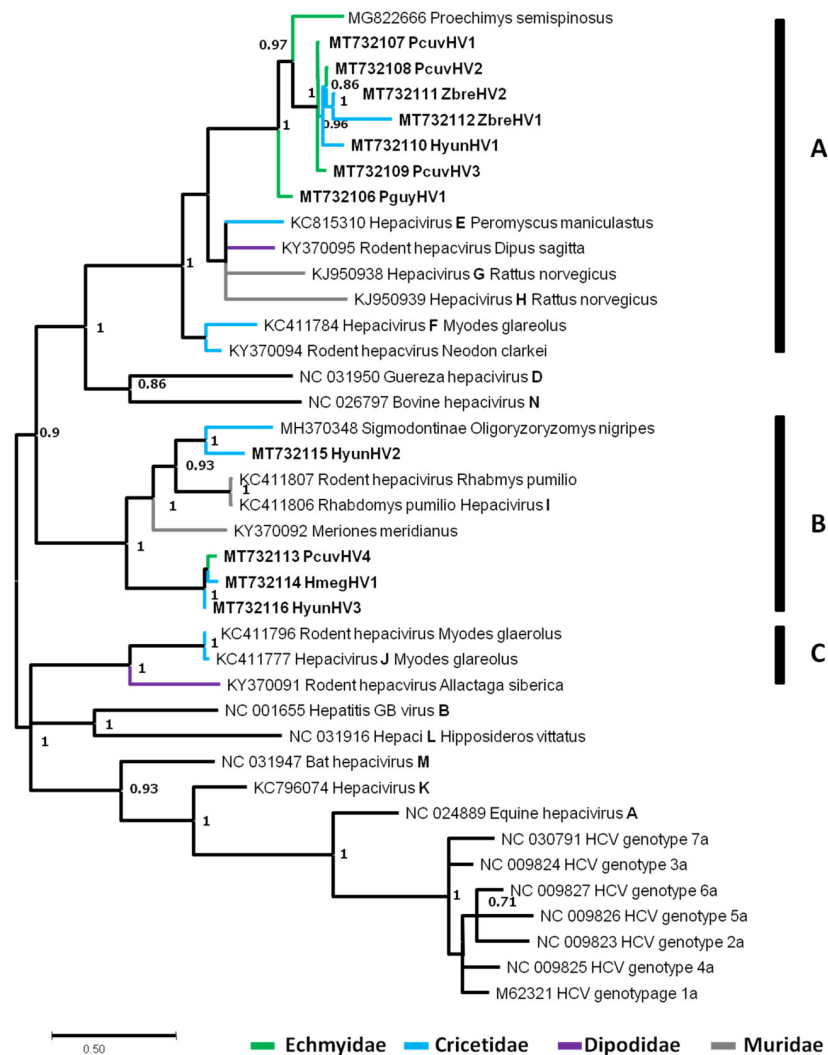
Hepaciviruses (HVs) constitute a genus that belongs to the ssRNA+ family of *Flaviviridae*. Their genome is approximately 10 kb long and encodes a single ORF translated into a polyprotein, which is processed by viral and cellular proteases, giving mature proteins. According to Smith and colleagues [79], there are 14 HV species in mammals (HV-A to HV-N), among which six are hosted by rodents (HV-E to HV-J). The main mode of transmission of HVs is vertical, but the fecal route for cross-species transmissions has been suggested.

A total of 16,686 contigs were assigned to HVs (Table S4). These sequences were found in 18 of 36 pools (seven sera, six kidneys, and five spleens) and in all species included in the study. After re-examination of contigs and progressive multiple alignments, an *NS5B* (i.e., *RDRP*) fragment of 384 nucleotides was chosen for subsequent phylogenetic analyses. This region included the main known mammal HVs and rodent HVs (RHVs) identified in five of the seven species analyzed (i.e., *P. cuvieri* with four contigs, *P. guyannensis* with one contig, *Z. brevicauda* with two contigs, *H. megacephalus* with one contig, *H. yunganus* with three contigs). We attributed arbitrary names by contracting rodent species names with “HV” for Hepacivirus with an incremental number to differentiate sequences from the same host species.

The percentages of identity between HV sequences obtained in this study ranged from 36.67% to 97.58% in nucleotides and from 27.27% to 97.27% in amino acids, showing their great diversity (Supplemental data, Table S6). Briefly, PcuHV4 (GenBank accession number MT732113), HmegHV1 (GenBank accession number MT732114), and HyunHV3 (GenBank accession number MT732116) sequences shared the highest percentages of identity (97.27–97.58% nucleotide identity and 94.55–96.36% amino acid identity) while PguyHV1 (GenBank accession number MT732106), PcuHV1-3 (GenBank accession number MT732107–MT732109), ZbreHV1-2 (GenBank accession number MT732112, MT732112), and HyunHV1 (GenBank accession number MT732110), even if close, showed a greater divergence, with the percentage of identity ranging from 58.48% to 96.06% in nucleotides and from 64.55% to 97.27% in amino acids. These two groups of sequences were separated into two distinct clusters distantly related to each other. The HyunHV2 sequence (GenBank accession number MT732115) had the highest percentages of identity (64.24% in nucleotides and 71.82% in amino acids) with another Sigmodontinae rodent HV (*Oligoryzomys nigripes* Hepacivirus, accession number MH370348). Compared with the other RHVs, the first group of sequences (PcuHV4, HmegHV1, and HyunHV3) showed 54.64–54.45% and 60–64.55% nucleotide identity and amino acid identity, respectively, with *Meriones meridianus* RHV (Supplemental data, Table S6). This group is also related to *Rhabdomys pumilio* RHV-I (GenBank accession number KC411806), showing between 58.48% and 59.09% nucleotide identity and between 52.72% and 56.36% amino acid identity. The second group showed 58.79–69.39% and 56.36–77.27% nucleotide identity and amino acid identity, respectively, with *Proechimys semispinosus* HV (GenBank accession number MG822666).

Three groups (A, B, C) of sequences regrouping rodent HVs were identified, highly supported with posterior probabilities of 1 (Figure 6). Group A was subdivided into three subgroups. The first one comprised seven sequences identified here (PguyHV1, PcuHV1-3, ZbreHV1-2, HyunHV1) and the previously identified RHV characterized in *P. semispinosus*. This subgroup was supported with a posterior probability of 1 and represented a group comprising Cricetidae and Echimyidae. It was related, however, with low support to a second subgroup composed of HV sequences identified in *Dipus sagitta*, *Peromyscus maniculatus*, and *Rattus norvegicus*, with these two last species hosting RHVs species E, G, and H. The third subgroup was composed two RHV sequences (*Neodon clarkei* HV and *Myodes glareolus* RHV-F). Group B contained the four remaining sequences from our samples (PcuHV4, HmegHV1, HyunHV2-3), along with *Oligoryzomys*, *Meriones*, and *Rhabdomys* RHVs. PcuHV4, HmegHV1, and HyunHV3 were grouped together (posterior probability = 1). The HyunHV2 sequence was clustered with the RHV sequence of *Oligoryzomys nigripes* (posterior probability = 1). This last group is related to RHV-I species hosted by *Rhabdomys pumilio* and to a lesser extent to RHV hosted by *Meriones*

*meridianus*. Group C was composed of three RHVs sequences among which was the RHV-J species identified in *Myodes glareolus* (Figure 6). All these groups did not seem to follow a co-evolution model with their rodent hosts since HVs identified in different rodent families were found independently in all groups. In addition, for a given rodent species, diverse and distantly related HV sequences were identified.



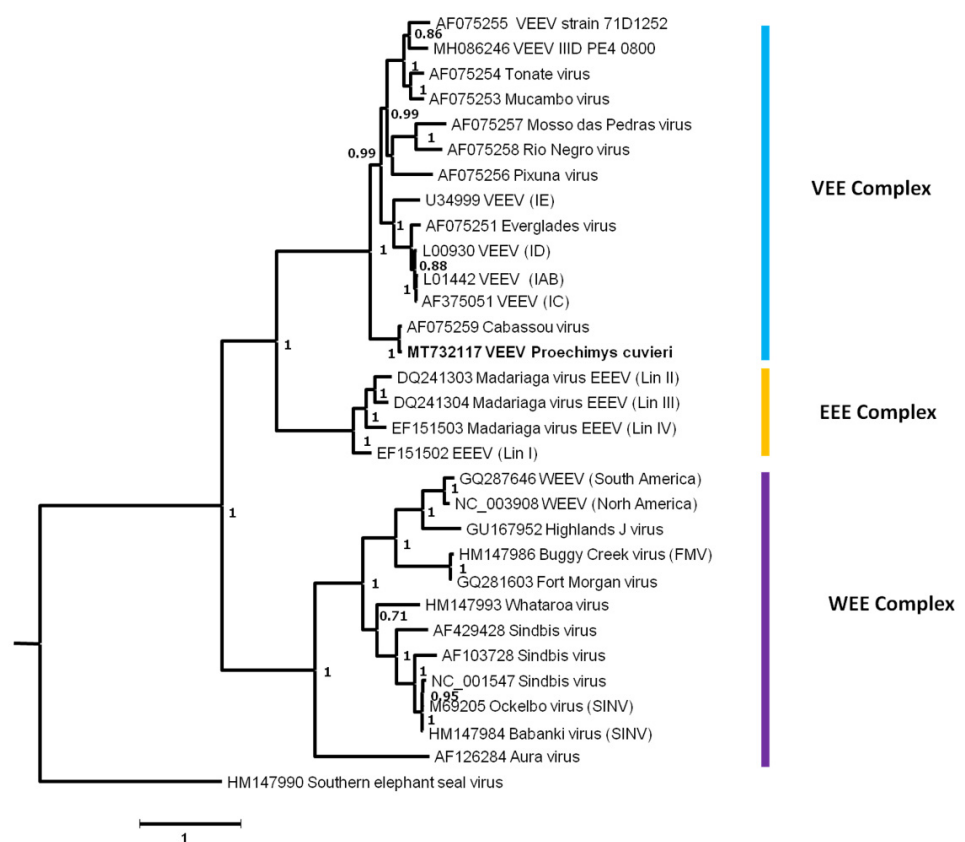
**Figure 6.** Unrooted phylogenetic tree based on the *ns5b* (alignment of 135 amino acid positions) of rodent hepaciviruses (RHVs). The tree was inferred from amino acid sequences using the Bayesian method with the WAG model. Sequence identifiers include the NCBI accession number and the isolate name. Posterior probabilities of the Bayesian analysis (>70%) are shown next to nodes. The scale bar indicates amino acid substitutions per site. The sequences identified in this study are shown in bold. The sequences identified in this study clustered in three well supported groups named **A**, **B** and **C** for ease of description.

### 3.4.3. Alphaviruses (*Togaviridae*)

The *Togaviridae* family is composed of ssRNA+ viruses. Their genome, 10–12 kb in size, is composed of nonstructural and structural parts. The *Togaviridae* family has recently become a monogenus family exclusively composed of alphaviruses since its former sibling genus *Rubivirus* was removed to its own family (*Matonaviridae*). Alphaviruses are classified as antigenic complexes, such as VEE (Venezuelan equine encephalitis), EEE (Eastern equine encephalitis), WEE (Western equine encephalitis), etc. Alphaviruses are mainly transmitted by mosquitoes.

A total of 14 contigs were identified, all from *P. cuvieri* sera from disturbed forest. They covered 7682 nucleotides and both the structural and non-structural parts of the genome. We used the structural part of the genome (1989 nucleotides) for phylogenetic analyses including representatives of the main antigenic complexes (VEE, WEE, Semliki Forest, etc.). *P. cuvieri* VEE showed the highest levels of nucleotide and amino acid identity (94.83% and 98.69%, respectively) with the VEE strain Cabassou, a member of the antigenic complex of the same name (Supplemental data, Table S7).

The phylogenetic tree showed that *P. cuvieri* VEE clustered together with VEE Cabassou (posterior probability = 1), forming a clade with a basal position of the VEE complex. The clade was highly supported (posterior probability = 1) (Figure 7). The other major alphavirus complexes, EEE and WEE, were also supported with posterior probability values of 1 (Figure 7).



**Figure 7.** A phylogenetic tree based on one fragment of the structural protein (alignment of 1989 nucleotide positions) of each well-known representative of an alphavirus antigenic complex. The tree was inferred from nucleotide sequences using the Bayesian method with the GTR + G (general time reversible) model. Sequence identifiers include the NCBI accession number and the isolate name. The posterior probability of the Bayesian analysis (>70%) is shown next to each node. The scale bar indicates the (nucleotide) substitution rate. The sequence of *P. cuvieri* VEE is highlighted in bold and the main antigenic complexes are indicated.

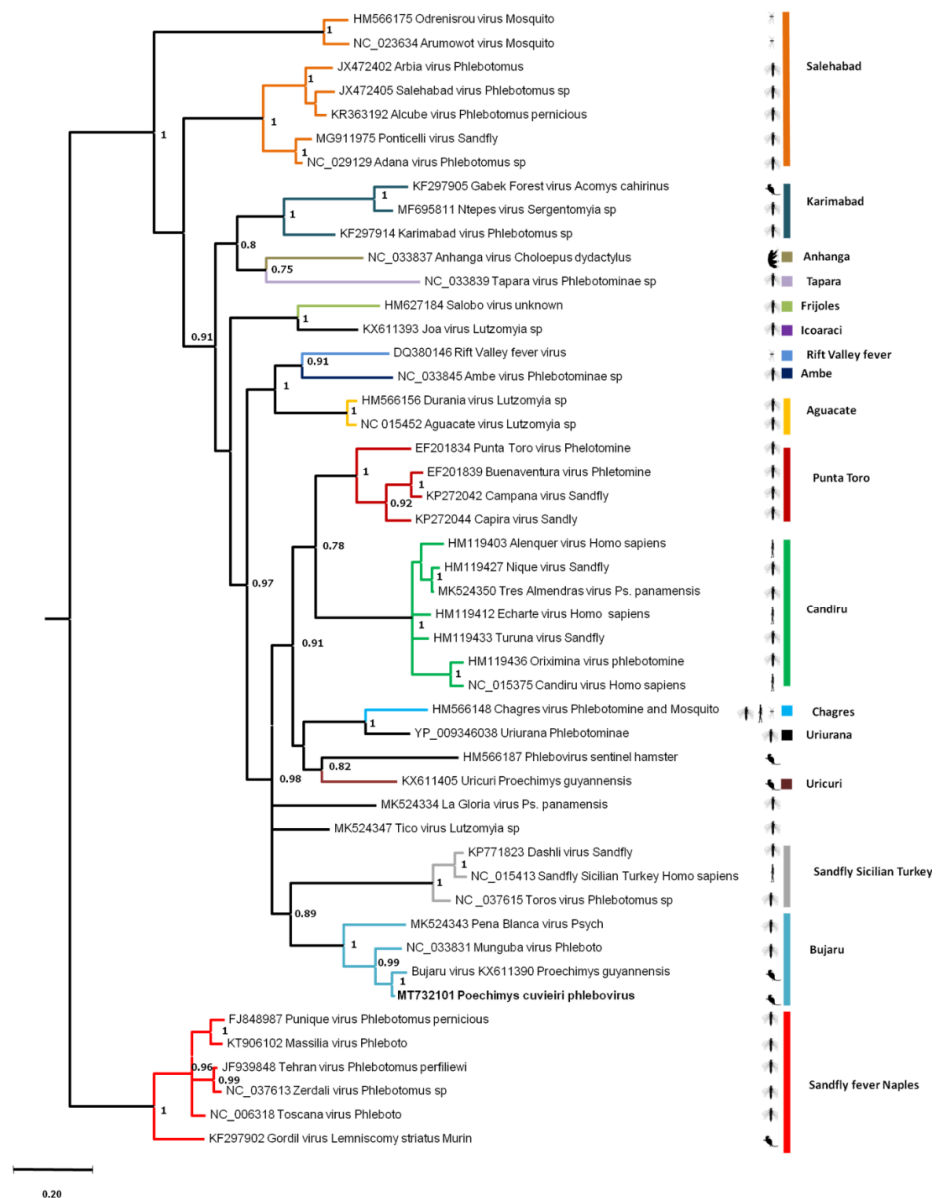
#### 3.4.4. Rodent Phleboviruses

Phleboviruses are members of the *Bunyavirales* order and belong to the *Phenuiviridae* family. They possess a segmented (three segments: small, medium, large) negative-sense ssRNA genome. Phleboviruses are arthropod-borne viruses frequently hosted by phlebotomin species (sandflies), but also by mosquitoes, ticks, and culicoides from which they are transmitted to humans and other vertebrates. Currently, 66 species have been

recognized by the ICTV [80] and other species remain to be described. Phleboviruses have been classified in two antigenic groups including ten different species complexes [81,82].

Phlebovirus-attributed sequences were found in *P. cuvieri* samples from disturbed forest only, one from sera and three from spleen samples. The phylogenetic analysis was based on a 179-aa segment of the nucleocapsid. The *P. cuvieri* phlebovirus sequence showed 79.80% and 95.78% nucleotide and amino acid identity, respectively, with Bujaru phlebovirus, which was identified from a *P. guyannensis* rodent in Brazil (Supplemental data, Table S8).

*P. cuvieri* phlebovirus were clustered phylogenetically with Bujaru virus with high support (posterior probability = 1). These *Proechimys*-originating sequences were grouped under the Bujaru serogroup with Munguba and Peña Blanca, both isolated from sandflies. The ancestral node of these four viruses was supported with a high posterior probability (pp = 1) (Figure 8). Nevertheless, phylogenetic relationships between all species-complexes remained unresolved with low support observed for basal nodes.



**Figure 8.** A phylogenetic tree based on one fragment of the nucleocapsid protein (alignment of 537 nucleotides/179 amino acid positions) of each well-known phlebovirus antigenic complex representative. The tree was inferred from amino acid sequences using the Bayesian method with the

WAG model. Sequence identifiers include the NCBI accession number and the isolate name. Posterior probabilities of the Bayesian analysis (>70%) are shown next to the nodes. The scale bar indicates amino acid substitutions per site. The sequence of *P. cuvieri* phlebovirus is indicated in red. The species–complexes (established by ICTV or suggested in [81,82]) are indicated with colored bars (if many members are present) and dots (if one member is represented). Names of established species–complexes are indicated in bold and italics. Animal pictures were downloaded from phylopic web site (<http://phylopic.org/>) [83].

#### 4. Discussion

Over the past decade, virome studies exploring the roles of wild species as reservoirs of infectious diseases have become more common thanks to the technological breakthrough of high-throughput sequencing. Considering that some species are reservoirs of numerous viruses, some of which have large impacts on human health, studies on viral diversity in rodents have recently increased [3,9,84–86]. Hence, 173 viral species belonging to more than 65 genera have been described in rodents to date, among which 53 are zoonotic, such as mammarenaviruses and hantaviruses [30,87]. However, few studies have explored the links among viral diversity, host ecology, and habitats [8,88,89]. Here, we presented the viral diversity identified in three different organs of seven rodent species from French Guiana, according to their natural hosts and habitats, and further explored the phylogenetic relationships of several viruses of interest for human health.

In order to ascertain the viral infection status in natural reservoirs and to identify a large number of vertebrate-related viruses, we chose to study three types of organs representing different tropisms of viruses. The kidney is the target organ of viruses that use the urinary tract to disseminate, such as hantavirus, arenaviruses, and paramyxoviruses, whereas viruses such as dengue or West-Nile have been detected in the spleen, a blood reservoir. Finally, serum is one of the most important media for the transmission of arboviruses. Furthermore, the analysis of such organ samples also limits the potential errors in the taxonomic assignment of new viruses compared to those that may be detected in respiratory or fecal samples. The latter viruses could indeed be from environment plants, insects, or fungi, and only incidentally found in rodents. Together, the use of organs should give a good representation of vertebrate viruses hosted by rodents [84].

Overall, this study identified 77,767 viral-associated contigs distributed within 27 viral families known to infect vertebrates, invertebrates, plants, and amoeba. The viromes were quantitatively dominated by vertebrate viral sequences (>99% of both contigs and reads were assigned to 11 viral families known to strictly infect vertebrates) and to a lesser extent to viral sequences from invertebrates, plants, and amoeba. Nevertheless, the smaller number of invertebrate and plant virus sequences indicates non-negligible diversity, accounting for 12 families.

The different viral families, whether originating from invertebrates, plants, or vertebrates, were not evenly distributed within the different species and habitats. Viruses from *Parvoviridae*, *Circoviridae*, *Astroviridae*, and *Anelloviridae* from vertebrates were found in most species and habitats and can be considered as generalists. These ubiquitous viruses were already reported in wild rodents in the United States where the *Circoviridae* family was the most abundant among the 24 families described [86], and in wild brown rats in Germany, with viruses of the *Parvoviridae* family [90]. On the other hand, viruses belonging to the *Caulimoviridae* (from plants), *Iflaviridae* (from invertebrates), or *Arteriviridae* (from vertebrates) families were rare and only present in some species and/or habitats. These differences in the distribution of viral families can be put in perspective by hypothesizing a rare biosphere for microbial communities in oceanic waters [91], with a portion of a few dominant microbial species and a second large, unexplored fraction with rare species. Accordingly, viromes in rodents could be dominated by a few dominant families, and a long distribution tail shaping a rare virosphere.

Such differences in virus abundance could be related to the ecology of the viruses (i.e., their ability to infect host cells and to persist and replicate) and to the ecology and behavior of their rodent hosts in a given habitat, such as a modified diet in a disturbed environment. The role of vectors in viral transmission and their diversity according to the environment can also have an impact on viral diversity. Indeed, for *P. cuvieri* and *H. megacephalus*, fourfold more viral families of invertebrate and vertebrate viruses have been detected in disturbed forest compared with pristine forest. In these two opportunistic species, diet can be supplemented by invertebrates when fruits and seeds are lacking [92], with subsequent impacts on their virome structures. On the other hand, a more specialized diet should restrict the range of viral diversity. Similar virome compositions were previously observed in house mice [3] and brown rats [85] in New York City, suggestive of an adaptive diet.

Viral diversity indices and the relative dominance levels of viral species were also impacted by the level of disturbance and the type of habitat. In this study, the highest viral diversity index values were mainly observed in pristine habitats where the highest diversity of hosts was also recorded. The viromes of *P. guyannensis* and *H. megacephalus* in pristine forest showed the highest diversities (mainly driven by viruses originating from plants) compared with their counterparts from disturbed forest. This trend was nevertheless not found in *P. cuvieri*, for which viral diversities were comparable between habitats (pristine vs. disturbed forest), but a higher number of rare viral entities were in pristine forest. In contrast, *H. megacephalus* presented a high number of rare viruses in disturbed forests. *Z. brevicauda*, the only species also sampled in the savannah, showed the highest viral diversities in this habitat, also reflecting the richness of the savannah ecosystem [93,94]. Peri-urban areas had the lowest viral diversity, which may be related to overall low biodiversity.

Among vertebrate hosts, rodents have been described as major reservoirs of arboviruses such as *Togaviridae*, *Flaviviridae*, and *Bunyaviridae* [95], and can serve as amplifiers of viruses that can be transmitted to humans. For instance, Cabassou virus (genus *Alphavirus*) was detected in *P. cuvieri* in disturbed forest. The circulation of arboviruses in disturbed habitats could be the result of increased contacts with vectors and may also reflect the lowest diversity of hosts available for arthropods to feed on.

The likelihood of disease emergence is indeed commonly accepted to increase in disturbed habitats [96]. The transmission of viruses from forest species to humans may result from two mechanisms. First, anthropic activities can increase contact between wildlife and humans and thereby the risk of infection [36] when humans enter slightly modified habitats and come into contact with a pristine viral cycle. Secondly, in more degraded forests, environmental changes may disrupt some ecological barriers and impact the structure and dynamics of rodent and arthropod communities, species richness, and ecological functions [97]. This may favor generalist over specialist species and ultimately the dominance of more synanthropic ones. Feeding networks between hosts and hematophagous vectors consequently change, influencing the transmission of viruses and potentially increasing cross-species transmission events.

From a theoretical point of view, the dilution effect hypothesis explores how the decrease of biodiversity may increase the amplification of zoonotic diseases. Briefly, the dilution effect proposes that a high diversity of putative hosts and vector species dilutes the more efficient carriers and amplifiers of viruses in a community of less efficient species, consequently reducing the circulation of the harmful ones and lowering the likelihood of infection [98]. The dilution effect may affect cycles involving a single animal host (i.e., reservoir) and those with two host compartments, i.e., reservoirs and vectors. In the latter case, a decrease in vertebrate diversity may concentrate blood meals taken by arthropods on a lower number of species, resulting in a higher viral circulation as soon as those resilient vertebrate species are also efficient carriers. The dilution effect can be suggested to illustrate the links between the diversity of rodent hosts and the spread of some zoonotic viruses. A higher probability of hepacivirus infection in *P. semispinosus* has been related to a loss of diversity in hosts due to land-use change [99]. Additionally, hantavirus outbreaks in

the Americas are related to environmental disturbances that result in a decrease in specific richness of non-murine rodents and in the dominance of a few *Muridae* species known to be more efficient reservoirs [28,100]. In French Guiana, all known human hantavirus cases occurred in agricultural and peri-urban areas, where rodent diversity is much lower than in forest habitats [33], likely favoring hantavirus circulation in most efficient reservoirs.

In this study, 14 viral families from *Rodentia* were detected of the 31 currently described. We established the phylogenetic relationships of viral sequences related to four viral families known to infect vertebrates including arthropod-borne viruses (*Polyomaviridae*, *Flaviviridae*, *Togaviridae*, and *Phenuiviridae*). Even if the sequences obtained were incomplete, their analysis added knowledge on viral evolution among rodent species in South America, a group of species with very few data available to date.

Most viral sequences were related to sequences previously detected in rodents from multiple geographic areas (Africa, Asia, and North and South America), suggesting common evolutionary processes. Nevertheless, cross-species transmission and spill-over events were also detected, emphasizing the importance of these mechanisms in their evolution. These events took place early during the evolution of mammals or could be linked to recent interactions between sympatric species, as suggested for polyomaviruses. Indeed, PyVs have been described in a wide range of hosts, including mammals, birds, amphibians, reptiles, fish, and invertebrates. Some PyVs are pathogenic for humans and animals [78]. In rodents, 45 PyVs have been described from 11 species originating from Europe, Asia, and Africa [30]. In South America, betapolyomaviruses were recently described in two Sigmodontinae (Cricetidae) species (*Akodon montensis* and *Calomys tener*), and an alphapolyomavirus in a Myocastoridae species (*Myocastor coypus*) [101,102]. We identified four PyV sequences in four species (Figure 4). The three Sigmodontinae PyV sequences (*ObicPyV1*, *ZbrePyV1*, and *HmegPyV1*) belong to the *Betapolyomavirus* genus, and the Echimyidae PyV sequence (*PguyPyV1*) to *Alphapolyomavirus*. The sequence identified in *P. guyannensis* did not cluster with the other alphaPyVs sequences from rodents and showed a basal position to a clade constituted of PyVs detected in primates and Chiroptera, suggesting duplication events [103]. Polyomaviruses were initially considered to be host-specific, with codivergence and lineage duplication being the main drivers of their diversification [6,104–106]. PyV cross-species transmissions were also identified, but they do not seem to play a major role in the diversification processes of PyVs [105–107]. In the present study, a duplication event can be suggested to explain the position of *PguyPyV1*, and host-switching events could explain the closeness of PyVs detected in the Sigmodontinae subfamily. Indeed, the high PyV sequence identity values observed between *O. bicolor*, *Z. brevicauda*, and *H. megacephalus* may reflect a geographical signature related to the sympatry of these three taxonomically related species, favoring PyV host-switching events [101,106]. Further work is needed to confirm this hypothesis so as to ascertain whether the novel PyVs detected in the present study show evidence of host-switching in the Sigmodontinae subfamily, and whether PyV host-switching is more common in rodents than in other mammalian orders.

We also identified a large number of hepacivirus sequences (HVs) in the seven species, suggesting a high prevalence of HVs in neotropical rodents. The *hepacivirus* genus prototype is the human-infecting *hepatitis C virus* (HCV). After its identification at the end of the 1980s, HCV remained, along with *GB-virus B* (GBV-B), the only known HV for years. Homologues were then described within a wide range of hosts, such as horses, bats, rodents, cows, dogs, and even sharks, thanks to high-throughput sequencing and extensive investigations [4,9,108–113]. Like polyomaviruses, HVs are considered to co-evolve with their hosts [4]. Nevertheless, their descriptions in a wide range of species, and more particularly in primates and rodents (RHVs), have demonstrated that they do not fully follow a co-speciation pattern [114]. The phylogenetic analysis of HVs, including those detected in this study, revealed the presence of genetically distinct RHVs in *P. cuvieri*, *H. yunganus*, and *Z. brevicauda*, and of two distantly related RHVs in *P. guyannensis* and *H. megacephalus* (Figure 5). All RHV clades did not seem to co-evolve with their hosts, since RHVs identified in *Muridae*, *Dipodidae*, *Echimyidae*, and *Cricetidae* rodents were

independently found in different clades. In addition, the identification of different RHVs in a single rodent species demonstrated the high level of genetic diversity among RHV species, reinforcing the idea that strict co-evolution is unlikely [114]. This high diversity reflects the idea that host shifts seem to be the main driver of RHV evolution [4,114,115]. The same evolutionary pattern was already suggested for RHVs from rodents in China [9]. Thus, the evolutionary history of the *Hepacivirus* genus remains to be deciphered along with the role of their rodent hosts as zoonotic transmitters, given their basal phylogenetic position in relation to other mammalian hepaciviruses [4,85,116].

Alphaviruses are arboviruses infecting a large number of vertebrates [5]. They are maintained in enzootic cycles involving arthropods as vectors and small mammals and/or birds as amplifier hosts. Occasionally, spill-overs into humans and domesticated animals lead to disease [117]. Among alphaviruses, VEEV have caused epizootics and epidemics in South America and the southern United States in the past few decades [118–120]. The VEEV complex is divided into six subtypes, some of which are known to cause diseases in humans and horses while others are enzootic, but can potentially be transmitted to humans [121–123]. In French Guiana, two viruses belonging to the VEEV complex were previously reported: *Cabassou virus* (CABV) isolated from *Culex portesi*, belonging to the subtype V, and *Tonate virus* (TONV), isolated from a bird and associated with the subtype IIIB [117,124]. Only a few cases of TONV infection were detected in French Guiana, Suriname, and North America, and most patients showed febrile illness. However, severe cases of encephalitis have been described [124,125]. To date, no human case of CABV infection has been reported.

Here, we identified sequences related to CABV (98.69% amino acid identity) in *P. cuvieri* samples in disturbed environments. Given that rodents belonging to the genera *Proechimys*, *Sigmodon*, *Oligoryzomys*, and *Oryzomys* have already been described as the main reservoirs for enzootic VEEV strains, CABV may circulate between mosquitoes and rodents, such as *Culex (melanoconion) portesi* (where it was previously isolated) and *Proechimys* [31,117,122,126,127]. VEEV strains are important candidates for future emergence in South America given that their reservoir hosts and vectors are facing an increasing number of anthropogenic disturbances. The potential circulation of CABV in the human population should be further investigated because its symptoms resemble those of dengue fever and it could therefore, be mis- or undiagnosed [128,129].

The *Phenuiviridae* family comprises 19 genera, including the *Phlebovirus* genus, which currently has ten recognized species [80] distributed worldwide. Among them, 14 phleboviruses were identified in rodents and 21 remain to be classified [30]. In South America, four species have been detected in rodents: *Icoaraci*, *Itaporanga*, *Jacunda* (Candiru complex), and *Bujaru*. The phylogenetic analysis of the sequence identified in *P. cuvieri* suggests that it belongs to the Bujaru complex composed of *Munguba*, *Peña blanca*, and *Bujaru* viruses [82]. Previous studies have already identified phleboviruses belonging to the Bujaru complex in *P. guyannensis* in Brazil and in sandflies (Phlebotominae spp.) [82,130], suggesting that viruses of the complex circulate in wild fauna and potentially in humans. Nevertheless, the pathogenicity of viruses of the Bujaru complex in humans is not known since most of the human cases described to date were due to viruses of the *Candiru* species complex. As for CABV, the fact that no human case has been reported for Bujaru viruses can be related to under-diagnosis given the similarity of the associated symptoms with other arboviral diseases [131,132]. Further studies are needed to investigate the ability of phlebotomine to spread the virus from their natural hosts and to clarify the real impact of phlebovirus infections on human health.

## 5. Conclusions

Only a few studies on the viral diversity in rodents have been conducted, even though they comprise the first order of mammals in terms of the number of species and are considered an important source of viral zoonotic pathogens. In addition, while Amazonia is considered a hotspot of diversity for hosts and pathogens [133], most virome studies have



been conducted in Asia and North America. In French Guiana, north of the Amazonian region, the description of the virome of seven neotropical rodent species allowed us to identify a large number of new viruses, most of which correspond to vertebrate viruses. These findings extend our knowledge on the host range and evolution of these viruses. We identified previously known viruses belonging to *Togaviridae* and *Phenuiviridae* in the spiny rat *P. cuvieri*, highlighting its role in the maintenance and circulation of these arthropod-borne viruses in disturbed areas. Further research is needed to better understand the transmission cycles and the ecology of the hosts and vectors involved. In addition, we showed that the diversity of rodent viromes varies according to the types of habitat, with higher viral diversity in pristine forests compared with disturbed forests for most rodent species. As well as the environment, the significance of species characteristics (including distribution, ecology, demography, and phylogenetic relationships), the importance of host switch throughout virus evolution, and the potential for local cross-species transmission should be studied to gain a better understanding of how viral diversity is shaped.

Environmental pressures on wild animal populations continue to grow, leading to increasing risks of contact between human and rodent populations. This could favor the emergence or re-emergence of viral diseases, including from viruses yet unknown or with undocumented roles on human health.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/v13091690/s1>, Figure S1: Rarefaction curves for the four species present in different types of habitats (nine species–habitat combinations), Table S1: Number of viral read of each viral genus/subfamily according rodent species and habitats, Table S2: Sampling variables and viral richness for the nine species-habitat at organ level (Organ level pool name, Number of rodents individuals, Number of viral genera/subfamilies detected, Sequencing platform, Organ type), Table S3: Assembly statistics of each sample by organ, species- and habitat, Table S4: Number of viral contigs by rodent species/habitats and viral families, Tables S5–S8: Pairwise sequence identities presents in nucleotides and amino-acids for respectively partial *LTA<sub>g</sub>* (*Polyomaviridae*), partial *ns5b* (*Flaviviridae*), *Hepacivirus*, partial *Structural protein* (*Togavirida*) and *nucleocapsid protein* (*Phenuiviridae*; *Phlebovirus*).

**Author Contributions:** S.T.: Data curation, Formal analysis, Software, Investigation, Writing—Original draft preparation. B.d.T.: Conceptualization, Investigation, Writing—Original draft preparation, Ressources. D.D.: Formal analysis, Writing—Review and Editing. C.B.: Formal analysis, Writing—Review and Editing. V.L.: Writing—Review and Editing. A.F.: Methodology, Writing—Review and Editing, Supervision. A.L.: Conceptualization, Data curation, Funding acquisition, Methodology; Project administration, Validation, Visualization, Writing—Original draft preparation, Writing—Review and Editing, Ressources, Supervision. All authors have read and agreed to the published version of the manuscript.

**Funding:** S. Tirera was funded by the RESERVOIRS program, supported by European funds (PO FSE 2014–2020), an “Investissement d’Avenir” grant managed by the Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01), a European Commission “REGPOT-CT-2011-285837-STRonGer” grant within the FP7, and the Institut Pasteur de la Guyane. This study was conducted within the “BioViRo” program, supported by an “Investissement d’Avenir” grant managed by the Agence Nationale de la Recherche (CEBA, reference ANR-10-LABX-25-01). Field work conducted by BdT was funded by the ViRUSES program, supported by European funds (FEDER) and assistance from Région Guyane and Direction Régionale pour la Recherche et la Technologie, the ZNIEFF Guyane (DEAL Guyane) program, the GUYAMAZON II program, and Réseau des Observatoires Hommes-Milieux (OHM-Oyapok APR 2013, François Catzefflis). High-throughput sequencing was performed on the Biomix Platform, C2RT, Institut Pasteur, Paris, France, supported by France Génomique (ANR-10-INBS-09-09) and IBISA. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** All in-house scripts used in this study are available at the github repository under the link: [github.com/stirera/rodentsvirome\\_filter1](https://github.com/stirera/rodentsvirome_filter1) (accessed on 9 August 2021).

**Acknowledgments:** We warmly acknowledge François Catzefflis for fruitful discussions on the biology, ecology, and dynamics of rodents. We acknowledge François Catzefflis and all field volunteers for their assistance in capturing. We thank Jean-Marc Frigerio of the UMR BIOGECO for his help in data management and analysis. We thank the C3BI teams at the Institut Pasteur and UMR BIOGECO for their technical support and access to the clusters. We acknowledge the support of IDRIS National Computing Center for access to Turing and Ada machines through the DARI project A0040307360. We thank L. Ma and M. Tichit (Biomics Platform) for technical assistance.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sanjuán, R.; Nebot, M.R.; Chirico, N.; Mansky, L.M.; Belshaw, R. Viral mutation rates. *J. Virol.* **2010**, *84*, 9733–9748. [CrossRef]
2. Houldcroft, C.J.; Beale, M.A.; Breuer, J. Clinical and biological insights from viral genome sequencing. *Nat. Rev. Microbiol.* **2017**, *15*, 183–192. [CrossRef]
3. Williams, S.H.; Che, X.; Garcia, J.A.; Klena, J.D.; Lee, B.; Muller, D.; Ulrich, W.; Corrigan, R.M.; Nichol, S.; Jain, K.; et al. Viral diversity of house mice in New York City. *MBio* **2018**, *9*, e02101-19. [CrossRef]
4. Hartlage, A.S.; Cullen, J.M.; Kapoor, A. The strange, expanding World of animal Hepaciviruses. *Annu. Rev. Virol.* **2016**, *3*, 53–75. [CrossRef]
5. Shi, M.; Lin, X.D.; Chen, X.; Tian, J.H.; Chen, L.J.; Li, K.; Wang, W.; Eden, J.S.; Shen, J.J.; Liu, L.; et al. The evolutionary history of vertebrate RNA viruses. *Nature* **2018**, *556*, 197–202. [CrossRef]
6. Buck, C.B.; Van Doorslaer, K.; Peretti, A.; Geoghegan, E.M.; Tisza, M.J.; An, P.; Katz, J.P.; Pipas, J.M.; McBride, A.A.; Camus, A.C.; et al. The ancient evolutionary history of Polyomaviruses. *PLoS Pathog.* **2016**, *12*, e1000863. [CrossRef]
7. Roux, S.; Brum, J.R.; Dutilh, B.E.; Sunagawa, S.; Duhaime, M.B.; Loy, A.; Poulos, B.T.; Solonenko, N.; Lara, E.; Poulain, J.; et al. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* **2016**, *537*, 689–693. [CrossRef] [PubMed]
8. Bergner, L.M.; Orton, R.J.; Benavides, J.A.; Becker, D.J.; Tello, C.; Biek, R.; Streicker, D.G. Demographic and environmental drivers of metagenomic viral diversity in vampire bats. *Mol. Ecol.* **2020**, *29*, 26–39. [CrossRef] [PubMed]
9. Wu, Z.; Lu, L.; Du, J.; Yang, L.; Ren, X.; Liu, B.; Jiang, J.; Yang, J.; Dong, J.; Sun, L.; et al. Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. *Microbiome* **2018**, *6*, 1–14. [CrossRef]
10. Cobián Güemes, A.G.; Youle, M.; Cantú, V.A.; Felts, B.; Nulton, J.; Rohwer, F. Viruses as winners in the game of life. *Annu. Rev. Virol.* **2016**, *3*, 197–214. [CrossRef] [PubMed]
11. Kreuder Johnson, C.; Hitchens, P.L.; Smiley Evans, T.; Goldstein, T.; Thomas, K.; Clements, A.; Joly, D.O.; Wolfe, N.D.; Daszak, P.; Karesh, W.B.; et al. Spillover and pandemic properties of zoonotic viruses with high host plasticity. *Sci. Rep.* **2015**, *5*, 1–8. [CrossRef] [PubMed]
12. Dennehy, J.J. Evolutionary ecology of virus emergence. *Ann. N. Y. Acad. Sci.* **2017**, *1389*, 124–146. [CrossRef]
13. McMahon, B.J.; Morand, S.; Gray, J.S. Ecosystem change and zoonoses in the Anthropocene. *Zoonoses Public Health* **2018**, *65*, 755–765. [CrossRef]
14. Jones, K.E.; Patel, N.G.; Levy, M.A.; Storeygard, A.; Balk, D.; Gittleman, J.L.; Daszak, P. Global trends in emerging infectious diseases. *Nature* **2008**, *451*, 990–993. [CrossRef]
15. Siu, A.; Wong, Y.C.R. Economic Impact of SARS: The case of Hong Kong. *Asian Econ. Pap.* **2004**, *3*, 62–83. [CrossRef]
16. Huber, C.; Finelli, L.; Stevens, W. The economic and social burden of the 2014 Ebola outbreak in West Africa. *J. Infect. Dis.* **2018**, *218*, S698–S704. [CrossRef] [PubMed]
17. European Centre for Disease Prevention and Control COVID-19 Situation Dashboard. Available online: <https://gap.ecdc.europa.eu/public/extensions/COVID-19/COVID-19.html#global-overview-tabs> (accessed on 20 July 2021).
18. Bogich, T.L.; Chunara, R.; Scales, D.; Chan, E.; Pinheiro, L.C.; Chmura, A.A.; Carroll, D.; Daszak, P.; Brownstein, J.S. Preventing pandemics via international development: A systems approach. *PLoS Med.* **2012**, *9*, e1001354. [CrossRef]
19. PREDICT. Available online: <https://ohi.vetmed.ucdavis.edu/programs-projects/predict-project/about> (accessed on 28 June 2021).
20. Daszak, P.; Carroll, D.; Wolfe, N.; Mazet, J. The global virome project. *Int. J. Infect. Dis.* **2016**, *53*, 36. [CrossRef]
21. Carroll, D.; Daszak, P.; Wolfe, N.D.; Gao, G.F.; Morel, C.M.; Morzaria, S.; Pablos-Méndez, A.; Tomori, O.; Mazet, J.A.K. The Global Virome Project. *Science (80)* **2018**, *359*, 872–874. [CrossRef]
22. Kruse, H.; Kirkemo, A.M.; Handeland, K. Wildlife as source of zoonotic infections. *Emerg. Infect. Dis.* **2004**, *10*, 2067–2072. [CrossRef]
23. Mollentze, N.; Streicker, D.G. Viral zoonotic risk is homogenous among taxonomic orders of mammalian and avian reservoir hosts. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 9423–9430. [CrossRef]
24. Corbet, G.B.; Hill, J.E.; Wilson, D.E.; Reeder, D.M. (Eds.) *Mammal Species Of the World: A Taxonomic and Geographic Reference*, 2nd ed.; Smithsonian Institution Press: Washington, DC, USA, 1993; ISBN 1-56098-217-9.
25. Huchon, D.; Douzery, E.J.P. From the Old World to the New World: A molecular chronicle of the phylogeny and biogeography of hystricognath rodents. *Mol. Phylogenetics Evol.* **2001**, *20*, 238–251. [CrossRef] [PubMed]

26. Gravinatti, M.L.; Barbosa, C.M.; Soares, R.M.; Gregori, F. Synanthropic rodents as virus reservoirs and transmitters. *Rev. Soc. Bras. Med. Trop.* **2020**, *53*, 1–11. [[CrossRef](#)]
27. Han, B.A.; Schmidt, J.P.; Bowden, S.E.; Drake, J.M. Rodent reservoirs of future zoonotic diseases. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7039–7044. [[CrossRef](#)] [[PubMed](#)]
28. Ruedas, L.A.; Salazar-Bravo, J.; Tinnin, D.S.; Armien, B.; Cáceres, L.; García, A.; Díaz, M.A.; Gracia, F.; Suzán, G.; Peters, C.J.; et al. Community ecology of small mammal populations in Panamá following an outbreak of Hantavirus pulmonary syndrome. *J. Vector Ecol.* **2004**, *29*, 177–191. [[PubMed](#)]
29. Chen, L.; Liu, B.; Wu, Z.; Jin, Q.; Yang, J. DRodVir: A resource for exploring the virome diversity in rodents. *J. Genet. Genomics* **2017**, *44*, 259–264. [[CrossRef](#)]
30. Database of Rodent-Associated Viruses. Available online: <http://www.mgc.ac.cn/DRodVir/> (accessed on 20 July 2021).
31. Meerburg, B.G.; Singleton, G.R.; Kijlstra, A. Rodent-borne diseases and their risks for public health. *Crit. Rev. Microbiol.* **2009**, *35*, 221–270. [[CrossRef](#)] [[PubMed](#)]
32. De Thoisy, B.; Matheus, S.; Catzefflis, F.; Clément, L.; Barrioz, S.; Guidez, A.; Donato, D.; Cornu, J.F.; Brunaux, O.; Guitet, S.; et al. Maripa Hantavirus in French Guiana: Phylogenetic position and predicted spatial distribution of rodent hosts. *Am. J. Trop. Med. Hyg.* **2014**, *90*, 988–992. [[CrossRef](#)]
33. Lavergne, A.; Matheus, S.; Catzefflis, F.; Donato, D.; Lacoste, V.; de Thoisy, B. Rodent-borne viruses in French Guiana. *Virologie* **2017**, *21*, E12–E27. [[CrossRef](#)]
34. Liu, J.; Liu, D.Y.; Chen, W.; Li, J.L.; Luo, F.; Li, Q.; Ling, J.X.; Liu, Y.Y.; Xiong, H.R.; Ding, X.H.; et al. Genetic analysis of hantaviruses and their rodent hosts in central-south China. *Virus Res.* **2012**, *163*, 439–447. [[CrossRef](#)]
35. Monath, T.P.; Newhouse, V.F.; Kemp, G.E.; Setzer, H.W.; Cacciapuoti, A. Lassa virus isolation from *Mastomys natalensis* rodents during an epidemic in Sierra Leone. *Science* **1974**, *185*, 263–265. [[CrossRef](#)]
36. Guégan, J.-F.F.; Ayoub, A.; Cappelle, J.; de Thoisy, B. Forests and emerging infectious diseases: Unleashing the beast within. *Environ. Res. Lett.* **2020**, *15*, 083007. [[CrossRef](#)]
37. Gentry, A.H. Tree species richness of upper Amazonian forests. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 156–159. [[CrossRef](#)] [[PubMed](#)]
38. Ellwanger, J.H.; Kulmann-Leal, B.; Kaminski, V.L.; Valverde-Villegas, J.M.; DA VEIGA, A.B.G.; Spilki, F.R.; Fearnside, P.M.; Caesar, L.; Giatti, L.L.; Wallau, G.L.; et al. Beyond diversity loss and climate change: Impacts of Amazon deforestation on infectious diseases and public health. *Anais da Academia Brasileira de Ciências* **2020**, *92*, 1–33. [[CrossRef](#)] [[PubMed](#)]
39. Han, B.A.; Kramer, A.M.; Drake, J.M. Global patterns of zoonotic disease in mammals. *Trends Parasitol.* **2016**, *32*, 565–577. [[CrossRef](#)] [[PubMed](#)]
40. Allen, T.; Murray, K.A.; Zambrana-Torrel, C.; Morse, S.S.; Rondinini, C.; Di Marco, M.; Breit, N.; Olival, K.J.; Daszak, P. Global hotspots and correlates of emerging zoonotic diseases. *Nat. Commun.* **2017**, *8*, 1–10. [[CrossRef](#)]
41. Salmier, A.; Tirera, S.; de Thoisy, B.; Franc, A.; Darcissac, E.; Donato, D.; Bouchier, C.; Lacoste, V.; Lavergne, A. Virome analysis of two sympatric bat species (*Desmodus rotundus* and *Molossus molossus*) in French Guiana. *PLoS ONE* **2017**, *12*, e0186943. [[CrossRef](#)]
42. Bolatti, E.M.; Zorec, T.M.; Montani, E.; Hošnjak, L.; Chouhy, D.; Casal, P.E.; Barquez, M. American Free-Tailed Bats (*Tadarida brasiliensis*) and identification of two novel mammalian viruses. *Viruses* **2020**, *12*, 422. [[CrossRef](#)]
43. Matheus, S.; Lavergne, A.; de Thoisy, B.; Dussart, P.; Lacoste, V. Complete genome sequence of a novel Hantavirus variant of Rio Mamore Virus, Maripa Virus, from French Guiana. *J. Virol.* **2012**, *86*, 5399. [[CrossRef](#)]
44. Matheus, S.; Djossou, F.; Moua, D.; Bourbigot, A.M.; Hommel, D.; Lacoste, V.; Dussart, P.; Lavergne, A. Hantavirus pulmonary syndrome, French Guiana. *Emerg. Infect. Dis.* **2010**, *16*, 739–741. [[CrossRef](#)]
45. Matheus, S.; Kalle, H.; Mayence, C.; Bremand, L.; Houcke, S.; Rousset, D.; Lacoste, V.; de Thoisy, B.; Hommel, D.; Lavergne, A. Hantavirus pulmonary syndrome caused by Maripa virus in French Guiana, 2008–2016. *Emerg. Infect. Dis.* **2017**, *23*, 1722–1725. [[CrossRef](#)] [[PubMed](#)]
46. Lavergne, A.; de Thoisy, B.; Donato, D.; Guidez, A.; Matheus, S.; Catzefflis, F.; Lacoste, V. Patawa Virus, a new Arenavirus hosted by forest rodents in French Guiana. *Ecohealth* **2015**, *12*, 339–346. [[CrossRef](#)] [[PubMed](#)]
47. Lavergne, A.; de Thoisy, B.; Tirera, S.; Donato, D.; Bouchier, C.; Catzefflis, F.; Lacoste, V. Identification of lymphocytic choriomeningitis mammarenavirus in house mouse (*Mus musculus*, Rodentia) in French Guiana. *Infect. Genet. Evol.* **2016**, *37*, 225–230. [[CrossRef](#)]
48. Sikes, R.S. 2016 Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. *J. Mammal.* **2016**, *97*, 663–688. [[CrossRef](#)] [[PubMed](#)]
49. Direction régionale de l'ONF Guyane. *Occupation du sol en 2015 sur la Bande Littorale de la Guyane et son Évolution Entre 2005 et 2015*; Office National des Forêts: Biscarrosse, France, 2017.
50. European Environment Agency. *European Environment Agency, Copenhagen Corine Land Cover*; EEA: København, Denmark, 1994.
51. Ivanova, N.V.; Dewaard, J.R.; Hebert, P.D.N. An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Mol. Ecol. Notes* **2006**, *6*, 998–1002. [[CrossRef](#)]
52. Borisenko, A.V.; Lim, B.K.; Ivanova, N.V.; Hanner, R.H.; Hebert, P.D.N. DNA barcoding in surveys of small mammal communities: A field study in Suriname. *Mol. Ecol. Resour.* **2008**, *8*, 471–479. [[CrossRef](#)]
53. Allander, T.; Emerson, S.U.; Engle, R.E.; Purcell, R.H.; Bukh, J. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 11609–11614. [[CrossRef](#)]

54. Lo, C.-C.; Chain, P.S.G. Rapid evaluation and quality control of next generation sequencing data with FaQCs. *BMC Bioinform.* **2014**, *15*, 366. [CrossRef]
55. Li, D.; Liu, C.M.; Luo, R.; Sadakane, K.; Lam, T.W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **2014**, *31*, 1674–1676. [CrossRef]
56. Li, D.; Luo, R.; Liu, C.M.; Leung, C.M.; Ting, H.F.; Sadakane, K.; Yamashita, H.; Lam, T.W. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **2016**, *102*, 3–11. [CrossRef]
57. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.
58. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef] [PubMed]
59. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
60. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 1–9. [CrossRef] [PubMed]
61. Index of/pub/taxonomy/new\_taxdump. Available online: [ftp.ncbi.nlm.nih.gov/pub/taxonomy/new\\_taxdump](ftp.ncbi.nlm.nih.gov/pub/taxonomy/new_taxdump) (accessed on 13 November 2019).
62. International Committee on Taxonomy of Viruses (ICTV). Available online: <https://talk.ictvonline.org/> (accessed on 9 August 2020).
63. ViralZone. Available online: <https://viralzone.expasy.org/5576> (accessed on 9 August 2020).
64. Research, Laboratory of Chemical Life Science, Kyoto University Bioinformatics Center, I. for C. Virus-Host DB. Available online: <https://www.genome.jp/virushostdb/> (accessed on 9 August 2020).
65. Nayfach, S.; Camargo, A.P.; Schulz, F.; Elie-Fadrosh, E.; Roux, S.; Kyrpides, N.C. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat. Biotechnol.* **2020**, *39*, 578–585. [CrossRef]
66. Oksanen, J.; Kindt, R.; Legendre, P.; O'Hara, B.; Simpson, G.L.; Solymos, P.M.; Stevens, M.H.H.; Wagner, H. The vegan package. *Community Ecol. Packag.* **2008**, *3*, 190.
67. Hill, M.O. Diversity and evenness: An unifying notation and its consequences. *Ecology* **1973**, *54*, 427–432. [CrossRef]
68. Rényi, A. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*; Contributions to the Theory of Statistics; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 547–561.
69. Margalef, R. Información y diversidad específica en las comunidades de organismos. *Investigación Pesquera* **1956**, *3*, 99–106.
70. Magurran, A.E.; Henderson, P.A. Explaining the excess of rare species in natural species abundance distributions. *Nature* **2003**, *422*, 714–716. [CrossRef]
71. Tothmeresz, B. Comparison of different methods for diversity ordering. *J. Veg. Sci.* **1995**, *6*, 283–290. [CrossRef]
72. Kumar, S.; Stecher, G.; Li, M.; Niyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]
73. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef] [PubMed]
74. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics* **2011**, *27*, 1164–1165. [CrossRef]
75. Ronquist, F.; Huelsenbeck, J.P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **2003**, *19*, 1572–1574. [CrossRef]
76. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [CrossRef] [PubMed]
77. Suchard, M.A.; Lemey, P.; Baele, G.; Ayres, D.L.; Drummond, A.J.; Rambaut, A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **2018**, *4*, 1–5. [CrossRef] [PubMed]
78. Moens, U.; Calvignac-Spencer, S.; Lauber, C.; Ramqvist, T.; Feltkamp, M.C.W.; Daugherty, M.D.; Verschoor, E.J.; Ehlers, B. ICTV virus taxonomy profile: Polyomaviridae. *J. Gen. Virol.* **2017**, *98*, 1159–1160. [CrossRef]
79. Smith, D.B.; Becher, P.; Bukh, J.; Gould, E.A.; Meyers, G.; Monath, T.; Muerhoff, A.S.; Pletnev, A.; Rico-Hesse, R.; Stapleton, J.T.; et al. Proposed update to the taxonomy of the genera Hepacivirus and Pegivirus within the Flaviviridae family. *J. Gen. Virol.* **2016**, *97*, 2894–2907. [CrossRef] [PubMed]
80. Abudurexiti, A.; Adkins, S.; Alioto, D.; Alkhovskiy, S.V.; Avšič-Županc, T.; Ballinger, M.J.; Bente, D.A.; Beer, M.; Bergeron, É.; Blair, C.D.; et al. Taxonomy of the order Bunyavirales: Update 2019. *Arch. Virol.* **2019**, *164*, 1949–1965. [CrossRef]
81. Marklewitz, M.; Dutari, L.C.; Paraskevopoulou, S.; Page, R.A.; Loaiza, J.R.; Junglen, S. Diverse novel phleboviruses in sandflies from the Panama Canal area, Central Panama. *J. Gen. Virol.* **2019**, *100*, 938–949. [CrossRef]
82. Nunes-Neto, J.P.; De Souza, W.M.; Acrani, G.O.; Romeiro, M.F.; Fumagalli, M.; Vieira, L.C.; De Almeida Medeiros, D.B.; Lima, J.A.; De Lima, C.P.S.; Cardoso, J.F.; et al. Characterization of the bujaru, frijoles and tapara antigenic complexes into the sandfly fever group and two unclassified phleboviruses from Brazil. *J. Gen. Virol.* **2017**, *98*, 585–594. [CrossRef] [PubMed]
83. PhyloPic. Available online: <http://phylopic.org> (accessed on 27 November 2020).

84. Wu, Z.; Han, Y.; Liu, B.; Li, H.; Zhu, G.; Latinne, A.; Dong, J.; Sun, L.; Du, J.; Zhou, S.; et al. Decoding the RNA viromes of rodent lungs provides new visions into the origin and evolution pattern of rodent-borne diseases in Mainland Southeast Asia. *Microbiome* **2020**, *9*, 1–19. [[CrossRef](#)]
85. 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.; et al. Detection of zoonotic pathogens and characterization of novel viruses carried by commensal *Rattus norvegicus* in New York city. *MBio* **2014**, *5*, e01933–14. [[CrossRef](#)] [[PubMed](#)]
86. Phan, T.G.; Kapusinszky, B.; Wang, C.; Rose, R.K.; Lipton, H.L.; Delwart, E.L. The fecal viral flora of wild rodents. *PLoS Pathog.* **2011**, *7*, e1002218. [[CrossRef](#)] [[PubMed](#)]
87. Nieto-Rabiela, F.; Suzán, G.; Wiratsudakul, A.; Rico-Chávez, O. Viral metacommunities associated to bats and rodents at different spatial scales. *Community Ecol.* **2018**, *19*, 168–175. [[CrossRef](#)] [[PubMed](#)]
88. Campbell, S.J.; Ashley, W.; Gil-Fernandez, M.; Newsome, T.M.; Di Giallonardo, F.; Ortiz-Baez, A.S.; Mahar, J.E.; Towerton, A.L.; Gillings, M.; Holmes, E.C.; et al. Red fox viromes in urban and rural landscapes. *Virus Evol.* **2020**, *6*, veaa065. [[CrossRef](#)]
89. Geoghegan, J.L.; Giallonardo, F.D.; Wille, M.; Ortiz-Baez, A.S.; Costa, V.A.; Ghaly, T.; Mifsud, J.C.O.; Turnbull, O.M.H.; Bellwood, D.R.; Williamson, J.E.; et al. Host evolutionary history and ecology shape virome composition in fishes. *bioRxiv* **2020**. [[CrossRef](#)]
90. Sachsenröder, J.; Braun, A.; Machnowska, P.; Ng, T.F.F.; Deng, X.; Guenther, S.; Bernstein, S.; Ulrich, R.G.; Delwart, E.; Johne, R. Metagenomic identification of novel enteric viruses in urban wild rats and genome characterization of a group A rotavirus. *J. Gen. Virol.* **2014**, *95*, 2734–2747. [[CrossRef](#)]
91. Sogin, M.L.; Morrison, H.G.; Huber, J.A.; Welch, D.B.M.M.; Huse, S.M.; Neal, P.R.; Arrieta, J.M.; Herndl, G.J. Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12115–12120. [[CrossRef](#)]
92. Guillotin, M. Rythmes d’activité et régimes alimentaires de *Proechimys cuvieri* et d’*Oryzomys capito velutinus* (Rodentia) en forêt guyanaise. *Rev. Ecol. (Terre Vie)* **1982**, *36*, 337–371.
93. Tier, A.; de Carvalho, W.D.; Rostain, S.; Catzefflis, F.; Claessens, O.; Dewynter, M.; McKey, D.; Mustin, K.; Palisse, M.; de Thoisy, B. The amazonian savannas of French Guiana: Cultural and social importance, biodiversity, and conservation challenges. *Trop. Conserv. Sci.* **2020**, *13*, 1940082919900471. [[CrossRef](#)]
94. Solbrig, O.T. The diversity of the savanna ecosystem. *Ecol. Stud.* **1996**, *121*, 1–27. [[CrossRef](#)]
95. Gubler, D.J. The global emergence/resurgence of arboviral diseases as public health problems. *Arch. Med. Res.* **2002**, *33*, 330–342. [[CrossRef](#)]
96. Machalaba, C.; Karesh, W.B. Emerging infectious disease risk: Shared drivers with environmental change. *OIE Rev. Sci. Technol.* **2017**, *36*, 435–444. [[CrossRef](#)]
97. Tilman, D.; May, R.M.; Lehman, C.L.; Nowack, M.A. Habitat destruction and the extinction debt revisited. *Nature* **1994**, *371*, 65–66. [[CrossRef](#)]
98. Ostfeld, R.S.; Keesing, F. Biodiversity and disease risk: The case of Lyme disease. *Conserv. Biol.* **2000**, *14*, 722–728. [[CrossRef](#)]
99. Schmid, J.; Rasche, A.; Eibner, G.; Jeworowski, L.; Page, R.A.; Corman, V.M.; Drosten, C.; Sommer, S. Ecological drivers of Hepacivirus infection in a neotropical rodent inhabiting landscapes with various degrees of human environmental change. *Oecologia* **2018**, *188*, 289–302. [[CrossRef](#)] [[PubMed](#)]
100. Suzán, G.; Marcé, E.; Giermakowski, J.T.; Mills, J.N.; Ceballos, G.; Ostfeld, R.S.; Armien, B.; Pascale, J.M.; Yates, T.L. Experimental evidence for reduced rodent diversity causing increased hantavirus prevalence. *PLoS ONE* **2009**, *4*, e5461. [[CrossRef](#)] [[PubMed](#)]
101. Maia, F.G.M.; de Souza, W.M.; Sabino-Santos, G.; Fumagalli, M.J.; Modha, S.; Murcia, P.R.; Figueiredo, L.T.M. A novel polyomavirus in sigmodontine rodents from São Paulo State, Brazil. *Arch. Virol.* **2018**, *163*, 2913–2915. [[CrossRef](#)] [[PubMed](#)]
102. da Silva, M.S.; Cibulski, S.P.; Alves, C.D.B.T.; Weber, M.N.; Budaszewski, R.F.; Silveira, S.; Mósena, A.C.S.; Mayer, F.Q.; Goltz, L.V.; Campos, R.; et al. New polyomavirus species identified in nutria, *Myocastor coypus* polyomavirus 1. *Arch. Virol.* **2018**, *163*, 3203–3206. [[CrossRef](#)]
103. Moens, U.; Krumbholz, A.; Ehlers, B.; Zell, R.; Johne, R.; Calvignac-Spencer, S.; Lauber, C. Biology, evolution, and medical importance of polyomaviruses: An update. *Infect. Genet. Evol.* **2017**, *54*, 18–38. [[CrossRef](#)]
104. Calvignac-Spencer, S.; Feltkamp, M.C.W.; Daugherty, M.D.; Moens, U.; Ramqvist, T.; Johne, R.; Ehlers, B. A taxonomy update for the family *Polyomaviridae*. *Arch. Virol.* **2016**, *161*, 1739–1750. [[CrossRef](#)]
105. Ehlers, B.; Anoh, A.E.; Salem, N.B.; Broll, S.; Couacy-Hymann, E.; Fischer, D.; Gedvilaite, A.; Ingenhütt, N.; Liebmann, S.; Martin, M.; et al. Novel polyomaviruses in mammals from multiple orders and reassessment of polyomavirus evolution and taxonomy. *Viruses* **2019**, *11*, 930. [[CrossRef](#)]
106. Tan, Z.; Gonzalez, G.; Sheng, J.; Wu, J.; Zhang, F.; Xu, L.; Zhang, P.; Zhu, A.; Qu, Y.; Tu, C.; et al. Extensive genetic diversity of Polyomaviruses in sympatric bat communities: Host switching versus coevolution. *J. Virol.* **2020**, *94*, e02101–19. [[CrossRef](#)]
107. Carr, M.; Gonzalez, G.; Sasaki, M.; Dool, S.E.; Ito, K.; Ishii, A.; Hang’ombe, B.M.; Mweene, A.S.; Teeling, E.C.; Hall, W.W.; et al. Identification of the same polyomavirus species in different African horseshoe bat species is indicative of short-range host-switching events. *J. Gen. Virol.* **2017**, *98*, 2771–2785. [[CrossRef](#)] [[PubMed](#)]
108. Rasche, A.; Sander, A.L.; Corman, V.M.; Drexler, J.F. Evolutionary biology of human hepatitis viruses. *J. Hepatol.* **2019**, *70*, 501–520. [[CrossRef](#)] [[PubMed](#)]
109. Shi, M.; Lin, X.D.; Tian, J.H.; Chen, L.J.; Chen, X.; Li, C.X.; Qin, X.C.; Li, J.; Cao, J.P.; Eden, J.S.; et al. Redefining the invertebrate RNA virosphere. *Nature* **2016**, *540*, 539–543. [[CrossRef](#)] [[PubMed](#)]

110. Li, L.L.; Liu, M.M.; Shen, S.; Zhang, Y.J.; Xu, Y.L.; Deng, H.Y.; Deng, F.; Duan, Z.J. Detection and characterization of a novel hepacivirus in long-tailed ground squirrels (*Spermophilus undulatus*) in China. *Arch. Virol.* **2019**, *164*, 2401–2410. [[CrossRef](#)] [[PubMed](#)]
111. Kapoor, A.; Simmonds, P.; Scheel, T.K.H.; Hjelle, B.; Cullen, J.M.; Burbelo, P.D.; Chauhan, L.V.; Duraisamy, R.; Sanchez Leon, M.; Jain, K.; et al. Identification of rodent homologs of hepatitis C virus and pegiviruses. *MBio* **2013**, *4*, e00216–13. [[CrossRef](#)] [[PubMed](#)]
112. Drexler, J.F.; Corman, V.M.; Müller, M.A.; Lukashev, A.N.; Gmyl, A.; Coutard, B.; Adam, A.; Ritz, D.; Leijten, L.M.; van Riel, D.; et al. Evidence for Novel Hepaciviruses in Rodents. *PLoS Pathog.* **2013**, *9*, e1003438. [[CrossRef](#)]
113. de Souza, W.; Fumagalli, M.; Sabino-Santos, G.; Motta Maia, F.; Modha, S.; Teixeira Nunes, M.; Murcia, P.; Moraes Figueiredo, L. A novel Hepacivirus in wild rodents from South America. *Viruses* **2019**, *11*, 297. [[CrossRef](#)]
114. Scheel, T.K.H.; Simmonds, P.; Kapoor, A. Surveying the global virome: Identification and characterization of HCV-related animal hepaciviruses. *Antivir. Res.* **2015**, *115*, 83–93. [[CrossRef](#)] [[PubMed](#)]
115. Pfaender, S.; Brown, R.J.P.; Pietschmann, T.; Steinmann, E. Natural reservoirs for homologs of hepatitis C virus. *Emerg. Microbes Infect.* **2014**, *3*, 1–9. [[CrossRef](#)] [[PubMed](#)]
116. Pybus, O.G.; Gray, R.R. The virus whose family expanded. *Nature* **2013**, *498*, 310–311. [[CrossRef](#)]
117. Weaver, S.C.; Winegar, R.; Manger, I.D.; Forrester, N.L. Alphaviruses: Population genetics and determinants of emergence. *Antivir. Res.* **2012**, *94*, 242–257. [[CrossRef](#)] [[PubMed](#)]
118. Aguilar, P.V.; Greene, I.P.; Coffey, L.L.; Medina, G.; Moncayo, A.C.; Anishchenko, M.; Ludwig, G.V.; Turell, M.J.; O’Guinn, M.L.; Lee, J.; et al. Endemic Venezuelan Equine Encephalitis in northern Peru. *Emerg. Infect. Dis.* **2004**, *10*, 880–888. [[CrossRef](#)] [[PubMed](#)]
119. Rico-Hesse, R.; Weaver, S.C.; De Siger, J.; Medina, G.; Salas, R.A. Emergence of a new epidemic/epizootic Venezuelan Equine Encephalitis virus in South America. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5278–5281. [[CrossRef](#)]
120. Rivas, F.; Diaz, L.A.; Cardenas, V.M.; Daza, E.; Bruzon, L.; Alcala, A.; De La Hoz, O.; Caceres, F.M.; Aristizabal, G.; Martinez, J.W.; et al. Epidemic Venezuelan Equine Encephalitis in La Guajira, Colombia, 1995. *J. Infect. Dis.* **1997**, *175*, 828–832. [[CrossRef](#)]
121. Taylor, K.G.; Paessler, S. Pathogenesis of Venezuelan Equine Encephalitis. *Vet. Microbiol.* **2013**, *167*, 145–150. [[CrossRef](#)] [[PubMed](#)]
122. Weaver, S.C.; Ferro, C.; Barrera, R.; Boshell, J.; Navarro, J.C. Venezuelan Equine Encephalitis. *Annu. Rev. Entomol.* **2004**, *49*, 141–174. [[CrossRef](#)]
123. Phillipotts, R.J.; O’Brien, L.; Appleton, R.E.; Carr, S.; Bennett, A. Intranasal immunisation with defective adenovirus serotype 5 expressing the Venezuelan Equine Encephalitis virus E2 glycoprotein protects against airborne challenge with virulent virus. *Vaccine* **2005**, *23*, 1615–1623. [[CrossRef](#)]
124. Hommel, D.; Heraud, J.M.; Hulin, A.; Talarmin, A. Association of Tonate virus (subtype IIIB of the Venezuelan Equine Encephalitis complex) with encephalitis in a human. *Clin. Infect. Dis.* **2000**, *30*, 188–190. [[CrossRef](#)]
125. Monath, T.P.; Lazuick, J.S.; Cropp, C.B.; Rush, W.A.; Calisher, C.H.; Kinney, R.M.; Trent, D.W.; Kemp, G.E.; Bowen, G.S.; Francy, D.B. Recovery of Tonate virus (“bijou bridge” strain), a member of the Venezuelan Equine Encephalomyelitis virus complex, from Cliff Swallow nest bugs (*Oeciacus vicarius*) and nestling birds in North America. *Am. J. Trop. Med. Hyg.* **1980**, *29*, 969–983. [[CrossRef](#)] [[PubMed](#)]
126. Carrara, A.S.; Gonzales, M.; Ferro, C.; Tamayo, M.; Aronson, J.; Paessler, S.; Anishchenko, M.; Boshell, J.; Weaver, S.C. Venezuelan Equine Encephalitis virus infection of spiny rats. *Emerg. Infect. Dis.* **2005**, *11*, 663–669. [[CrossRef](#)]
127. Deardorff, E.R.; Forrester, N.L.; Travassos Da Rosa, A.P.; Estrada-Franco, J.G.; Navarro-Lopez, R.; Tesh, R.B.; Weaver, S.C. Experimental infection of potential reservoir hosts with Venezuelan Equine Encephalitis virus, Mexico. *Emerg. Infect. Dis.* **2009**, *15*, 519–525. [[CrossRef](#)]
128. Nutricy, R.; Djossou, F.; Matheus, S.; Lorenzi-Martinez, E.; De Laval, F.; Demar, M.; Nacher, M.; Rousset, D.; Epelboin, L. Discriminating Tonate Virus from Dengue virus infection: A matched case–control study in French Guiana, 2003–2016. *Am. J. Trop. Med. Hyg.* **2020**, *102*, 195–201. [[CrossRef](#)] [[PubMed](#)]
129. Aguilar, P.V.; Estrada-Franco, J.G.; Navarro-Lopez, R.; Ferro, C.; Haddow, A.D.; Weaver, S.C. Endemic Venezuelan Equine Encephalitis in the Americas: Hidden under the dengue umbrella. *Future Virol.* **2011**, *6*, 721–740. [[CrossRef](#)] [[PubMed](#)]
130. De Carvalho, M.S.; De Lara Pinto, A.Z.; Pinheiro, A.; Rodrigues, J.S.V.; Melo, F.L.; Da Silva, L.A.; Ribeiro, B.M.; Dezengrini-Slhessarenko, R. Viola Phlebovirus is a novel Phlebotomus fever serogroup member identified in *Lutzomyia (Lutzomyia) longipalpis* from Brazilian Pantanal. *Parasites Vectors* **2018**, *11*, 1–10. [[CrossRef](#)] [[PubMed](#)]
131. Palacios, G.; Wiley, M.R.; Travassos da Rosa, A.P.A.; Guzman, H.; Quiroz, E.; Savji, N.; Carrera, J.P.; Bussetti, A.V.; Ladner, J.T.; Ian Lipkin, W.; et al. Characterization of the Punta Toro species complex (genus *Phlebovirus*, family *Bunyaviridae*). *J. Gen. Virol.* **2015**, *96*, 2079–2085. [[CrossRef](#)]
132. Gundacker, N.D.; Carrera, J.P.; Castillo, M.; Diaz, Y.; Valenzuela, J.; Tamhane, A.; Moreno, B.; Pascale, J.M.; Tesh, R.B.; López-Vergès, S. Clinical manifestations of Punta Toro virus species complex infections, Panama, 2009. *Emerg. Infect. Dis.* **2017**, *23*, 872–874. [[CrossRef](#)]
133. Lorenz, C.; de Oliveira Lage, M.; Chiaravalloti-Neto, F. Deforestation hotspots, climate crisis, and the perfect scenario for the next epidemic: The Amazon time bomb. *Sci. Total Environ.* **2021**, *783*, 147090. [[CrossRef](#)] [[PubMed](#)]