

OPEN

# Distribution and genetic diversity of adeno-associated viruses in bats from coastal areas of Southeast China

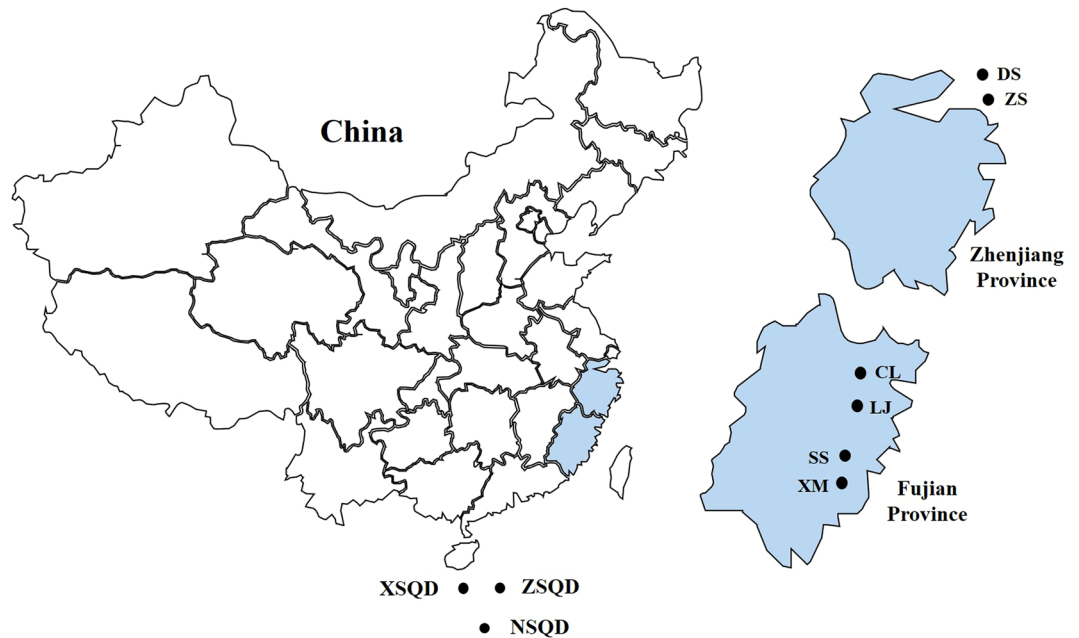
Changqiang Zhu<sup>1,5</sup>, Chunhui Wang<sup>1,5</sup>, Jiahong Wu<sup>3,5</sup>, Fuqiang Ye<sup>1,5</sup>, Ruichen Lv<sup>1</sup>, Dan Hu<sup>1</sup>, Lele Ai<sup>1</sup>, Lu Yang<sup>1</sup>, Ting Wu<sup>4</sup>, Bo Li<sup>1</sup>, Chenxi Ding<sup>1</sup>, Bin Zhang<sup>1</sup>, Heng Lv<sup>1</sup>, Changjun Wang<sup>2\*</sup> & Weilong Tan<sup>1\*</sup>

Bats are associated with several important zoonotic viruses from different families. One example includes adeno-associated viruses (AAVs), that are extensively detected in several animals, especially primates. To understand AAVs distribution and genetic diversity in the coastal areas of Southeast China, a total of 415 intestine samples were mostly collected from two provinces of southeast China, i.e., Zhejiang and Fujian province. Intestine samples from five bat species were collected for AAVs detection. The average prevalence rate for AAV detection among these samples was 18.6% (77 positives out of 415 samples) and ranged from 11.8 to 28.9% between the five bat species. This suggests that AAVs are widely distributed in diverse bat populations in southeast coastal areas of China. Based on the genome sequence of bat adeno-associated virus-CXC1 (BtAAV-CXC1) from one AAV-positive sample, the genetic diversity of the detected AAVs were assessed and analyzed. Phylogenetic analysis revealed that BtAAV-CXC1 was comparatively distant to other major AAVs from mammals and non-mammals, with only a 52.9–64.7% nucleotide identity. However, they were phylogenetically closer to *Rhinolophus sinicus* bat adeno-associated virus (Rs-BtAAV1), with a 74.5% nt similarity. Partial analysis of the rep and cap overlapping open reading frame (ORF) sequences from bat AAV samples revealed 48 partial rep sequences and 23 partial cap sequences from positive samples shared 86.9 to 100% and 72.3 to 98.8% nucleotide identities among themselves, respectively. This suggests that the detected AAVs had a distinctly high genetic diversity. These findings led us to conclude that diverse AAVs may be widely distributed in bat populations from the southeast regions of China.

Adeno-associated viruses (AAVs) were discovered in 1965 as a small contaminant. AAVs are a group of small single-stranded DNA (ssDNA) viruses without envelopes that belong to the parvovirus family. The entire genome of AAV is approximately 4.7 kb and contains two major ORFs that encode for Rep and Cap proteins. Under normal circumstances, the AAV virus can coexist within the mammalian host for an extended period without apparent pathogenicity. However, productive infection only occurs in the presence of a helper virus, either adenovirus or herpesvirus<sup>1–7</sup>. Due to this, AAV has become an important gene therapy vector for its non-pathogenic and conditional infectious life cycle<sup>8</sup>.

The genetic diversity of AAV is very high among its natural hosts. To date, thirteen AAV serotypes (designated AAV1–AAV13) and more than 100 unique variants have been isolated and identified from primate and non-primate hosts<sup>9–25</sup>. Hosts can be as diverse as Chiropteras, Porcine, Serpentes, Bovines, Caprines and Chickens. Among these natural hosts, bats are more geographically dispersed. In addition, Bats host a greater number of emerging viruses per species compared to any other mammalian clade<sup>26,27</sup>. Bats have been increasingly recognized as a natural host for AAVs in recent years, in addition to the many other important zoonotic

<sup>1</sup>Eastern Theater Command Centers for Disease Control and Prevention, 293 Zhongshan East Rd, Nanjing, 210002, P. R. China. <sup>2</sup>Institute for Disease Prevention and Control of PLA, Beijing, 100071, China. <sup>3</sup>Key Laboratory of Environmental Pollution Monitoring and Disease Control, Ministry of Education, Guizhou Medical University, Guiyang, 550025, China. <sup>4</sup>Jinling Hospital Nanjing, Nanjing, 210002, China. <sup>5</sup>These authors contributed equally: Changqiang Zhu, Chunhui Wang, Jiahong Wu and Fuqiang Ye. \*email: [science2008@hotmail.com](mailto:science2008@hotmail.com); [njcdc@163.com](mailto:njcdc@163.com)



**Figure 1.** Sampling locations of bats in the two southeast coastal provinces of China. (D.S.: Daishan, Z.S.: Zhoushan, C.L.: Changle, L.J.: Lianjiang, S.S.: Shishi, X.M.: Xiamen, X.S.Q.D.: Xishaqundao, Z.S.Q.D.: Zhongshaqundao, N.S.Q.D.: nanshaqundao).

viruses, such as SARS, MERS, and Ebola virus<sup>28–31</sup>. With the development of molecular and sequencing technologies, numerous novel AAVs are being discovered. These include BtAAV-YNM<sup>14</sup>, Rp-BtAAV1 and Rs-BtAAV1<sup>32</sup>. Yan Li identified several adenoviruses and AAVs in a wide host range of bat species with remarkable genetic diversity<sup>14</sup>. Lau SKP investigated the presence of parvoviruses among different bat species in Hong Kong and mainland China<sup>32</sup>. Their work broadened the current knowledge of AAV diversity in Bats. Several studies also have investigated the origins of viral diversity in bats or the maintenance of bat-associated viruses in nature. In addition, considerable emphasis has been placed on understanding the direct epidemiological interface from bats to humans<sup>33–35</sup>.

In our previous study, several novel bat-associated viruses in southeast China were annotated through metagenomic sequencing technology<sup>36</sup>, including coronavirus, norovirus, adenovirus, bocavirus, astrovirus, circovirus, and adeno-associated viruses. To date, there have been no adeno-associated virus studies in the coastal areas of Southeast China. The main aim of this study was to understand AAV distribution and genetic diversity in these areas. Our study could enrich the understanding of bats as a pathogen library and provided new insights into the transmission of AAV from bats to humans.

## Results

### Identification and prevalence of AAVs in bats in the coastal regions of Southeast China.

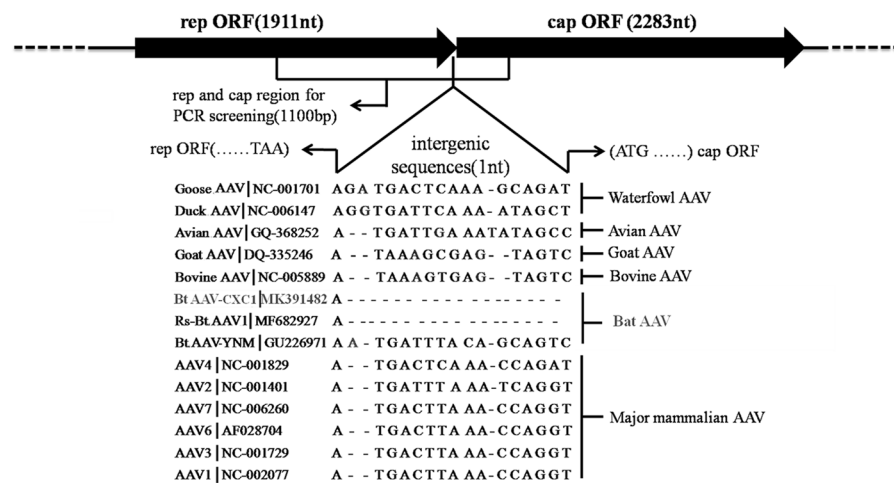
A total of 415 intestinal tissue samples were collected from five bat species in Zhejiang and Fujian provinces from Southeast China from 2015 to 2018 (Fig. 1). These samples were then used to determine and understand the presence of AAVs. A 1100 bp fragment encompassing the rep and cap regions of known AAV genome was successfully amplified from 77 out of 415 bat samples (18.6%). The prevalence rate ranged from 11.8 to 28.9% and was detected from four insectivorous bat species. AAVs in the study were detected at nearly all sampling locations except in the SS and LJ regions. Among all four bat species, the virus detection rate for *Rhinolophus pusillus* was the highest (4 locations out of 6), followed by *Myotis davidii*, *Rhinolophus ferrumequinum* and *Scotophilus kuhli*. These results demonstrated that AAVs were widely distributed in diverse bat populations in southeast coastal regions of China (Table 1).

**Nearly complete genome analysis of BtAAV-CXC1.** We sequenced one bat AAV strain that was isolated from *Rhinolophus pusilla* sample in Fujian province and designated it BtAAV-CXC1 (GenBank accession number: MK391482). The nearly complete genome of BtAAV-CXC1 was 4263 nt in length with a G + C content of 54.51%, with a 4194 nt coding region and a 1 nt intergenic region. The putative ORFs of rep and cap were predicted to be 1911 and 2283 nt in length, respectively. The Rep and Cap proteins of BtAAV-CXC1 shared an average of 61.0% (ranging from 46.4 to 85.7%) and 27.7% (ranging from 10.7 to 35.5%) pairwise amino acid identity, respectively to the reference AAV genome (Fig. 2). This demonstrated that the genome of BtAAV-CXC1 was obviously different from the reference AAV genome in this study.

**Phylogenetic origin of BtAAV-CXC1.** In order to investigate the phylogenetic origin of BtAAV-CXC1, a large-scale maximum likelihood (ML) phylogenetic tree was constructed with 26 sequences from members of the *Dependovirus* genus (both full length and partial genomes). The ML tree demonstrated that BtAAV-CXC1 was

Bat species	No. positive/no. tested (%)						
	ZS	DS	XM	CL	SS	LJ	All locations
<i>Rhinolophus pusillus</i>	30/102(29.4)	20/91(22.0)	2/19(10.5)	8/44(18.2)	0/35	0/24	60/315(11.8)
<i>Rhinolophus ferrumequinum</i>	—	5/29(17.2)	—	—	—	—	5/29(17.2)
<i>Scotophilus kuhli</i>	—	—	1/8(12.5)	—	—	—	1/8(12.5)
<i>Myotis davidii</i>	7/27(25.9)	—	—	4/11(36.4)	—	—	11/38(28.9)
<i>Myotis formosus</i>	—	0/25	—	—	—	—	0/25
All species	37/129(28.7)	25/145(17.2)	3/27(11.1)	12/55(21.8)	0/35	0/24	77/415(18.6)

**Table 1.** Prevalence of AAVs among the five bat species from the two provinces in China. \*ZS: Dinghai; DS: Daishan; XM: Xiamen; CL: Changle; SS: Shishi; LJ: Lianjiang.



**Figure 2.** The genome of BtAAV-CXC1. Schematic diagram demonstrating the relative scale of rep and cap ORFs, and the intergenic sequence of BtAAV-CXC1. The region selected for PCR is indicated. Intergenic sequence alignment of BtAAV-CXC1 to those of other AAVs. AAV species are shown left of the alignment and the rough classification is shown on the right of the alignment.

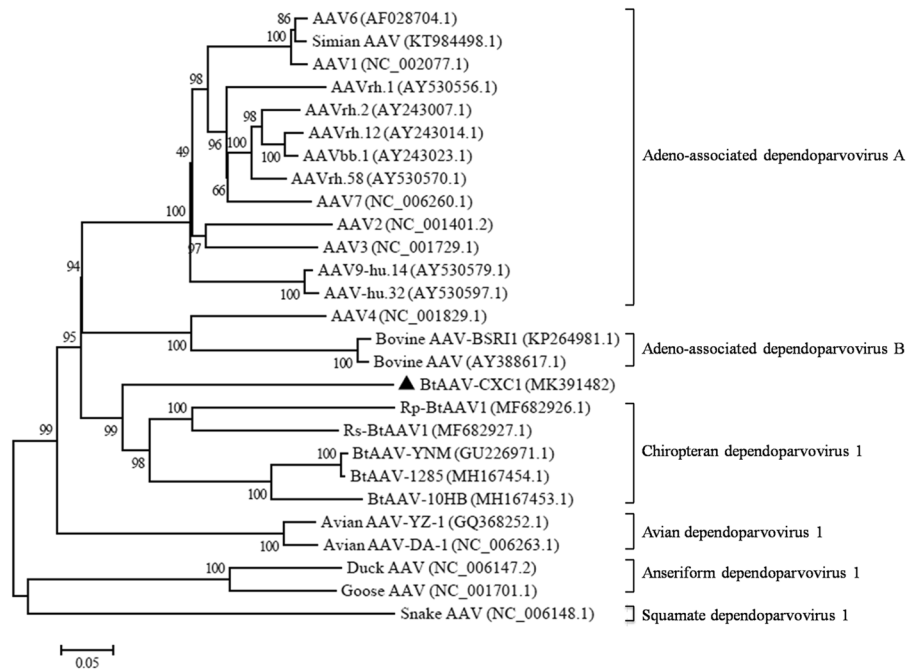
most closely related to Rs-BtAAV1, a previously reported BtAAV from a *Rhinolophus sinicus* bat<sup>32</sup>, with a 25.5% nucleotide difference (Fig. 3). BtAAV-CXC1 was also related to other bat AAVs with a 58.4–77.5% nt identity and was relatively divergent from the majority of AAVs from primates and other animals (i.e., bovines, chicken, waterfowl and snakes), with a 52.9–64.7% nucleotide similarity. These findings indicated that BtAAV-CXC1 was a novel BtAAV species.

**Partial rep and cap sequence analysis of bat AAVs.** To investigate the genetic diversity of bat AAVs, partial rep and cap ORF sequences were obtained through direct sequencing of the PCR products ( $n = 48$  and  $n = 23$ ; selected from positive samples during the PCR screening assay). The amplified regions covered 1100 nt of the cap region and 841 nt of the rep region based on Rs-BtAAV1 (GenBank accession number: MF682927). As shown in Fig. 4, 48 amplicons of the rep region were grouped into bat clade with 86.9 to 100% nucleotide identities between them. In addition, they shared 71.5 to 84.2% identities with other AAVs (Table 2). Further analysis of the 23 partial cap sequences demonstrated that 23 bat AAVs shared 72.3 to 98.8% nucleotide identities among themselves and 59.4 to 72.9% identities with other AAVs (Fig. 5). These results demonstrated that diverse AAVs were widely distributed in bat populations from the southeast coastal regions of China.

## Discussion

Adeno-associated viruses were detected from the majority of bats collected from the southeastern coastal regions of China. However, no adeno-associated virus was successfully isolated from Vero E6 cells and no genomic sequences of adeno-associated virus were detected in challenge experiments using suckling mice. This may be due to the viral replication of adeno-associated viruses that rely on adenovirus proliferation and expansion.

Four out of the five insectivorous bat species, including *Rhinolophus pusillus*, *Rhinolophus ferrumequinum*, *Scotophilus kuhli*, and *Myotis davidii*, have been reported to be native carriers of diverse AAVs around the southeastern coastal regions of China. The average prevalence rate is as high as 18.6%, which suggests that AAVs are considerably prevalent among various bat species. However, the absence of *Myotis formosus* AAV in this study may be due to the limited sampling size. Partial rep and cap ORF sequences analysis of the detectable bat AAVs revealed a relatively large genetic diversity among the bat samples. This suggests that diverse AAVs may be widely distributed in bat populations as previous studies have reported<sup>14,32</sup>. However, further extensive work has to be



**Figure 3.** Phylogenetic origin of BtAAV-CXC1. The phylogenetic tree was constructed using the maximum likelihood method and the best evolutionary model (GTR + G + I). Bootstrap values were calculated from 1000 trees. The triangles in the phylogenetic trees denote the sequences derived in this study.

performed to discover new AAV strains and their native bat species around the southeastern coastal regions of China.

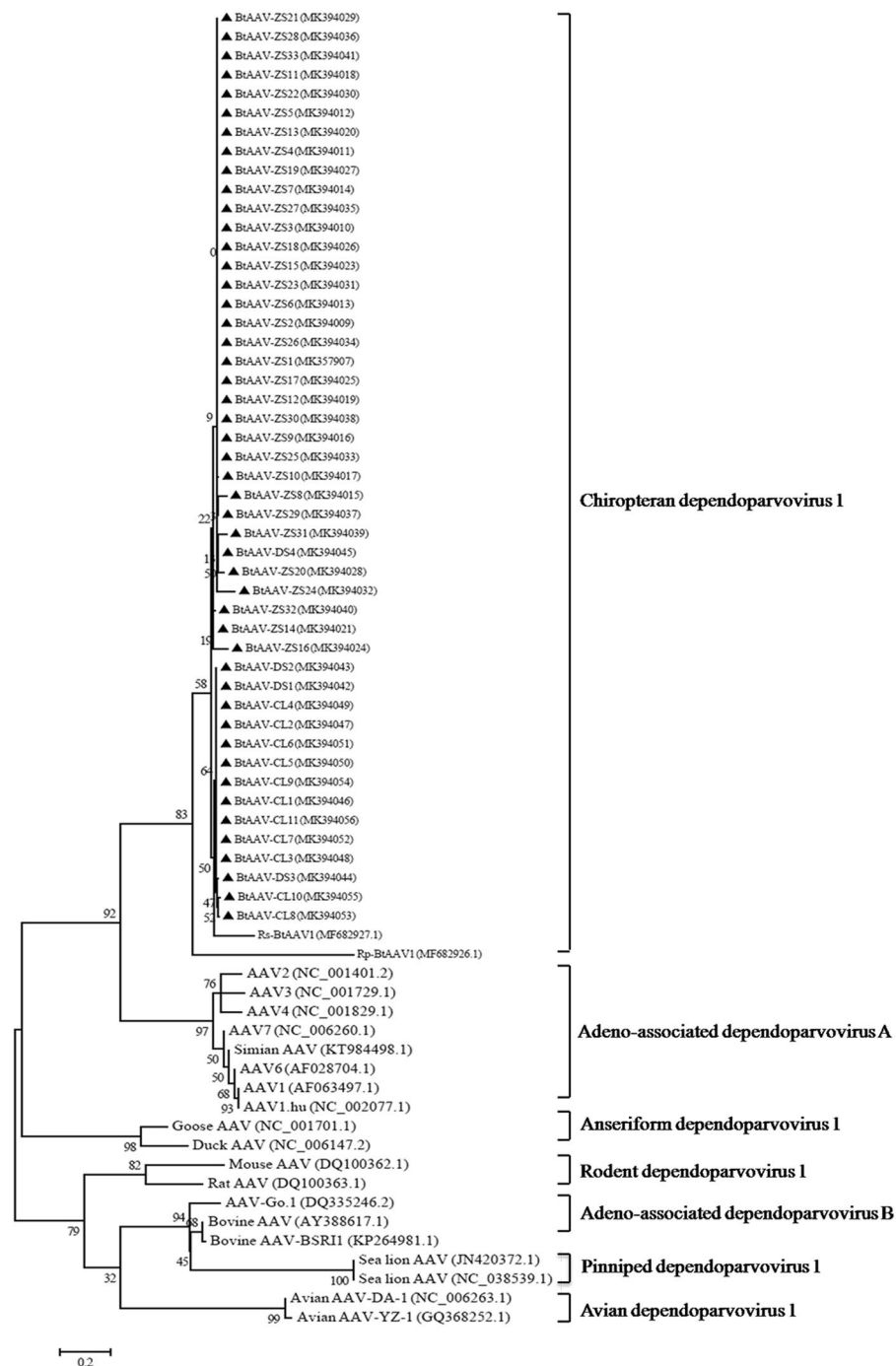
The phylogenetic study of BtAAV-CXC1 with 26 other AAVs suggested that it was a novel AAV member of the genus Dependovirus. It was phylogenetically closer to mammalian AAVs compared to avian, waterfowl or snake AAVs. BtAAV-CXC1 shared a large branch with the previously reported Rp-BtAAV1, BtAAV1-YNM, and other known human or non-human primates AAVs. This included Adeno-associated virus A, Adeno-associated virus B, and all proven human AAVs (Fig. 3). Amino acid sequence analysis of the rep gene of this novel BtAAV-CXC1 showed 64.8% similarity with human AAV2. Similar results were obtained in a previous study<sup>32</sup>. In addition, several previous studies have indicated that mammalian and non-mammalian AAVs had a separate evolutionary path, except for a few avian, waterfowl and reptilian AAVs that were distantly related to mammalian AAVs. Bats are thought to be the original host for alpha coronaviruses and beta coronaviruses, while birds for gamma coronaviruses and delta coronaviruses<sup>32,37,38</sup>. The discovery of additional AAVs from bats and other mammals will contribute to a better understanding of the evolutionary origin and pathways of mammalian AAVs. Our study found that the co-occurrence of bat adenovirus and AAV in bat samples was as high as 33.7%, which was similar to the findings in Li *et al.*'s study<sup>14</sup>. This may be because other viruses such as herpesvirus and vaccinia virus are essential for the replication of AAV<sup>39</sup>.

Adeno-associated viruses (AAVs) are frequently used viral vectors in biomedical research and human gene therapy. The clinical use of AAV vectors have been studied in detail. Over 160 phase I, II, and III clinical trials have been implemented since 1994, and have noticeably increased since 2011<sup>40</sup>. To date, substantial efforts have been dedicated on engineering strategies to develop novel AAV capsid variants with increased tropism for precise cell types with lower seroreactivity, and increased manufacturability. Our results revealed that the deduced Cap protein of BtAAV-CXC1 shared an average of 27.7% (ranging from 10.7 to 35.5%) pairwise amino acid identity with other mammalian AAVs. This suggests that the Cap protein of BtAAV-CXC1 may have a low cross-serological reactivity to known mammalian AAVs and hence could likely be considered to an substitutable vector for gene therapy.

In conclusion, our study of viruses in bat samples in southeast China identified novel AAVs and their distinct genetic diversity. This extends our knowledge of the diversity of bat AAVs and its evolution and emergence in humans and animals. However additional studies are warranted using larger samplings from different locations to further understand the global diversity of bat viruses.

## Materials and Methods

**Ethical approval.** All procedures and protocols for sample collection and processing were approved by the Administrative Committee on Animal Welfare of the Institute of Zhejiang and Fujian CDC Veterinary and the Ethics Committee of the CDC of Easten Theater. All methods were performed in accordance with the relevant guidelines and regulations (Approval number: 2015009).



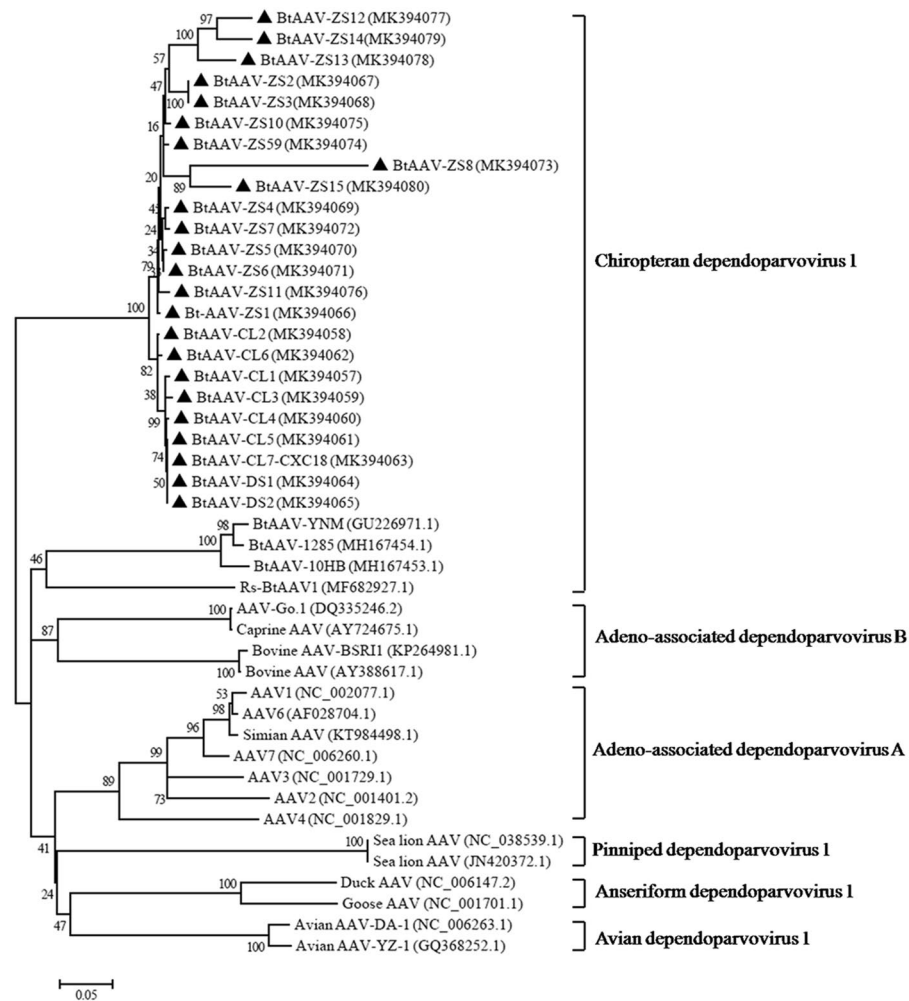
**Figure 4.** The phylogenetic origin derived from partial rep sequence analysis of bat AAVs. The phylogenetic tree was constructed using the maximum likelihood method and the best evolutionary model (GTR + G + I). Bootstrap values were calculated from 1000 trees. The triangles in the phylogenetic trees denote sequences derived in the study.

**Sample collection.** Bat samples were collected from July 2015 to May 2018 from Zhejiang Province (including Dinghai and Daishan) and Fujian Province (including Changle, Lianjiang, Shishi and Xiamen), China. A total of 415 adult bats were captured alive at mountain caves with mist nets. Captured bats were carefully removed from the nets from each sampling location and were immediately euthanized and dissected. Bat details are shown in Table 1. Each sample (approximately 1 g of intestine tissue) was immediately transferred into viral transport medium and stored in liquid nitrogen prior to transportation to the laboratory<sup>35</sup>. Samples were stored at  $-80^{\circ}\text{C}$  until analyzed.

**RNA extraction and RT-PCR analysis.** All samples were pooled and then underwent reverse transcription PCR analysis as previously described<sup>36</sup>. Briefly, each intestine sample ( $\sim 0.1$  g) was homogenized

Location/ species	Minimum value/Maximum value (Mean value)										
	CL(batAAVs)	DS(batAAVs)	ZS(batAAVs)	Bat	Human	Caprine	Bovine	Sea_lion	Avian	Duck	Goose
CL(batAAVs)	96.9/100.0(99.3)	95.2/100.0(98.2)	87.0/98.8(95.9)	78.3/83.0(81.9)	73.2/8.16(81.7)	74.8/78.0(76.4)	71.9/78.3(75.7)	74.7/82.0(80.3)	78.4/80.8(79.8)	76.5/77.9(77.4)	74.8/76.3(75.7)
DS(batAAVs)	95.2/100.0(98.2)	94.3/100.0(97.0)	87.6/99.9(96.2)	78.6/83.0(81.7)	76.2/84.1(81.7)	75.5/78.6(76.8)	74.1/77.3(75.7)	77.3/81.2(80.1)	75.3/80.8(78.6)	77.9/78.6(78.2)	76.3
ZS(batAAVs)	87.0/98.8(96.2)	87.6/99.9(96.2)	86.9/100.0(98.2)	75.9/82.5(81.8)	73.4/84.1(81.2)	73.3/78.8(77.9)	71.6/77.9(76.6)	75.2/81.9(80.5)	71.5/78.6(77.7)	73.4/78.6(77.9)	76.3

**Table 2.** Comparison of AAV genetic diversity in different hosts (pairwise nucleotide similarity based on the cap region). \*ZS: Dinghai; DS: Daishan; CL: Changle.



**Figure 5.** The phylogenetic origin derived from partial cap sequence analysis of bat AAVs. The tree was constructed using the maximum likelihood method and the best evolutionary model (GTR + G + I). Bootstrap values were calculated from 1000 trees. The triangles in the phylogenetic trees denote sequences derived in the study.

using a glass grinder with 10 volumes of SM buffer (50 mM Tris, 10 mM MgSO<sub>4</sub>, 0.1 M NaCl, pH7.5). The homogenates were centrifuged at 12,000 g for 10 min at 4 °C. The supernatants from each sample were passed through a 0.22 μm Pellicon II filter (Millipore, Billerica, MA). Viral RNA was extracted using the Viral RNA Mini Kit (Qiagen) following the manufacturer's instructions. RNA was eluted in 30 μl RNase-free H<sub>2</sub>O and stored at -80 °C. Reverse transcription was performed using the 1st cDNA synthesis kit (TaKaRa), according to the manufacturer's protocol. PCR was performed using cDNA with the following primers; forward primer 5'-AAGACCAACATCGCGGACATC-3' and reverse primer 5'-TAGTTCTTGTGWRGTRTT-3'. The PCR reaction generated a 1000 nt fragment of the partial rep and cap gene. Primers were designed based on the nucleotide sequence of Rs-BtAAV1 (MF682927). The PCR amplification was performed in a 50 μl reaction mix containing 2 μl extracted cDNA, 10 μl PCR Buffer, 20 pmol of each primer, 4 μl dNTP (2.5 mM) and 1 μl Taq DNA polymerase (PrimeSTAR GXL, Takara). PCR conditions included 30 cycles of amplification, consisting of denaturation at 98 °C for 10 s, annealing at 65 °C for 15 s, extension at 68 °C for 1 min. Standard precautions were taken to avoid PCR contamination, and a negative control was included for each PCR assay. Amplified PCR products

Primer	Sequences (5'-3')	Amplification length
AAV-Rep1	GTCCCATTGACGTGGAGGAACATC	841 bp
AAV-Rep2	GGGAGCCGCTTAGTCAGTTCAAAT	
AAV-Cap1	AAGACCAACATCGCGGAGGCCATC	1100 bp
AAS- Cap2	GCGTGGTTGACTTGAGGTA	

**Table 3.** Primers used for PCR amplification of the partial rep and cap sequences.

were sequenced in both directions using an ABI 3730 DNA Analyzer (Invitrogen, Beijing, China). Bat AAV identified from bat species *Rhinolophus pusillus* in the Fujiang province (designated BtAAV-CXC1) was selected for additional full-length genome sequencing (Supplementary Table S1).

**Sequencing of bat AAVs partial rep and cap genome regions.** After PCR screening, all positive samples were PCR amplified that targeted 841 bp and 1100 bp fragments of rep and cap ORFs. Primer sequences are shown in Table 3. PCR amplification was performed in a 50  $\mu$ l reaction mix containing 2  $\mu$ l extracted cDNA, 10  $\mu$ l PCR Buffer, 20 pmol of each primer, 4  $\mu$ l dNTP (2.5 mM) and 1  $\mu$ l *Taq* DNA polymerase (PrimeSTAR GXL, Takara). PCR conditions were; 30 cycles of amplification consisting of denaturation at 98 °C for 10 s, annealing at 65 °C for 15 s, extension at 68 °C for 1 or 2 min. Standard precautions were taken to avoid PCR contamination, and a negative control was included in each PCR assay. Amplified PCR products were sequenced in both directions using the ABI 3730 DNA Analyzer (Invitrogen, Beijing, China).

**Sequence and phylogenetic analysis.** Viral nucleotide and derived amino acid sequences were analyzed using the DNASTAR software package (Lasergene). Putative ORFs with a minimum size of 150 aa and a coding capacity were predicted using the NCBI ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). Amplified sequences obtained by PCR amplification were aligned using BLAST in GenBank. Representative sequences of other known AAVs were downloaded from GenBank. MEGA7.0 software was used to construct the maximum-likelihood evolutionary tree of the nearly complete genome sequence of BtAAV-CXC1 or partial rep and cap sequences of bat AAVs. Bootstrap values were estimated using 1000 replicates on the ML substitution model<sup>41</sup>.

**Nucleotide sequence accession numbers.** All sequences obtained in this study were deposited in GenBank under accession numbers MK391482 for BtAAV-CXC1 complete genome, MK357907, MK394009–MK394021 and MK394023–MK394056 for the 48 partial rep sequences of AAVs, MK394057–MK394080 for the 24 partial cap sequences of AAVs<sup>42</sup>.

Received: 1 November 2019; Accepted: 13 February 2020;

Published online: 28 February 2020

## References

- Atchison, R. W., Casto, B. C. & Hammon, W. M. Adeno associated virus defective virus particles. *Sci.* **149**, 754–756 (1965).
- Hastie, E. & Samulski, R. J. Adeno-associated virus at 50: a golden anniversary of discovery, research, and gene therapy success—a personal perspective. *Hum. Gene Ther.* **26**, 257–65 (2015).
- Rose, J. A., Hoggan, M. D. & Shatkin, A. J. Nucleic acid from an adenoassociated virus: chemical and physical studies. *Proc. Natl Acad. Sci. USA* **56**, 86–92 (1966).
- Michael, F., Naso, B., Tomkowicz, Perry, William L. III & Strohl, William R. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *BioDrugs* **31**, 317–334 (2017).
- TRL, F. & Berns, K. I. Adeno-Associated Virus: A Ubiquitous Commensal of Mammals. *Hum. Gene Ther.* **16**, 401–407 (2005).
- Samulski, R. J. & Muzyczka, N. AAV-mediated gene therapy for research and therapeutic purposes. *Annu. Rev. Virol.* **1**, 427–451 (2014).
- Sandoval, I. M., Kuhn, N. M. & Manfredsson, F. P. Multimodal Production of Adeno-Associated Virus. *Methods Mol. Biol.* **1937**, 101–124 (2019).
- He, X., Xie, H., Liu, X. & Gu, F. Basic and Clinical Application of Adeno-Associated Virus-Mediated Genome Editing. *Hum. Gene Ther.* **30**(6), 673–681 (2019).
- Bantel-Schaal, U. & zur Hausen, H. Characterization of the DNA of a defective human parvovirus isolated from a genital site. *Virology* **134**, 52–63 (1984).
- Chiorini, J. A., Kim, F., Yang, L. & Kotin, R. M. Cloning and characterization of adeno-associated virus type 5. *J. Virol.* **73**, 1309–1319 (1999).
- Zhang R, Cao L, Cui M, Sun Z, Hu M. Adeno-associated virus 2 bound to its cellular receptor AAVR. *Nat. Microbiol.*, <https://doi.org/10.1038/s41564-018-0356-7>(2019).
- Schmidt, M. *et al.* Adeno-associated virus type 12 (AAV12): a novel AAV serotype with sialic acid- and heparan sulfate proteoglycan-independent transduction activity. *J. Virol.* **82**, 1399–1406 (2008).
- Schnepf, B. C., Jensen, R. L., Clark, K. R. & Johnson, P. R. Infectious molecular clones of adeno-associated virus isolated directly from human tissues. *J. Virol.* **83**, 1456–1464 (2009).
- Li, Y. *et al.* Prevalence and genetic diversity of adeno-associated viruses in bats from China. *J. Gen. Virology* **91**, 2601–2609 (2010).
- Salganik, M., Hirsch, M. L. & Samulski, R. J. Adeno-associated Virus as a Mammalian DNA Vector. *Microbiol. Spectr.* **3**, 1–32 (2015).
- Schmidt, M., Katano, H., Bossis, I. & Chiorini, J. A. Cloning and characterization of a bovine adeno-associated virus. *J. Virol.* **78**, 6509–6516 (2004).
- Ng, T. F. *et al.* A metagenomics and case-control study to identify virus associated with bovine respiratory disease. *J. Virol.* **89**, 5340–5349 (2015).
- Puppo, A. *et al.* Recombinant vectors based on porcine adeno-associated viral serotypes transduce the murine and pig retina. *PLoS ONE* **8**, 1–12 (2013).
- Bello, A. *et al.* Novel Adeno-associated Viruses Derived From Pig Tissues Transduce Most Major Organs in Mice. *Sci. Rep.* **4**, 1–11 (2014).

20. Arbetman, A. E. *et al.* Novel caprine adeno-associated virus (AAV) capsid (AAV-Go.1) is closely related to the primate AAV-5 and has unique tropism and neutralization properties. *J. Virol.* **79**, 15238–15245 (2005).
21. Bossis, I. & Chiorini, J. A. Cloning of an avian adenoassociated virus (AAAV) and generation of recombinant AAAV particles. *J. Virol.* **77**, 6799–6810 (2003).
22. Wang, J., Zhu, L., Zhu, J., Sun, H. & Zhu, G. Molecular characterization and phylogenetic analysis of an avian adeno-associated virus originating from a chicken in China. *Arch. Virol.* **156**, 71–77 (2011).
23. Zádori, Z., Stefancsik, R., Rauch, T. & Kisary, J. Analysis of the complete nucleotide sequences of goose and Muscovy duck parvoviruses indicates common ancestral origin with adeno-associated virus 2. *Virology* **212**, 562–573 (1995).
24. Farkas, S. L. *et al.* A parvovirus isolated from royal python (*Python regius*) is a member of the genus Dependovirus. *J. Gen. Virol.* **85**, 555–561 (2004).
25. Li, L. *et al.* The fecal viral flora of California sea lions. *J. Virol.* **85**, 9909–9917 (2011).
26. Smith, I. & Wang, L. F. Bats and their virome: an important source of emerging viruses capable of infecting humans. *Curr. Opin. Virol.* **3**, 84–91 (2013).
27. Wang, L. F. & Anderson, D. E. Viruses in bats and potential spillover to animals and humans. *Curr. Opin. Virol.* **34**, 79–89 (2019).
28. Li, W. *et al.* Bats are natural reservoirs of SARS-like coronaviruses. *Sci.* **310**, 676–679 (2005).
29. Leroy, E. M. *et al.* Fruit bats as reservoirs of Ebola virus. *Nat.* **438**, 575–576 (2005).
30. Banerjee, A., Kulcsar, K., Misra, V., Frieman, M. & Mossman, K. Bats and Coronaviruses. *Viruses* **11**, 1–15 (2019).
31. Halpin, K., Young, P. L., Field, H. E. & Mackenzie, J. S. Isolation of Hendra virus from pteropid bats: a natural reservoir of Hendra virus. *J. Gen. Virol.* **81**, 1927–1932 (2000).
32. Lau, S. K. P. *et al.* Bats host diverse parvoviruses as possible origin of mammalian dependoparvoviruses and source for bat–swine interspecies transmission. *J. Gen. Virology* **98**, 3046–3059 (2017).
33. Brierley, L., Vonhof, M. J., Olival, K. J., Daszak, P. & Jones, K. E. Quantifying global drivers of zoonotic bat viruses: a process-based perspective. *Am. Nat.* **187**, E53–64 (2016).
34. Mari Saéz, A. *et al.* Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Mol. Med.* **7**, 17–23 (2015).
35. Plowright, R. K. *et al.* Ecological dynamics of emerging bat virus spillover. *Proc. Biol. Sci.* **282**, 1–9 (2015).
36. Hu, D. *et al.* Virome analysis for identification of novel mammalian viruses in bats from southeast China. *Sci. Rep.* **7**, 1–7 (2017).
37. Lau, S. K. *et al.* Complete genome sequence of bat coronavirus HKU2 from Chinese horseshoe bats revealed a much smaller spike gene with a different evolutionary lineage from the rest of the genome. *Virology* **367**, 428–439 (2007).
38. Woo, P. C. *et al.* Discovery of seven novel mammalian and avian coronaviruses in the genus Deltacoronavirus supports bat coronaviruses as the gene source of Alphacoronavirus and Betacoronavirus and avian coronaviruses as the gene source of Gammaparvovirus and Deltacoronavirus. *J. Virol.* **86**, 3995–4008 (2012).
39. Wibbelt, G., Kurth, A., Yasmum, N., Bannert, M. & Nagel, S. Discovery of herpesviruses in bats. *J. Gen. Virol.* **88**, 2651–2655 (2007).
40. Cabanes-Creus, M. *et al.* Codon-Optimization of Wild-Type Adeno-Associated Virus Capsid Sequences Enhances DNA Family Shuffling while Conserving Functionality. *Mol. Ther. Methods Clin. Dev.* **12**, 71–84 (2018).
41. Hu, D. *et al.* Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerg. Microbes Infect.* **7**(1), 154 (2018).
42. Xu, Z. *et al.* Isolation and Identification of a Highly Divergent Kaeng Khoi Virus from Bat Flies (*Eucampsipoda sunndaica*) in China. *Vector Borne Zoonotic Dis.* **19**(1), 73–80 (2019).

## Acknowledgements

We thank Weng Zhang for valuable technical support and Boyu Shen for providing additional bat samples. This study was funded by grants from the National Major Infectious Diseases (2017ZX10303401-007), National Natural Science Foundation of China (U1602223), National Postdoctoral Special Aid (2016T91011), Opening Project of Key Laboratory of Ministry of Education for Environmental Pollution and Disease Monitoring (GMU-2018-HJZ-05), and Medical Research Topics of Jiangsu Provincial Health Planning Commission (Z2018006).

## Author contributions

C.Q.Z., C.H.W., C.J.W., and W.L.T. conceived and designed the study. F.Q.Y., R.C.L. and D.H. performed the sequencing, phylogenetic and molecular analysis. L.L.A., L.Y. and B.L. performed field and lab studies. T.W. and C.X.D. provided bat samples for analysis. J.H.W., B.Z. and H.L. analyzed the results. C.Q.Z. wrote the draft manuscript. C.H.W., C.J.W. and W.L.T. edited the manuscript. All authors read and approved the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-60721-z>.

**Correspondence** and requests for materials should be addressed to C.W. or W.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020