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

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

Audrey Lacroix, Veasna Duong, Vibol Hul, Sorn San, Holl Davun, Keo Omaliss, Sokha Chea, Alexandre Hassanine, Watthana Theppangnag, Soubanh Silithammavongh, Kongsy Khammavongh, Damien O. Joly, Lucy Keatts, Philippe Dussart, Roger Frutos, Sinpakone Singhalath, Aneta Afelt, Zoe Greatorex, Amanda E. Fine, Tracey Goldstein, Sarah Olson, and are divided into two genera:

Astroviruses (astVs) are non-enveloped positive-sense single-stranded RNA viruses, with a genome size ranging from 6.4 to 7.3 kb. They infect birds and mammals, respectively, despite the association between host species and astV genus has been recently evidenced to be permeable (Pankovics et al., 2015). AstVs are distributed widely and some have been identified as a cause of gastroenteritis in humans and other mammals (Koci et al., 2000; Kurtz and Lee, 1987; Matsui and Herrmann, 2003). Previous studies reported a high prevalence and genetic diversity of astroviruses in bats and rodents. AstVs have been detected in a wide variety of bat species, mostly Yangochiroptera and Yinzpterocirhoptera bats. Evidence for varying degrees of host restriction for astroviruses in bats was found. Furthermore, additional Pteropodid hosts were detected. The astroviruses formed distinct phylogenetic clusters within the genus Mamastrovirus, most closely related to other known bat astroviruses. The astrovirus sequences tested them for astroviruses. Our study revealed a high diversity of astroviruses among various Yangochiroptera and Yinzpterocirhoptera bats. Evidence for varying degrees of host restriction for astroviruses in bats was found. Furthermore, additional Pteropodid hosts were detected. The astroviruses formed distinct phylogenetic clusters within the genus Mamastrovirus, most closely related to other known bat astroviruses. The astrovirus sequences were found to be highly saturated indicating that phylogenetic relationships should be interpreted carefully. An astrovirus clustering in a group with other viruses from diverse hosts, including from ungulates and porcupines, was found in a Rousettus bat. These findings suggest that diverse astroviruses can be found in many species of mammals, including bats.

1. Introduction

Astroviruses (astVs) are non-enveloped positive-sense single-stranded RNA viruses, with a genome size ranging from 6.4 to 7.3 kb (De Benedictis et al., 2011). They belong to the family of Astroviridae and are divided into two genera: Avastrovirus and Mamastrovirus. They infect birds and mammals, respectively, despite the association...
Cambodia and Lao PDR both have a high diversity of bat species, respectively 70 and 80 species of chiroptera are present including Yangochiroptera and Yinpterochiroptera bats. Many bat species roost close to human dwellings and several of them are also hunted for food and trade (Lee et al., 2014), leading to opportunities for close contact between humans and bats which may present a risk for the transmission of potential pathogens. Circulation of several viruses has been reported in Cambodian bats including flavivirus, bunyavirus, lyssavirus, and Nipah virus (Salain et al., 1974, Olson et al., 2002; Osborne et al., 2003; Reynes et al., 2004, 2005). Information on the presence and diversity of astVs in bats was lacking in in Lao PDR and Cambodia. We therefore undertook this work in order to investigate the presence of astroviruses in bats in these two countries.

2. Material and methods

2.1. Ethics

The study was approved by the National Veterinary Research Institute and the Forest Administration Department of the Ministry of Agriculture Forestry and Fisheries in Cambodia, as well as by the National Animal Health Laboratory, Department of Livestock and Fisheries, Ministry of Agriculture and Forestry in Lao PDR, and under the Institutional Animal Care and Use Committee at the University of California, Davis (protocol number: 16,048). As Cambodia and Lao PDR have no ethics committee overseeing animal experimentation, animals were treated in accordance with the guidelines of the American Society of Mammalogists, and within the European Union legislation guidelines (Directive 86/609/EEC).

2.2. Collection of bat samples

Bat samples from 17 genera were collected from 45 locations in Lao PDR and Cambodia, over a 3-year period (from November 2010 to December 2013). Bat genera, or species when possible, were identified, based on morphological characteristics by field biologists and veterinarians. The sampling was carried out in two phases: Phase 1 was performed in 2010 by the Institut Pasteur in Cambodia (IPC) in collaboration with the Muséum National d’Histoire Naturelle (MNHN; Paris, France). Bats were captured using harp traps, stored in Mist Net bags (Ecotone, Gdynia, Poland) and humanely euthanized under supervision of the National Veterinary Research Institute in full compliance with local ethical and legal guidelines at the sampling sites in the Cambodian provinces of Ratanakiri, Stung Treng and Preah Vihear. Rectal swabs were stored in viral transport medium (VTM; containing tryptose phosphate Broth 2.95%, 145 mM of NaCl, 5% gelatin, 54 mM Amphotericin B, 106 U of penicillin-streptomycin per liter, 80 mg of gentamicin per liter (Sigma-Aldrich, Irvine, UK)). Lung and other tissue specimens (i.e. liver, spleen, kidney, heart) were placed in separate cryotubes. All specimens were immediately transferred into liquid nitrogen containers before being transported to the Institut Pasteur laboratory where they were stored at −80 °C prior to testing. Phase 2 of sampling was performed by the Wildlife Conservation Society (WCS), as part of USAID PREDICT project, from 2011 to 2013. Bats on Cambodian guano farms are wild and free-ranging, with farmers constructing artificial roosts adjacent to their homes in order to attract bats. Farmers, who wear no protective equipment, collect guano voided onto tarpaulins or nets laid beneath the roosts. Sterile swabs were used to collect freshly voided fecal samples from tarpaulins placed under the roosts. Rectal and oral swabs and tissue samples were also collected from individual animals that had died of natural causes and were found fresh beneath bat guano farm roosts. Each collected fecal sample was considered as representative of one animal. Additionally, oral and rectal swab were collected from freshly dead bats presented by traders or for sale in restaurants or food markets. The freshness of the carcass was determined based on the appreciation of veterinarians (<48 h after being killed), and direct information from the sellers or hunters. Phase 2 samples were immediately placed in VTM or RNA stabilization solution (RNAlater®, Qiagen). Specimens collected in VTM were immediately stored in liquid nitrogen before being transported to the laboratory and stored at −80 °C prior to testing.

2.3. RNA extraction and astrovirus detection

Viral RNA from swabs, feces and from lung specimens were extracted by the QIaamp viral mini kit (Qiagen, Hilden, Germany) and RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the supplier’s instructions. For samples collected during phase 1, if a rectal swab and a lung were both collected from the same animal, the two samples were pooled for RNA extraction. Similarly, during phase 2, if an oral and rectal swab were both collected from one animal, the samples were pooled for RNA extraction. Pooling was implemented to facilitate the detection of other viruses than astroviruses displaying different tropism. The RNA was eluted in 60 μl of AVE buffer (Qiagen) and was used as template for reverse transcription (RT). Reverse transcription (RT) was performed using a SuperScript III kit (Invitrogen, San Diego, CA). Astrovirus RNA detection was performed by a semi-nested PCR targeting a fragment of the RNA-dependent RNA polymerase (RdRp) gene, using broadly reactive primers and the amplification conditions described previously by Chu et al. (2008). The visualization of a 436 bp and a 421 bp fragment after the first and second PCR round was performed by agarose gel electrophoresis. To limit the risk of contamination, RNA extraction, reverse transcription-PCR (RT-PCR), nested-PCR and gel electrophoresis were carried out in separate rooms. Negative (water) and positive (plasmids prepared by cloning the gene of interest) controls were included in each PCR run. When pooled samples tested positive, RNA from the corresponding oral, rectal or lung specimens were re-extracted and tested separately to identify the positive specimen in the pool. DNA products amplified by nested-RT-PCR were sequenced in both directions by direct Sanger sequencing in commercial facilities (Macrogen, Inc., Seoul, Korea). Amplification of longer fragments of the RdRp (ORF1b, ~620 bp) was performed on RNA of the positive samples, by a nested RT-PCR using the same forward primers as for the assay for astVs detection and a gene-specific reverse primer targeting the 3′ end of ORF1b previously described (Chu et al., 2008). When sequencing indicated coinfections may be present (multiple nucleotide peaks seen in the chromatogram), PCR products were cloned by Macrogen, Inc., using a TOPcloner™ TA kit (Enzymics), and 5 individual clones were selected for sequencing in order to assess clonal sequence diversity. Longer fragments of the ORF2 generated by PCR using gene-specific (Supplementary Table 1) and oligo (dT) primers, flanking the ORF2 region at the 3′ end. Fragments were then sequenced by primer walking (Macrogen, Inc., Seoul, Korea).

2.4. Sequence analyses

Sequences were analyzed by CLC Genomics Workbench 3.6.1 and aligned by BioEdit 7.0.9.1. (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Sequences were deposited in GenBank and accession numbers are provided in Supplementary Table 2. Alignments of the detected astrovirus sequences with a representative set of reference sequences retrieved from GenBank were performed using Seaview 4.5.4 (Gouy et al., 2010). The best tree construction model was calculated using MEGA 6 (Tamura et al., 2013); Nucleic acid and protein phylogenetic trees were constructed by the Maximum Likelihood method with the LG + G model in Seaview 4.5.4 (Gouy et al., 2010) and bootstrap values were determined by 1000 replicates. For the purpose of classification, mean amino acid genetic distances of the complete capsid sequences shared by the samples that were successfully sequenced and other known astVs were calculated using the p-dist method of the MEGA 6 software. All positions containing alignment gaps were eliminated using the pairwise deletion option. To assess the robustness of the
phylogenetic analysis, nucleotide polymorphism and genome saturation were analyzed. The ratio (Ts/Tv) of observed transition (mutation from a purine nucleotide to another or a pyrimidine nucleotide to another) versus transversion (mutation for a purine nucleotide to pyrimidine one or from a pyrimidine nucleotide to purine one) was calculated as an indicator of genome saturation. Ts/Tv was calculated for indel-free regions of ORF1b and ORF2 sequences using Seaview 4.5.4 (Gouy et al., 2010). Sequence polymorphism was investigated using the DnaSP 5.10.01 package (Librado and Rozas, 2009).

2.5. Geographic data

Land cover data was obtained from GlobeLand30 service operated by the National Geomatics Center of China (NGCC, 2014). Initial data was produced in 2010 with an update in 2014. Images used for GlobeLand30 (GLC30) classification were multispectral images with a 30-meter resolution. Six classes of land cover were considered: crop land, forest, grassland, wetland, water bodies and human settlement area. Land cover structure for each sampling location of sampling is described in Supplementary Table 3. Data mapping was conducted with Quantum GIS, version 2.8.2.

3. Results

3.1. Sampling areas and sample collection

A total of 1876 animals were sampled in 14 districts in Cambodia (n = 1247 bats) and 13 districts in Lao PDR (n = 629 bats). A total of 1211 oral swabs, 1684 rectal swabs, 187 feces and 328 lung specimens (n = 1247 bats) and 13 districts in Lao PDR (n = 629 bats). A total of 1876 animals were sampled in 14 districts in Cambodia (n = 1247 bats) and 13 districts in Lao PDR (n = 629 bats). A total of 1211 oral swabs, 1684 rectal swabs, 187 feces and 328 lung specimens were obtained from 7 genera belonging to six families: 4 families from the Yinpterochiroptera suborder and 2 families belonging to the Yangochiroptera suborder (Table 1). For samples collected during phase 2, apart from the samples collected in bat guano farms, the exact original location of animals remained unknown as bats were hunted for market trade. However, bats were always captured by hunters in areas close to the location where animals were sampled, i.e. mostly in areas at the border of deep forests, in mixed agricultural zones with sparse forests, in suburban zones close to sparse forest, in natural forests, in places close to water surfaces or in limestone karst areas with mountain forests (Supplementary Fig. 1, Supplementary Table 3). Landscape analysis around each sampling location led to a typology comprising four main groups of land cover (Supplementary Table 3, Supplementary Fig. 1): 1) Deep forest area well isolated from human settlements (sites: A5, A17, A19, A26, A30, A31, A33, A34, A35, A42, A43 and A44), 2) Isolated pockets of cropslands surrounded by forest very often under fragmentation (sites: A4, A14, A18, A20, A21, A22, A23, A24, A25, A27, A28, A29, A32, A36, A37, A38, A39, A40, A41 and A45), 3) Forest edge, mixed agricultural zones with sparse forests (sites: A7, A9, A15 and A16) and 4) typical agricultural zones, villages and suburban zones, often close to highly fragmented forests (sites: A1, A2, A3, A6, A8, A10, A11, A12 and A13). For sites A1, A6, A8, A10 and A11 a key parameter was the presence of wetlands in the immediate vicinity. Almost 70% of specimens collected in Cambodia were from 4 of the 14 districts at sampling sites located in Chom Khsant, Kean Svay, Muong Russei and Kang Meas. Diversity of bats varied between sites as bats from 7 genera belonging to 5 families including Yangochiroptera and Yinpterochiroptera were sampled in the Thala Barivat district, whereas at other sampling sites, only one single genus of bat was sampled, including the lesser Asian house bat (genus Scotophilus) and horsfield’s bat (Myotis horsfieldii) in the Kang Meas and the Kampong Chnang districts, respectively.

3.2. Detection of astroviruses in bats

The number of samples from each bat genus, as well as the detection of astroviruses by semi-nested RT-PCR are shown in Table 1. Of 1876 animals tested, 100 (5.5%) were positive for an astV. Overall, 21 fecal specimens, 4 oral swabs and 76 rectal swabs tested positive (Table 1). In one

| Table 1 | RT-PCR detection of astroviruses in bats from Cambodia and Lao PDR. |
|---------|---------------------------------------------------------------|
| Location | Total No. of bats (percentage (%) of positive) | Cambodia No. of bats (percentage (%) of positive) | Lao PDR No. of bats (percentage (%) of positive) |
|---------|---------------------------------------------|------------------------|---------------------------------------------|
| Total   | 1876 (5.3)                                   | 1247 (68)              | 629 (32)                                    |
| Yangochiroptera | 781 (8.2)                                   | 749 (63)               | 32 (1)                                      |
| Emballonuridae   | 147 (2.7)                                   | 147 (4)                | 0                                           |
| Vesperilionidae  | 634 (9.5)                                   | 602 (59)               | 32 (1)                                      |
| Hesperopterus sp | 0.1 (0)                                     | 0                      | 32 (1)                                      |
| lio       | 32 (3.1)                                     | 0                      | 0                                           |
| Myotis horsfieldii | 47 (42.6)                                   | 47 (20)                | 0                                           |
| Pipistrellus sp | 29 (0)                                      | 29 (0)                 | 0                                           |
| Scotophilus sp  | 524 (7.4)                                   | 524 (39)               | 0                                           |
| Tylonycteris sp | 1 (0)                                       | 1 (0)                  | 0                                           |
| Yinpterochiroptera | 1095 (3.3)                                   | 498 (5)                | 597 (31)                                    |
| Hipposideridae   | 37 (5.4)                                     | 4 (1)                  | 33 (1)                                      |
| Aegyptius sp    | 7 (0)                                       | 0                      | 7 (0)                                       |
| Hipposideros sp | 30 (6.7)                                     | 4 (1)                  | 26 (1)                                      |
| Megadermatidae  | 21 (9.5)                                     | 21 (2)                 | A26                                         |
| Megaderma lyra  | 21 (9.5)                                     | 21 (2)                 | A21, A22, A24, A25                         |
| Pteropodidae    | 882 (3)                                      | 420 (1)                | 422 (26)                                    |
| Cynopterus sp   | 340 (0)                                      | 321 (0)                | 19 (0)                                      |
| Eonycteris sp   | 79 (3.8)                                     | 28 (0)                 | A18, A22, A24, A25                         |
| Macrogrisssus sp| 22 (0)                                       | 21 (0)                 | A25                                         |
| Megacreeps sp   | 98 (0)                                       | 29 (0)                 | A15, A16, A17, A18, A24, A25, A27, A28     |
| Pteropus sp     | 10 (0)                                       | 10 (0)                 | 0                                           |
| Rousettus sp    | 333 (7.2)                                    | 11 (1)                 | A16, A17, A18, A19                         |
| Rhinolophidae   | 155 (3.2)                                    | 53 (1)                 | A42                                         |
| Rhinolophus sp  | 155 (3.2)                                    | 53 (1)                 | A31, A33                                    |

Bat families are shown in bold.
Sites where bats tested positive for astroviruses are underlined and bold.
A. Lacroix et al. / Infection, Genetics and Evolution 47 (2017) 41–50

44

bat, an astV was found in both oral and rectal samples. Astrovirus-positive bats were detected in 13 of the 45 sampling sites in 11 districts: 6 provinces in Lao and 5 in Cambodia. Nine of 17 bat genera were found to be positive for astroviruses. Among the positive bats sampled in phase 2 (n = 90/1547), 76% (n = 68) were destined for human consumption and 24% (n = 22) were from bat guano farms, i.e., artificial roosts erected to facilitate bat guano collection for use as agricultural fertilizer. The proportion positive among genera varied from 2.7% to 42.5%. The highest number of positives was found in Myotis bats. Differences were found within one bat genus when they were sampled at different sites, e.g., the proportion of positives from the genus Scotophilus in a guano farm from Bakan district (site reference A8) was 26.3% (15/70), whereas only 5.3% (4/80) were positive from Scotophilus bats from Kang Meas district (site reference A3). Some genera (e.g., Cynopterus), with large sample size (n = 340 respectively) were found to be negative, while other genera (e.g., Myotis, Megaderma) with small sampling numbers (n = 47 and n = 21 respectively) tested positive (Table 1). In Cambodia, most of the astrovirus-positive animals (65/68) were Yangochiroptera bats, whereas in Lao PDR 27 out of the 32 astVs were found in Yimp terochiroptera, including bats from genera Rousettus and Eonycteris. Differences in the proportion of positives were seen both in host genera and between sampling sites. For example, in Thala Barivat District (reference site A43, Supplementary Fig. 1), 7 astVs were detected in 4 different bat genera (Taphozous, Hipposideros, Rousettus and Rhinolophus) whereas at the Kampong Chhnang site (reference site A6, Supplementary Fig. 1), 23 astVs were detected, all in the Myotis genus (Table 1).

3.3. Phylogenetic analyses

Phylogenetic analyses of 112 amino acid fragments of the RdRp gene revealed that all the astrovirus sequences belonged to the Mamastrovirus genus (Fig. 1). A similar topology was observed using alignments of the corresponding 336 nucleic acid fragments (Supplementary Fig. 2). The virus sequences fell into 10 monophyletic groups, most related to astV sequences previously detected in bats, with the exception of groups 7 and 8 which included astroviruses detected in multiple host and murine species (Fig. 1). A characteristic of this phylogenetic analysis is the very low bootstrap values (Fig. 1) indicative of a lack of population structure, possibly due to the short sequences, and therefore must be interpreted cautiously.

3.4. Host/astrovirus specificity

Astrovirus sequences showed varying degrees of host specificity. For example, astV sequences clustering in groups 1, 3, 4, 8, 9 and 10 were restricted to only one bat genus, i.e., Rousettus for group 1, Myotis for groups 4 and 10, Megaderma for group 9, Scotophilus for groups 3 and 8, whereas sequences from groups 2, 5, 6 and 7 were each detected in several bat genera (Fig. 1), whereas in groups 2, 5, 6 and 7, sequences belong to insectivorous bat genera only (Ia, Taphozous, Rhinolophus, Hipposideros). Group 3 contained 36 astV sequences all detected in lesser Asian house bats (Scotophilus sp.) from three Cambodian provinces (site references A1, A2, A3, A8) and closely related to bat astVs previously found in the same genus in Hainan Island, China (Hu et al., 2014). Astroviruses detected in Scotophilus (n = 1), Myotis (n = 3) and Rhinolophus (n = 1) fell into group 6, which also contained sequences detected previously in insectivorous bats (Zhu et al., 2009; Hu et al., 2014). In this group, sequences detected Scotophilus genus (BatAstV-Anhui, KOC04, KC29, KC37) formed their own branches, closely related to the strain BatAstV-Anhui, also detected in Myotis, showing 83.5 to 91.8% and 98.2 to 99.1% of nucleotide and amino acid identities. The sequences BatAstV-HKD3_2016 and MAstV-22_PA190, detected in a lesser Asian bat and a horseshoe bat respectively, were distant from the other group 6 bat astVs. BatAstV-HKD3_2016 showed only 64% of amino acid similarities with the BatAstV-NX28 sequence detected in a Miniopterus bat in China, and MAstV-22_PA190 displayed 68% of amino acid similarities with the BatAstV-AFCD11, found in a pipistrelle from Hong Kong. Group 9 also presented a high divergence with 77.4% of identity in the amino acid sequence with the closest astV of the tree (i.e. MAstV-24_P039_Sco). Interestingly, 19 sequences from viruses identified from 3 genera of Yimp terochiroptera bats (Rousettus, Eonycteris and Rhinolophus) in Lao PDR and Cambodia, fell into group 7, showing 78.3%, 86.8% of amino acid identity with astVs previously detected in ungulate and porcupine hosts (Hu et al., 2014; Reuter et al., 2011; Shan et al., 2012, 2011) (Fig. 2). Moreover, two sequences identified in lesser Asian bats (Scotophilus sp.) in Cambodia (MAstV-20_TP104 and MAstV-20_TP089) clustered with murine astVs, showing 84 to 87% of amino acid sequence identity. The average pairwise similarities of amino acids within/among bat astV sequences from the groups 1 and 6 were the lowest, with 73.6% and 71.7% (Supplementary Table 4). Additionally, bat astV sequences from monophyletic group 1 related, displayed 50% to 63.6% of amino acid similarities to other MAstVs in the partial RdRp sequence (Supplementary Table 4). Astrovirus sequences were detected in bats sampled in various environments. All the sequences detected in bats from Cambodian guano farms from Bakan (site reference A8) and Kang Meas (site reference A2) districts, belong to the group 3 and were all detected in fecal samples. Heterogeneities were found in terms of location, date of bat collection and host genus, depending on the group of astV sequences. Sequences from 4 different groups (2, 6, 7 and 9) were detected in bats roosting in the same environment in the Thala Barivat District (site references A33, A34) in a suburban environment close to sparse forest. In Kampong Chhnang Province, Horsefield's bats (genus Myotis) sampled from the same colony (A6 on Fig. 1) at the same time, were found to carry astVs belonging to groups 4, 6 and 10. Moreover, astV sequences detected in different sampling locations at different times and/or in different host genus, were shown to be identical or highly similar. For example, in the group 7, 2 sequences from the genus Rousettus bats collected in December 2010 and January 2012 (i.e. MAstV-13_PA178 and MAstV-13_A11) at two sites 530 km apart (i.e., site references A34 (Cambodia) and A24 (Lao PDR)), had 100% identity in nucleotide and amino acid sequences. Within the same group, the sequences MAstV-13 × 60 and MAstV-13 A09, detected both in a Rhinolophus and a Rousettus from the Lao districts of Vieng Thong and Yang Vieng, in December 2011 and January 2012 respectively, also displayed 100% sequences identity.

3.5. Polymorphism and genome saturation assessment

After additional RNA extraction, only 55 out of the 100 positive samples showed a successful amplification using the RT-PCR detection assay (Chu et al., 2008), and were used for additional sequencing. For 29 samples, longer ORF1b fragments were obtained, ranging from 609 bp and 701 bp. Given the lack of structure in the phylogenetic trees, nucleotide polymorphism and saturation assessment were conducted on indel-free fragments from ORF1b and ORF2 sequences. Sequences of ORF1b(2) fragments, ranging from 609 bp and 701 bp were included in the analysis (n = 29). Thus, the three batches of sequences (named ORF1b(1), ORF1b(2) and ORF2) contained fragments of respectively 246, 582 and 716 nucleotides in length. Polymorphism analysis indicated the

---

**Fig. 1.** Relationship of astroviruses constructed using a multiple alignment of 112 amino acids of the partial RdRp gene of detected astroviruses and selected members of Mamastrovirus. The tree was constructed using the Neighbor-Joining method. Bootstrap values were determined by 1000 replicates. PoAstV: Porcine astrovirus (astV); BoAstV: bovine astVs; HAstV: human astV; DromAstV: dromadery astV; PpcAstV: porcine astV; AvAstV: avian astV. The genus of bat host is given for Bat astroviruses, i.e. Min: Miniopterus; Hip: Hipposideros; Tph: Taphozous; Pip: Pipistrellus; Tyl: Tylonycteris; Sco: Scotophilus; Rs: Rousettus; Mgy: Myotis; Mgd: Megaderma; Eon: Eonycteris; Ia: Ia; Rhin: Rhinolophus. For each detected sequence, the strain, sample name and bat genus were included in the nomenclature. For clarity, strains named “PREDICT-MAstV-xx” are abbreviated as “MAstV-xx” in the tree. Accession numbers of the sequences are listed in the Supplementary Table 2. An avian astrovirus (AB033998) is used as outgroup.
Fig. 2. Relationship of the partial ORF1b and ORF2 sequences of the bat astrovirus PREDICT-MAstV-13_A91_Rs to other species of astroviruses. a) Relationship of the partial ORF1b (~230 aa) sequence of the bat astrovirus PREDICT-MAstV-13_A91_Rs to other species of astroviruses. b) Relationship of the partial ORF2 (~380 aa) sequence of the bat astrovirus PREDICT-MAstV-13_A91_Rs to other species of astroviruses. Alignments were based on the encoded amino acid sequences. Trees were constructed using the Neighbor-Joining method. Five other bat astrovirus sequences (BatAstV-LD38, BatAstV-LS11, BatAstV-AFCD57, BatAstV-LD77) corresponding to groups 2, 5, and 6 in Fig. 1, are included. The strain PREDICT-MAstV-13_A91_Rs is marked with a bullet. Astroviruses are abbreviated as follows: PoAstV: Porcine astrovirus (astV); BatAstV: bat astV; BovAstV: bovine astV; HAstV: human astV; DromAstV: dromadery astV; PcpAstV: porcupine astV; AvAstV: avian astV. The genus of bat host is given for Bat astroviruses, i.e. Miniopterus; Hip: Hipposideros; Tph: Taphozous; Pip: Pipistrellus. An avian astrovirus (AB033998) was used as outgroup.
presence of differing trends. A very high level of polymorphism was detected on both ORF1b and ORF2 (Table 2). The number of mutated sites was high and represented 80.5%, 70% and 84.2% of the nucleotides analyzed for ORF1b1, ORF1b2 and ORF2, respectively. The number of mutations observed (\(\eta_i\)) represented more than twice the number of polymorphic sites (Table 2). This indicates that the populations analyzed are not growing or emerging populations. Mutated codons represented 42.7%, 60.3% and 34.9% of the codons analyzed for ORF1b1, ORF1b2 and ORF2, respectively. However, in face of this high level of polymorphism, the number of synonymous mutations was also very high, indicative of a negative selective pressure. Accordingly, the Ka/Ks ratio was found to be low for both genes. The transition/transversion ratio (Ts/Tv) was also very low (<0.8) indicating a highly saturated genome (Table 2). No bias was observed in the percentage of A, T, C and G bases. All these data indicate that some astVs are under high mutational dynamics leading to full genome saturation and to a strong negative, or purifying, selective pressure. Thirty of the 100 sequences of partial ORF1b had multiple nucleotides at the same position. The amplicons were therefore cloned to further investigate the nucleic acid polymorphism. A total of 17 amplicons were successfully cloned. For 12 positive samples, a single cloned daughter sequence was obtained whereas for 5 samples, more than one differing sequence was obtained after cloning. The sequences obtained after cloning showed varying degrees of distance with the consensus sequences generated by direct sequencing of the amplicons, ranging from 0% to 44.7% for nucleotide sequences and from 0% and 43.5% for amino acid sequences. When more than one different clone was found, the average sequence difference among them varied from 5.2% to 43.1% and from 4.8% to 43.5%, for nucleotides and amino acids sequences, respectively (Supplementary Table 5).

### 3.6. Genomic features of the strain PREDICT-MAstV-13_A91

Partial genome sequencing was successful for only one sample, corresponding to the sequence PREDICT-MAstV-13_A91. A longer sequence of 3123 nucleotides of the ORF1b gene and the complete ORF2 region was obtained for one newly detected virus (PREDICT-MAstV-13_A91, group 7) (Supplementary Fig. 2). ORF2 contained a region encoding a 774 amino acid capsid protein, a 3′ untranslated region (UTR) followed by a poly-A tail. An overlap of 8 nucleotides (ATGGCCTAG) between the stop codon of ORF1b and the start codon of ORF2, specific to mamastrovirus genomes (De Benedictis et al., 2011) was found. To investigate the relationship of PREDICT-MAstV-13 to other mamastroviruses, the region covering 699 nt of ORF1b and the complete capsid gene on ORF2 were compared to sequences from various representative mamastroviruses retrieved from GenBank (HMO strains, human astVs, porcine astV, bovine astVs, porcine astV, rat astV, bat astVs and murine astVs) (Fig. 2). ORF2 sequences of BatAstV available in GenBank were limited, nevertheless some sequences from group 2 (BatAstV-LD38), group 5 (BatAstV-LS11) and group 6 (BatAstV-AFCD11, BatAstV-AFCD57 and BatAstV-LD77) were included in the analysis. Both ORF1b and ORF2 sequences’ analyses confirmed the presence of astrovirus PREDICT-MAstV-13_A91 within group 7 along with ungulate and porcupine astVs (Fig. 2), showing a relative distance from sequences with the other astVs from group 7 (68.5 to 78% amino acid similarity on partial ORF1b sequences). Based on complete capsid gene sequence (ORF2), A91 shared 41.2% to 49.4% of amino acid similarity with other group 7 members, and displayed <32.1% with the other mamastroviruses from the tree, which is the criterion for astrovirus species definition, considering the ICTV taxonomic proposal on Mamastrovirus (2010.018v).

### 4. Discussion

This work highlights the apparent high diversity of astroviruses circulating among bat populations in Lao PDR and Cambodia. This finding was consistent with previous reports (Chu et al., 2008; Zhu et al., 2009; Xiao et al., 2011; Wu et al., 2012; Hu et al., 2014; Kememesi et al., 2014a, 2014b; Dufkova et al., 2015). In some sites, multiple viruses were found in bats from a single genus and collected during one sampling session, within one habitat, whereas in other cases bats from different genera, sampled in different habitats at various times were found to have just one astV. Our results were consistent with previous studies in China, revealing that the circulation of astVs is widely spread among Chiroptera and does not show strict host restriction (Chu et al., 2008; Zhu et al., 2009; Xiao et al., 2011). These results vary from those reported in Germany where host-specific viruses have been described in insectivorous bats (Fischer et al., 2015). Host specificity in insectivorous bats and variation between regions and species is therefore an aspect to further investigate. Some insectivorous bat species are leaving close to human habitat and may thus represent a risk for the

### Table 2

|               | Hp  | Nt  | S(%) | \(\eta\) | Pa(%) | S(%) | MC(%) | Na(%) | Nc(%) | Ka/Ks | Ts/Tv | Pi   | %A  | %C  | %G  | %T  |
|---------------|-----|-----|------|---------|-------|------|-------|-------|-------|-------|-------|-----|-----|-----|-----|
| ORF1b1a       | 178 | 144 | 246  | 196(85%)| 448   | 183(92.4)| 15(7.6)| 35(42.7)| 5(4.1)| 30(85.7)| 0.097 | 0.728| 0.73 | 29  | 19.4 | 21.3| 30.4|
| ORF1b2b       | 28  | 28  | 582  | 407(70%)| 746   | 360(90%)| 41(10.1)| 11(58.3)| 23(19.7)| 94(60.4)| 0.187 | 0.789| 0.32 | 27  | 19.8 | 21.8| 31.3|
| ORF2c         | 32  | 30  | 716  | 603(84.2)| 1328  | 570(94.5)| 33(5.5)| 83(34.9)| 92(22.9)| 64(77.1)| 0.147 | 0.662| 0.44 | 23.9| 23.1 | 25.4| 27.6|

N: Number of sequences.
Hp: Number of haplotypes.
Nt: Sequence size in nucleotides.
S: Number of mutated sites.
\(\eta\): Number of mutations.
Pa: Number of parsimony informative sites.
S: Number of singleton.
MC: Number of mutated codons.
Na: Number of non-synonymous mutations.
Nc: Number of synonymous mutations.
Ka/Ks: Ka/Ks ratio.
Ts/Tv: Transition/Transversion ratio.
Pi: Average number of nucleotide differences per site between two sequences.
%A: Overall percentage of adenine.
%C: Overall percentage of cytosine.
%G: Overall percentage of guanosine.
%T: Overall percentage of thymidine.

a Owing to the presence of indels in the ORF sequences, statistics were calculated on regions free of indels.
b Due to the important frequency of indels in the protein, polymorphism parameters were calculated based on 716 nucleotides corresponding to the first cleaved protein at the N terminal position only (homologous to the VP34 protein for Human astVs).
transmission of viruses to humans and domestic animals. One astV strain was repeatedly detected during different sampling periods, the group 7 astV sequences detected in bats from different Cambodian and Lao districts between 2010 and 2012. Drexler et al. (2011) also reported one strain in one bat colony at different times over a three-year period. Seasonality may be related to the breeding season, correlating with more numerous contacts between bats during the mating period, as well as a lower immunocompetence in females during gestation and lactating stages (Christe et al., 2000; Plowright et al., 2008; Baker et al., 2013), which could be important for viral shedding. Unfortunately, the lack of longitudinal sampling in our study did not allow us to explore the effect of seasonality on viral shedding, an aspect that would be valuable in future studies.

A high polymorphism was observed among astV sequences within a single group while displaying a level of divergence with other mammal astVs, i.e. astV from groups 1 and 6. Furthermore, twenty-two identified sequences fell into groups 7 and 8, which comprise unguulate and murine astVs. To date, this is the first report of astV strains related to diverse hosts including ungulate and porcine, and murine astVs, in frugivorous and insectivorous bats. This suggests that bats can host many diverse astroviruses that may be found in more virus groups than previously thought. However, given that only short sequences were analyzed from a conserved region of the RdRp gene, interpretations regarding relatedness to other viruses must be considered with caution. Bosch et al. (2010) proposed new classification criteria for astrovirus species in the latest ICTV proposal (2010.018aV), based on the ORF2 region encoding the capsid protein. ICTV recognises a new species of astrovirus if the complete ORF2 sequence shows >32.1% of difference in the amino acid with the closest sequence (Bosch et al., 2010). Although several sequences were original, only one sample. i.e. PREDICT-MastV-13_A91, met the dual requirement of a full length ORF2 sequence displaying >32.1% difference at the amino acid level. Considering these ICTV criteria for astrovirus systematics, molecular analyses on the ORF2 region suggest that astrovirus PREDICT-MastV-13_A91 is a new species (Pringle, 2014). Following ICTV guidelines, we thus propose to name this astrovirus: PREDICT-MastV-13/3AP11-A0091/2011.

As suggested by previous reports (Lukashov and Goudsmit, 2002; van Hemert et al., 2007b), the genome of astroviruses is highly saturated making phylogenetic and genetic diversity analyses difficult to interpret. A genome is saturated when all possibilities of mutations have been exploited and there is therefore no further linear correlation between the accumulation of mutations and time. As a consequence, phylogenies are blurred and weak, trees are poorly structured and genotyping is not reliable as there is an overall loss of significance. A similar situation was demonstrated with human rhinoviruses (Naughtin et al., 2015). Saturation also causes the molecular clock to be overestimated and leads to estimated segregation events to have occurred far earlier than they effectively did (Bromham and Penny, 2003). Astroviruses seem to be exposed to two different trends: a high level of mutations which leads to extensive genome saturation and, also a strong negative selection pressure leading to high conservation of the amino acid sequences, and an evolution of the codon usage which was shown to be host-driven (van Hemert et al., 2007b).

Such an evolution could also be explained with a quasi-species model (Lauring and Andino, 2010). One can thus expect to simultaneously see a high mutation level, conserved proteins and host-specific variations which are all the traits displayed by astVs described in this work. In such a model, the consensus sequence is imposed by the host through selection imposed on variants present in the sequence (van Hemert et al., 2007a). The simultaneous presence of different clone sequences could be due to the presence of a co-infection, hyperpolymorphism within the same astV strain or a result of quasi species model of evolution. Host specificity in the quasi species model is not host restriction or limitation, but instead the mark of the host replication system on the viral progeny.

Apart from one astrovirus belonging to group 7, sequencing attempts of ORF2 regions of other positive samples were unsuccessful. Several hypotheses could explain the difficulties encountered. First, the quality of viral genetic material in PCR-positive samples may be questionable. Since we found numerous positive samples which restated negative after an additional RNA extraction procedure, we can assume the viral load was low or the viral RNA too damaged, and additional thaw may have further damaged the RNA, to obtain better and longer sequence fragments. Human and several mammal (porcine, bovine) and avian astVs have been isolated and amplified in cell culture in order to provide a larger amount of high quality viral genetic material (Lee and Kurtz, 1981; Shimizu et al., 1990; Woode et al., 1985). To date, this technique has remained unsuccessful for bat astVs (De Benedictis et al., 2011), which may explain the limited number of available ORF2 references for bat astVs in GenBank. ORF2 is known to be highly divergent among astroviruses, even among bat astVs from the same cluster (Chu et al., 2008; Zhu et al., 2009), thus the development of the consensus or specific primers to amplify regions of interest may be difficult.

Bat genera where astV sequences were detected, roost in various environments, such as natural caves, old buildings, old disused wells, (Eonycteris, Rousettus, Taphozous, Megaderma, Rhinolophus), karst environment (ia), nearby forest or orchards (Eonycteris, Rousettus) or in water areas, where food resources can be found (Scotophilus, Myotis) (Kunz and Fenton, 2005). Despite the fact that they do not share same direct food sources, the spatial overlap of feeding areas and temporary night roosting between bats from different genera may facilitate the exchange of viruses. This was already suggested with coronavirus infection among bats that were displaying different feeding and roosting behaviors, while sharing the same habitat, i.e. in hipposiderid and rousette in China, and in Cynopterus and Scotophilus bats in Thailand (Lau et al., 2012; Tang et al., 2006; Wacharaphusadee et al., 2015). As a result, more investigations are needed to understand the correlation between interspecific bat behavior and the potential transmission of astroviruses.

Besides bat hosts, the human-bat interface may be important to consider when examining astrovirus infection in bats. Cambodia and Lao PDR have undergone substantial environmental change associated with agricultural development and suburban growth in the last fifty years (WWF, 2013). These changes affect natural habitats and therefore the distribution of bats, leading them to be closer to humans. This can be exacerbated by the high tolerance for habitat modification of certain bat genera that easily adapt to disturbed agricultural lands (Rousettus, Eonycteris, Hipposideros, Rhinolophus) or even adopt roosting in human dwellings (Myotis, Scotophilus) (Kunz and Fenton, 2005; Bates et al., 2008a, 2008b; Bates and Helgen, 2008; Csorba et al., 2008; Francis et al., 2008; Rosell-Ambal et al., 2008; Walston et al., 2008). This might facilitate increased contact between humans and bats and therefore, increase the risk of potential transmission of astVs to humans.

Most of the positive samples were from rectal swabs and feces, which was consistent with previous reports (Chu et al., 2008). Moreover, most of the positive bats were sampled either in guano farms or bats hunted and sold for food. These results suggest that the risk of transmission to humans through the fecal route may exist, especially through guano farming or when bats are handled after being hunted (Thi et al., 2014) in areas where they are sold in markets or provide food resources (Lee et al., 2014). In this study, the health status of the bats was unknown, and no study to date has reported clinical signs associated with astV infection in bats (Chu et al., 2008; Xiao et al., 2011). This is also the case for bat species that may carry zoonotic viruses such as henipa, filo or lyssa viruses (Mackenzie and Field, 2004; Leroy et al., 2005; Banyard et al., 2014). Thus, without a better understanding of the risks to humans associated with bat astrovirus infection, there is a need for rural communities exposed to bats to be aware of the general potential risk of virus transmission associated with close contact, and improved guidance on how to handle the animals and reduce such
Conflicts of interest

Philippe Buchy is currently an employee of GSK Vaccines.

Acknowledgements

This study was made possible by the generous support of the American people through the United States Agency for International Development (USAID) Emerging Pandemic Threats PREDICT project (co-operative agreement number GHN-A-00-09-00010-00). We thank the governments of Cambodia and Laos PDR for permission to conduct this study. We thank WCS teams and all the students who helped collect field samples. We also thank Dara Kong and Serey Roth Long for their technical support in the laboratory diagnostics. Thanks also to Neil Furey, from the Center for Biodiversity Conservation of the Royal Phnom Penh, for his help and advice on species identification.

We also thank Sarah Olson for her help in extracting, organizing, and providing the sample data (species, sample type, sampling location) from the phase 2 sample collection performed by WCS as part of the PREDICT project.

References

Baker, M.L., Schouten, T., Wang, L.F., 2013. Antiviral Immune Responses of Bats: A Review. Zoonoses Public Health 60, 104–116.

Banyard, A.C., Evens, J.S., Luo, T.R., Fooks, A.R., 2014. Lyssaviruses and bats: emergence and risks. There is also a need for continued surveillance for the circulation of viruses in bats and to monitor for the risk of transmission to humans.

We thank WCS teams and all the students who helped collect field samples. We also thank Dara Kong and Serey Roth Long for their technical support in the laboratory diagnostics. Thanks also to Neil Furey, from the Center for Biodiversity Conservation of the Royal Phnom Penh, for his help and advice on species identification.

We also thank Sarah Olson for her help in extracting, organizing, and providing the sample data (species, sample type, sampling location) from the phase 2 sample collection performed by WCS as part of the PREDICT project.
Tang, X.C., Zhang, J.X., Zhang, S.Y., Wang, P., Fan, X.H., Li, L.F., Li, G., Dong, B.Q., Liu, W., Cheung, C.L., Xu, K.M., Song, W.J., Vijaykrishna, D., Poon, L.L.M., Peiris, J.S.M., Smith, G.J.D., Chen, H., Guan, Y. 2006. Prevalence and genetic diversity of coronaviruses in bats from China. J. Virol. 80, 7481–7490.

Thi, S., Furey, N.M., Jurgens, J.A., 2014. Effect of bat guano on the growth of five economically important plant species. J. Trop. Agric. 52, 169–173.

van Hemert, F.J., Berkhout, B., Lukashov, V.V., 2007a. Host-related nucleotide composition and codon usage as driving forces in the recent evolution of the Astroviridae. Virology 361, 447–454.

van Hemert, F.J., Lukashov, V.V., Berkhout, B., 2007b. Different rates of (non-)synonymous mutations in astrovirus genes; correlation with gene function. Virology J. 4, 25.

Wacharapluesadee, S., Duengkae, P., Rodpan, A., Kaevpon, T., Maneenorn, P., Kanchanasaka, B., Yingsakmongkon, S., Sittidethoripat, N., Charreesaan, C., Khlangsap, N., Pidthongh, A., Leadprathom, K., Ghai, S., Epstein, J.H., Daszak, P., Olival, K.J., Blair, P.J., Callahan, M.V., Henrichduhua, T., 2015. Diversity of coronavirus in bats from eastern Thailand. Virol. J. 12, 57.

Walston, J., Kingston, T., Hutson, A.M., 2008. Rhinolophus Affinis [WWW Document]. IUCN Red List Threat. Species. URL The IUCN Red List of Threatened Species 2008: e.T19522A8952553 (accessed 9.21.16).

Wang, L.-F., Shi, Z., Zhang, S., Field, H., Daszak, P., Eaton, B.T., 2006. Review of bats and SARS. Emerg. Infect. Dis. 12, 1834–1840.

Wild, T.F., 2009. Henipaviruses: a new family of emerging Paramyxoviruses. Pathol. Biol. 57, 188–196.

Woode, G.N., Gourley, N.E., Pohlenz, J.F., Liebler, E.M., Mathews, S.L., Hutchinson, M.P., 1985. Serotypes of bovine astrovirus. J. Clin. Microbiol. 22, 668–670.

Wu, Z., Ren, X., Yang, L., Hu, Y., Yang, J., He, G., Zhang, J., Dong, J., Sun, L., Du, J., Liu, L., Xue, Y., Wang, J., Yang, F., Zhang, S., Jin, Q., 2012. Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces. J. Virol. 86, 10999–11012.

WWF, 2013. Ecosystems in the Greater Mekong: Past Trends, Current Status, Possible Futures. World Wide Fund For Nature, Greater Mekong, Bangkok, Thailand.

Xiao, J., Li, J., Hu, G., Chen, Z., Wu, Y., Chen, Y., Chen, Z., Liao, Y., Zhou, J., Ke, X., Ma, L., Liu, S., Zhou, J., Dai, Y., Chen, H., Yu, S., Chen, Q., 2011. Isolation and phylogenetic characterization of bat astroviruses in southern China. Arch. Virol. 156, 1415–1423.

Zhu, H.C., Chu, D.K.W., Liu, W., Dong, B.Q., Zhang, S.Y., Zhang, J.X., Li, L.F., Vijaykrishna, D., Smith, G.J.D., Chen, H.L., Poon, L.L.M., Peiris, J.S.M., Guan, Y., 2009. Detection of diverse astroviruses from bats in China. J. Gen. Virol. 90, 883–887.