Taylor & Francis
Taylor & Francis Group

EMI: Animal & Environment

ISSN: 2998-6990 (Online) Journal homepage: www.tandfonline.com/journals/tema20

Identification and Genomic Characterization of Known and Novel Highly Divergent Sapoviruses in Frugivorous and Insectivorous Bats in Nigeria

Uwem E. George, Lander De Coninck, Oluwadamilola A. George, Taiye Adeyanju, Arthur Oragwa, Joshua Kamani, Joseph Igbokwe, Andrew Adamu, Temitope Faleye, Richard Adeleke, Tomiwa Adesoji, Timothy K. Soh, Oluyomi Sowemimo, Philomena Eromon, Olubusuyi M. Adewumi, Johnson A. Adeniji, Onikepe Folarin, Scott C. Weaver, Anise Happi, Jens B. Bosse, Robert W. Cross, Isaac Komolafe, Jelle Matthijnssens & Christian Happi

To cite this article: Uwem E. George, Lander De Coninck, Oluwadamilola A. George, Taiye Adeyanju, Arthur Oragwa, Joshua Kamani, Joseph Igbokwe, Andrew Adamu, Temitope Faleye, Richard Adeleke, Tomiwa Adesoji, Timothy K. Soh, Oluyomi Sowemimo, Philomena Eromon, Olubusuyi M. Adewumi, Johnson A. Adeniji, Onikepe Folarin, Scott C. Weaver, Anise Happi, Jens B. Bosse, Robert W. Cross, Isaac Komolafe, Jelle Matthijnssens & Christian Happi (08 May 2025): Identification and Genomic Characterization of Known and Novel Highly Divergent Sapoviruses in Frugivorous and Insectivorous Bats in Nigeria, EMI: Animal & Environment, DOI: 10.1080/29986990.2025.2503155

To link to this article: https://doi.org/10.1080/29986990.2025.2503155

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Accepted author version posted online: 08 May 2025.

Submit your article to this journal

Article views: 447

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=

Publisher: Taylor & Francis

Journal: *EMI: Animal & Environment*

DOI: 10.1080/29986990.2025.2503155

Check for updates

Identification and Genomic Characterization of Known and Novel Highly Divergent Sapoviruses in Frugivorous and Insectivorous Bats in Nigeria

Uwem E. George[1,2], Lander De Coninck[3], Oluwadamilola A. George[4], Taiye Adeyanju[5], Arthur Oragwa[6], Joshua Kamani[7], Joseph Igbokwe[8], Andrew Adamu[9,10,11], Temitope Faleye[12], Richard Adeleke[13], Tomiwa Adesoji[14], Timothy K. Soh[15,16,17,18], Oluyomi Sowemimo[8], Philomena Eromon[1], Olubusuyi M. Adewumi [19,20], Johnson A. Adeniji [19,20], Onikepe Folarin[1,21], Scott C. Weaver[2], Anise Happi[1], Jens B. Bosse[15,16,17,18] Robert W. Cross[2,±], Isaac Komolafe[21,±,*], Jelle Matthijnssens[3,±*], and Christian Happi[1,21,22,±,*]

[1]Institute of Genomics and Global Health (formerly ACEGID), Redeemer's University, Ede, Osun State, Nigeria
[2]Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston, 77555, Texas, USA
[3]KU Leuven, Department of Microbiology, Immunology and Transplantation, Rega Institute, Division of Clinical and Epidemiological Virology, Laboratory of Viral Metagenomics, Leuven, Belgium
[4]Ibadan Diagnostic and Epidemiology Laboratory, National Veterinary Research Institute, Mokola, Ibadan, 200212, Oyo State, Nigeria
[5]Department of Wildlife and Ecotourism Management, Faculty of Renewable Natural Resources, University of Ibadan, Oyo State, Nigeria
[6]Department of Veterinary Microbiology, Faculty of Veterinary Medicine, University of Jos, Jos 930003, Plateau State, Nigeria
[7]Parasitology Division, National Veterinary Research Institute (NVRI), PMB 01, Vom 930103, Plateau State, Nigeria
[8]Department of Zoology, Faculty of Science, Obafemi Awolowo University, Ile Ife 220005, Osun State, Nigeria
[9]Australian Institute of Tropical Health and Medicine, Division of Tropical Health and Medicine, James Cook University, Townsville, QLD 4811, Australia
[10]College of Public Health, Medical and Veterinary Sciences, James Cook University, 1 James Cook Drive, Bebegu Yumba Campus, Douglas, QLD 4811, Australia
[11]Department of Veterinary Public Health and Preventive Medicine, University of Abuja, Abuja 900105, Federal Capital Territory, Nigeria
[12]Center for Environmental Health Engineering, Biodesign Institute, Arizona State University, Tempe, Arizona, USA

ACCEPTED MANUSCRIPT

[13]Department of Immunology and Infectious Diseases, College of Veterinary Medicine, Cornell University, New York, NY 14853, USA
[14]Department of Microbiology, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria
[15]Hannover Medical School (MHH), Institute of Virology, Hannover, Germany
[16]Leibniz Institute of Virology (LIV), Hamburg, Germany
[17]Centre for Structural Systems Biology (CSSB), Hamburg, Germany
[18]Cluster of Excellence RESIST (EXC 2155), Hannover Medical School, Hannover, Germany
[19]Department of Virology, College of Medicine, University of Ibadan, Ibadan, Nigeria
[20]Infectious Disease Institute, College of Medicine, University of Ibadan, Ibadan, Nigeria
[21]Department of Biological Sciences, Faculty of Natural Sciences, Redeemer's University, Ede, Osun State, Nigeria
[22]Department of Immunology and Infectious Diseases, Harvard T H Chan School of Public Health, Boston, MA, USA

[±] These senior authors contributed equally

*Corresponding Authors: toshokomolafe@gmail.com (Isaac Komolafe), jelle.matthijnssens@kuleuven.be (Jelle Matthijnssens) happic@run.edu.ng (Christian Happi)

ACCEPTED MANUSCRIPT

Abstract

Sapovirus (SaV) infections have been linked with moderate-to-severe acute gastroenteritis (AGE) in animals and humans and represent a significant risk to public health. SaVs from animals including pigs, chimpanzees, and rodents have been reported to be closely related with human SaVs, indicating the possibility of cross-species transmission. Divergent SaVs have been reported in various bat species across various continents including Asia, Europe, Oceania and Africa. However, little is known about the evolutionary history of SaVs across various bat species and their zoonotic potential. In this report, we describe the findings of a surveillance study across various bat species in Nigeria. Samples were pooled and subjected to metagenomics sequencing and analyses. Nine of 57 sample pools (containing 223 rectal swabs from five bat species) had SaV reads from which we assembled a total of four complete and three near-complete (having complete coding sequences) genomes. The bat SaV (BtSaV) strains from this study formed five distinct lineages of which four represented novel genogroups. BtSaV lineages clustered mainly according to bat families, which might suggest a likely virus-host-specific evolution. The BtSaV VP1 capsid protein structure prediction confirmed

three main domains (S, P1, and P2) as reported for Human SaV (HuSaV). We found that the P2 subdomain of the VP1 protein contains a degree of homology to known immunoreactive epitopes suggesting these conserved regions may be valuable for diagnostics or medical countermeasure development. This study expands our understanding of reservoir hosts, provides information on the genetic diversity and continuous evolution of SaVs in bats.

Keywords: Sapovirus, Bats, metagenomic sequencing, protein modelling, VP1 protein, new genogroups, Nigeria

Introduction

Sapovirus alongside *Norovirus* is one of the 11 genera in the family *Caliciviridae*, having only one species (*Sapporo virus*) [1]. Sapoviruses (SaVs) were initially identified in faecal samples collected from infants as part of a stool survey in Glasgow children using electron microscopy [2]. SaV infections have been associated with moderate-to-severe acute gastroenteritis (AGE) in animals and humans and represent a significant risk to public health [3,4]. Currently, there are no specific therapeutics or vaccines available for the management and prevention of SaV diseases.

The SaV virion consists of a non-enveloped icosahedral capsid (typically 30-38 nm in diameter) containing the positive-sense, single-stranded RNA genome [5]. The RNA genome has two or three open reading frames (ORFs), depending on the SaV genogroup. The large ORF1 polyprotein encodes both non-structural proteins (NS1-7) and the major capsid protein (VP1) while the small ORF 2 encodes VP2 (a minor structural protein) [6]. The VP1 protein of SaV, like that of other caliciviruses, is the main capsid component of the SaV virion, and it is critical for immunological response as well as determining SaV genetic variation and genotypes [7]. The VP1 protein consists of approximately 560 amino acids (aa) organized into two main domains: the protruding (P) domain and the shell (S) domain, plus an N-terminal arm (NTA) which connects the S domain in the capsid shell [8,9]. The P domain is made up of two subdomains: P1 and P2, with P2 being the outermost virion structure. Hypervariable regions on the P2 subdomain have been shown to serve as the primary targets of neutralizing antibodies [10].

ACCEPTED MANUSCRIPT

SaVs are classified into genogroups based on amino acid (aa) sequence identity analysis of the complete VP1 amino acid sequences [11]. Recently, a dual classification criterion has been proposed, with cutoff values for genogroup clusters defined as <0.503 (VP1) and <0.531 (RdRp). Based on these criteria, SaVs have been classified into 34 genogroups (GI to GXXI and GNA1 - GNA3). SaVs can be further classified into genotypes, with cutoff values for genotype clusters for VP1 and RdRp sequence identity defined as <0.161 and <0.266. Currently, at least 52 genotypes have been reported within the genus Sapovirus [12]. SaVs have been reported in a wide range of hosts, including domestic pigs [13], chimpanzees [14], foxes, hyenas, lions [15], rats [16], and humans [17,18].

Human SaVs fall into four genogroups (GI, GII, GIV and GV) and some SaVs detected in animal hosts have been found to cluster closely with human SaVs especially members in genogroups GI, GV, and GII [14,16,19]. While this suggests that SaVs are likely to cross host barriers, it does not indicate the direction of transmission.

Bats are the second largest group of mammals after the order Rodentia, accounting for more than 20% of all known mammalian species worldwide [20]. They are extensively distributed in nature and contribute significantly to both the biological and ecological diversity of numerous habitats. In Africa, anthropogenic factors especially land-use change and overexploitation have resulted in bats being heavily hunted [21]. This has resulted in a rise in the number of bat colonies located close to or inside urban environments, as well as more frequent encounters with humans, pet animals, and livestock, which may increase the risk for zoonotic disease outbreaks. Recently, divergent SaVs have been reported in various bat species across Asia [22], Europe [23], Oceania [24] and Africa [25]. However, little is known about the evolutionary history of SaV across various bat species and their zoonotic potential.

Most of the recent discoveries of SaVs in bats were based on the use of enhanced pathogen discovery techniques including metagenomics. This approach has been pivotal not just in the discovery of distinct strains of SaV in humans and animals but also in our improved understanding of the disease burden of SaVs [17,18,24,25]. In this study, we report the identification and genomic characterization of highly divergent SaVs in both fructivorous and insectivorous bats in Nigeria using next-generation sequencing. We also provide further insight into the genetic diversity of SaVs across various bat species. Furthermore, using molecular modelling, we predicted the VP1 and VP2 structures of the detected bat SaVs (BtSaV), their unique antigenic sites and B cell epitopes in the

ACCEPTED MANUSCRIPT

immunodominant VP1 domain. Overall, these data are vital for improving SaV classification, viral taxonomy, developing effective diagnostic tests, and gaining a better knowledge of SaV evolution and host species adaptability.

Materials and Methods

Ethical Authorization and Sample Collection

The bat rectal swabs analysed in this study were collected as part of the Nigerian bat virome project from insectivorous bats (*Mops condylurus*, *Chaerephon spp*. and *Hipposideros ruber*) and fruit-eating bats (*Rousettus aegyptiacus* and *Eidolon helvum*) from 6 states (Benue, Bauchi, Ondo, Niger, Osun and Plateau) in Nigeria between 2019 and 2022 (**Figure 1 and Table S1**). Bats were captured around caves and fruit trees at night using harp traps and mist nets as previously described [26]. Morphological features were assessed for each captured bat to determine species, and molecular confirmation of bat species was done as previously described [27] using primers targeting cytochrome b (Cyt b) and mitochondrial cytochrome oxidase subunit 1 (COI). Briefly, rectal swabs were collected and placed in tubes containing 1 mL of virus transport medium. The bat sample collection was done in a microbiological safety station, with personal protective equipment on. The samples were transported in -20°C containers to the African Centre of Excellence for Genomics of Infectious Diseases (ACEGID) laboratory at Redeemer's University in Nigeria, where they were stored at -80°C until processed. A total of four hundred and twenty samples were collected (409 between 2019 and 2021 [26] and 11 samples in 2022). The National Veterinary Research Institute (NVRI) Nigeria's Animal Care and Use Committee approved the study design and sampling technique (approval number AEC/03/65/19). We were also given authorization by the Plateau State Health Research Ethics Committee (approval number PSSH/ADM/ETH.CO/2019/005).

In this study, 223 archived (stored at −20°C) rectal swabs suspended in a virus transport medium were randomly selected and combined into 57 pools based on bat species and the state of collection and subsequently analysed. Approximately 200 µL of the rectal swab suspensions were combined to create each pool, and each contained one to eight faecal suspensions (**Table S1**).

ACCEPTED MANUSCRIPT

Sample preparation and Sequencing

Virus enrichment and library preparation were performed using either the unbiased next-generation sequencing technique as previously reported [28] at Institute of Genomics and Global Health (formerly ACEGID), Redeemer's University, Ede, Nigeria or the NetoVIR protocol [29] at KU Leuven, Rega Institute, Laboratory of Clinical and Epidemiological Virology in Belgium. At KU Leuven, 40 sample pools were analysed using the NetoVIR protocol. Briefly, Fecal suspensions were filtered through a 0.8 µm PES filter and free-floating nucleic acids digested using a combination of Micrococcal Nuclease (New England Biolabs, Ipswich, MA, USA) and Benzonase (Millipore, Billerica, MA, USA). Nucleic acid was then extracted using the QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany). For first- and second-strand synthesis, a slightly modified Whole Transcriptome Amplification (WTA2) Kit procedure (Sigma-Aldrich, St Louis, MO, USA) was used, followed by 17 cycles of random PCR amplification. The Nextera XT Library Preparation Kit (Illumina, San Diego, CA, USA) was used to prepare the libraries after which sequencing was performed on the Illumina NextSeq 500 platform (300 cycles, 2 x 150 bp paired ends). At ACEGID, 17 individual samples were processed and sequenced. Briefly, after elution, the RNA was treated with turbo DNase to remove any contaminating DNA. The cDNA was then synthesized using a Superscript III Synthesis kit (Invitrogen) and random primers. Sequencing libraries were prepared using the Illumina Nextera XT kit. Subsequently, paired-end sequencing was carried out using the Illumina MiSeq Reagent Kit v2 (500 cycles) on an Illumina Miseq platform (**Table S1**).

Read Processing, Genome Assembly and Annotation

We processed all raw reads using the Virome Paired-End Reads (ViPER) pipeline (https://github.com/Matthijnssenslab/ViPER), regardless of sequencing platform, to identify reads containing SaVs. In brief, we used Trimmomatic [30] to remove sequencing adapters and improve the quality of the raw reads, and Bowtie2 to eliminate reads that mapped to the host genome [31]. All the trimmed and filtered reads were then de novo assembled into contigs using metaSPAdes [32]. We also performed de novo

ACCEPTED MANUSCRIPT

assembly with MEGAHIT [33] to validate the BtSaV genomes assembled using metaSPAdes. We annotated contigs with DIAMOND (sensitive option) [34] and used BLASTn to further validate the assembled SaV contigs. Using Bowtie2 [31], trimmed reads were mapped against the SaV contigs to assess coverage depth. We identified SaV Open Reading Frames (ORFs) using the ORF finder analysis tool and MetaGeneMark (https://genemark.bme.gatech.edu/genemark/meta_gmhmmp.cgi version 3.25) [35].

Phylogenetic and Recombination analysis

Following a BLASTn search with SaV contigs from this study as queries, we retrieved five sequences from GenBank for each contig with highest identity and coverage. We also retrieved all available SaV reference sequences from the thirty-four documented SaV genogroups from NCBI virus database, along with high-quality BtSaV genomes meeting criteria of <10% ambiguous nucleotides and 80% alignment length. Post-deduplication, we created two sets of data: one comprising the alignment of SaVs found in our research alongside all reference SaV sequences (GI - GXXI and GNA1 - GNA3) from both humans and animals (including complete genome and VP1-only reference sequences for comparison) and another dataset consisting of all BtSaVs found worldwide with a minimum of 80% genome coverage (to assess the genetic diversity of SaVs in different bat species/families). We utilized MAFFTv7.505 [36] to align the data and reconstructed phylogenies using IQTREE2 [37] with ModelFinder [38]. The phylogenetic trees were visualized using Interactive Tree of Life (iTOL) v6 [39]. To further validate the divergence and pairwise identity of our BtSaV sequences and published reference sequences, we aligned each distinct pair using the Sequence Demarcation Tool [40]. We examined all the BtSaV sequences for recombination using RDP5 [41].

Molecular modelling and structural prediction of Bat Sapovirus VP1 proteins

We aligned the VP1 domain of Nigerian BtSaV with known VP1 sequences from human SaVs (AB455803.1 and AJ606694.2) to identify conserved regions and amino acid substitutions related to potential binding to human receptors. The S, P1, and P2 domains in the VP1 protein were the focus of the analysis. Physical and general biological

ACCEPTED MANUSCRIPT

properties of the bat SaV VP1 proteins were calculated using ProtParam and ProtScale tools on the ExPASy Server (accessible at https://web.expasy.org/cgi-bin/protparam/protparam). To predict regions in the BtSaV VP1 protein sequences that are likely to be antigenic (antigenic epitopes), a previously described approach [51] with an antigen prediction tool (accessible at http://imed.med.ucm.es/Tools/antigenic.pl) was used.

To predict potential B-cell epitopes present within the BtSaV VP1 domain, we employed the Bepipred Linear Epitope Prediction 3.0 tool [52]. Protein structure prediction was performed with LocalColabFold v1.5.2[53] (https://github.com/YoshitakaMo/localcolabfold) [54] to generate 5 models using 20 recycles and a stop-at-score of 100. AlphaFold 3 [55] was used through the AlphaFold Server (https://alphafoldserver.com). Visualization was performed with ChimeraX version 1.8 (https://www.cgl.ucsf.edu/chimerax) [56].

Results

Bat Sample Demographic, Sapovirus Diversity and Phylogenetic analysis

We screened a total of 223 rectal swabs from 5 bat species (*Mops condylurus*, *Eidolon helvum*, *Chaerephon spp*, *Hipposideros ruber* and *Rousettus aegyptiacus*) collected from six Nigerian states (Benue, Bauchi, Ondo, Niger, Osun and Plateau) (**Figure 1**), which were divided over 57 pools (**Table 1 and Table S1**).

ACCEPTED MANUSCRIPT

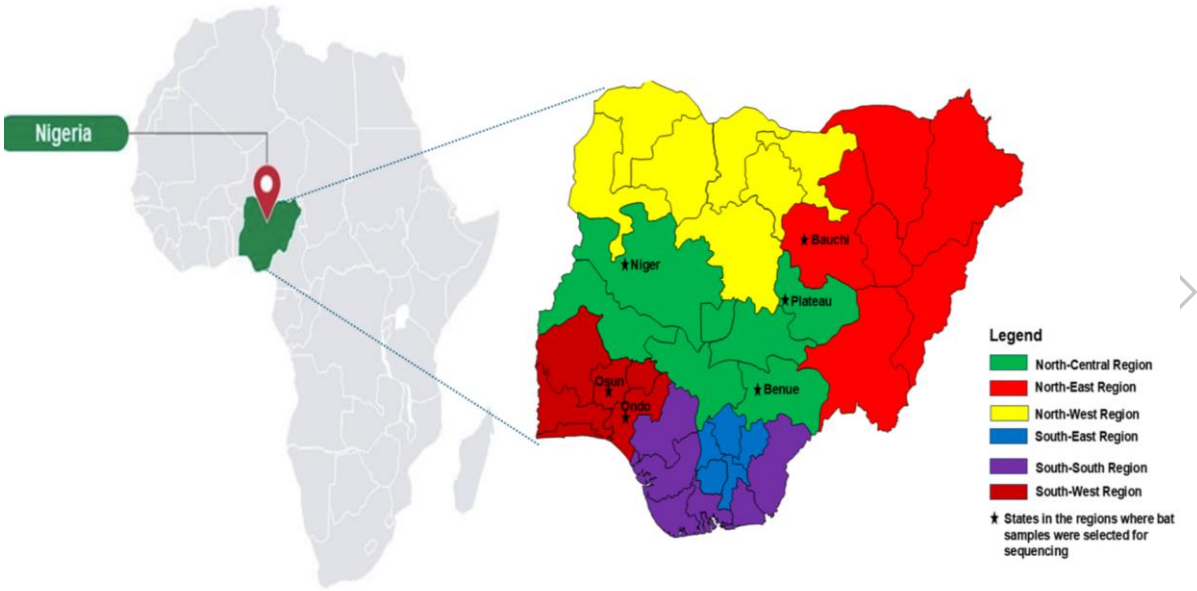

Figure 1. Map of Africa highlighting Nigeria and the various geographical regions where bat samples were selected for sequencing in this study.

Of the 57 pools, we detected SaVs reads in 9 pools (15.8%). Out of the 9 pools with SaV reads, 6 pools were derived from straw-coloured fruit bat (*Eidolon helvum*) (Osun State), 2 from *Rousettus aegyptiacus* (Ondo State), and 1 from *Mops condylurus* (Benue State). No SaV reads were detected in samples from the Bauchi, Niger and Plateau states. We assembled a total of 4 complete, 3 near-complete (having complete coding sequences) and 11 partial genomes ranging in length from 7611 – 898 (**Table S2)**.

Table 1. Demographic description and summary of 9 pools containing Sapovirus reads

S/N	Pool ID	Year of sample collection	Location	Bat family	Bat species	Total reads	SaV Mapped reads	Number of Sapovirus contigs

ACCEPTED MANUSCRIPT

1.	B8/BT/OAU/2	2020	Osun	*Pteropodidae*	*E. helvum*	12,731,940	155,425	2
2.	B12/BT/OAU/6	2020	Osun	*Pteropodidae*	*E. helvum*	14,935,074	248	2
3.	B14/BT/OAU/8	2020	Osun	*Pteropodidae*	*E. helvum*	15,225,166	27,224	5
4.	B15/BT/OAU/9	2020	Osun	*Pteropodidae*	*E. helvum*	1,447,204	46	1
5.	B17/BT/OAU/11	2020	Osun	*Pteropodidae*	*E. helvum*	10,948,508	429	1
6.	B23/BT/OAU/17	2020	Osun	*Pteropodidae*	*E. helvum*	3,975,982	153	3
7.	GB09	2020	Benue	*Molossidae*	*Mops condylurus*	381,938	10,853	2
8.	G10	2022	Ondo	*Pteropodidae*	R. aegyptiacus	4,001,878	22,270	1
9.	G11	2022	Ondo	*Pteropodidae*	R. aegyptiacus	3,314,134	586	1

To determine the divergence of the BtSaV strains detected in this study, BtSaV sequences (we focused mostly on the four complete, three near-complete and one partial genome

ACCEPTED MANUSCRIPT

having over 70% genome coverage with complete VP1 and VP2) and all previously described reference SaV genogroup sequences (GI to GXXI and GNA1 - GNA3) were subjected to Sequence Demarcation tool analysis. The analysis was based on cutoff values for genogroup clusters defined as <0.503 and cutoff values for genotype clusters defined as <0.161) [12]. We found that the Nigerian BtSaV sequences formed 5 distinct lineages (L1 - L5) and only one of the BtSaV from this study (BtSaV/A12B8/OAU/NGR/2020 – L2) belonged to a previously described genogroup (GXXII), which consisted of BtSaV detected in Eidolon helvum bats from Cameroon. The remaining BtSaV from this study were likely members of four novel genogroups, tentatively named GXXXII to GXXXV (**Figure 2**).

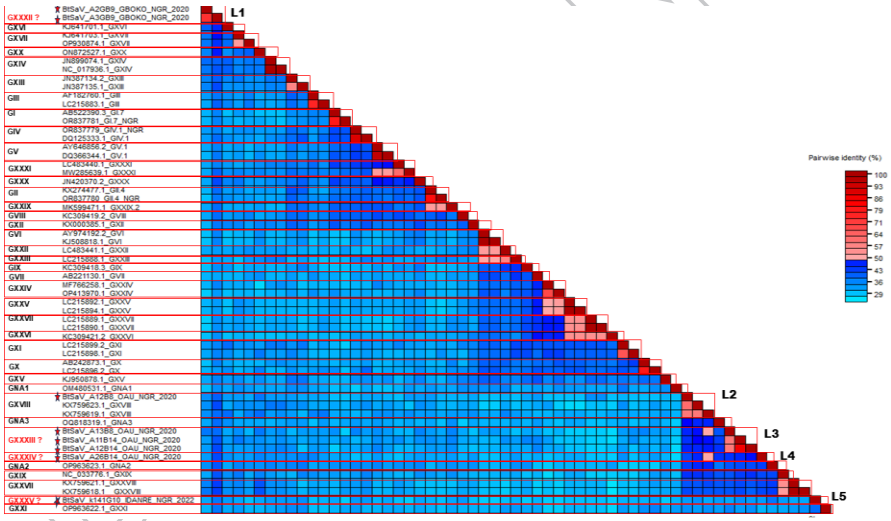

Figure 2. Classification of SaV (GI to GXXI and GNA1 - GNA3) based on cutoff values for genogroup clusters defined as <0.503 and cutoff values for genotype clusters defined as <0.161 using SDT showing evidence of several tentative new genogroups. SaV contigs detected in this study are highlighted in red font with a question mark and a star while previously described SaVs are in black font.

To further investigate the genetic diversity of the Nigerian BtSaV lineages in relation to SaV from humans and other animals, we generated two phylogenetic trees using available dataset of SaV reference sequences (GI to GXXI and GNA1 - GNA3). The first tree was based on complete VP1 regions, since it is the most representative region and often used for typing of SaV and the second tree was based on complete genomes sequences (where available). In our phylogenetic analyses, we found that the BtSaV strains from this study formed 5 distinct lineages (like the SDT analysis) with high

ACCEPTED MANUSCRIPT

support (bootstrap = 100) in both phylogenies, suggesting the co-circulation of divergent BtSaVs in Nigerian bats (**Figure 3A, B**). The BtSaVs from this study were all distantly related to human SaVs (HuSaVs) including the recently reported HuSaVs from Nigeria (**Figure 3A**) [17].

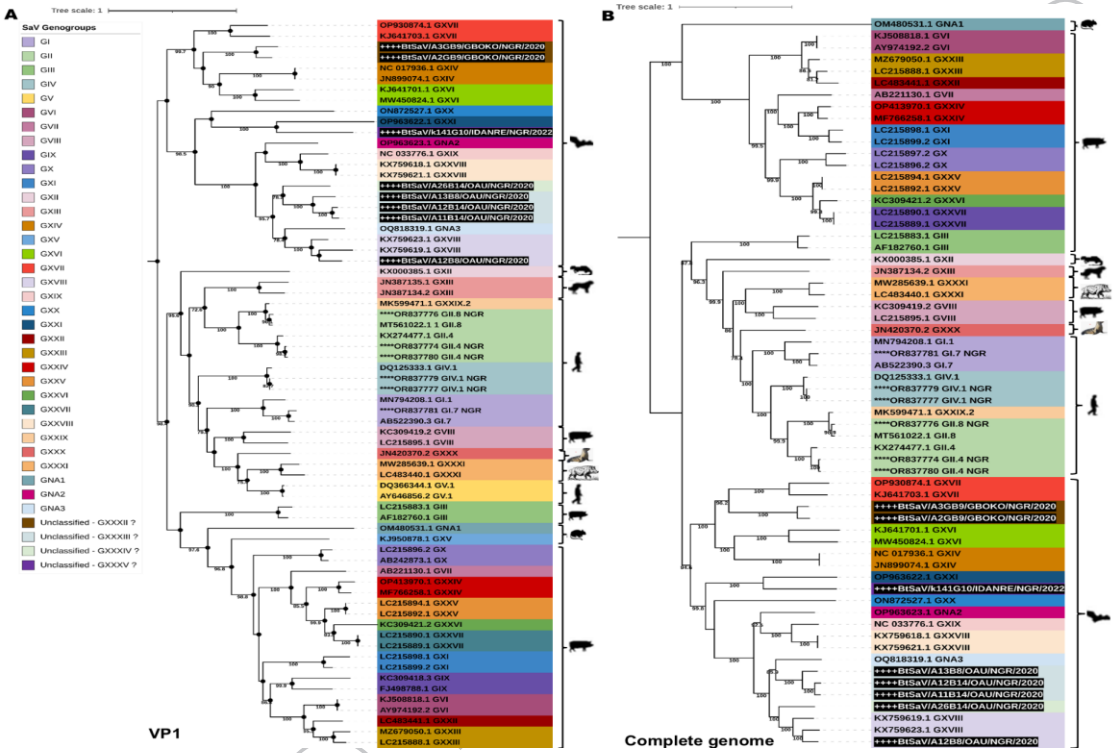

Figure 3. Maximum likelihood trees of SaV GI - GXXI and GNA1 - GNA3 based on **A.** complete VP1 gene and **B.** Complete genome sequences. All SaV genogroups were assigned a specific colour according to the legend provided. The BtSaV strains reported in this research are marked with an asterisk and highlighted in white font and black background while previously reported HuSaV from Nigeria are highlighted in black asterisk.

Next, we focused our attention specifically to bat SaVs. We constructed a phylogenetic tree of the complete VP1 and RdRp gene to assess the genetic diversity of SaVs in different bat species/families using BtSaV sequences from this study and BtSaVs found worldwide (sequences with a minimum of 80% genome coverage). Our maximum likelihood (ML) tree results showed that a total of 17 bat species in six different bat families were found to be infected with SaVs. Based on SaV genogroup demarcation criteria using SDT, we found that at least 16 distinct SaV lineages may be circulating in various bat species globally with species members in the *Pteropodidae* family harbouring the highest number of divergent SaVs (**Figure 4, 5 and Figure S1**). We also found that various BtSaV lineages clustered mainly according to bat families, which suggests virus-host-specific evolution. Furthermore, we observed that species within some bat families

ACCEPTED

(especially species members in the families *Molossidae* and *Vespertilionidae*), harboured genetically related BtSaV irrespective of location where bats were sampled. For instance, SaVs detected in *Mops condylurus* bats (species member in *Molossidae* family) in this study from Nigeria were genetically related to SaVs detected in *Chaerephon leucogaster* bats sampled in Kenya (**Figure S1**).

Based on boot scan analysis in RDP5, the BtSaV sequences found in this study did not contain any evidence for recombination events.

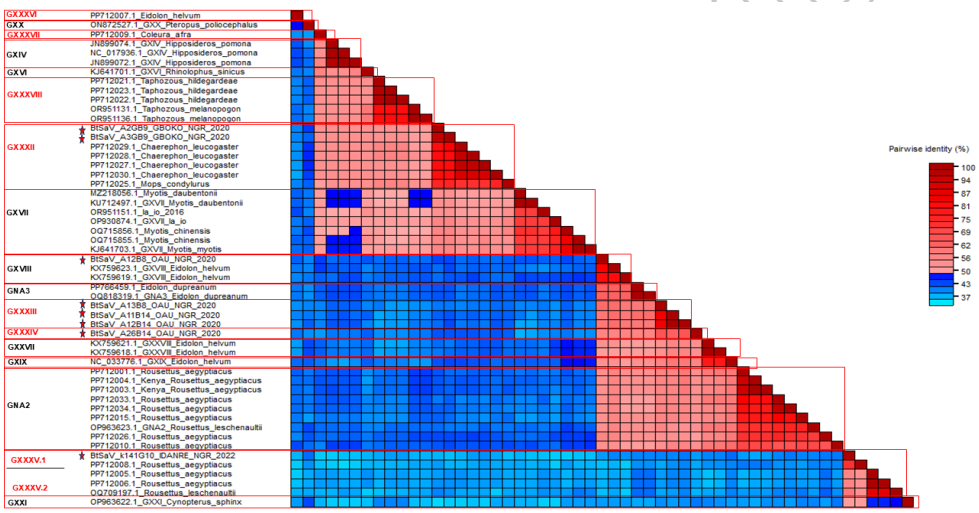

Figure 4. Classification of BtSaV based on cutoff values for genogroup clusters defined as <0.503 and cutoff values for genotype clusters defined as <0.161 using SDT showing evidence of 17 distinct BtSaV lineages detected in various bat species. Tentative new genogroups are highlighted in red font with a question mark and red star while previously described BtSaVs are in black font.

ACCEPTED MANUSCRIPT

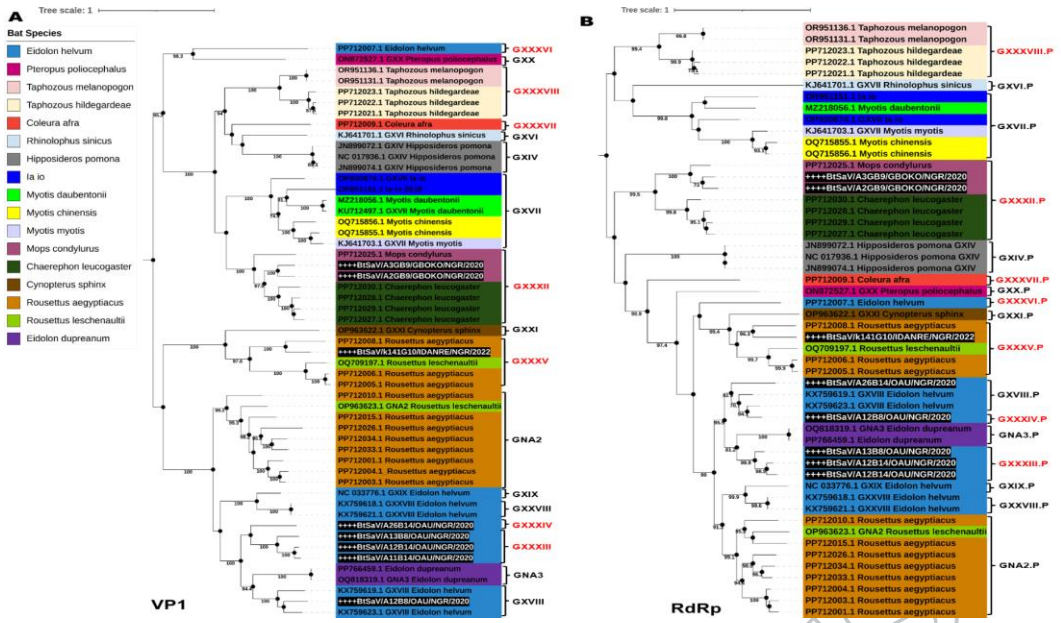

Figure 5. Maximum likelihood tree of All BtSaV based on **A.** complete VP1 gene and **B.** Complete RdRp gene, with 1000 bootstrap replications. Bat species are assigned a specific colour according to the legend provided. The BtSaV strains reported in this research are marked with an asterisk and highlighted in white.

Physicochemical Characterization and B-cell Epitope Prediction of Bat Sapovirus VP1 Protein

To determine the physicochemical characteristics of the BtSaV VP1 protein detected in this study, the sequences were submitted to the ProtParam database for analysis. We found that the BtSaV VP1 ranged from 534 to 544 amino acids while their molecular weights were between 55.8 and 57.6 kDa (**Table S3**).

To determine epitopes in the VP1 domain that may serve as targets for immune recognition, the BtSaV sequences along with a reference human SaV (HuSaV) sequence (AB455803) were subjected to the Predicting Antigenic Peptide software. Overall, we found that the BtSaV had antigenic determinants across the VP1 domain ranging from 23 - 27 epitopes (**Figure S2**).

Based on the Bepipred Linear Epitope Prediction analysis of potential B-cell epitopes in the BtSaV VP1 domain, we observed a preponderance of B-cell epitopes located in the P2 subdomain; mostly in the hypervariable regions (HVR) 1 - 4 of both HuSaVs and BtSaVs (**Figure 6**). Determining the specific regions of highly probable B

ACCEPTED MANUSCRIPT

cell epitopes, is of great practical interest and could be crucial for the design and development of potent and highly effective anti-SaV vaccines.

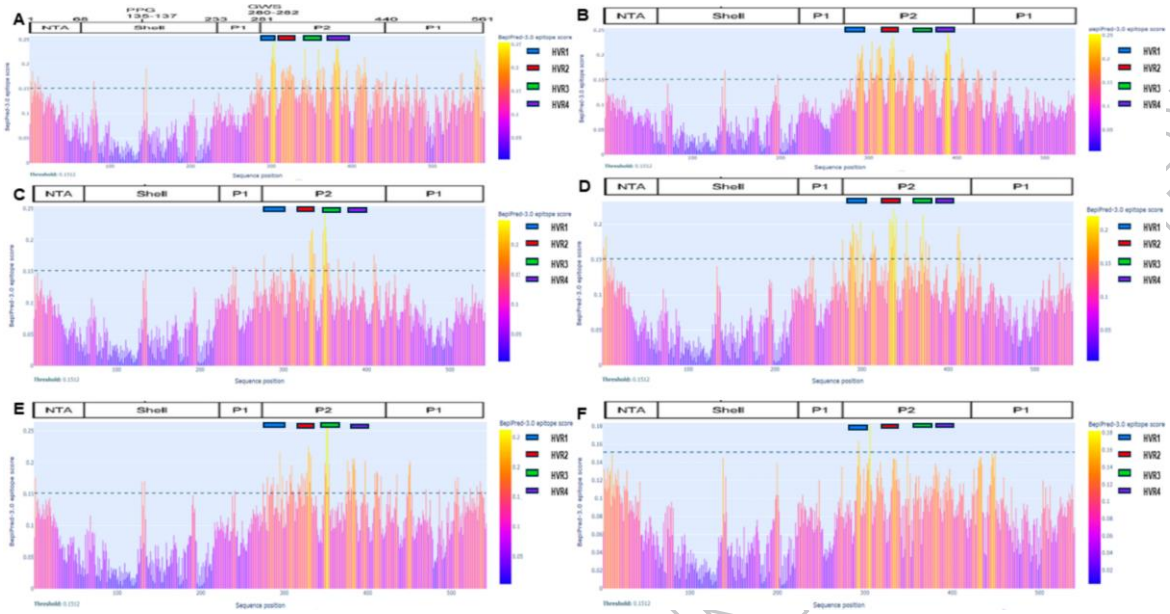

Figure 6. Predicted target sites for B cell binding across the VP1 domain with threshold set at 0.15. (A). Human SaV AB455803, (B) BtSaV/A2GB9/GBOKO/NGR/2020, (C) BtSaV/A11B14/OAU/NGR/2020, (D) BtSaV/A13B8/OAU/NGR/2020, (E) BtSaV/A26B14/OAU/NGR/2020 and (F) BtSaV/k141G10/IDANRE/NGR/2022. The hypervariable regions (HVR) located within the P2 subdomain in the VP1 are colour-coded as shown in the legend.

The virion architecture is conserved within the *Caliciviridae* family

Our analysis to gain insight into the conserved regions and amino acid substitutions related to potential binding to human receptors in the S, P1, and P2 domains in the VP1 protein of the BtSaV in comparison with HuSaV showed minor conserved motifs in the S and P1 domains. The P2 domain alignment (containing the hypervariable regions), known to be the target of neutralizing antibodies [10], differed greatly within both HuSaVs and BtSaVs (**Figure S3**). To examine the mean distance between the Cα atoms of HuSaV and BtSaV and ascertain the structural similarity of immunogenic and antigenic regions in HuSaV VP1 protein in relation to BtSaVs, the predicted VP1 protein structure of BtSaV/A2GB9/GBOKO/NGR/2020 was compared to that of HuSaV (pdb: 7dod). It was discovered that the P1 and a section of the P2 subdomain (crucial for antigen binding) showed moderate-to-high conservation (**Figure S5**).

To further characterize the structural proteins, structure predictions were generated of oligomers of VP1 and/or VP2 from HuSaV AJ606694.2 and

ACCEPTED MANUSCRIPT

BtSaV/A2GB9/GBOKO/NGR/2020. The VP1 pentamer aligns to the 5-fold axis of the solved structure of a HuSaV VLP [9] (**Figure 7. A-B**), the VP1 trimer aligns to the asymmetric unit in the T=3 VLP [9] (**Figure 7. C-D**), and the VP2 dodecamer aligns to the receptor-induced portal dodecamer from feline calicivirus (FCV) [48] (**Figure 7. E-F**). While the human virus generates a high confidence dodecamer, the bat virus prediction is less confident but still exhibits the same organization. These recapitulated higher-order macromolecular structures suggest a conserved virion structure between caliciviruses.

Structure predictions of VP1-VP2 implicates a conserved α-helix in human and bat VP2 in stabilizing the heterodimer (**Figure 7. G-H**). The prediction of the human virus has higher confidence than the bat virus. Since the bat virus still shows the same pattern, collectively it suggests a common feature. Aligning the heterodimer into each of the VP1 positions in the VLP asymmetric unit [48] displays clashes with VP2 (**Figure 7. I**). This suggests that VP2 does not fit into the icosahedral packing of VP1 and perhaps incorporation of VP2 disrupts the organization. This could explain why VP2 is a minor capsid protein and that receptor-induced conformation changes are not coordinated [48]. In contrast, aligning the heterodimer into the portal complex [48] points this α-helix towards VP1 (**Figure 7. J**). Collectively, this suggests that a C-terminal α-helix in VP2 interacts with VP1 without complete occupancy in the virion.

ACCEPTED MANUSCRIPT

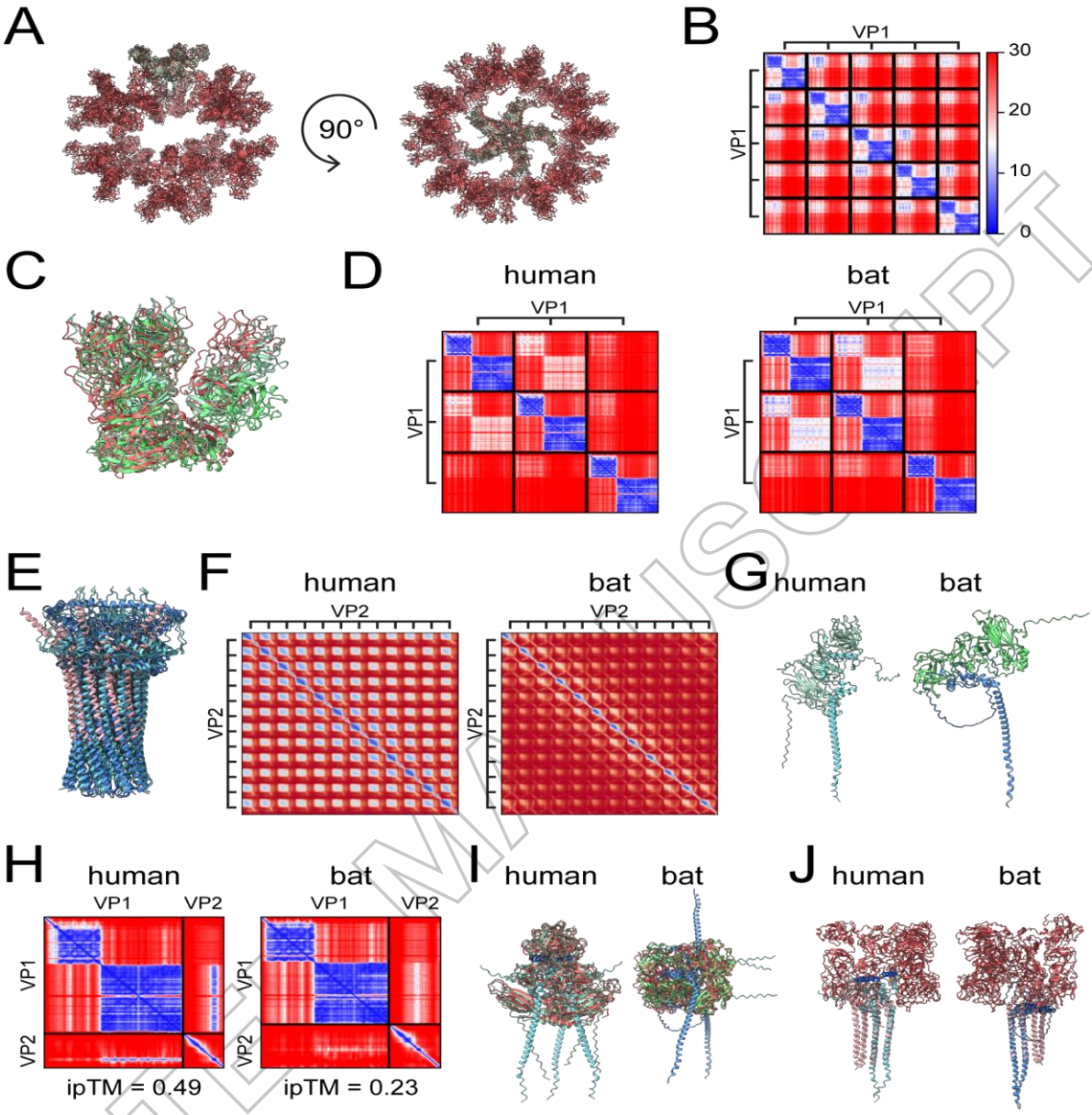

Figure 7. Structure predictions of VP1 and VP2 propose conserved virion architecture and an important α-helix in VP2. (A) LocalColabFold prediction of a pentamer of VP1 from HuSaV AJ606694.2 (light green) reproduces the 5-fold axis from the HuSaV VLP structure[4] (red). (B) The PAE plot of the pentamer structure in (A) shows protein-protein interactions. This heat map scheme is used for all subsequent PAE plots. (C) LocalColabFold predictions of a trimer of VP1 from HuSaV AJ606694.2 (light green) and BtSaV/A2GB9/GBOKO/NGR/2020 (green) reproduce the asymmetric unit from the HuSaV VLP structure[4] (red). (D) The PAE plots of the trimers in (C) show protein-protein interactions. (E) AlphaFold 3 predictions of a dodecamer of VP2 from HuSaV AJ606694.2 (light blue) or BtSaV/A2GB9/GBOKO/NGR/2020 (blue) reproduces the VP2 portal barrel of cat FCV[5] (pink). (F) The PAE plot of the dodecamers in (E) shows protein-protein interactions. (G) LocalColabFold prediction of the VP1-VP2 heterodimer (green and blue, respectively) from HuSaV AJ606694.2 and BtSaV/A2GB9/GBOKO/NGR/2020. (H) The PAE plots of the VP1-VP2 heterodimers in (G) predict an interaction of VP1 with an α-helix in VP2. This heterodimer was aligned to all 3 conformations of VP1 in the HuSaV VLP[9] (red) (I) or both conformations of VP2 in cat FCV[48] (pink) (J). The interacting α-helix in VP2 is shown in dark blue.

ACCEPTED MANUSCRIPT

Discussion

We report the detection of 4 complete, 3 near-complete (having complete coding sequences) and 11 partial genomes of SaVs from rectal swabs collected from frugivorous and insectivorous bats in Nigeria using metagenomics. This is the first report of SaV in Nigerian bats, and our analysis revealed the co-circulation of highly divergent novel SaV lineages in Nigerian bats. Specifically, we document the presence of five distinct lineages, representing probably 4 novel genogroups with one lineage belonging to a previously reported genogroup (GXVIII) (**Figure 2 and Figure 3**). SaV infections in both humans and animals have been associated with cases of AGE, posing a significant public health issue [4]. SaVs from animals including pigs, chimpanzees, and rodents have been reported to be closely related with human SaVs, indicating the possibility of cross-species transmission [14,16]. There is a need for continuous surveillance to generate full-length or partial genome sequences of SaVs to guide the development of therapeutics and diagnostic tests, enhance viral taxonomy and classification, and provide insight into the molecular mechanisms influencing the host adaptation and evolution of SaVs in humans and animals.

While PCR surveillance of SaV has shown success in identifying circulating genotypes in humans and animals, this method may not be effective in detecting novel or highly divergent lineages of SaV, resulting in partial knowledge of SaV circulation. The utilization of primer-independent deep sequencing has led to the identification of divergent SaVs in both human and animal populations [17,22-25].

We also found that a high proportion (66%) of the distinct SaV lineages were in straw-coloured fruit bat species (**Table 1, Figure 3 and 5**). A similar preponderance of distinct BtSaV in straw-coloured fruit bat species was reported in Cameroon, which suggests that this bat species may play a vital role as a reservoir in the circulation and evolution of SaV [25]. The straw-coloured fruit bat is the most widespread and commonly hunted bat species in Africa and often roosts in urban areas, increasing its opportunities to encounter humans [21,549]. Therefore, it is crucial to comprehend the mechanisms of SaV persistence and transmission dynamics in this species to avoid future spillover.

The BtSaV identified in this study were not closely related to HuSaVs, including the SaVs recently reported in children from Nigeria [17] (**Figure 3**). However, in our global assessment of the genetic diversity of SaVs in different bat species/families, we

ACCEPTED MANUSCRIPT

discovered that distinct BtSaV lineages grouped primarily according to bat families (**Figure 5 and Figure S1**) with species members within the bat family harbouring genetically related viruses regardless of bat sampling location or country. Kemenesi et al. [30] reported a similar geographic distant evolutionary relationship between similar BtSaV lineage from European and Asian bats. This could reflect virus-host-specific evolution due to a predilection of certain SaV lineages for bat species in the respective families.

Our analysis of potential B-cell epitopes in the BtSaV VP1 domain showed that most B-cell epitopes are in the P2 subdomain, specifically in the hypervariable regions (HVR) 1 – 4 (**Figure 6**). The main goal of predicting B cell epitopes in protein sequences is to identify specific segments that can be used to generate specific antibodies instead of using the whole protein. Linear B cell epitope prediction is also essential for developing synthetic peptides to generate antibodies targeting specific antigens. Our comparison of the BtSaV VP1 structure to HuSaV revealed three main domains (S, P1, and P2), similar to the reported structure of mature HuSaV capsid protein [9]. The moderate to high conservation in the P1 and a section of the P2 subdomain in comparison with HuSaV (crucial domain for antigen binding and primary targets of neutralizing antibodies) [10] (**Figure 6**) suggests potential similarities in receptor motifs and antigenic determinants. The protein folding in the P1 and P2 subdomains is noted to be conserved across caliciviruses [7]. However, a key distinction in the sequence of BtSaVs and HuSaV was the presence of various minor insertions and deletions in the P2 subdomain (**Figure S4**). Comparable changes in the P2 region have been noted in both HuSaVs and BtSaV [9,30] and could result in longer loops on the exteriors. Nevertheless, experimental research is required to comprehend the interactions between these novel BtSaVs and receptors on the surface of the bat host cells and to clarify the impact of this interaction on viral pathogenesis and regulation of morbidity and mortality in the bat host.

Some of our study's limitations included nucleic acid degradation caused by previous freeze-thaw procedures, which may have influenced the quantity and quality of genomes recovered in the present investigation. We were also unable to assess the true prevalence of SaV in individual samples because of the small sample volume of the archived samples.

In conclusion, we report the detection and genomic characterization of novel SaVs from frugivorous and insectivorous bats in Nigeria. This study extends our knowledge about

ACCEPTED MANUSCRIPT

the genetic diversity and geographical distribution of SaVs as well as the diversity of host species and ongoing evolution of SaV in bats. Our findings showed that the VP1 protein, specifically the P2 subdomain, contains immunoreactive epitopes which can serve as targets for development of an anti-SaV vaccine and antiviral medications for treating SaV infections.

Acknowledgement

This research was supported by the Laboratory of Viral Metagenomics, Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, sequencing and computational resources. This research was partly supported by the Maxwell computational resources operated at Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany.

Funding details

This work was supported partially by the [International Society for Infectious Diseases] under Grant [2020 ISID grant]; [National Institute of Health] under Grant [Grant Number D43TW012246 and U01AI151801 to UEG, SCW and RWC]; [National Institute of Health] under Grant [Grant Number U54HG007480 to CH]; [World Bank] under Grant [Grant Number project ACE-019 and ACE-IMPACT to CH]; [Rockefeller Foundation] under Grant [Grant Number #2021 HTH to CH]; [Deutsche Forschungsgemeinschaft] under Grant [390874280, 453548970, 49735088 and 4158/5-1 to JBB]; [Wellcome Trust] under Grant [209250/Z/17/Z to JBB].

Conflict of interest declaration

Authors have no conflict of interest to declare.

Author contributions

UEG, RWC, IK, JM and CH designed the study. UEG, OG, TA (Adeyanju), AO, JK, JI, AA (Adamu), RA, TA (Tomiwa) and OS collected the samples; UEG, LDC, PE and JM performed the molecular assays and metagenomic sequencing; UEG, LDC, TF, TKS and JM conducted the bioinformatics analysis and protein modelling; UEG and TKS wrote the initial draft manuscript; SCW, AH, OF, JM, OMA, JAA, JBB, RWC, IK and CH reviewed and edited the manuscript; JBB, RWC, IK, JM and CH supervised the work; OMA, JAA, JBB, SCW, RWC, IK, JM and CH provided mentorship. UEG, SCW, JBB,

ACCEPTED MANUSCRIPT

RWC, JM and CH facilitated funding acquisition. All the authors read and approved the final manuscript before submission.

Disclosure statement:

The authors declare that the study was carried out without any commercial or financial ties that could be seen as a possible conflict of interest.

Data Availability

The bat sapovirus sequenced raw reads for this study and the generated genomes have been deposited in NCBI under BioProject **PRJNA1185806** with accession numbers **PQ623340- PQ623357.**

References

1. Vinjé J, Estes MK, Esteves P, et al. ICTV Virus Taxonomy Profile: Caliciviridae. J Gen Virol. 2019;100(11):1469-1470. doi:10.1099/jgv.0.001332

2. Madeley CR, Cosgrove BP. Letter: Caliciviruses in man. Lancet. 1976;1(7952):199-200. doi:10.1016/s0140-6736(76)91309-x

3. Shen H, Zhang J, Gauger PC, et al. Genetic characterization of porcine sapoviruses identified from pigs during a diarrhoea outbreak in Iowa, 2019. Transbound Emerg Dis. 2022;69(3):1246-1255. doi:10.1111/tbed.14087

4. Becker-Dreps S, Bucardo F, Vinjé J. Sapovirus: an important cause of acute gastroenteritis in children. Lancet Child Adolesc Health. 2019;3(11):758-759. doi:10.1016/S2352-4642(19)30270-6

5. Oka, T.; Wang, Q.; Katayama, K.; Saif, L.J. Comprehensive review of human sapoviruses. Clin. Microbiol. Rev. 2015, 28, 32–53.

6. Li J, Shen Q, Zhang W, et al. Genomic organization and recombination analysis of a porcine sapovirus identified from a piglet with diarrhea in China. Virol J. 2017;14(1):57.doi:10.1186/s12985-017-0729-1

7. Chen R, Neill JD, Noel JS, et al. Inter- and intragenus structural variations in caliciviruses and their functional implications. J Virol. 2004;78(12):6469-6479. doi:10.1128/JVI.78.12.6469-6479.2004

8. Miyazaki N, Taylor DW, Hansman GS, Murata K. Antigenic and Cryo-Electron Microscopy Structure Analysis of a Chimeric Sapovirus Capsid. J Virol. 2015;90(5):2664-2675. doi:10.1128/JVI.02916-15

9. Miyazaki N, Song C, Oka T, et al. Atomic Structure of the Human Sapovirus Capsid

ACCEPTED MANUSCRIPT

Reveals a Unique Capsid Protein Conformation in Caliciviruses. J Virol. 2022;96(9):e0029822. doi:10.1128/jvi.00298-22

10. Ossiboff RJ, Zhou Y, Lightfoot PJ, Prasad BV, Parker JS. Conformational changes in the capsid of a calicivirus upon interaction with its functional receptor. J Virol. 2010;84(11):5550-5564. doi:10.1128/JVI.02371-09

11. Oka T, Mori K, Iritani N, et al. Human sapovirus classification based on complete capsid nucleotide sequences. Arch Virol. 2012;157(2):349-352. doi:10.1007/s00705-011-1161-2

12. Zhao W, Gao Z, Guo C, et al. A dual typing system establishment and global diversity analysis for sapoviruses. BMC Genomics. 2024;25(1):1131. Published 2024 Nov 22. doi:10.1186/s12864-024-11048-w

13. Liu X, Song C, Liu Y, et al. High Genetic Diversity of Porcine Sapovirus From Diarrheic Piglets in Yunnan Province, China. Front Vet Sci. 2022;9:854905.doi:10.3389/fvets.2022.854905

14. Mombo IM, Berthet N, Bouchier C, et al. Characterization of a genogroup I sapovirus isolated from chimpanzees in the republic of congo. Genome Announc. 2014;2(4):e00680-14.doi:10.1128/genomeA.00680-14

15. Olarte-Castillo XA, Hofer H, Goller KV, Martella V, Moehlman PD, East ML. Divergent Sapovirus Strains and Infection Prevalence in Wild Carnivores in the Serengeti Ecosystem: A Long-Term Study. PLoS One. 2016;11(9):e0163548.doi:10.1371/journal.pone.0163548

16. Firth C, Bhat M, Firth MA, et al. Detection of zoonotic pathogens and characterization of novel viruses carried by commensal Rattus norvegicus in New York City. mBio. 2014;5(5):e01933-14.doi:10.1128/mBio.01933-14

17. George UE, Faleye TOC, De Coninck L, et al. Metagenomic Detection and Genetic Characterization of Human Sapoviruses among Children with Acute Flaccid Paralysis in Nigeria. Pathogens. 2024;13(3):264. doi:10.3390/pathogens13030264

18. Diez-Valcarce M, Castro CJ, Marine RL, et al. Genetic diversity of human sapovirus across the Americas. J Clin Virol. 2018;104:65-72. doi:10.1016/j.jcv.2018.05.003

19. Martella V, Lorusso E, Banyai K, et al. Identification of a porcine calicivirus related genetically to human sapoviruses. J Clin Microbiol. 2008;46(6):1907-1913. doi:10.1128/JCM.00341-08.

20. Kunz TH, Braun de Torrez E, Bauer D, Lobova T, Fleming TH. Ecosystem services provided by bats. Ann N Y Acad Sci. 2011;1223:1-38. doi:10.1111/j.1749-6632.2011.06004.x

21. Kamins AO, Restif O, Ntiamoa-Baidu Y, et al. Uncovering the fruit bat bushmeat commodity chain and the true extent of fruit bat hunting in Ghana, West Africa. Biol Conserv. 2011;144(12):3000-3008. doi:10.1016/j.biocon.2011.09.003

22. Wu Z, Yang L, Ren X, et al. Deciphering the bat virome catalog to better understand

ACCEPTED MANUSCRIPT

the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. ISME J. 2016;10(3):609-620. doi:10.1038/ismej.2015.138

23. Kemenesi G, Gellért Á, Dallos B, et al. Sequencing and molecular modeling identifies candidate members of Caliciviridae family in bats. Infect Genet Evol. 2016;41:227-232. doi:10.1016/j.meegid.2016.04.004

24. Van Brussel K, Mahar JE, Ortiz-Baez AS, et al. Faecal virome of the Australian grey-headed flying fox from urban/suburban environments contains novel coronaviruses, retroviruses and sapoviruses. Virology. 2022;576:42-51. doi:10.1016/j.virol.2022.09.002

25. Yinda CK, Conceição-Neto N, Zeller M, et al. Novel highly divergent sapoviruses detected by metagenomics analysis in straw-colored fruit bats in Cameroon. Emerg Microbes Infect. 2017;6(5): e38.doi:10.1038/emi.2017.20

26. George U, George O, Oragwa A, et al. Detection of Alpha- and Betacoronaviruses in Frugivorous and Insectivorous Bats in Nigeria. Pathogens. 2022;11(9): 1017.doi:10.3390/pathogens11091017

27. Townzen JS, Brower AV, Judd DD. Identification of mosquito bloodmeals using mitochondrial cytochrome oxidase subunit I and cytochrome b gene sequences. Med Vet Entomol. 2008;22(4):386-393. doi:10.1111/j.1365-2915.2008. 00760.x

28. Matranga CB, Gladden-Young A, Qu J, et al. Unbiased Deep Sequencing of RNA Viruses from Clinical Samples. J Vis Exp. 2016;(113):54117. Published 2016 Jul 2. doi:10.3791/54117

29. Conceição-Neto N, Yinda KC, Van Ranst M, Matthijnssens J. NetoVIR: Modular Approach to Customize Sample Preparation Procedures for Viral Metagenomics. Methods Mol Biol. 2018; 1838:85-95. doi:10.1007/978-1-4939-8682-8_7

30. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-2120. doi:10.1093/bioinformatics/btu170

31. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-359. Published 2012 Mar 4. doi:10.1038/nmeth.1923

32. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 2017;27(5):824-834. doi:10.1101/gr.213959.116

33. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31(10):1674-1676. doi:10.1093/bioinformatics/btv033

34. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12(1):59-60. doi:10.1038/nmeth.3176

35. Zhu W, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences. Nucleic Acids Res. 2010;38(12):e132. doi:10.1093/nar/gkq275

ACCEPTED MANUSCRIPT

36. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-780. doi:10.1093/molbev/mst010

37. Minh BQ, Schmidt HA, Chernomor O, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era [published correction appears in Mol Biol Evol. 2020;37(5):1530-1534. doi:10.1093/molbev/msaa015

38. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587-589. doi:10.1038/nmeth.4285

39. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. 2024;52(W1):W78-W82. doi:10.1093/nar/gkae268

40. Muhire BM, Varsani A, Martin DP. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One. 2014;9(9):e108277. Published 2014 Sep 26. doi:10.1371/journal.pone.0108277

41. Martin DP, Varsani A, Roumagnac P, et al. RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus Evol. 2020;7(1):veaa087. Published 2020 Apr 12. doi:10.1093/ve/veaa087

42. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 1990;276(1-2):172-174. doi:10.1016/0014-5793(90)80535-q

43. Clifford JN, Høie MH, Deleuran S, Peters B, Nielsen M, Marcatili P. BepiPred-3.0: Improved B-cell epitope prediction using protein language models. Protein Sci. 2022;31(12):e4497. doi:10.1002/pro.4497

44. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2

45. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Methods. 2022;19(6):679-682. doi:10.1038/s41592-022-01488-1

46. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630(8016):493-500. doi:10.1038/s41586-024-07487-w

47. Meng, E.C.; Goddard, T.D.; Pettersen, E.F.; Couch, G.S.; Pearson, Z.J.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Tools for structure building and analysis. Protein Sci. 2023, 32, e4792.

48. Conley MJ, McElwee M, Azmi L, et al. Calicivirus VP2 forms a portal-like assembly following receptor engagement. Nature. 2019;565(7739):377-381. doi:10.1038/s41586-018-0852-1

ACCEPTED MANUSCRIPT

49. Hayman DT, Peel AJ. Can survival analyses detect hunting pressure in a highly connected species? Lessons from straw-coloured fruit bats. Biol Conserv. 2016;200:131-139. doi:10.1016/j.biocon.2016.06.003

Identification and Genomic Characterization of Known and Novel Highly Divergent Sapoviruses in Frugivorous and Insectivorous Bats in Nigeria

Table S1: Demographic description and summary of bat samples analysed in this study

S/N	Sample ID	Site of sample collection	Bat Family	Bat Species	Year	Individual or pool
1	BA1	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2019	Pool of 5 bats
2	BA2	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2019	Pool of 5 bats
3	BA3	Mkar-Benue state	*Pteropodidae*	*Eidolon helvum*	2019	Pool of 5 bats
4	BA4	Lim-Bauchi state	*Pteropodidae*	*Eidolon helvum*	2019	Pool of 5 bats
5	BA5[a]	Lim, Bauchi state	*Pteropodidae*	*Eidolon helvum*	2019	Pool of 5 bats
6	BA6	Vom-Jos, Plateau state	*Molossidae*	Chaerephon spp	2019	Individual
7	B7/BT/OAU/1	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 5 bats
8	B8/BT/OAU/2 [b, c]	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 5 bats

ACCEPTED MANUSCRIPT

9	B9/BT/OAU/3	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 5 bats
10	B10/BT/OAU/4	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 5 bats
11	B11/BT/OAU/5	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 5 bats
12	B12/BT/OAU/6 [b, c]	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 5 bats
13	B13/BT/OAU/7	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 5 bats
14	B14/BT/OAU/8 [b, c]	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 5 bats
15	B15/BT/OAU/9 [b, c]	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 6 bats
16	B16/BT/OAU/10	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 4 bats
17	B17/BT/OAU/11 [b, c]	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 4 bats
18	B18/BT/OAU/12	OAU-Osun state	*Hipposideridae*	*Hipposideros ruber*	2020	Pool of 5 bats
19	B19/BT/OAU/13	OAU-Osun state	*Hipposideridae*	*Hipposideros ruber*	2020	Pool of 6 bats
20	B20/BT/OAU/14	OAU-Osun state	*Hipposideridae*	*Hipposideros ruber*	2020	Pool of 5 bats
21	B21/BT/OAU/15	OAU-Osun state	*Hipposideridae*	*Hipposideros ruber*	2020	Pool of 5 bats
22	B22/BT/OAU/16	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 4 bats
23	B23/BT/OAU/17 [b, c]	OAU-Osun state	*Pteropodidae*	*Eidolon helvum*	2020	Pool of 2 bats

ACCEPTED MANUSCRIPT

24	B24/BT/PL/01	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 3 bats
25	B25/BT/PL/02	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 3 bats
26	B26/BT/PL/03	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 3 bats
27	B27/BT/PL/04	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 8 bats
28	B28/BT/PL/05	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 8 bats
29	B29/BT/PL/06	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 8 bats
30	B30/BT/PL/07	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 8 bats
31	B31/BT/PL/08	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 8 bats
32	B32/BT/PL/09	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 7 bats
33	B33/BT/PL/010	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Pool of 7 bats
34	B34/BT/GB/01	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2020	Pool of 5 bats
35	B35/BT/GB/02	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2020	Pool of 5 bats
36	B36/BT/GB/03	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2020	Pool of 5 bats
37	B37/BT/BA/01	Lim, Bauchi state	*Pteropodidae*	*Eidolon helvum*	2019	Pool of 5 bats
38	B38/BT/BA/02	Lim, Bauchi state	*Pteropodidae*	*Eidolon helvum*	2019	Pool of 5 bats

ACCEPTED MANUSCRIPT

39	B39/BT/BA/03	Lim, Bauchi state	*Pteropodidae*	*Eidolon helvum*	2019	Pool of 5 bats
40	B40/BT/BA/04	Lim, Bauchi state	*Pteropodidae*	*Eidolon helvum*	2019	Pool of 5 bats
41	GB04	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2020	Individual
42	GB09 [b, d]	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2020	Individual
43	GB10	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2020	Individual
44	GB12	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2020	Individual
45	GB13	Gboko-Benue state	*Molossidae*	*Mops Condylurus*	2020	Individual
46	NG17	Paiko - Niger state	*Molossidae*	*Chaerephon spp*	2021	Individual
47	NG19	Paiko - Niger state	*Molossidae*	*Chaerephon spp*	2021	Individual
48	NG22	Paiko - Niger state	*Molossidae*	*Chaerephon spp*	2021	Individual
49	NG24	Paiko - Niger state	*Molossidae*	*Chaerephon spp*	2021	Individual
50	NG33	Paiko - Niger state	*Molossidae*	*Chaerephon spp*	2021	Individual
51	NG34	Paiko - Niger state	*Molossidae*	*Chaerephon spp*	2021	Individual
52	NG35	Paiko - Niger state	*Molossidae*	*Chaerephon spp*	2021	Individual
53	CER24	OAU-Osun state	*Hipposideridae*	*Hipposideros ruber*	2021	Individual

ACCEPTED MANUSCRIPT

54	CER43	OAU-Osun state	*Hipposiderida e*	*Hipposideros ruber*	2021	Individual
55	PL66	Jos Zoo, Plateau state	*Pteropodidae*	*Eidolon helvum*	2021	Individual
56	G10 [b, d]	Idanre – Ondo state	*Pteropodidae*	*R. aegyptiacus*	2022	Individual
57	G11 [b, d]	Idanre – Ondo state	*Pteropodidae*	*R. aegyptiacus*	2022	Individual

Table S2: Summary of SaV contigs detected in the pooled bat samples

S/N	Virus strain/Isolate name	Accession number	Nucleotide length	Mean Coverage	Genome completeness	SaV Genogroup
1.	BtSaV/A2GB9/GBOKO/NGR/2020	PQ623354	7587	214.3	Complete genome	Unclassified
2.	BtSaV/A3GB9/GBOKO/NGR/2020	PQ623355	7533	101	Complete genome	Unclassified
3.	BtSaV/A12B8/OAU/NGR/2020	PQ623340	7611	2,912	Near-complete genome with complete	GXVIII

ACCEPTED MANUSCRIPT

					coding sequence	
4.	BtSaV/A13B8/OAU/NGR/2020	PQ623341	7453	220	Near-complete genome with complete coding sequence	Unclassified
5.	BtSaV/A11B14/OAU/NGR/2020	PQ623344	7483	521	Near-complete genome with complete coding sequence	Unclassified
6.	BtSaV/A12B14/OAU/NGR/2020	PQ623345	7361	372	Near-complete genome with complete coding sequence	Unclassified

ACCEPTED MANUSCRIPT

7.	BtSaV/k141G10/IDANRE/NGR/2022	PQ623356	7110	467	Complete genome	Unclassified
8.	BtSaV/A26B14/OAU/NGR/2020	PQ623346	5573	32	Partial genome (complete VP1 and VP2)	Unclassified
9.	BtSaV/A128B12/OAU/NGR/2020	PQ623342	1176	9.147	Partial genome	GXVIII
10.	BtSaV/A141B12/OAU/NGR/2020	PQ623343	1715	15	Partial genome	GXVIII
11.	BtSaV/A97B14/OAU/NGR/2020	PQ623347	2735	521.4044	Partial genome	Unclassified
12.	BtSaV/A347B14/OAU/NGR/2020	PQ623348	1673	15	Partial genome	Unclassified
13.	BtSaV/A93B15/OAU/NGR/2020	PQ623349	898	6.8296	Partial genome	Unclassified
14.	BtSaV/A43B17/OAU/NGR/2020	PQ623350	2293	26.2957	Partial genome	GXVIII
15.	BtSaV/A277B23/OAU/NGR/2020	PQ623351	1883	11.3011	Partial genome	GXVIII
16.	BtSaV/A583B23/OAU/NGR/2020	PQ623353	1288	10	Partial genome	GXVIII

ACCEPTED MANUSCRIPT

17.	BtSaV/A904B23/OAU/NGR/2020	PQ623352	1036	11	Partial genome	GXVIII
18.	BtSaV/A25G11/IDANRE/NGR/2022	PQ623357	2103	40.3428	Partial genome	Unclassified

Abbreviations: Obafemi Awolowo University, OAU;

[a] This sample was excluded from sequencing after library preparation due to low DNA concentration and contamination.

[b] Samples with Sapovirus reads detected.

[c] Samples where Sapovirus reads were detected using the NetoVIR protocol [37].

[d] Samples where Sapovirus reads were detected using the Matranga et al. [36] protocol.

ACCEPTED MANUSCRIPT

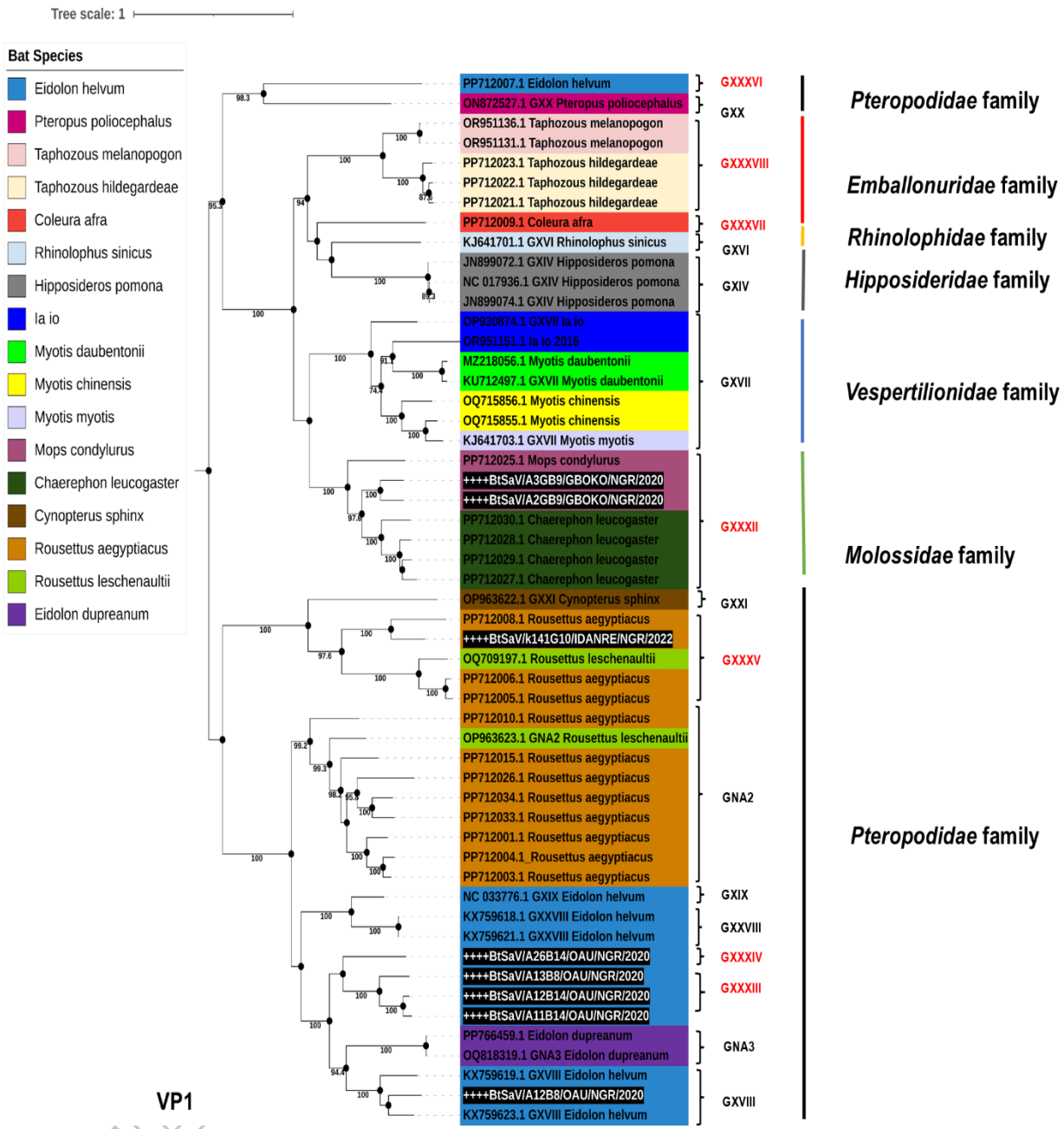

Figure S1. Maximum likelihood tree using ModelFinder of All BtSaV based on A. complete VP1 gene, with 1000 bootstrap replications. Bat species and families are assigned a specific colour according to the legend provided. The BtSaV strains reported in this research are marked with an asterisk and labelled in white. Tentative newly described genogroups are highlighted in red. The tree was visualized using the Interactive Tree of Life (iTOL) v6 with midpoint rooting.

Table S3: Physical and Chemical properties of BtSaV VP1 protein detected in this study.

Characteristic	BtSaV/A2GB9/GBOKO/NGR/2020	BtSaV/A3GB9/GBOKO/NGR/2020	BtSaV/A12B8/OAU/NGR/2020	BtSaV/A13B8/OAU/NGR/2020	BtSaV/A11B14/OAU/NGR/2020	BtSaV/A12B14/OAU/NGR/2020	BtSaV/k141G10/IDANRE/NGR/2022	BtSaV/A26B14/OAU/NGR/2020
Number of amino acids	534	536	544	544	542	541	540	542
Formula	$C_{2508}H_{3843}N_{677}O_{747}S_{24}$	$C_{2514}H_{3855}N_{677}O_{754}S_{24}$	$C_{2542}H_{3937}N_{687}O_{777}S_{17}$	$C_{2543}H_{3969}N_{705}O_{787}S_{17}$	$C_{2527}H_{3961}N_{693}O_{781}S_{19}$	$C_{2527}H_{3960}N_{677}O_{747}S_{19}$	$C_{2472}H_{3879}N_{659}O_{780}S_{15}$	$C_{2534}H_{3976}N_{694}O_{786}S_{13}$
Molecular weight	56200.63	56396.79	57099.38	57555.76	57155.56	57120.56	55791.84	57156.40
Theoretical isoelectric point (PI)	5.23	5.23	5.37	5.41	5.40	5.40	5.30	5.55
Number of negatively charged res	40	40	37	35	33	33	32	34

ACCEPTED MANUSCRIPT

idues								
Number of positively charged residues	25	25	28	25	23	23	23	25
Instability index (II)	26.84 (stable)	28.37 (stable)	35.56 (stable)	29.79 (stable)	32.51 (stable)	30.31 (stable)	27.70 (stable)	29.26 (stable)
Aliphatic index	78.88	78.92	90.04	85.09	87.97	88.30	86.98	88.49
Grand average of hydropathicity (GRAVY)	0.010	0.004	0.021	-0.025	0.047	0.040	0.155	0.045

ACCEPTED MANUSCRIPT

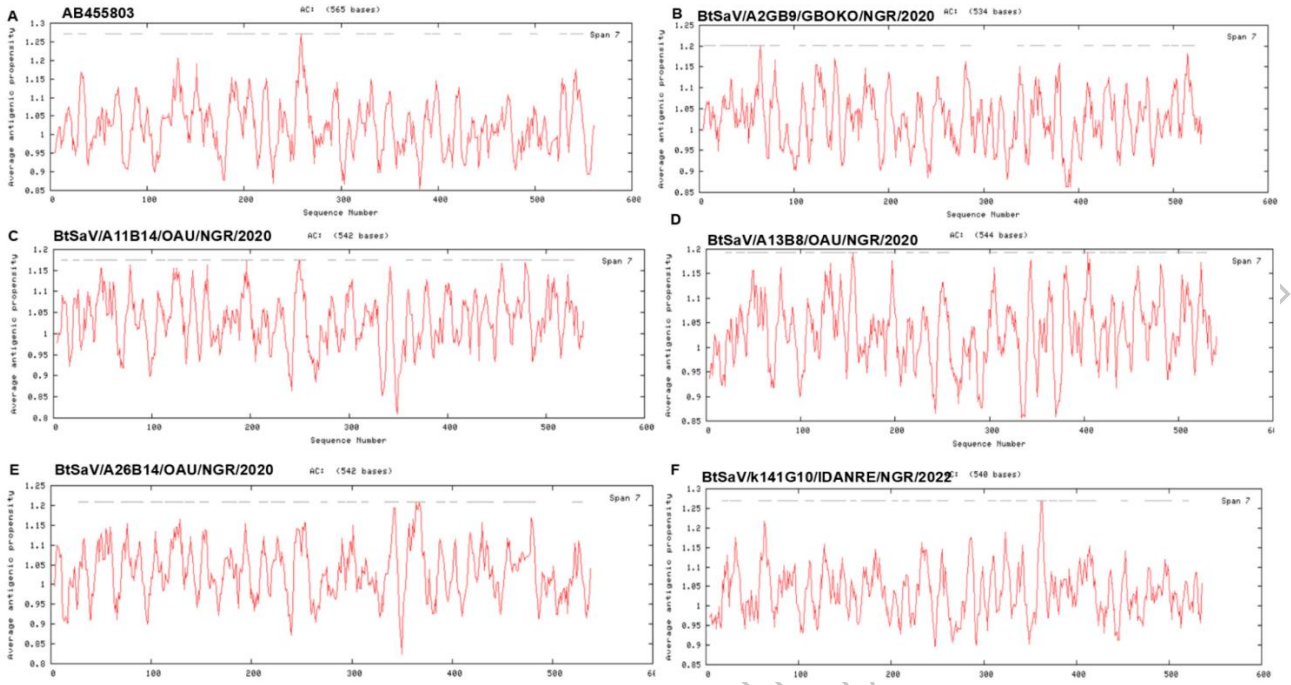

Figure S2. Analysis of antigenic determinants in the VP1 protein. (A). AB455803 VP1 protein showing 24 epitopes, (B). BtSaV/A2GB9/GBOKO/NGR/2020 VP1 protein showing 24 epitopes, (C). BtSaV/A11B14/OAU/NGR/2020 VP1 protein showing 27 epitopes, and (D). BtSaV/A13B8/OAU/NGR/2020 VP1 protein showing 25 epitopes, (E). BtSaV/A26B14/OAU/NGR/2020 VP1 protein showing 23 epitopes and (F). BtSaV/k141G10/IDANRE/NGR/2022 VP1 protein showing 23 epitopes respectively.

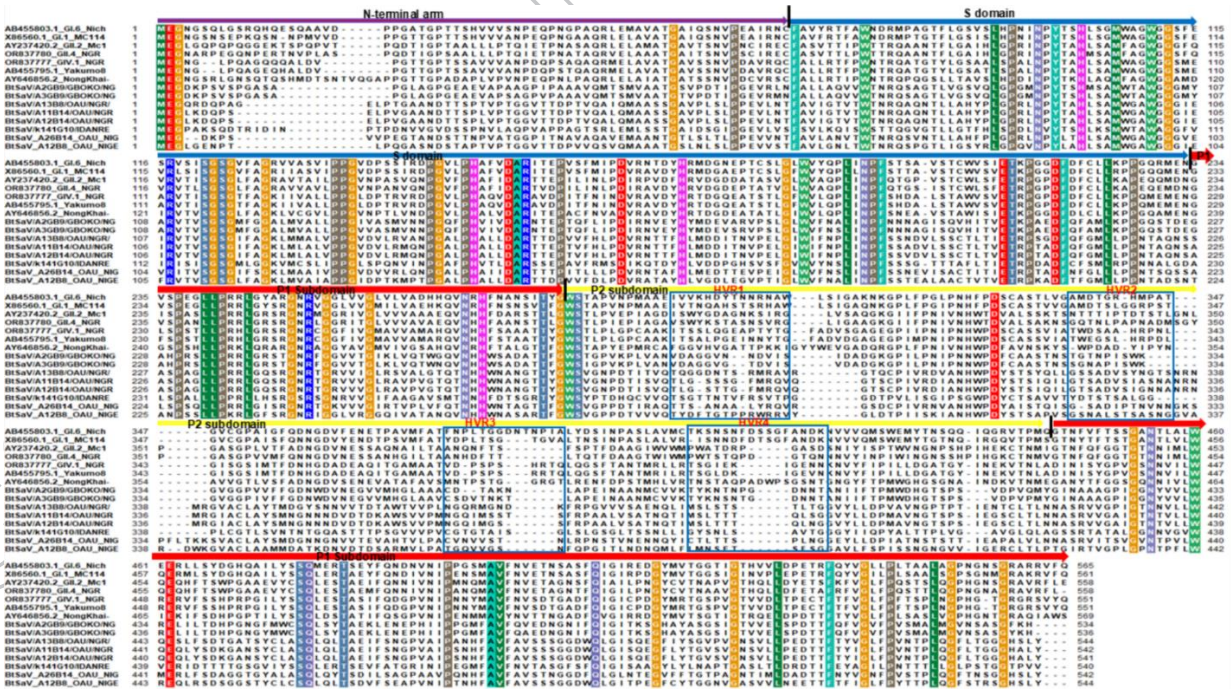

Figure S3: Sequence alignment of HuSaV and BtSaV complete VP1 protein. conserved motifs are coloured across the residue. The N-terminal arm is highlighted in purple, the S-domain in blue, the P1 subdomain in red and the P2 subdomain in yellow. Hypervariable regions (HVR) 1 - 4 are also highlighted in blue.

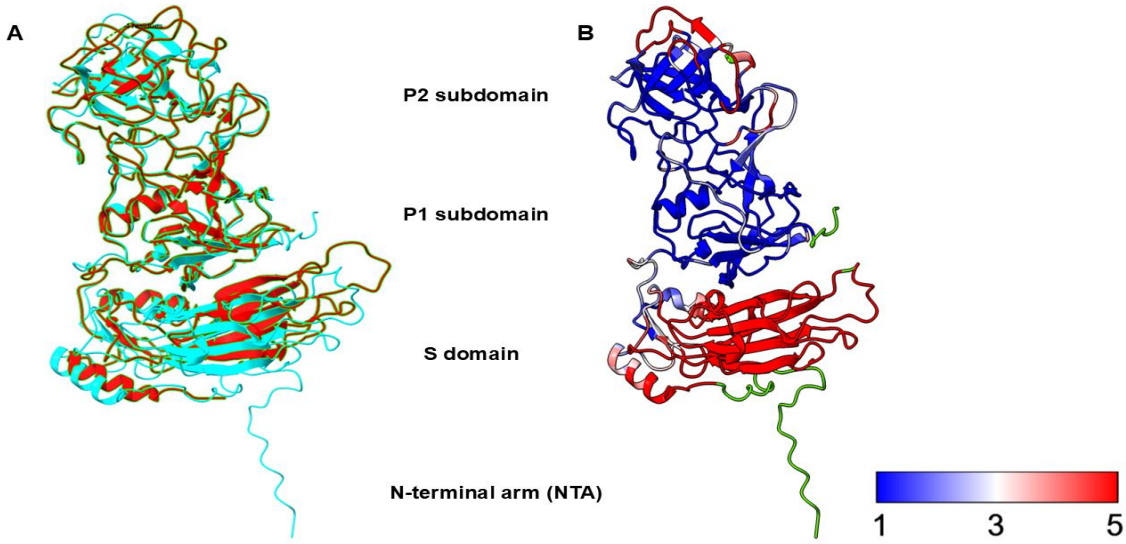

Figure S4. Comparing the similarity and average distance between the Cα atoms of superimposed HuSaV (pdb: 7dod) and BtSaV/A2GB9/GBOKO/NGR/2020. (**A**) The HuSaV VP1 protein template structure is shown in red cartoon, while BtSaV/A2GB9/GBOKO/NGR/2020 VP1 model structure is depicted in cyan cartoon. (**B**) Conservation of amino acid residues of BtSaV/A2GB9/GBOKO/NGR/2020 VP1 proteins mapped onto the molecular surface of HuSaV template to show the average distance between their Cα atoms. Blue indicates that the root mean square deviation (RMSD) of Cα atoms are within 1.0-Å between the two proteins while Red indicates that the RMSD of Cα atoms of the two proteins are 5.0-Å or more away from each other. The portion-coloured green indicates the region with no correspondence between the two proteins.

ACCEPTED MANUSCRIPT