Type of the Paper (Article)

Whole Genome Sequence Analysis of Porcine Astroviruses Reveals Novel Genetically Diverse Genotypes Circulating in East African Smallholder Pig Farms

Joshua O. Amimo[1,2,3*], Eunice M. Machuka[3], Edward O. Abworo[3], Anastasia N. Vlasova[2], Roger Pelle[3]

[1]Department of Animal Production, Faculty of Veterinary Medicine, University of Nairobi

[2]Food Animal Health Research Program, Department of Veterinary Preventive Medicine, The Ohio State University

[3]Biosciences eastern and central Africa-International Livestock Research Institute (BecA-ILRI) Hub, Kenya

*Corresponding Author - jamimo@uonbi.ac.ke

Abstract

Astroviruses (AstVs) are occurs globally and are common causes of gastroenteritis in human and animals. The genetic diversity and epidemiology of AstVs in Africa is not well known, hence, we aimed to genetically characterize astroviruses in asymptomatic smallholder piglets in East Africa. Twenty-four samples randomly selected from 446 piglets (<6 months old), initially collected for rotavirus study, was sequenced for metagenomic analysis. Thirteen (13/24) samples had contigs with high identity to genus *Mamastrovirus*. Analysis of 7 strains with complete (or near complete) genome revealed variable nucleotide and amino acid sequence identities with known PoAstV strains. The U083 and K321 strains had nucleotide sequence similarities ranging from 66.4 to 75.4 % to the known PoAstV2 strains, nucleotide sequence similarity of U460 strain with known PoAstV3 ranged 57.0 to 65.1 % to the, while K062, K366, K451, and K456 strains showed nucleotide sequence similarities of 63.5 to 80 % to the known PoAstV4 strains. The low sequence identities (<90 %) indicate that novel genotypes of PoAstVs are circulating in the study area. Multiple recombination events were detected in our PoAstV4 strains, indicating that the genetic diversity observed in these strains may be due to recombination. Importantly, we identified potential candidate epitopes with conserved peptides in our PoAstV strains that could aid in the design of immune diagnosis tools and subunit vaccines. Our data provide new intuitions into the genetic structure of porcine astroviruses in East African.

Keywords: porcine astroviruses; linear antigenic epitopes; recombination; glycosylation; whole genome sequences; East Africa

 doi:10.20944/preprints202009.0559.v1

1. Introduction

Porcine astroviruses (PoAstV) belong to the family *Astroviridae* consisting of *Avastrovirus* and *Mamastrovirus* genera [2]. Astroviruses (AstVs) have been reported in a wide variety of mammals and birds [2-5]. AstVs are the second most common cause (after rotaviruses) of viral gastroenteritis in infants and children [6], and they have been reported in samples from water and sewage [7, 8]. The first AstV were identified by electron microscopy in 1975 in Scotland in fecal samples from diarrheic infants [9]. AstVs is disseminated through the fecal–oral route, direct contact, or contaminated food and water [10]. In Africa the number of studies on AstV detection is relatively low, with only one in East Africa in children [11]. Rapid mutation of AstVs, and co-infection of multiple AstV strains in a single individual [3, 12], could result in a recombination event from which a zoonotic strains may emerge. Emergence of porcine-human AstV recombinants has been reported in regions where there is close interaction between pigs and humans, such as our study region, however, transmission from humans to pigs is suspected although the reverse has not been described [13].

AstVs are small, nonenveloped viruses containing a positive-sense single-stranded RNA of 6.4–7.3 kb in size. The positive-sense viral RNA genome consists of untranslated regions (UTR) at both 5'- and 3'-ends with a poly-A tail at the 3'-end. AstV genome contains 3 open reading frames (ORF1a, ORF1b, and ORF2) starting from the 5'-end, with a frame shift between ORF1a and ORF1b [2, 14]. The three ORFs encode for nonstructural protease (ORF1a), RNA-dependent RNA-polymerase (ORF1b), and structural capsid protein (ORF2), respectively [13]. Genome analysis has shown that ORF1b is the most conserved while ORF2 is highly divergent due to selective pressure in this region. Recently, the Astroviruses Study Group, International Committee on Taxonomy of Viruses (ICTV) (http://ictvonline.org/virusTaxonomy.asp), proposed a classification based on the amino acid sequence of ORF2 [15, 16], of which 19 genotype species of Mamastrovirus (MAstV 1–19) has been identifed. In pigs, the first porcine astrovirus (PoAstV) was identified by electron microscopy in 1980 [17]. To date, four more PoAstV types (PAstV2– PAstV5) have been identified [18-21]. Association of PoAstV with gastroenteritis has been reported [17, 19, 22, 23], however, PoAstVs have also been found in healthy pigs [21, 24-28]. Despite the clinical and agricultural significance of astroviruses, they are among the least studied enteric RNA viruses [29], which could be due to the lack of small animal models and fewer full genome sequences for many AstV species. Additionally, there are limited established culture systems for propagating AstVs [23, 30]. Studies of genetic and antigenic diversity among PoAstV strains in a given location may aid the design of accurate diagnostic assays and vaccine and hence improved disease prevention [25]. Currently, there are no PoAstV vaccines available commercially. In this study, the complete (or near complete) genomes

 doi:10.20944/preprints202009.0559.v1

of seven porcine astroviruses were analyzed and characterized, which revealed novel porcine astrovirus strains. Additionally, we determined potential linear antigenic epitope on capsid region of our PoAstV strains with predicted high antigenicity, which could aid in the design of immune diagnosis reagents and subunit vaccines.

2. Materials and methods

2.1 Ethics statement

This study was conducted according to the ethical guidelines of both Commonwealth Scientific Research and Industrial Organization (CSIRO)-Social Science Human Research Ethics Committee (CSSHREC), Australia (approval 059/11), and Institutional Animal Care and Use Committee (IACUC) of International Livestock Research Institute (ILRI), Kenya.

2.2 Selection of samples: Twenty four (24) samples were randomly selected from a total of 446 piglets aged below 6 months that were initially collected to study rotaviruses distribution and diversity in western Kenya and eastern Uganda [31].

2.3 Processing of fecal sample, RNA library preparation and sequencing: The detailed sample and library preparation and sequencing of samples used in this study is described by Amimo et al 2016 [3]. Briefly, Fecal suspensions were filtered (0.22µm membrane filters) to remove eukaryotic and bacterial cell debris, then treated with DNases and RNases (Takara, Japan). RNA was extracted using the QIAamp Viral RNA Mini Kit (Qiagen, USA) according to the manufacturer's instructions. Library construction was done using the TruSeq™ RNA library preparation kit v2 (Illumina®, USA) according to the manufacturer's instructions. The library size and concentration were estimated by Bioanalyzer (Agilent, Santa Clara, CA, USA) and the Qubit, respectively. A ten (10nM) library pool was prepared by combining the 24 samples to attain an equal molar concentration for each library. Pooled libraries were sequenced by the MiSeq (Illumina®) platform using sequencing runs of 2×150 paired-end reads at the BecA-ILRI Hub genomics laboratory facility in Nairobi, Kenya.

2.4 Data analysis: Comprehensive sequence data quality check, *de novo* contig assembly and contigs classification has been described by Amimo et al 2016 [3]. Using preliminary results, the contigs from 13 samples, having hits with high similarity to astroviruses, were combined and analyzed with the Metavir2 beta version, the web based tool for virome analysis [32], which compute taxonomic

 doi:10.20944/preprints202009.0559.v1

composition through a BLAST comparison with the Refseq complete viral genomes protein sequences database from NCBI.

2.4.1 Phylogenetic analysis: We carried out multiple sequence alignment using Clustal Omega (ClustalO) web server (http://www.ebi.ac.uk/Tools/msa/clustalo/) to evaluate the sequence percent similarity of our strains with other known AstV genomes at nucleotide and amino acid levels. Phylogenetic analysis of our assembled genomes and deduced amino acids were performed in MEGA X [1] after building alignment by ClustalW algorithm and phylogenetic tree created using the neighbour-joining method [33] with 1000 bootstraps replicates.

2.4.2 Recombination analysis: A nucleotide alignment was created by using ClustalW on the full genome sequences of 28 mamastroviruses, including the seven porcine astroviruses from this study and 21 known astroviruses from swine, bovine, camel and human. Using Recombination Detection Program (RDP4, Version 4.94), potential recombination patterns were screened by RDP, GENECONV, MaxChi, Chimaera, SiScan, 3Seq and Bootscan methods, following Instruction Manual [34], using the step-down correction for multiple comparisons and a p-value cutoff of 0.01. We checked regions of potential recombinant interest using the above methods. We considered recombination events only when an event was involving at least one of our 7 newly identified strains and supported by highest acceptable p-value of 0.05 with all the methods.

2.4.3 Linear antigen epitope prediction: We examined potential linear antigen epitopes on capsid protein of our PoAstV strains using SVMTriP web-based software [35] which predicts linear antigen epitope based on Support Vector Machine. The astrovirus uses capsid gene for attachment and entry into host cells, therefore, accurate prediction of antigenic epitopes in this gene could be useful in the design of immune diagnosis assays and/or subunit vaccines. The predicted epitopes by SVMTrip software was further analyzed by IEDB analysis resource and Immuno-medicine Group tool, web based programs which predicts protein segments that are probable antigenic and elicit antibody response [36, 37]. We examined the antigenicity of predicted candidate epitopes using VaxiJen v2.0 server [38]. This server uses auto cross covariance (ACC) transformation of selected protein sequences based on unique amino acid properties. Each sequence is used to find out 100 known antigen and 100 non-antigens. The identified sequences are tested for antigenicity by leave-one-out cross-validation and overall external validation, with prediction accuracy of up to 89%. Thereafter, we modelled the 3D structure of capsid proteins using Swiss model server which is fully automated and provides full stoichiometry and the whole structure of the complex as inferred by homology modelling [39]. Conformational B-cell epitopes

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2020 doi:10.20944/preprints202009.0559.v1

from the 3D model of our PoAstVs proteins were predicted by ElliPro, a webtool designed by Thornton's method together with MODELLER program of a residue clustering algorithm and Jmol viewer [40].

2.4.4 Glycosylation analysis: Glycosylation is an vital post-translational modification, which influences protein folding, localization and trafficking, protein solubility, antigenicity, biological activity and half-life, as well as cell-cell interactions. We investigated the spread of known and predicted N-glycosylation sites across the capsid protein of the our PoAstV field strains using NetNglyc software [41]. This software uses artificial neural networks that examine the sequence context of Asn-X-Ser/Thr sequons (Asn=Asparagine, Ser=Serine, Thr=Threonine and X=any other amino acid except Proline) and differentiates glycosylated sequences from non-glycosylated ones. The predictions are only shown on Asn-X-Ser/Thr sequences, since only asparagine residues within Asn-X-Ser/Thr (and in some cases, Asn-X-Cys) are N-glycosylated *in-vivo*.

3. Results and discussion

3.1 General features of whole genome sequences of East African porcine astroviruses:

The length of RNA genome sequences of the seven newly identified field strains (U083, K321, U460, K456, K451, K366, K062), excluding the 30 adenines [poly(A) tail)] at the 3' end, varied from 4,281 to 6,649 nt as shown in Table 1. Our strains had a characteristic AstV genomic organization having three ORFs (ORF1a, ORF1b and ORF2), preceded by a 5' untranslated region (UTR) and ending with a 3' UTR, with ribosomal slippage site between ORF1a and ORF1b. The 5' UTR region of K451 and U460 strains was not assembled. They had a frameshift heptamer (AAAAAAC) and a stem-loop structure near the 3' end of ORF1a. The occurrence of heptamer signifies a ribosomal frameshift during translation to create the ORF1ab (replicase polyprotein) [12]. PoAstVs detected in this study showed the conserved tyrosine residue within the TEEEY motif in the viral protein genome-linked (VPg) putative protein at the 3' end of ORF1a (PoAstV3 contained SEEEY). Additionally, a classical YGDD motif was conserved in the middle of the predicted RNA-dependent RNA-polymerase (RdRp) protein of all our strains [12, 42, 43]. All the strains also contained a conserved sequence located at the junction of RdRp and capsid region (UUUGGAGGGG(A/C)GGACCAAA(G/A)$_{8/11}$AUGGC), which is a regulatory element utilized as a promoter for sgRNA transcription [12, 16]. Finally, all our strains contained Trypsin-like peptidase domain in the nonstructural protein 1a and astrovirus capsid protein precursor domain.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2020 doi:10.20944/preprints202009.0559.v1

153 **Table 1:** Comparison of the genomic organization of porcine astroviruses circulating in smallholder
154 swine farms in East Africa (- = region missing)

Genotype	Strains	Accession No.	5'UTR	ORF1a	ORF1ab	ORF2	3'UTR	Total Length	Source
PoAstV2	U083	KY940077	18	2475	4053	2325	162	6434	Samia, Uganda
	K321	KY940076	27	2475	4056	2328	63	6347	Budalangi, Kenya
PoAstV3	U460 (Partial, ORF1a)	KY933399	-	1616	3208	2148	-	5281	Budama, Uganda
PoAstV4	K456	KY933398	85	2550	3995	2511	67	6649	Funyula, Kenya
	K451	KY940075	-	2602	4138	2541	55	6634	Funyula, Kenya
	K366	MT451917	87	2550	3995	2469	76	6618	Funyula, Kenya
	K062	MT451918	87	2550	3995	2481	75	6629	Amukura, Kenya

155

156 **3.2 Genetic diversity and phylogenetic analysis**: Genome sequence comparison revealed that the
157 complete genomes of K456, K451, K366 and K062 strains shared relatively moderate to high nucleotide
158 sequence identities among themselves (64.5 to 88.4%) and with the known sequences for PoAstV4 in
159 the GenBank (63.5 to 80 %), while low identities (41 to 48 %) were noted with other AstV types (Table
160 S1). The U083 and K321 strains had nucleotide sequence identities of 82% among themselves and 66.4
161 to 75.4 % to the known sequences for PoAstV2 in the GenBank, while U460 strain showed nucleotide
162 sequence similarities ranging from 57.0 to 65.5 % with the known sequences for PoAstV3. These data
163 demonstrate a broad genetic divergence among the PoAstV strains circulating in East Africa region, and
164 therefore, we expect significant serological differences among them [13, 44]. Further analysis of the
165 nucleotide and deduced amino acid sequences of the capsid region revealed significant variation among
166 the our strains. Analysis of the nucleotide and deduced amino acid sequences of the capsid region
167 revealed significant variations among East African field strains ranging from 51.7 to 77.8 % and 49 to
168 76.4 %, respectively, to other PoAstVs in the same group (Tables S2 and S3), suggesting that they are
169 novel PoAstV subtype, similar to reports in Japan [45]. We carried out phylogenetic analysis using the
170 nucleotide and/or amino acid sequences of the complete genome and capsid region (ORF2) of the
171 PoAstV reported in this study and those available from GenBank, together with selected AstV sequences
172 from other species to establish genetic relatedness (Figures 1 and 2). In the phylogenetic trees
173 constructed, East African strains clustered with astroviruses of the PoAstV2 (U083 and K321), PoAstV3
174 (U460) and PoAstV4 (K456, K451, K366 and K062) lineages, indicating that they are similar to those
175 strains.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2020 doi:10.20944/preprints202009.0559.v1

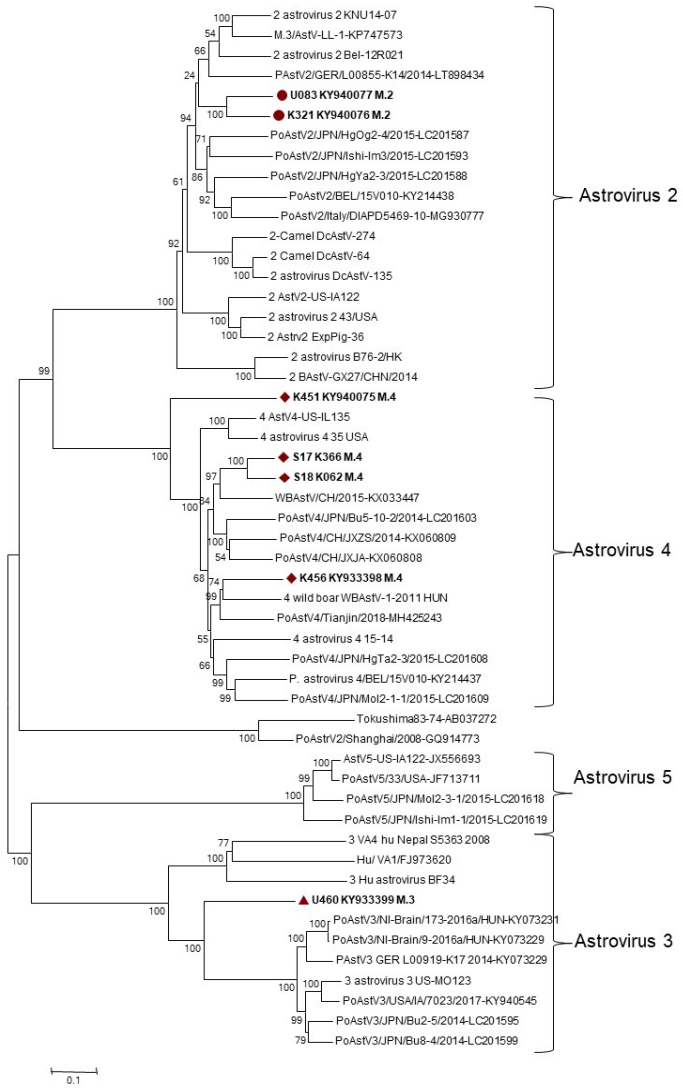


Figure 1. Phylogenetic analysis based on the complete (nearly complete, U460) nucleotide sequences of the East African PoAstVs (**bold**) and known humans and other species AstVs in the GenBank. Multiple sequence alignments were performed using the ClustalO program. The evolutionary history was inferred by using the Maximum Likelihood method based on the General Time Reversible model in MEGA X [1]. The scale bar is given in numbers of substitutions per site. Phylogeny was inferred following 1,000 bootstrap replications, and the node values show percentage bootstrap support

The phylogenetic analysis showed our PoAstV2 were closely related at the nucleotide level, while the PoAstV4 were very diverse, consistent with our comparative sequence analysis results. The sequence identity between our identified novel viruses was mostly greater in the RdRp region than in the NSP1a capsid region (data not shown). Capsid proteins are naturally under intense positive selective pressure from the host immune reaction [46], hence is likely to be more diverse, as shown in this study. According to the ICTV, the capsid protein encoded by the ORF2 is used to distinguish genotypes and species of astroviruses. They described an amino acid sequence diversity in the capsid gene product of <0.312 and >0.378 within and between astrovirus species, respectively. Analysis of evolutionary divergence using the capsid region (ORF2) of our strains (Table S4) showed that PoAstV2 were 0.512 divergent among themselves and 0.418 to 0.579 with the known PoAstV2 strains in the same group.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2020 doi:10.20944/preprints202009.0559.v1

Similarly, PoAstV4 were diverse with 0.309-0.624 and 0.243-0.661 among themselves and with other known PoAstV4 strains, respectively. PoAstV3 was also diverse with 0.499-0.730 with known PoAstV3 strains. Thus, these results further confirm that our strains were novel.

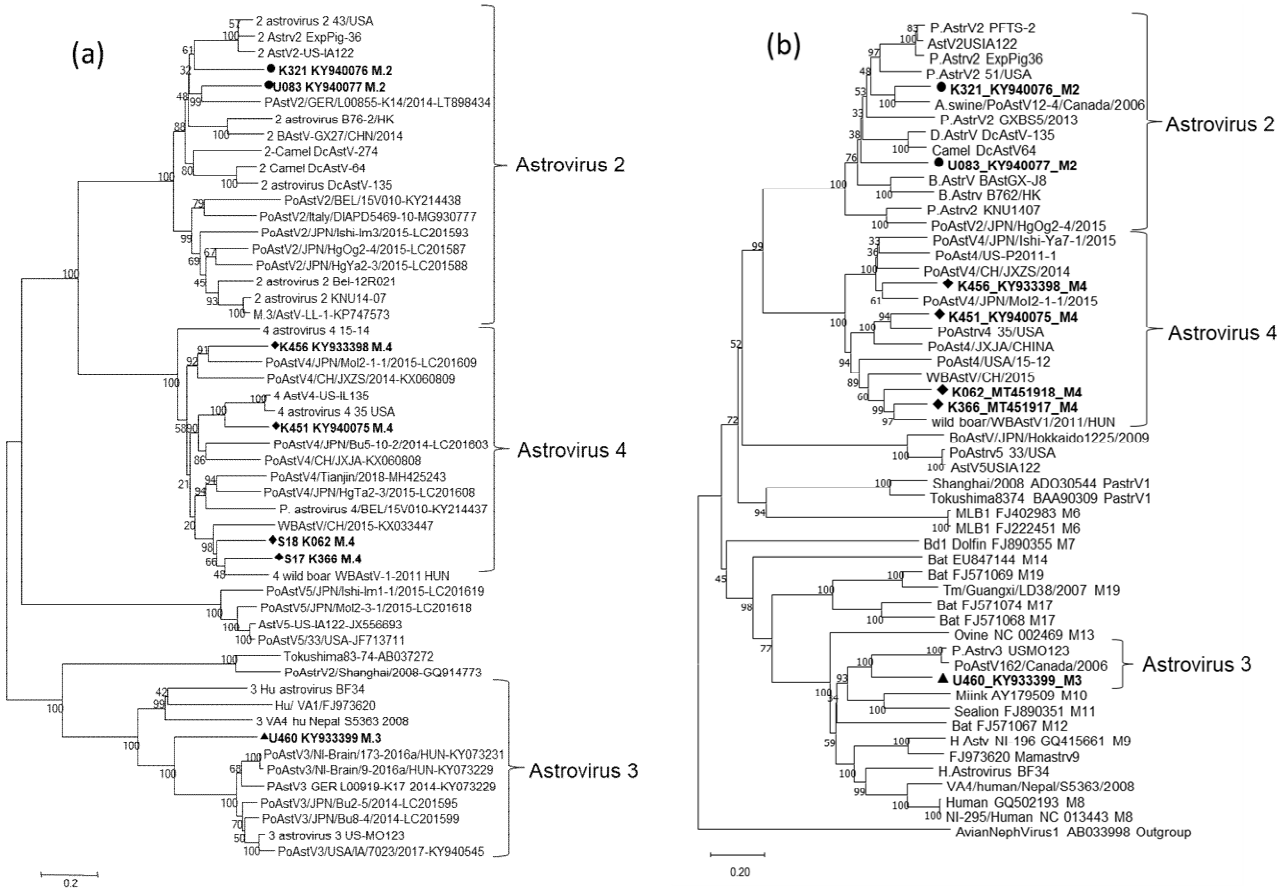


Figure 2. Phylogenetic tree based on the nucleotide (a) and amino acid (b) sequences of the Capsid protein (ORF2) of the East African astrovirus field strains and the known astroviruses in the GenBank. Multiple sequence alignments were performed using the ClustalO program. The evolutionary history was inferred by using the Maximum Likelihood method in MEGA X [1]. The scale bar is given in numbers of substitutions per site. Phylogeny was inferred following 1,000 bootstrap replications, and the node values show percentage bootstrap support

Here we described seven novel and genetically distinct PoAstVs from apparently clinically healthy pig populations in East Africa. The phylogenetic analysis of their full-length genome and capsid gene revealed broader phylogenetic diversity than previously understood for the respective genus. Our data confirmed that pigs are reservoir of genetically diverse PoAstVs presumably derived from distinct ancestors as has been report elsewhere [21].

 doi:10.20944/preprints202009.0559.v1

3.3 Recombination Analysis: Virus recombination has been reported to affect phylogenetic categories, increase the virulence of the virus, confound molecular epidemiological investigations, and have effects in vaccine design [47, 48]. We further analyzed our strains for potential recombination, since our previous study identified multiple genotypes of PoAstV in same pigs and/or same farms [3, 24]. To determine probable recombination events among our strains, we carried out recombination analysis of complete sequences of our seven new strains of astrovirus and selected known AstV strains from the GenBank. Between the four PoAstV4 strains (K456, K451, K366 and K062), three recombination events were recognized, event 1 (2713 -4089nt position), event 2 (4244 -6305nt position), and the last event (event 3) starting at 6346 and ends at 6649nt position in K062, K366 and K456 respectively (Figure 3). Based on the recombination analysis, we concluded that all the four PoAstV4 strains may be of recombinant origin. Event 1, where K062 is recombinant, was predicted at the ORF1a-ORF1b overlap, whereas event 2 (K366 is recombinant) was predicted from the ORF1b-ORF2 junction and covering almost the entire ORF2. The event 3 (K456 recombinant) was predicted at the 3' end of capsid region (ORF2). These recombination patterns were supported by RDP, GENECONV, MaxChi, Chimaera, SiScan, 3Seq, LARD, Phylpro and Bootscan programs as shown in Table 2. The potential recombination event 2 reported in this study supports the previous suggestion that the ORF1b/ORF2 junction region is prone to the recombination region in AstVs [49, 50]. The co-existence of different PoAstVs within swine farms is absolutely possible due to the high prevalence of PoAstV in swine farms [3, 12], and this may promote co-infections of one animal with more than one AstV strain at the same time as observed in this study. Therefore, recombination events between different strains have been observed. As an immune-escape mechanism, virus recombination events may lead to the generation of novel virus strains, to which the affected host may have a lower immunity than to the parent strains. These novel viruses may potentially cross species barriers at some point. Many prior studies have reported evidence for recombination amongst astrovirus types in human, pigs, marine mammals, turkey, small ruminants and dogs [49, 51-54].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2020 doi:10.20944/preprints202009.0559.v1

Table 2. Summary of p-values of different recombination methods using the step-down correction for multiple comparisons between our field strains and a P-value cutoff of 0.01 in the Recombination Detection Program (RDP4, Version 4.94)

Method	Event 1 in K062	Event 2 in K366	Event 3 in K456
RDP	$2.811X10^{-13}$	$1.522X10^{-13}$	$2.744X10^{-16}$
GENECONV	$2.747X10^{-17}$	$1.099X10^{-02}$	$1.816X10^{-03}$
BootScan	$2.199X10^{-15}$	$4.858X10^{-05}$	$1.518X10^{-16}$
MaxChi	$2.377X10^{-21}$	$9.780X10^{-13}$	$4,161X10^{-05}$
Chimaera	$2.805X10^{-13}$	$2.604X10^{-12}$	$5.516X10^{-09}$
Siscan	$4.626X10^{-02}$	$1.434X10^{-07}$	$2.183X10^{-13}$
Phylpro	$4.612X10^{-09}$	$1.165X10^{-14}$	$2.642X10^{-13}$
LARD	$1.219X10^{-71}$	$1.267X10^{-76}$	$2.947X10^{-22}$
3Seq	$4.612X10^{-05}$	$1.165X10^{-14}$	$2.642X10^{-12}$

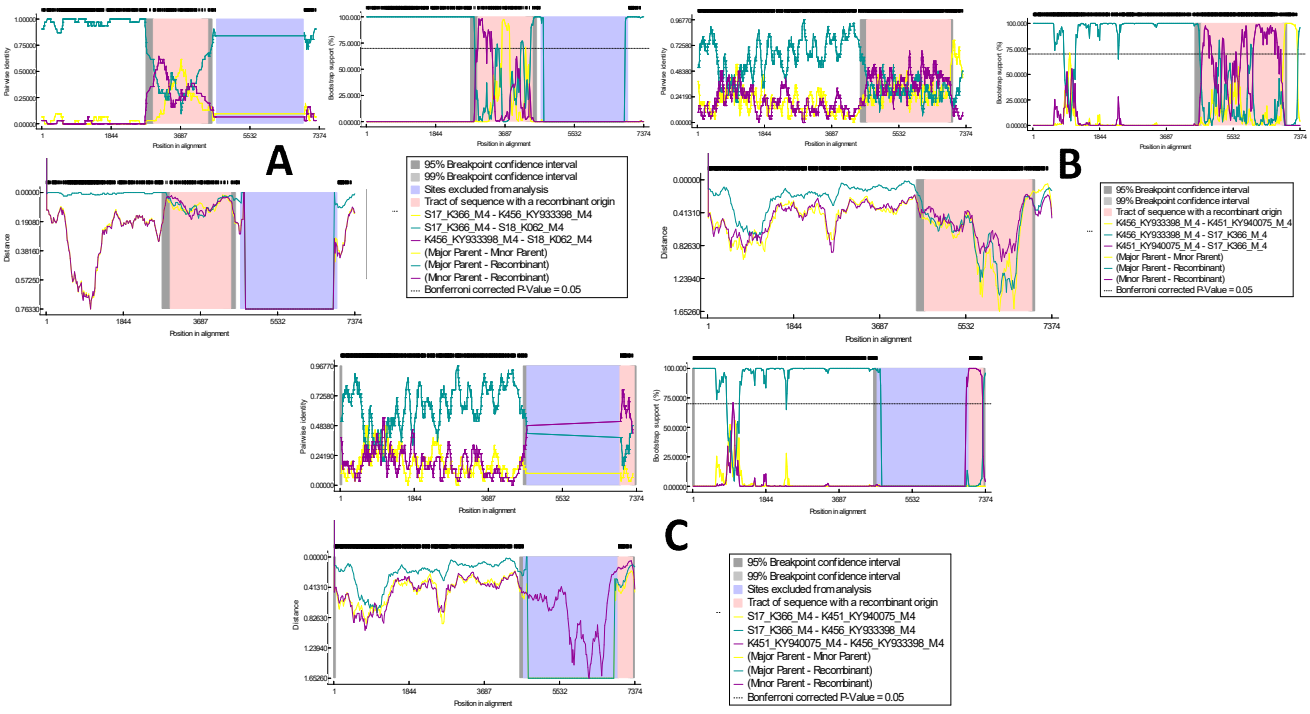


Figure 3. Recombination analysis of complete genomes of newly discovered astrovirus in East Africa: (A) predicted recombination event has likely occurred in the ORF1a-ORF1b junction and covering nearly the entire RdRp region. K062 strain is recombinant. (B) predicted recombination event has likely occurred in RdRp-ORF2 overlap and covering almost entire capsid region (ORF2), K366 is the recombinant. (C)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2020 doi:10.20944/preprints202009.0559.v1

predicted recombination event has occurred in the 3' end of capsid region ORF2, K456 strain is recombinant

3.4 Prediction of potential linear antigenic epitopes: AstVs are difficult to grow in cell culture, consequently evaluation of their antigenic property is difficult. However, since the capsid protein stimulates host immunity, serological property is hypothesized based on sequence similarity of the capsid protein. Since AstVs uses capsid protein for attachment and entry into host cells, we used the region in our strains to predict linear antigenic epitopes to generate information that could aid in the design of subunit vaccines and immunological diagnostic tools. A total of 10 linear antigenic epitopes were detected in the capsid gene products of each of the seven field strains by SVMTrip web-based software. The program recommended antigenic linear epitopes predicted by SVMTrip tool for PoAstV4 strains were all antigenic in nature when analyzed by VaxiJen software, however, some of the recommended predicted epitopes by SVMTrip for PoAstV2 and PoAstV3 strains were non-antigenic in nature when analyzed by VaxiJen (Table S5). Therefore, to narrow down on the potential linear antigenic epitope candidates, which could be used as immunological targets, we further analyzed our sequences using IEDB analysis resource software and immune-medicine group tool which used different algorithms to predict antigenic epitopes. Using threshold value of 0.4, the antigenic property of recognized target sequences from these epitopes was predicted by VaxiJen software. The sequences with values below the threshold value were regarded as non-antigenic in nature while sequences with values above threshold were deemed antigenic. Importantly, using the three analytic software above (SVMTrip, IEDB and Immune-medicine), among all the epitopes, we determine three potential candidate motifs at the surface of the structure and were present at same position on all the capsid proteins of PoAstV2, PoAstV3 and PoAstV4 strains as shown in Table 3. The antigenicity of these predicted epitopes was further analyzed with VaxiJen software with threshold of 0.4, which means the segment greater than the threshold were potentially antigenic in nature. Based on the results of the VaxiJen, we are proposing that the epitope at the amino acid position 126-161 is the best potential candidate epitope since it contained a conserved motif in each genotype (Table 3).

After predicting the potential linear B cell epitope, we constructed a 3D model of capsid protein of all our strains using Swiss model [39], to be able to predict potential conformational B-cell epitopes. We then predicted the conformational B-cell epitopes of our PoAstVs capsid protein models using Ellipro server ([40]. The sequence identity between capsid protein of our PoAstV strains and the selected template (Q82452-Human AstV 1 strain) ranged from 39.45% – 42.86%, a value greater than the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2020 doi:10.20944/preprints202009.0559.v1

273 required 30% sequence similarity for creating suitable models [55]. Our potential candidate epitopes
274 predicted were identified in the capsid protein model by Ellipro software and visualized in Jmol to show
275 their 3D structures and comparative orientation of protein and peptide molecule (Table 4). The amino
276 acid positions of each epitope were also verified by Jmol viewer. The peptide on these sites was
277 predicted to be highly antigenic (>1) that could be considered for effective vaccination or immune
278 diagnosis. Furthermore, all the selected epitopes were on surface of the capsid protein structure, hence,
279 suitable as the immunological targets for diagnosis of PoAstVs in the study region.
280
281 **Table 3**. Predicted antigenic epitopes within capsid protein (ORF2) of our field strains using three
282 different software and antigenicity of the predicted epitopes determined

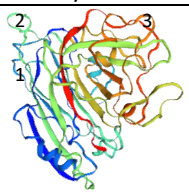
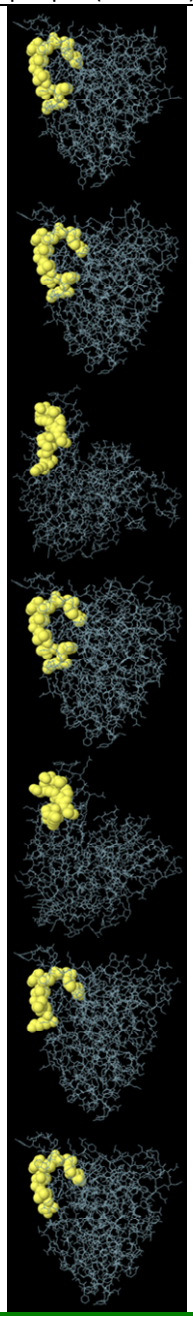
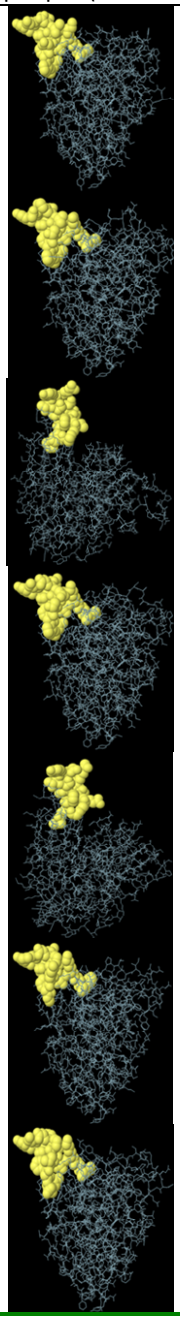
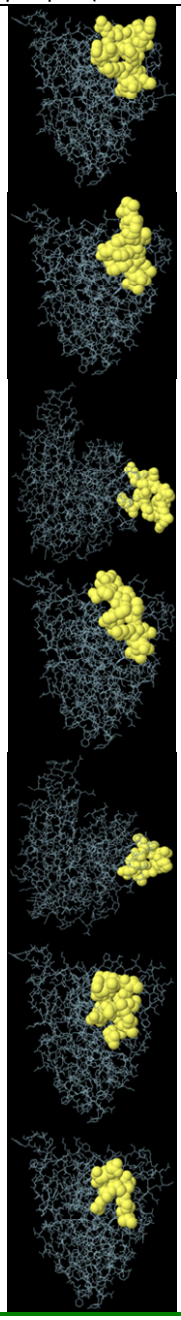
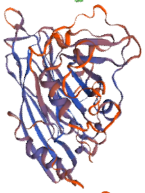
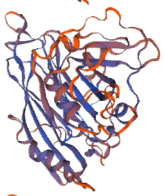
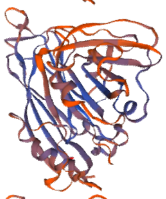
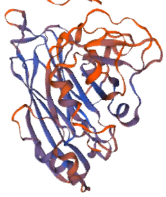
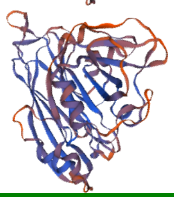
Strains	Amino acid position 126-161	Amino acid position 219-241	Amino acid position 332 – 363
U083	FKMTKCELVLKPLVGDSAVSGTVV**RASWNPTA**T	IGKTMSTYQSRAFEGGLFLAELTT	**RAANAPVR**TGETTFDIYASISDARSDSPCVST
K321	YKMTRCVVTLKPIVGDSAVAGTVT**RVSWNPTS**S	CHTFGKTTSTYRNEPFKGGLFLAE	VK**RAAGAPVR**ANDNEIRFDIYASISDARSNTP
U460	WRLTNLK**IKCTPLVG**PSAVTGSVYRVSLNLTQS	MIEIHGLGKTSSTYKDEPWVGDLF	PFQWLIKG**GWWFVKKA**LGRSMNSDEVYYVYAS
K456	WRVQYLD**IKLTPLVG**ASAVSGTVIRTSLNLAAQ	TLGKTMSTYKSDIFDGPLFLAEVT	QWLIKA**GWWFLKR**IANKKKSGDHIDGQPDANE
K451	WRVDNIL**IKLTPLVG**ASAVSGTAVRVSLNNAAT	TLGQTMSTYQAKVFTGPLFLCEMT	LFQA**GWWFVKR**IANKKKVGGSVDGEPDPGEVT
K366	WRVKNMII**KLTPLVG**GSAVSGTAVRTSLNLSGQ	TYGKTVSTYRNDPFTGPLFLAELT	LFKA**GWWFVKK**IANKSQNRNRPGEPDPGELTF
K062	WRARDII**VKLTPLVG**GSAVSGTAIRTSLNLSAQ	TLGKTLSTYKNEDFTGPLFLAELT	LFKA**GWWFVKK**IANKKTSGNAPGEPAPGELTF

		SVMTriP web-based tool		immune-medicine group tool		IEDB: Ellipro software		VaxiJen software	
Strains		Position	Score	Position	Score	Position	Score	Position	Antigenicity*
U083	PoAstV2	127 - 146	1.000	129-151	1.215	136-148	0.804	126-158	0.6357
K321		130 - 149	0.777	130-153	1.200	139-151	0.810	129-161	0.7207
U460	PoAstV3	-	-	135-161	1.150	142-154	0.809	132-164	1.0511
K456	PoAstV4	127 - 146	1.000	125-159	1.150	135-149	0.765	127-159	1.1882
K451		128 - 147	0.994	129-154	1.150	135-149	0.765	127-159	0.7063
K366		128 - 147	0.836	134-156	1.150	135-149	0.751	127-159	0.7977
Ko62		128 - 147	0.836	130-147	1.175	135-149	0.767	127-159	0.9112

283 *Antigenicity of the predicted epitope was analyzed by VaxiJen Software with threshold of 0.4 (values
284 above threshold are probable antigens); **Bold** colored amino acid sequence represent a motif conserved
285 in each PoAstV types for position 126-161 and 332-363, with high antigenicity (>1)
286
287

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2020 doi:10.20944/preprints202009.0559.v1

288 **Table 4.** Capsid protein structures and B-cell epitopes conformation of PoAstV protein predicted by
289 Ellipro. Yellow balls are the residues of predicted peptides and white sticks are the non-epitope residues
290 of protein. Each epitope is predicted with residue number and position mentioned

Strains	Capsid protein tertiary structure	Conformational B-cell epitopes (126-161)	Conformational B-cell epitopes (219-241)	Conformational B-cell epitopes (332-363)
K456 PoAstV4				
K451 PoAstV4				
K366 PoAstrV4				
K062 PoAstV4				
U083 PoAstV2				
K321 PoAstV2				
U460 PoAstV3				

291
292

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2020 doi:10.20944/preprints202009.0559.v1

3.5 Glycosylation analysis: Glycosylation is normally required for progeny formation and infectivity of many viruses [56]. High levels of glycosylation serve as a protective shield from the host's immune system, where during virus entry into host cells, host cell glycans are viral receptors interacting with carbohydrate binding proteins on the viral surface [57, 58]. We observed glycosylation on the capsid protein of PoAstV2, PoAstV3 and PoAstV4 with higher glycosylation sites in PoAstV2 compared to PoAstV3 and PoAstV4 (Table 5). Studies have shown that, N- and O-linked glycans shield immunodominant epitopes from immune recognition [57-59]. Our analysis of the predicted antigenic epitopes at position 126-161 common between PoAstV2, PoAstV3 and PoAstV4 showed that they had at least one glycosylation sites. Therefore, in-depth studies of the glycosylation in AstVs would be an important step in designing suitable antigens for diagnostic tools and vaccine development. Based on these results, we suggest that any approach that is based on inhibition of the host mechanism to glycosylate astrovirus proteins may offer the best potential approach to develop therapeutics to astrovirus infection.

Table 5. Glycosylation analysis of the capsid proteins (ORF2) of astrovirus field strains using NetNGlyc software

Name	Position	Sequence	Potential†	Jury agreement	N Glycosylated
U083	12	NTTN	0.734	(9/9)	++
	20	NGSS	0.513	(5/9)	+
	55	NKTV	0.764	(9/9)	+++
	86	NGSE	0.691	(9/9)	++
	154	NPTA	0.697	(9/9)	++*
	297	NKTI	0.752	(9/9)	+++
	439	NYTT	0.648	(9/9)	++
	542	NGTG	0.735	(9/9)	++
	557	NRTN	0.618	(7/9)	+
	611	NNTM	0.406	(8/9)	-
K321	13	NTTN	0.746	(9/9)	++
	21	NGSS	0.505	(6/9)	+
	41	NRTR	0.749	(9/9)	++
	56	NQSQ	0.560	(6/9)	+
	80	NTTL	0.625	(9/9)	++
	89	NESG	0.557	(6/9)	+
	157	NPTS	0.714	(9/9)	++*
	412	NPTR	0.729	(9/9)	++*
	457	NGTK	0.685	(9/9)	++
	493	NNTT	0.547	(6/9)	+
	494	NTTA	0.609	(7/9)	+
	511	NESP	0.129	(9/9)	---
	560	NNSN	0.377	(9/9)	--
U460	160	NLTQ	0.750	(9/9)	+++
	306	NATT	0.597	(6/9)	+
	625	NYTF	0.646	(8/9)	+
K451	274	NATP	0.110	(9/9)	---
	398	NITQ	0.681	(9/9)	++
	573	NYTM	0.731	(9/9)	++
	658	NTTP	0.110	(9/9)	---
K456	251	NPTP	0.271	(8/9)	--
	391	NITG	0.631	(9/9)	++
	521	NPTL	0.643	(8/9)	+*
	601	NGTL	0.698	(9/9)	++
	655	NLTA	0.644	(9/9)	++
K366	123	NYSL	0.738	(9/9)	++
	155	NLSG	0.670	(9/9)	++
	287	NGSL	0.573	(7/9)	+
	345	NKSQ	0.677	(8/9)	+
	502	NYTP	0.197	(9/9)	---
	559	NPTR	0.580	(8/9)	+*
	565	NFTQ	0.584	(7/9)	+
K062	123	NYSL	0.709	(9/9)	++
	155	NLSA	0.627	(8/9)	+
	287	NSSS	0.497	(4/9)	-
	615	NQTV	0.607	(7/9)	+

†=Any potential crossing the default threshold of 0.5 (predicted glycosylated site); +=N glycosylated, -=a negative site; * = Proline occurs just after the Asparagine residue (unlikely to be glycosylated); the **jury agreement** column indicates how many of the nine networks support the prediction; N=Asparagine, S=Serine, T=Threonine. For picking up N-glycosylation sites with **high specificity** (Asparagine residues very likely to be glycosylated), use only (++) predictions (and better) for Asparagines that occur within the Asn-X-Ser/Thr triplet (no Proline at the X position)

 doi:10.20944/preprints202009.0559.v1

4. Conclusions

To our knowledge, we report for the first time a detailed genetic characterization of whole genome of three PoAstV strains from African region. Our findings give intuitions into the epidemiology and evolution of PoAstV in the region and would facilitate investigations on the genetic diversity of PoAstV globally. The finding of these novel PoAstV strains provides reveals how diverse AstVs are in smallholder pig population in the study region where there is close contact between pigs and human. Our data taken together would be helpful in the development of PoAstV immunodiagnostic tools and vaccines for African region. However, the high genetic diversity among PoAstV strains (especially diversity at predicted potential antigenic epitopes) and high levels of glycosylation reported in the PoAstVs reported here may pose difficulties for development of virus detection methods, subunit vaccines, as wel as implications on epidemiological investigations. Importantly, this study identified three potential linear antigenic epitopes occurring at the surface of the capsid protein structure, which could be used as the immunological targets for the proper diagnosis and treatment of PoAstVs in the study region. The generated information can be used to test the predicted function of these peptides by conducting *in vitro* and *in vivo* experiments to confirm the immunogenicity and ultimately the vaccine properties to prevent PoAstV infections. Finally, understanding the genetic differences of these novel PoAstV variants, that may emerge locally or globally through genetic drift or shift, in wild and domestic animals could lead to early identification of the source of an emerging outbreak leading to faster and more targeted interventions to control and/or limit the spread of such outbreaks.

Data availability: The sequences of the strains discussed on this study has been deposited in GenBank under accession numbers KY940075-KY940077, KY933398, KY933399, MT451917 and MT451918. Additional data are presented in the supplementary files.

Supplementary Materials:

Table S1: Pairwise comparison of nucleotide sequence identities of the complete (near complete, U460) genomes of the seven (7) astrovirus field strains (**bold**) and with sequences of other astroviruses available in GenBank

Table S2. Summary of nucleotide sequence identity matrix of the capsid protein (ORF2) among the seven (7) astroviruses field strains (**bold**) and the known reference strains in the GenBank using Clustal Omega

Table S3. Summary of amino acid sequence identity matrix of the capsid protein (ORF2) among the 7 astroviruses field strains (**bold**) and the known reference strains in the GenBank using Clustal Omega

Table S4: Estimates of evolutionary divergence between the East African PoAstVs and selected known AstV in the GenBank based on the amino acid sequences of complete ORF2 protein. The number of amino acid differences per site from between sequences is shown. Standard error estimate(s) are shown above the diagonal for our strains.

Table S5. Recommended potential linear antigenic epitopes predicted inside capsid protein (ORF2) of our field strains by SVMTriP web-based tool and corresponding antigenicity predicted by VaxiJen software

Author Contributions: Conceptualize and design the work: JOA, EAO. Methodology: JAO, EMM. Analysis and interpretation of data: JOA, ANV and RP. Write and review the paper: JOA, EMM, EAO, ANV and RP. All the authors read and approved the final version of the paper.

Conflict of interest: The authors declare no conflict of interest.

Acknowledgement BecA-ILRI Hub supported this research through the Africa Biosciences Challenge Fund (ABCF) program. The ABCF Program is funded by the Australian Department for Foreign Affairs and Trade (DFAT) through the BecA-CSIRO partnership; the Syngenta Foundation for Sustainable Agriculture (SFSA); the Bill & Melinda Gates Foundation (BMGF); the UK Department for International Development (DFID) and; the Swedish International Development Cooperation Agency (Sida). We thank smallholder pig farmers in the study region for allowing us to sample their pigs for this study.

Reference

1. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K., MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **2018,** 35, (6), 1547-1549.
2. Bosch, A.; Guix, S.; Krishna, N. K.; Mendez, E.; Monroe, S. S.; Pantin-Jackwood, M.; Schultz-Cherry, S., Family Astroviridae. In *Virus Taxonomy: Classification and Nomenclature of Viruses (Ninth Report of the International Committee on the Taxonomy of Viruses)*, King, A. M. Q.; Lefkowitz, E.; Adams, M. J.; Carstens, E. B., Eds. Elsevier Academic Press: New York, 2011; pp 953-959.
3. Amimo, J. O.; El Zowalaty, M. E.; Githae, D.; Wamalwa, M.; Djikeng, A.; Nasrallah, G. K., Metagenomic analysis demonstrates the diversity of the fecal virome in asymptomatic pigs in East Africa. *Archives of virology* **2016,** 161, (4), 887-897.
4. Chu, D. K.; Poon, L. L.; Guan, Y.; Peiris, J. S., Novel astroviruses in insectivorous bats. *Journal of virology* **2008,** 82, (18), 9107-14.
5. Tse, H.; Chan, W. M.; Tsoi, H. W.; Fan, R. Y.; Lau, C. C.; Lau, S. K.; Woo, P. C.; Yuen, K. Y., Rediscovery and genomic characterization of bovine astroviruses. *JOURNAL OF GENERAL VIROLOGY* **2011,** 92, (Pt 8), 1888-98.

6. Glass, R. I.; Noel, J.; Mitchell, D.; Herrmann, J. E.; Blacklow, N. R.; Pickering, L. K.; Dennehy, P.; Ruiz-Palacios, G.; de Guerrero, M. L.; Monroe, S. S., The changing epidemiology of astrovirus-associated gastroenteritis: a review. *Archives of Virology Supplement* **1996,** 12, 287-300.
7. Meleg, E.; Jakab, F.; Kocsis, B.; Banyai, K.; Melegh, B.; Szucs, G., Human astroviruses in raw sewage samples in Hungary. *Journal of Applied Microbiology* **2006,** 101, (5), 1123-9.
8. Taylor, M. B.; Cox, N.; Vrey, M. A.; Grabow, W. O., The occurrence of hepatitis A and astroviruses in selected river and dam waters in South Africa. *Water Res* **2001,** 35, (11), 2653-60.
9. Madeley, C. R.; Cosgrove, B. P., Letter: 28 nm particles in faeces in infantile gastroenteritis. *Lancet* **1975,** 2, 451-452.
10. Espinosa, A. C.; Mazari-Hiriart, M.; Espinosa, R.; Maruri-Avidal, L.; Mendez, E.; Arias, C. F., Infectivity and genome persistence of rotavirus and astrovirus in groundwater and surface water. *Water Research* **2008,** 42, (10-11), 2618-28.
11. Kiulia, N. M.; Mwenda, J. M.; Nyachieo, A.; Nyaundi, J. K.; Steele, A. D.; Taylor, M. B., Astrovirus infection in young Kenyan children with diarrhoea. *Journal of Tropical Pedriatics* **2007,** 53, (3), 206-9.
12. Xiao, C.-T.; Giménez-Lirola, L. G.; Gerber, P. F.; Jiang, Y.-H.; Halbur, P. G.; Opriessnig, T., Identification and characterization of novel porcine astroviruses (PAstVs) with high prevalence and frequent co-infection of individual pigs with multiple PAstV types. *Journal of General Virology* **2013,** 94, 570-82.
13. De Benedictis, P.; Schultz-Cherry, S.; Burnham, A.; Cattoli, G., Astrovirus infections in humans and animals - molecular biology, genetic diversity, and interspecies transmissions. *Infection Genetics and Evolution* **2011,** 11, (7), 1529-44.
14. King, A. M. Q.; Adams, M. J.; Lefkowitz, E. J., *Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses*. Elsevier, London, UK.: 2011; Vol. 9.
15. Bosch, A.; Pintó, R. M.; Guix, S., Human Astroviruses. *Clinical Microbiology Reviews* **2014,** 27, (4), 1048-1074.
16. Mendez, E.; Arias, C. F., Astroviruses. In *Fields's virology*, 6th ed.; Knipe, D. M.; Howley, P. M., Eds. Philadelphia, PA: Lippincott Williams and Wilkins, 2013; pp 609–628
17. Bridger, J. C., Detection by electron microscopy of caliciviruses, astroviruses and rotavirus-like particles in the faeces of piglets with diarrhoea. *Veterinary Record* **1980,** 107, 532-533.
18. Brnic, D.; Jemersic, L.; Keros, T.; Prpic, J., High prevalence and genetic heterogeneity of porcine astroviruses in domestic pigs. *Vet J* **2014,** 202, (2), 390-2.
19. Lan, D.; Ji, W.; Shan, T.; Cui, L.; Yang, Z.; Yuan, C.; Hua, X., Molecular characterization of a porcine astrovirus strain in China. *Archives of virology* **2011,** 156, 1869-75.
20. Laurin, M. A.; Dastor, M.; L'Homme, Y., Detection and genetic characterization of a novel pig astrovirus: relationship to other astroviruses. *Archive of Virology* **2011,** 156, 2095-2099.
21. Luo, Z.; Roi, S. p.; Dastor, M.; Gallice, E.; Laurin, M.-A.; Lâ€™Homme, Y., Multiple novel and prevalent astroviruses in pigs. *Veterinary Microbiology* **2011,** 149, (3), 316-323.
22. Geyer, A.; Steele, A. D.; Peenze, I.; Lecatsas, G., Astrovirus-like particles, adenoviruses and rotaviruses associated with diarrhoea in piglets. *Journal of the South Africa Vetetrinary Association* **1994,** 65, (4), 164-6.
23. Shimizu, M.; Shirai, J.; Narita, M.; Yamane, T., Cytopathic astrovirus isolated from porcine acute gastroenteritis in an established cell line derived from porcine embryonic kidney. *Journal of clinical microbiology* **1990,** 28, 201-206.
24. Amimo, J. O.; Okoth, E.; Junga, J. O.; Ogara, W. O.; Njahira, M. N.; Wang, Q.; Vlasova, A. N.; Saif, L. J.; Djikeng, A., Molecular detection and genetic characterization of kobuviruses and astroviruses in asymptomatic local pigs in East Africa. *Archives of virology* **2014,** 159, (6), 1313-9.

25. Mor, S. K.; Chander, Y.; Marthaler, D.; Patnayak, D. P.; Goyal, S. M., Detection and molecular characterization of Porcine astrovirus strains associated with swine diarrhea. *Journal of veterinary diagnostic investigation* **2012,** 24, (6), 1064-7.
26. Reuter, G.; Nemes, C.; Boros, A.; Kapusinszky, B.; Delwart, E.; Pankovics, P., Astrovirus in wild boars (Sus scrofa) in Hungary. *Archives of virology* **2012,** 157, 1143-7.
27. Reuter, G.; Pankovics, P.; Boros, A., Identification of a novel astrovirus in a domestic pig in Hungary. *Archives of virology* **2011,** 156, 125-8.
28. Shan, T.; Li, L.; Simmonds, P.; Wang, C.; Moeser, A.; Delwart, E., The Fecal Virome of Pigs on a High-Density Farm. *Journal of virology* **2011,** 85, (22), 11697-11708.
29. Cortez, V.; Freiden, P.; Gu, Z.; Adderson, E.; Hayden, R.; Schultz-Cherry, S., Persistent Infections with Diverse Co-Circulating Astroviruses in Pediatric Oncology Patients, Memphis, Tennessee, USA. *Emerging infectious diseases* **2017,** 23, (2), 288-290.
30. Janowski, A. B.; Bauer, I. K.; Holtz, L. R.; Wang, D., Propagation of Astrovirus VA1, a Neurotropic Human Astrovirus, in Cell Culture. *Journal of virology* **2017,** 91, (19).
31. Amimo, J. O.; Junga, J. O.; Ogara, W. O.; Vlasova, A. N.; Njahira, M. N.; Maina, S.; Okoth, E. A.; Bishop, R. P.; Saif, L. J.; Djikeng, A., Detection and genetic characterization of porcine group A rotaviruses in asymptomatic pigs in smallholder farms in East Africa: Predominance of P[8] genotype resembling human strains. *Veterinary Microbiology* **2015,** 175, (2-4), 195-210.
32. Roux, S.; Tournayre, J.; Mahul, A.; Debroas, D.; Enault, F., Metavir 2: new tools for viral metagenome comparison and assembled virome analysis. *BMC Bioinformatics* **2014,** 15, (1), 76.
33. Saitou, N.; Nei, M., The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **1987,** 4, 406-425.
34. Martin, D. P.; Murrell, B.; Golden, M.; Khoosal, A.; Muhire, B., RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol* **2015,** 1, (1), vev003.
35. Yao, B.; Zhang, L.; Liang, S.; Zhang, C., SVMTriP: a method to predict antigenic epitopes using support vector machine to integrate tri-peptide similarity and propensity. *PLoS One* **2012,** 7, (9), e45152.
36. Ingale, S.; Gach, J. S.; Zwick, M. B.; Dawson, P. E., Synthesis and analysis of the membrane proximal external region epitopes of HIV-1. *J Pept Sci* **2010,** 16, (12), 716-22.
37. Kolaskar, A. S.; Tongaonkar, P. C., A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* **1990,** 276, (1-2), 172-4.
38. Doytchinova, I. A.; Flower, D. R., VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* **2007,** 8, 4.
39. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research* **2018,** 46, (W1), W296-W303.
40. Ponomarenko, J.; Bui, H. H.; Li, W.; Fusseder, N.; Bourne, P. E.; Sette, A.; Peters, B., ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics* **2008,** 9, 514.
41. Gupta, R.; Jung, E.; Brunak, S., Prediction of N-glycosylation sites in human proteins. *In Preparation* **2004**.
42. Kamer, G.; Argos, P., Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic acids research* **1984,** 12, (18), 7269-82.
43. Poch, O.; Sauvaget, I.; Delarue, M.; Tordo, N., Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *The EMBO journal* **1989,** 8, (12), 3867-74.
44. Walter, J. E.; Briggs, J.; Guerrero, M. L.; Matson, D. O.; Pickering, L. K.; Ruiz-Palacios, G.; Berke, T.; Mitchell, D. K., Molecular characterization of a novel recombinant strain of human astrovirus associated with gastroenteritis in children. *Archives of virology* **2001,** 146, (12), 2357-67.

45. Ito, M.; Kuroda, M.; Masuda, T.; Akagami, M.; Haga, K.; Tsuchiaka, S.; Kishimoto, M.; Naoi, Y.; Sano, K.; Omatsu, T.; Katayama, Y.; Oba, M.; Aoki, H.; Ichimaru, T.; Mukono, I.; Ouchi, Y.; Yamasato, H.; Shirai, J.; Katayama, K.; Mizutani, T.; Nagai, M., Whole genome analysis of porcine astroviruses detected in Japanese pigs reveals genetic diversity and possible intra-genotypic recombination. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **2017,** 50, 38-48.
46. van Hemert, F. J.; Berkhout, B.; Lukashov, V. V., Host-related nucleotide composition and codon usage as driving forces in the recent evolution of the Astroviridae. *Virology* **2007,** 361, (2), 447-54.
47. Bull, R. A.; Tanaka, M. M.; White, P. A., Norovirus recombination. *The Journal of general virology* **2007,** 88, (Pt 12), 3347-59.
48. Medici, M. C.; Tummolo, F.; Martella, V.; Banyai, K.; Bonerba, E.; Chezzi, C.; Arcangeletti, M. C.; De Conto, F.; Calderaro, A., Genetic heterogeneity and recombination in type-3 human astroviruses. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **2015,** 32, 156-60.
49. Li, M.; Yan, N.; Ji, C.; Wang, M.; Zhang, B.; Yue, H.; Tang, C., Prevalence and genome characteristics of canine astrovirus in southwest China. *The Journal of general virology* **2018,** 99, (7), 880-889.
50. Martella, V.; Medici, M. C.; Terio, V.; Catella, C.; Bozzo, G.; Tummolo, F.; Calderaro, A.; Bonura, F.; Di Franco, M.; Banyai, K.; Giammanco, G. M.; De Grazia, S., Lineage diversification and recombination in type-4 human astroviruses. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **2013,** 20, 330-5.
51. Kauer, R. V.; Koch, M. C.; Hierweger, M. M.; Werder, S.; Boujon, C. L.; Seuberlich, T., Discovery of novel astrovirus genotype species in small ruminants. *PeerJ* **2019,** 7, e7338.
52. Pantin-Jackwood, M. J.; Spackman, E.; Woolcock, P. R., Phylogenetic analysis of Turkey astroviruses reveals evidence of recombination. *Virus Genes* **2006,** 32, (2), 187-92.
53. Ulloa, J. C.; Gutiérrez, M. F., Genomic analysis of two ORF2 segments of new porcine astrovirus isolates and their close relationship with human astroviruses. *Canadian Journal of Microbiology* **2010,** 56, 569-577.
54. Wolfaardt, M.; Kiulia, N. M.; Mwenda, J. M.; Taylor, M. B., Evidence of a recombinant wild-type human astrovirus strain from a Kenyan child with gastroenteritis. *Journal of clinical microbiology* **2011,** 49, (2), 728-31.
55. Xiang, Z., Advances in homology protein structure modeling. *Current protein & peptide science* **2006,** 7, (3), 217-27.
56. Mossenta, M.; Marchese, S.; Poggianella, M.; Slon Campos, J. L.; Burrone, O. R., Role of N-glycosylation on Zika virus E protein secretion, viral assembly and infectivity. *Biochem Biophys Res Commun* **2017,** 492, (4), 579-586.
57. Behrens, A. J.; Vasiljevic, S.; Pritchard, L. K.; Harvey, D. J.; Andev, R. S.; Krumm, S. A.; Struwe, W. B.; Cupo, A.; Kumar, A.; Zitzmann, N.; Seabright, G. E.; Kramer, H. B.; Spencer, D. I.; Royle, L.; Lee, J. H.; Klasse, P. J.; Burton, D. R.; Wilson, I. A.; Ward, A. B.; Sanders, R. W.; Moore, J. P.; Doores, K. J.; Crispin, M., Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein. *Cell Rep* **2016,** 14, (11), 2695-706.
58. Walls, A. C.; Tortorici, M. A.; Frenz, B.; Snijder, J.; Li, W.; Rey, F. A.; DiMaio, F.; Bosch, B. J.; Veesler, D., Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol* **2016,** 23, (10), 899-905.
59. Gram, G. J.; Hemming, A.; Bolmstedt, A.; Jansson, B.; Olofsson, S.; Akerblom, L.; Nielsen, J. O.; Hansen, J. E., Identification of an N-linked glycan in the V1-loop of HIV-1 gp120 influencing

neutralization by anti-V3 antibodies and soluble CD4. *Archives of virology* **1994,** 139, (3-4), 253-61.