Molecular Characterization of Human Enteroviruses Detected in Children Under Five Years Old in Kenya 2009 - 2015.

* Mwasi S. Lydia[1], Wallace D. Bulimo[3], Jennifer R. Verani [2], Omballa Victora[1], Alice Ouma [1] Samuel Kiplangat [1], Silvanos Opanda[3], Imbuga Mabel [4], Bonventure Juma [2], Bigogo Godfrey [1], Elizabeth Hunsperger [2].

1. *Kenya Medical Research Institute (CGHR)*
2. *Centers for Disease Control and Prevention - Kenya (CDC-Kenya), Center for Global Health, Division of Global Health Protection*
3 *USA Medical Research Directorate - Kenya (USAMRD-K)*
4. *Jomo Kenyatta University of Science and Technology (JKUAT)*

Correspondence Author: Lydia Mwasi, Email: Lmwasi@kemricdc.org

Summary

INTRODUCTION

Human enterovirus (HEVs) infection is common, with an extensive array of clinical displays ranging from asymptomatic to life-threatening. Presentation include nonspecific febrile illness often accompanied by muscle pain, sore throat, abdominal discomfort, rash, headache, encephalitis, aseptic meningitis and acute flaccid paralysis [2].

OBJECTIVES

The study objective was to investigate the natural selection and genetic variability of HEVs and to identify HEV serotypes in circulation among children below 5 years old with diarrhea in an informal settlement(Kibera) in Kenya.

METHODOLOGY

Specimens (n=628) from a prospective cohort study assessing the incidence and *etiology* of diarrhea from 2009-2015 were analyzed. Enteric Taqman array cards (TAC) were used for initial screening where two hundred and nine (78%) tested positive for HEVs.

Of these specimens, 72 (42%) had a cycle threshold (Ct) ≤30 and were tested by conventional PCR targeting the 3' regions of the viral protein 1 (VP1) gene. A total of 48 (67%) underwent sequencing; 11 (23%) of which yielded *nucleotide* sequences. *Phylogenetic* analyses clustered the Kenyan serotypes to HEVs groups C, B and A. Evaluation of the VP1 amino acid sequences revealed numerous amino acid substitutions in relation to reference strains, which were confirmed to be due to natural selection by negative or positive selection.

CONCLUSION

The *Heterogeneous* nature of stool samples is known to influence disparities in viral nucleic acid yields. TAC detected 209 of which 171 (82%) were confirmed positive for HEVs by real-time reverse transcription polymerase chain reaction (RRT-PCR), targeting the 5' NTR regions. Therefore, the results may not be a representative of all circulating HEVs in the study area. Since this was a retrospective study of previously collected samples, it is possible that some HEVs strains may have failed to amplify.

[***Afr. J.*** Health Sci. 2019 ; 32(2) : 15 - 33]

Introduction

Human enteroviruses (HEVs) belong to the *Picornaviridae* family and are associated with a wide spectrum of clinical disease ranging from asymptomatic to severe life threatening [1].

Clinical presentations include nonspecific febrile illness often accompanied by muscle pain, sore throat, abdominal discomfort, rash, headache, encephalitis, aseptic meningitis and acute flaccid paralysis [2].

Most infections are mild or asymptomatic, although some infections may progress to severe diseases. Particularly in immuno - compromised individuals and children under five years this is possible[3, 4, 5].

The *Picornaviruses* are small in size (22 to 30 *nm)* non-enveloped, single-stranded, positive-sense RNA viruses with a *genome* of about 7,500 *nucleotides* and an icosahedral symmetry [6].

The virus capsid consists of four *polypeptides viral* proteins (VP1-4). HEVs serotypes are defined by neutralizing *epitopes* of the virus capsid proteins, predominantly VP1 as well as VP2 and VP3 [7].

The VP1 protein is the outermost viral capsid protein and contains key neutralization sites thus an ideal target for vaccine development [2, 7].

Besides being a major neutralization site, several other determinants are mapped to the VP1 including determinants of virulence, virion thermostability, host range alteration and obstinate infection [8].

The VP1 partial *genome* sequence is sufficient to identify all serotypes with routine use in enterovirus *serotype* characterization;

(a.) A nucleotide identity of $\geq 75\%$

(b.) Amino acid identity of $\geq 85\%$ is considered for serotype determination [2, 9, 10].

More than 100 *serotypes* of non - polio enteroviruses are currently recognized worldwide with additional antigenic variants within serotypes defined by reduced nonreciprocal cross - neutralization between prototype and variant strains [7].

HEVs have a high rate of mutation and recombination because they co - circulate in human and animal hosts, rapidly evolving to create new variants [9, 10].

Natural selection due to both negative and positive pressures has been known to shape patterns of *nucleotide polymorphism* as observed herein where *nucleotide* substitutions were confirmed to be as a result of negative and positive selection [7].

Whereas negative or purifying selection deletes mutations that are disadvantageous to the virus, positive selection increases the frequency of advantageous mutations. These two forces contribute to the fitness of the virus within its environment increasing its chances of survival and dominance [7, 9, 10].

Following the infection, HEVs are transmitted via respiratory droplets, feces, and in some cases via conjunctival secretions and skin lesion exudates [8, 9].

Detection of HEVs from the gastrointestinal tract provides evidence of infection, although viral shedding at these sites may also happen in the absence of symptoms [5].

Stool specimens of infected HEVs individuals have higher viral titers with persistent viral shedding for several weeks following the onset of infection [6, 11].

This study was to determine the serotypes of HEVs circulating in Kibera.

During the study period; establishment of *phylogenetic* relationships of HEVs, determination of the genetic evolution and natural selection of the Kenyan HEVs was important.

We report on the molecular characteristics of HEVs types detected in stool samples from children ≤5 years with diarrhea in Kibera, Kenya between 2009 and 2015 by partially sequencing the VP1 gene.

Materials and Methodology

Sample Selection Criteria

The stool samples were collected through the Population-Based Infectious Disease Surveillance

(PBIDS) platform, carried out by the Kenya Medical Research Institute (KEMRI) and the US Centers for Disease Control and Prevention (CDC) in Kibera.

Kibera is a large urban informal settlement located 7 kilometers Southwest of Nairobi's Central Business. It is characterized by congested population with deprived hygiene standards [12].

PBIDS participants received free medical care for acute illness at a centrally located clinic. Stool samples Which were collected from those with diarrhea, were equal or grater than three (≥ 3) stools in 24-hours or clinician diagnosis of diarrhea).

A subset of archived specimen collected from children aged <5 years from 2009 to 2015 were tested using enteric TAC in 2016.

Criteria for selection for TAC testing included:

(a) Children less than two (<2) years with diarrhea

(b) Children less than(<5) years with two or more diarrhea illness for which a stool was provided

(C) Children less than(<5) years with diarrhea

(d) Children from a household with Human Immunodeficiency Virus-infected member.

Of 831 specimens from children meeting those criteria, 631 (76%) were available for TAC testing [1-3].

Extraction of Nucleic Acid Material

From stool samples, 180 – 220 mg or 200ul of liquid were extracted using a modified QIAamp Fast Stool Mini Kit procedure where they underwent a lysate preparation process that included a mechanical disruption step (bead beating), removal of inhibitors, purification and elution of nucleic material using spin columns.

The lysate was transferred into the spin columns and centrifuged at 15,600Xg for 1 minute. The columns were washed twice according to the manufacturer's instructions by adding 500µl wash solution, centrifuging at 15,600Xg for 1 minute and eluted with 100µl total nucleic acid (TNA) [13].

RT-PCR

The nucleic material was analyzed using an entero-pathogen TAC card pre-coated with multiple sets of primers and probe concordant to the conserved 5′-untranslated region of the enterovirus genome [14, 15].

A master mix was prepared using the Ag-Path-ID One-Step RT-PCR Kit (Life Technologies, Foster City, CA) as follows;

(a) 50 µL of the buffer,

(b) 4 µL enzyme mix

(c) 26µL nuclease-free water per reaction.

A total volume of 100 µL: 80 µL of master mix and 20 µL of TNA were loaded per reaction.

All TAC tests were run on the Applied Biosystems ViiA7 Real-Time PCR system (Life Technologies). Thermo cycling conditions of 10 min at 45°C for reverse transcription; 10 min at 94°C activation of Taq polymerase.

Subsequent 45 cycles consisting of cDNA denaturation step for 15s at 95°C and a combined annealing-extension step of 1 min at 60°C were employed.

Samples with a Cycle threshold (Ct) value of ≤35 were confirmed by RRT-PCR [15].

Briefly, 25µl reaction mixtures were made containing:

a. 12.5µL of 2x Taq man one-step RRT-PCR master mix buffer,

b. 1µl 40x Multi Scribe and RNase inhibitor mixture,

c. 1µM forward primer,

d. 1µM reverse primer,

e. 1µM probe (Table 7),

f. 5.75µl nuclease-free water

g. 5µL of purified RNA (or nuclease-free water for controls).

The reaction was incubated at 50°C for 30 min followed by 94°C for 3 min. Thermo cycling was performed for 45 cycles at 94°C for 30 s, 42°C for 30 s and 72°C for 30 s in a 7500F model thermocycler (Applied Biosystems). Ct values of ≤30 were considered for sequencing.

***Table** 7: Primers used for RT-PCR*

Primer or probe	Sequence	Nucleotide position
Forward primer	5'-TCCTCCGGCCCCTGA	452-466
Reverse primer 1	5'-AATTGTCACCATAAGCAGCCA	607-587
Reverse primer 2	5'-GATTGTCACCATAAGCAGCCA	607-587
Probe 1	5'-CGGAACCGACTACTTTGGGTGTCCGT	541-566
Probe 2	5'-CGGAACCGACTACTTTGGGTGACCGT	541-566

cDNA Synthesis and Amplification

A standardized in-house protocol was used for cDNA synthesis and amplification obtain amplicons for sequencing.

Briefly, a master mix containing 2.0 µL RNase free water; 16 µL of 2X reaction mix; 1 µL of 20 pico moles/µL forward and reverse primers (292 and 222 in ***Table 8***); 1 µL of SuperScript III RT/Platinum Taq mix and 4 µL of the RNA template was used for each reaction.

The reaction mix were incubated at 50°C for 45 min; 94°C for 3min then thermo cycling was performed for 40 cycles of 94°C for 30s; 42°C for 1 min and 60°C for 2 min followed by a final extension at 60°C for 10 min.in a model 9700 thermocycler (Applied Biosystems, Foster City, CA) [2].

The reaction products were analyzed by electrophoresis in a 1% agarose gel and staining with 0.5 mg/ml ethidium bromide.

Table 8: *Primers Used For Amplification and Sequencing*

Target	Primer/Probe Name	Sequence (5'-3')	Gene	Location (nt)	Specificity
HEV (ALL-VP1)	HEV292	MIGCIGYIGARACNGG	VP1	2612–2627	All EV
	HEV222	CICCIGGIGGIAYRWACAT	VP1	2969–2951	All EV

Nucleotide Sequencing

Gel bands of approximately 350bp were excised and purified using the Promega gel purification kit (Wizard® SV Gel and PCR Clean-Up System, Madison, Wisconsin). Sequencing of the purified PCR amplicons was performed using the BigDye Terminator *version 3.1* Cycle Sequencing kit (Applied Bio systems, Foster City, CA).

The product was further purified using sephadex-50 (Sigma Aldrich Co; USA). PCR products were finally loaded in an automated 3500 Genetic Analyzer (Applied Bio systems Foster City, CA).

Sequence Assembly and Data Analysis

The sequences were aligned using the ClustalW algorithm implemented in MEGA6. The Bayesian *phylogenetic* trees were based on partial VP1 nucleotide gene sequences of HEV. Phylogenetic analysis was performed using MrBayes 3.2 [16] employing the general time-reversible (GTR + Gamma) substitution model estimated in MEGA 6. A posterior probability was shown in percentages on each node.

The identity of each specimen was determined to be the serotype of the highest scoring strain in Gen Bank using the Basic Local Alignment Search Tool (BLAST) followed by the pair-wise comparison of the partial VP1 sequences to VP1 sequences of prototypes and variant strains from Gen Bank

Homologous strains from various African countries were selected according to proximity to study strains upon applying the BLAST tool. Nucleotide sequences displaying a percent identity of ≥75% or ≥85% amino acid homology and above were considered sufficient for serotype assignment [2].

Natural selection signatures were detected from the DNA sequences using SLAC/FEL/BUSTED methods implemented in the Data monkey [17, 18].

Natural selective pressure on VP1 was surveyed by appraising the ratio of non-synonymous (dN) to synonymous (dS) substitutions (ω=dN/dS) on a codon by codon basis through the lineages.

To determine whether any particular codon location on VP1 was under positive selection, both the FEL and SLAC methods were applied [17].

For those that failed on SLAC and FEL analyses, BUSTED analysis was applied to determine evidence of diversifying selection at both the site and test branch levels [19].

Results

TAC and RT-PCR Results

Of the 628 stool specimens tested by the enteric Taqman array cards (TAC) assay on the ViiA 7 instrument (Life Technologies), 209 (33%) tested positive for HEV with a cut-off Cycle threshold (Ct) of ≤ 30 recommended for the enteric TAC (12).

Of these, 171 (82%) were confirmed positive for HEVs by a subsequent real-time RT-PCR assay, targeting the 5' NTR region performed on the ABI 7500 F instrument (Ct ≤40). Specimens with a Ct ≤30 by RRT-PCR (n=72; 42%), were tested by conventional PCR targeting the 3' region of the VP1 gene.

Of these, 48 (67%) specimens gave a band of approximately 350bp, and their nucleotide sequences were determined.

628 stool samples available for TAC Testing

419 (67%) negative for Enterovirus by TAC

209 (33%) positive for Enterovirus by TAC underwent rRT-PCR Testing

38 (18%) negative for Enterovirus by RRT-PCR

99 (47%) Ct>30 by RRT-PCR

72 (34%) positive for Enterovirus by RRT-PCR (with a Ct ≤30) underwent conventional PCR

24(33%) unexpected band length or smear

48 (67%) samples with expected band length (350bp) underwent sequencing

37 (77%) no readable sequence

11 (23%) readable sequences analyzed

Figure 1: *Sample Testing Algorithm For Human Enterovirus Stool From Children Below Five Years In Kibera Kenya*

Sequencing Results

Nucleotide and Amino Acid Homology Results

Eleven (23%) of the sequenced samples yielded usable nucleotide sequences amongst which, two sequences ≥200nt were deposited in Gen Bank with accession numbers MG838739 and MG838740 ***(Table1).***

Table 1: *All Human Enterovirus Serotypes and Clade/Species Identified In Stool Specimens From Kibera.* Kenya *2009 - 2015. Nine different serotypes were identified showing HEV serotype diversity within this population.*

Sample ID	HEV Serotype	Year of Collection	Clade/ Species group	Accession Number
KEN_KIBERA_03	*Enterovirus A76*	2009	A	MG838740
KEN_KIBERA_10	*Echovirus E12*	2009	B	-
KEN_KIBERA_16	*Echovirus E13*	2010	B	-
KEN_KIBERA_15	*Echovirus E25*	2011	B	-
KEN_KIBERA_17	*Echovirus E9*	2012	B	-
KEN_KIBERA_01	*Coxsackie A24*	2013	C	-
KEN_KIBERA_04	*Coxsackie A13*	2013	C	-
KEN_KIBERA_06	*Coxsackie A13*	2013	C	-
KEN_KIBERA_13	*Coxsackie A13*	2014	C	-
KEN_KIBERA_05	*Coxsackie A20*	2014	C	MG838739
KEN_KIBERA_11	*Enterovirus C99*	2015	C	-

These two belong to HEV-A76 and CV-A20 serotypes. *Phylogenic* analyses of the 11 nucleotide sequences, revealed a total of 9 HEV serotypes. Interestingly, an *Echovirus E13* from year 2010 clustered together with group C instead of group B ***(Fig 2)***.

The HEV-A76 strain had an 88% nucleotide identity and 99% amino acid identity to the prototype (FRA91) and an 84% - 86% nucleotide identity and 97% to 99% amino acid identity to other African *homologous* strains ***(Table 2).*** CV-A20 had a 95% nucleotide identity and 91% amino acid identity to the prototype (IH35).

This strain had a 91% to 94% nucleotide identity and 85% to 91% amino acid identity to other African *homologous* strains ***(Table 2).***

***Figure** 2: Bayesian Phylogenetic Tree Of VP1 Region Showing Kenyan Serotypes In Relation To The HEV Prototypes.*

Table 2: *Nucleotide and Amino Acid Homology Of Kenyan HEV-A76 and Cv-A20 In Relation To Prototypes and Homologous African Strains.*

HEV - A76	Nucleotide % Identity	Amino acid % Identity
	KEN_KIBERA_03	KEN_KIBERA_03
(Prototype) AY697458_HEV76_FRA91	87.56%	98.61%
KY433796_HEV_A76_Niger	84.33%	98.61%
JX307650_HEV_A76_Cameroon	86.11%	97.22%
EU481516_HEV_A76_Egypt	86.11%	98.61%
CV - A20		
	KEN_KIBERA_05	KEN_KIBERA_05
(Prototype) AF499642_CVA20_IH35	94.57%	90.69%
KY433692_CV_A20_Niger	94.40%	90.57%
JX417858_CV_A20_Cameroon	90.60%	85.00%

Phylogeny Results

Phylogenetic analysis of the VP1 region of the Kenyan HEV serotypes with HEV prototypes strains retrieved from Gen Bank revealed that these viruses separated into three main clusters: A, B and C ***(Fig. 2).*** Most of the Kenyan strains belonged to group C (6/11, 55%) and group B (4/11, 36%) while one (9%) belonged to group A. The HEV serotypes showed diversity when compared to homologous strains from Kenya and other African countries ***(Fig. 3).***

***Figure 3:** Bayesian Phylogenetic Tree Of The VP1 Region Showing Kenyan Serotypes In Relation To Homologous Strains From Other African Countries.*

Amino acid Substitution Amongst the Kenyan HEVs

The HEV-A76 strain in this study showed an amino-acid substitution (Q103M) in VP1 when compared to prototype strain, HEV76_FRA91 ***(Fig 4; Table 3).***

The CV-A20 had nine- amino-acid substitutions compared to the prototype strain, CVA20_IH35. The amino acid substitutions included: A58G; R83C; A85G; C86W; Y95N; I101T; E103T; T108A and T115S ***(Fig 5; Table 3)***. Amino acid substitutions at codon I101T

concurred with Niger and Cameroon strains and varied with some homologous strains from the two countries. Amino acid substitutions at codon E103T differed in the homologous strains under comparison; Niger and Cameroon strains had E103K while Cameroon strains had I103Q.

Amino acid substitutions at codon 108 had a consensus in all homologous strains selected for comparison *(Fig 5; Table 3)*. CV-A13 study serotypes had 22 inter-serotype amino acid substitutions observed among the study CV-A13 and a deletion at codon site154 *(Fig 6)*

```
                                                  60        70        80        90        100
                                         ....I....I....I....I....I....I....I....I....I....I
AY697458_HEV76_FRA91                     ETGATSNTTDENMIETRAVINKHSTEETCIEHFFSRAGLVSMISLLVSEE
KEN_KIBERA_03                            ..................................................
JX307650.1   Human enterovirus 7         ..................................................
KY433796.1   Enterovirus A76 iso         ..................................................
EU481516.1   Human enterovirus 7         ----------........................................
EU481517.1   Human enterovirus 7         --................................................
EU481518.1   Human enterovirus 7         --------------N...................................
JX307650.1   Human enterovirus 7         ..................................................
KY433796.1   Entrovirus A76    iso       ..................................................

                                                  110       120       130       140       150
                                         ....I....I....I....I....I....I....I....I....I....I
AY697458_HEV76_FRA91                     TTQAYANWEIDTMGFVQMRRKLEMFTYMRFNAEFTFVATGLNGTAPQTML
KEN_KIBERA_03                            ..M...............................................
JX307650.1   Human enterovirus 7         .........................................X--------
KY433796.1   Enterovirus A76 iso         .............................................PP...
EU481516.1   Human enterovirus 7         ..................................................
EU481517.1   Human enterovirus 7         ..................................................
EU481518.1   Human enterovirus 7         ..................................................
JX307650.1   Human enterovirus 7         ..................................................
KY433796.1   Entrovirus A76    iso       ..........................................X-------
```

Figure 4: *Vp1 Amino Acid Sequence Alignment Of The Kenyan HEV-A76 Strain Relative To The Prototype FRA91 Strain and Other Homologous African Strains.*

Table 3: *Nucleotide Substitutions That Translated In Amino Acid Changes In Kenyan HEV-A76 and CV-A20.*

HEV	Codon Changes				
Serotype	Codon Positions	From		To	
		Codon	Amino acid	Codon	Amino acid
HEV - A76	103	CAG	Glutamine	ATG	Methionine
	58	GCT	Alanine	GGT	Glycine
	83	CGG	Arginine	TGT	Cysteine
	85	GCG	Alanine	GGG	Glycine
	86	TGT	Cysteine	TGG	Tryptophan
CV - A20	95	TAC	Tyrosine	AAC	Asparagine
	101	ATC	Isoleucine	ACT	Threonine
	103	GAG	Glutamic acid	ACA	Threonine
	108	ACC	Threonine	GCA	Alanine
	115	ACA	Threonine	AGC	Serine

		60 70 80 90 100
		....I....I....I....I....I....I....I....I....I....I
AF499642_CVA20_IH35		ATSSVVPADLVQTRSRSESTVESFFARGACVTFFARGACVTIMSVENYNETXX
KEN_KIBERA_05		.........G..C.GW.............N............
KY433693.1	CV_A20_Niger	........................K...AG
KY433692.1	CV_A20_Niger	....
JX417858.1	CV_A20_Cameroon	...E.I.S..........................M...AG
JX417855.1	CV_A20_Cameroon	....AG
JX417856.1	CV_A20_Cameroon	...T................V...
JX417857.1	CV_A20_Cameroon	....VG
JX417859.1	CV_A20_Cameroon	.......S................S...
		110 120 130 140 150
		....I....I....I....I....I....I....I....I....I....I
AF499642_CVA20_IH35		IAESKLFTKWNITYTDTVQLRRKLEMFTYSRFDIEEFTEVVTERYH
KIBERA_05		T.T....A......S.....X........................C.GW.............N............
KY433693.1	CV_A20_Niger	VSK....A...........X..AG
KY433692.1	CV_A20_Niger	.SK....A...........X..
JX417858.1	CV_A20_Cameroon	T-K....A...........X..AG
JX417855.1	CV_A20_Cameroon	T.K...AG
JX417856.1	CV_A20_Cameroon	VSQ...A...........X..
JX417857.1	CV_A20_Cameroon	T.K....A...........X..VG
JX417859.1	CV_A20_Cameroon	VTK...A...........X..

Figure 5 : VP1 Amino Acid Sequence Alignment Of The Kenyan CV-A20 Strain Relative To The Prototype IH35 Strain And Other Homologous African Strains.

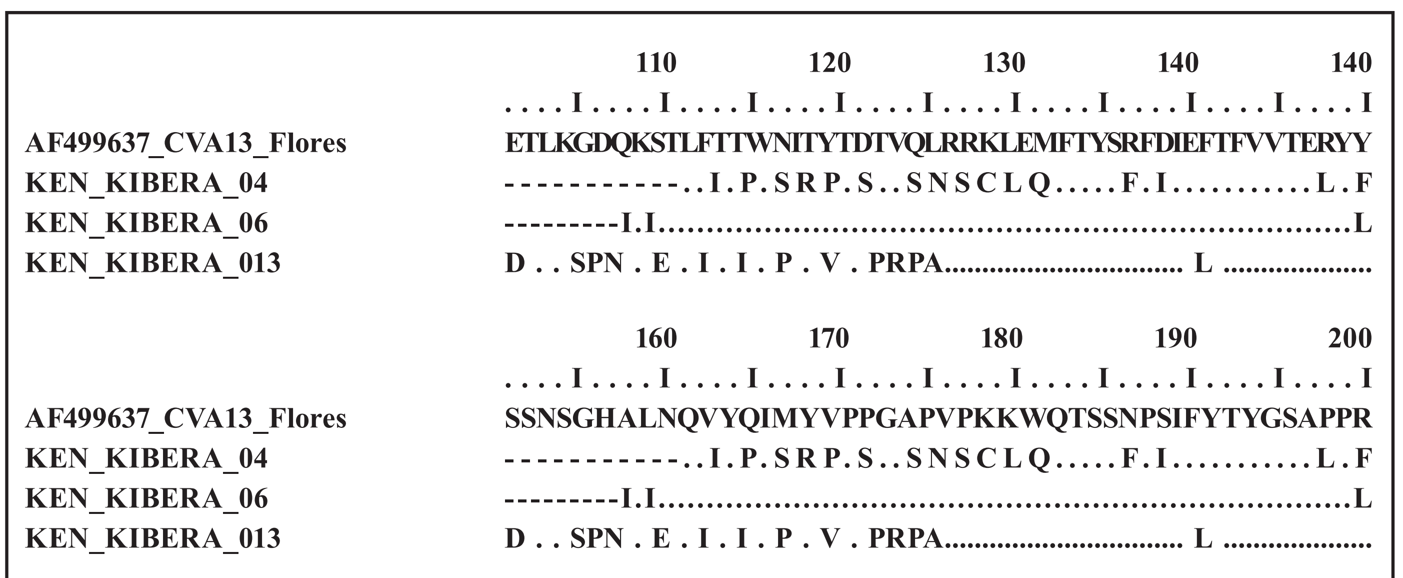

Figure 6: *Inter-serotype amino acid sequence alignment of the Kenyan CV-A13 serotypes [h].*
h Amino acid inter-serotype substitutions in study strain CV-A13. A deletion was also observed at codon sites 154 compared to the prototype Flores strain.

Evolutionary Pressure amongst the Kenyan HEVs

Both FEL and SLAC analyses indicated a significant (p = 0.05) negative selection/pressure on HEV-A76 study strain. Negative selection in the HEV-A76 strain were observed at codon sites 78 (p = 0.025); 88 (p = 0.034); 111 (p = 0.033); 112 (p = 0.025) and at 116 (p = 0.047) by FEL and at codon site 111 (p = 0.024) by SLAC analyses ***(Table 4).***

FEL analysis also determined significant positive and negative selection in study strain CV - A20 (p = 0.05). Positive selection was observed at codon 171 (p = 0.046) and negative selection at codon 181 (p = 0.039) ***(Table 5)***.

SLAC and FEL analyses did not find any positive or negative selection site to the other study strains which could be attributed to their short sequences.

However, Branch-Site Unrestricted Statistical Test for Episodic Diversification (BUSTED) found significant evidence (Likelihood Ratio Test (LRT, p-value = ≤ 0.05) of positive selection in the selected test branches of ECHO-E25 and other study serotypes ***(Table 6) .***

Table 4: *Negative selection of HEV-A76 serotype by FEL and SLAC analysis. Significant negative selection was observed at p values ≤0.05 by both analyses e*

Analysis	*Codon Site*	*Alpha(Ds)*	*Beta (Dn)*	*Alpha = Beta*	*P-Value*
FEL	78	22.23	0.00	1.79	0.03
FEL	88	8.37	0.00	2.34	0.03
FEL	111	23.04	0.00	7.38	0.03
SLAC	111	3.47	0.00	-5.58	0.02
FEL	112	20.99	0.00	1.75	0.03
FEL	116	17.23	0.00	1.76	0.05

e (dN) non-synonymous; (dS) synonymous substitutions

Table 5: *Negative and Positive Selection Sites In Kenyan CV - A20 by FEL analysis f*

Codon Site	(dS)	(dN)	dS = dN	p-value	Selection
171	0.00	296.54	34.48	0.05	Pos. Selection
181	10000	0.00	10.79	0.04	Neg. Selection

f Significant pervasive (p= 0.046)
and purifying (p=0.039) selection in CV-A20 serotype at codon 171 and 181 respectively.
(dN) non-synonymous;
(dS) synonymous substitutions

Table 6: *ECHO-E25 BUSTED analysis. Significant gene-wide episodic positive selection observed at codon sites 49 and 68 in ECHO-E25 Kenyan strain; p-value 0.01 and 0.03 respectively*

Codon Site	Unconstrained likelihood	Constrained likelihood	Optimized Null Likelihood	Constrained Statistic	Optimized Null Statistic (p)
49	-3.90	-3.82	-3.90	-0.16	0.01
68	-5.20	-5.12	-5.22	-0.16	0.03

Discussion

Partial sequencing of the HEVs VP1 gene revealed that, strains belonging to groups A, B and C were in co-circulation in Kibera, Kenya during the study period (2009 - 2015).

The HEVs serotypes reported in this study include:

a. CV - A24, A13 and A20
b. *Echovirus* E12, E13, E25, and E9
c. HEVs A76 and C99.

Previous studies of HEVs serotypes in stool and respiratory samples from Kenyan children had some overlap with our findings [20, 21].

However, this study was the first to report HEVs A76; E13; E9, CV A13 and CV A20 amongst Kenyan children.

CV-A20 has been associated with aseptic *meningitis, herpangina, myocarditis* and Bornholm diseases [6].

HEV - A76 is a recently reported HEVs serotype classified genetically as group A EV (EV-A) and is associated with viral *encephalitis* [22].

Although HEV-76 is not among the most commonly reported HEVs serotypes [3]. It has caused various enterovirus outbreaks including the viral encephalitis outbreak that occurred in northern India in 2006 [22].

Nucleotide and amino acid homology results confirmed the HEV serotypes through amino acid identity to prototypes and variant strains [2].

Evolution and genomic diversity of the various HEVs circulating strains are important in identifying recombination and mutations that may contribute to the *pathogenesis* [11].

Serotype multiplicity presents an environment for intra-serotype recombination that would determine the molecular characteristics of the virus population.

Additionally, multi-serotype infections as well as the variations in viral antigenicity are responsible for epidemics [23].

In the present study, ECHO-E13 clustered with group C HEVs despite being a group B HEVs suggesting that ECHO-E13 had genetic substitutions from group C species.

The VP1 gene harbors the BC and DE - loops which are important immune-provoking regions connected with HEVs infectivity [20].

Amino acid substitutions within these loops can considerably modify the host's neutralizing response to the virus.

An example is the *Echovirus 9* where *cytolytic* proficiencies of *lytic* clone-derived *Echovirus 9* in human pancreatic islets have been reported to be altered by a single amino acid substitution (T81A) in the VP1 capsid protein [24].

We found that the ECHO-E9 strain identified in this study had the T81A amino acid substitution suggesting it could result in *necrosis* of the human pancreatic islets.

Amino acid substitutions have in some cases resulted in changes in viral antigenicity. For example, an amino acid alteration from *alanine* to *threonine* at position 107 in VP1 protein dramatically diminishes replication of EV-71 virus due to impaired maturation [25].

A sole amino acid alteration at codon position 103 observed in HEV-A76 translated to a variation from *glutamine* to *methionine*.

Although further, in vitro studies could elucidate the impact of this amino acid substitution to host. Negative selection plays an imperative role in maintaining an enduring stability of a virus by eliminating *lethal* mutations [26].

The numerous amino acid substitutions in study serotypes were confirmed as negative selection by FEL and SLAC analysis, suggesting purifying selection. These findings correlated well with previous findings by Opanda *et al.,* 2016 that also suggested negative selection of circulating enteroviruses in the same age group in Kenya [7].

Strong ($p \leq 0.05$) negative selection as observed in the HEVs serotype suggests elimination of highly deleterious variants whose genes have vital roles in the host and withstand limited amino acid substitutions to maintain stability [26].

Evidence of negative and positive selection was observed in CV-A20 suggesting that both purifying and pervasive pressures are at play in this serotype strain.

Positive selection in CV-A20 and ECHO-E25 suggested a tendency of the virus to deviate from stabilizing selection by positively mutating residues in the antigenic epitope.

Particularly, those located on BC and DE loops thus promote subsequent divergence from other *homologous* HEVs serotypes. Positive selection on the VP1 region has been reported to mostly occur within the BC and DE regions [17].

Consequently, positive selection identified in CV-A20 and other strains in this study could have antigenic effects.

Conclusion

This study had several limitations. First, the *heterogeneous* nature of stool samples is known to influence disparities in viral nucleic acids yields and could have contributed to the few samples analyzed in this study [27].

TAC PCR primers were designed to capture a broad range of enteroviruses. Thus TAC detected 209 specimens however only 171 could be confirmed with specific HEV RRT-PCR test [28].

Sequencing was limited to specimen with RRT-PCR Ct values ≤30 and not all sequenced samples

yielded nucleotide sequences. The results therefore, may not be representative of all circulating HEVs in the study area.

Since this was a retrospective study of previously collected samples, it is possible that some HEVs strains may have failed to amplify; thus specimens from the earlier years of the study period may have been under represented.

The sole use of the VP1 gene to analyze the viruses was limiting; more analysis could have been done if we had performed whole genome sequencing (WGS) of the virus.

Notwithstanding these shortcomings, we established some of the HEVs serotypes in circulation in Kibera and their genetic diversity. We also found serotypes belonging to Enterovirus A, B, and C not previous reported in Kenya [7, 29].

Findings from this study have also indicated pervasive and purifying pressure in some enterovirus serotypes which have been shown to contribute to patterns of nucleotide *polymorphism* in HEV [30].

Inter-serotype differences of VP1 gene as observed among the study CV-A13 serotypes could be as a result of insertions or deletions of distinct amino acids that result in alterations of the virus capsid structure associated with divergent strains [18, 31].

Future WGS studies will provide more insight into the genetic and evolutionary characteristics of human enteroviruses in Kenya.

Conflict of Interest

None declared.

Ethical Approval

The study protocol was submitted and approved by the KEMRI's Scientific and Ethics Review Unit (SERU) under protocol number 3381.

Stool samples were collected under the protocol "Active population-based surveillance of SERU, major infectious disease syndromes in Western Kenya and Nairobi" (PBIDS).

This was approved by KEMRI under protocol numbers 932, 1899 and CDC Institutional Review Board Number 4566.

Samples collected from the field were given unique Identification(IDs) with no patient's identification.

Acknowledgment

We wish to appreciate the assistance of colleagues from Diagnostic and Laboratory Systems Program (DLSP) for their support during the entire study period.

Specifically, we appreciate contributions by Derrick Amon, for helping us with sample retrieval and Nucleic Material extraction. Moses Baraza for helping us with the initial sample screening by TAC assay.

We also appreciate the PBIDS program for funding the initial screening of study samples by TAC.

References

1. **Maël Bessaud, Sophie Jegouic, Marie-Line Joffret, Christelle Barge, Jean Balanant, Ionela Gouandjika-Vasilache,** and **Francis Delpeyroux..** Characterization of the genome of human enteroviruses: design of generic primers for amplification and sequencing of different regions of the viral genome. *Journal of Virological Methods*, Elsevier. 2008; 149 (**2**):277-84. [JviroMet]. [Pasteur].

2. **Oberste MS, M K, Williams AJ, Dybdahl-Sissoko N, Brown BA, Gookin MS, Peñaranda S, Mishrik N, Uddin M, Pallansch MA.** Species-specific RT-PCR amplification of human enteroviruses: a tool for rapid species identification of uncharacterized enteroviruses. ***Journal of General Virology.*** 2006.87:119-128. [JGV].

3. **Nasri D, Bouslama L, Omar S, Saoudin H, Bourlet T, Aouni M, Pillet S**. Typing of Human Enterovirus by Partial Sequencing of VP2. *Journal of Clinical Microbiology. 2007.* 45(8): 2370–2379. [JCM].

4. **Noor A, Krilov LR**. Enterovirus infections. Pediatr Rev. 2016. 37:505–515. [PUBMED].

5. **Juliana Attoh, Theophilus Adiku, John Kofi Odoom.** Prevalence of human enteroviruses among apparently healthy nursery school children in Accra. ***Pan Afr Med. 201***4.18(66). [doi];http://www.panafrican-med-journal.com/content/article/18/66/full/

6. **Yin-Murphy M, Almond JW,** eds. Picornaviruses. In: Baron S, editor. *Medical Microbiology 4th edition*. Galveston (TX): U*niversity of Texas Medical Branch at Galveston. 1996*.Chapter 53. https://www.ncbi.nlm.nih.gov/books/NBK7687/

7. **Opanda SM, Wamunyokoli F Khamadi, S Coldren R,** and **Bulimo W D**. Genotyping of enteroviruses isolated in Kenya from pediatric patients using partial VP1 region. *SpringerPlus. 2016*. 5 (158). http://doi.org/10.1186/s40064-016-1834-0

8. **Zhang S, Wang J, Yan Q, He S, Zhou W,** Ge S, Xia N. A One-Step, Triplex, Real-Time RT-PCR Assay for the Simultaneous Detection of Enterovirus 71, Coxsackie A16 and Pan-Enterovirus in a Single Tube. PLoS ONE. 2014. 9(7). https://doi.org/10.1371/journal.pone.0102724

9. **Muir P, Kämmerer U, Korn K, Mulders M. N, Pöyry T, Weissbrich B, van Loon, AM**. Molecular Typing of Enteroviruses: Current Status and Future Requirements. *Clinical Microbiology Reviews*. 1998.11(1):202–227. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC121380/

10. **Hawkes M T, Vaudry W**. Nonpolio enterovirus infection in the neonate and young infant. Paediatrics & Child Health. 2005.10(7): 383–388. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2722558/

11. **Oberste MS, Maher K, Kilpatrick DR, Flemister MR, Brown BA, Pallansch M A.**. Typing of human enteroviruses by partial sequencing of VP1. *J. Clin. Microbiol.* 1999a. 37 :1288–1293 [PMC free article] [PubMed]

12. **Daniel R Feikin Allan Audi Beatrice Olack Godfrey M Bigogo Christina Polyak Heather Burke John Williamson Robert F Breiman**. Evaluation of the optimal recall period for disease symptoms in home-based morbidity surveillance in rural and urban Kenya. *International Journal of Epidemiology. 2010*. 39: 450–458. https://doi.org/10.1093/ije/dyp374

13. **Zhou F, Kong F, McPhie K, Ratnamohan M, Donovan L, Zeng F, Dwyer D E.** Identification of 20 Common Human Enterovirus Serotypes by Use of a Reverse Transcription-PCR-Based Reverse Line Blot Hybridization Assay. *Journal of Clinical Microbiology.2009*. 47(9): 2737–2743. https://jcm.asm.org/content/47/9/2737

14. **Diaz, M. H, Waller, J. L, Napoliello, R. A, Islam, M. S, Wolff, B. J, Burken, D. J, Holden, R. L, Srinivasan, V, Arvay, M, McGee, L, Oberste, M. S, Whitney, C. G., Schrag, S. J, Winchell, J. M, Saha, S. K**. Optimization of Multiple Pathogen Detection Using the TaqMan Array Card: Application for a Population-Based *Study of Neonatal Infection. PLoS ONE. 2013*.8(6). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3689704/

15. **Liu J, Gratz J, Amour C, Kibiki G, Becker S, Janaki L, Houpt, E. R. A** Laboratory-Developed TaqMan Array Card for Simultaneous Detection of 19 Enteropathogens. *Journal of Clinical Microbiology. 2013*.51(2): 472–480. https://jcm.asm.org/content/51/2/472

16. **Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Huelsenbeck J P. Mr Bayes** 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. *Systematic Biology. 2012*. 61(3): 539–542. http://doi.org/10.1093/sysbio/sys029

17. **Kosakovsky Pond SL and Frost SD**. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics. 2005*a. 21:2531-2533. https://doi.org/10.1093/bioinformatics/bti320

18. **Kosakovsky Pond SL** and Frost SD.. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Molecular Biology and Evolution. 2005b. 22: 1208-1222. https://doi.org/10.1093/molbev/msi105

19. **Murrell B, Weaver S, Smith M D, Wertheim**

JO, Murrell S, Aylward A, … Kosakovsky Pond S L. Gene-Wide Identification of Episodic Selection. Molecular Biology and Evolution. 2015. 32(5):1365–1371. http://doi.org/10.1093/molbev/msv035

20. **Linsuwanon P, Puenpa J, Suwannakarn K, Auksornkitti V, Vichiwattana P, Korkong S, Apiradee Theamboonlers** and **Yong Poovorawan.** Molecular Epidemiology and Evolution of Human Enterovirus Serotype 68 in Thailand, 2006–2011. *PLoS ONE. 2012.*7(5). https://doi.org/10.1371/journal.pone.0035190

21. **Khetsuriani, N., Helfand, R., Pallansch, M., Kew, O., Fowlkes, A., Oberste, M. S., Tukei, P., Muli, J., Makokha, E and Gary, H.** Limited duration of vaccine poliovirus and other enterovirus excretion among human immunodeficiency virus infected children in Kenya. *BMC infectious diseases. 2009.* 9:136. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2739212/

22. **Sapkal G N, Bondre V P, Fulmali P V, Patil P, Gopalkrishna Dadhania V, Gore M M.** Enteroviruses in Patients with Acute Encephalitis, Uttar Pradesh, India. Emerging *Infectious Diseases. 2009.* 15(2): 295–298. http://doi.org/10.3201/eid1502.080865

23. **Behura S K, Severson D W.** Nucleotide substitutions in dengue virus serotypes from Asian and American countries: insights into intracodon recombination and purifying selection. *BMC Microbiology. 2013.* 13(37). http://doi.org/10.1186/1471-2180-13-37

24. **Paananen AYP, Smura T, Lempinen M, Galama J, Roivainen M.** A single amino acid substitution in viral VP1 protein alters the lytic potential of clone-derived variants of echovirus 9 DM strain in human pancreatic islets. *Journal of Medical Virology. 2013.*85: 1267-1273. https://doi.org/10.1002/jmv.23574

25. **Zhang Y-X, Huang Y-M, Li Q-J, Li X-Y, Zhou Y-D, Guo F, Zhou JM, Cen S.** A highly conserved amino acid in VP1 regulates maturation of enterovirus 71. *PLoS Pathog. 2017.*13(9) https://doi.org/10.1371/journal.ppat.1006625

26. **Wang, J, Street, N. R., Scofield, D. G,** and **Ingvarsson, P. K .** Natural Selection and Recombination Rate Variation Shape Nucleotide Polymorphism across the Genomes of Three Related Populus Species. Genetics: 2015. 202 (3) : 1185 - 200. http : // www.genetics . org /content / 202 / 3 / 1185.

27. **Bolanaki E, Kottaridi C, Dedepsidis E, Kyriakopoulou Z, Pliaka V, Pratti A, Levidiotou-Stefanou S, Markoulatos P.** Direct extraction and molecular characterization of enteroviruses genomes from human *faecal* samples. Mol Cell Probes. 2008. 22(3) : 156 - 161 https // www.ncbi.nlm.nih.gov/pubmed/18378420

28. **Harvey JJ, Chester S, Burke SA, Ansbro M, Aden T, Gose R, Sciulli R, Bai J, DesJardin L, Benfer JL, Hall J, Smole S, Doan K, Popowich MD, St George K, Quinlan T, Halse TA, Li Z, Pérez-Osorio AC, Glover WA, Russell D, Reisdorf E, Whyte T Jr, Whitaker B, Hatcher C, Srinivasan V, Tatti K, Tondella ML, Wang X, Winchell JM, Mayer LW, Jernigan D, Mawle AC. Co**mparative analytical evaluation of the respiratory TaqMan Array Card with real-time PCR and commercial Multi-Pathogen Assays. *Virol Methods. 2008*. 228:151-157. https://www.ncbi.nlm.nih.gov/pubmed/26640122

29. **L'Huillier AG, Kaiser L, Petty T J, Kilowoko M, Kyungu E, Hongoa P, Tapparel C.** Molecular Epidemiology of Human Rhinoviruses and Enteroviruses Highlights Their Diversity in Sub-Saharan Africa. *Viruses. 2015.*7(12) : 6412–6423. http://doi.org/10.3390/v7122948

30. **Vasseur E and Quintana-Murci L.** The impact of natural selection on health and disease: uses of the population genetics approach in humans. *Evol Appl. 2013.*6(4): 596-607. https://www.ncbi.nlm.nih.gov/pubmed/23789027

31. **Zhou F, O'Sullivan MVN, Iredell JR, Sintchenko V, Gilbert GL, Dwyer DE. 2016.** Molecular Analysis of Enteroviruses. *J Hum Virol Retrovirol* 3(2).https://medcraveonline.com/JHVRV/JHVRV-03-00083