Prevalence and molecular characterization of human noroviruses and sapoviruses in Ethiopia

Zufan Sisay1,2 · Appolinaire Djikeng3 · Nega Berhe1 · Gurja Belay2 · Wondwossen Gebreyes5 · Woldaregay Erku Abegaz1 · Moses N. Njahira3 · Q. H. Wang4 · Linda J. Saif4

Abstract Viral gastroenteritis is a major public health problem worldwide. In Ethiopia, very limited studies have been done on the epidemiology of enteropathogenic viruses. The aim of this study was to detect and characterize noroviruses (NoVs) and sapoviruses (SaVs) from acute gastroenteritis patients of all ages. Fecal samples were collected from diarrheic patients (n = 213) in five different health centers in Addis Ababa during June-September 2013. The samples were screened for caliciviruses by reverse transcription polymerase chain reaction (RT-PCR) using universal and genogroup-specific primer pairs. Phylogenetic analyses were conducted using the sequences of the PCR products. Of the clinical samples, 25.3 % and 4.2 % were positive for NoV and SaV RNA, respectively. Among the norovirus positives, 22 were sequenced further, and diverse norovirus strains were identified: GI (n = 4), GII (n = 17) and GIV (n = 1). Most strains were GII (n = 17/22: 77.2 %), which were further divided into three different genotypes (GII.4, GII.12/GII.g recombinant-like and GII.17), with GII.17 being the dominant (7/17) strain detected. GI noroviruses, in particular GL4 (n = 1), GL5 (n = 2) and GL8 (n = 1), were also detected and characterized. The GIV strain detected is the first from East Africa. The sapoviruses sequenced were also the first reported from Ethiopia. Collectively, this study showed the high burden and diversity of noroviruses and circulation of sapoviruses in diarrheic patients in Ethiopia. Continued surveillance to assess their association with diarrhea is needed to define their epidemiology, disease burden, and impact on public health.

Introduction

Acute gastroenteritis (AGE) remains one of the major causes of death in children under the age of five worldwide. It results in 1.8 million fatalities annually in developing countries alone [1] and is associated with about 14 % of deaths in children less than five years of age [2] in Ethiopia. Noroviruses and sapoviruses, members of the family Caliciviridae, are responsible for the majority of gastroenteritis cases. Norovirus, is the leading cause of gastroenteritis in people of all ages and is the second most common cause of diarrhea in children under the age of five, worldwide [3, 4]. In countries where universal rotavirus vaccination has been established, NoVs have become the leading cause of admission to hospital with acute gastroenteritis, also for children <5 years of age [5–8]. Noroviruses and sapoviruses are small, non-enveloped, single-stranded, positive-sense, polyadenylated RNA viruses with a genome of 7.4-7.7 kb in size [9]. Their genome is
organized into two to three open reading frames (ORFs), which encode both the structural and the non-structural protein genes, including the RNA-dependent RNA polymerase (RdRp) and the major capsid protein (VP1) [10, 11].

The NoVs are genetically highly diverse, which is attributed to the accumulation of point mutations associated with error-prone RNA replication and recombination. Currently, they are classified into at least five genogroups (GI-V) based on the complete sequence of the capsid protein VP1 [11–13]. Genogroup I (GI), GII, and GIV viruses are primarily human pathogens, although there are three porcine-specific genotypes within the GII genogroup [11, 12]. Comparison of full-length capsid sequences demonstrated that the five different genogroups share >60 % amino acid sequence identity in the VP1 region [12]. Each genogroup is subdivided further into distinct genetic clusters or genotypes that share >80 % identity in the amino acid sequence of VP1 [12]. Due to frequent recombination of NoVs, a dual classification/nomenclature system relating to the diversity of the RdRp and capsid genes has been proposed and appears to be the way forward [14, 15]. The other pathogenic enteric calicivirus, SaV, is classified into five genogroups based on capsid gene (ORF1) sequences, genogroups GI to GV [16]; however, potential new genogroups have been proposed [17–19]. Genogroups GI, GII, GIV, and GV contain the human SaVs. Each genogroup is further divided into genetically diverse genotypes.

The role of NoVs and SaVs as causative agents of gastroenteritis and their diversity in Africa is not well studied, except in some reports from a few African countries [20–36]. There has been only one published study from Ethiopia regarding the presence and genetic diversity of NoVs [37], and none for SaVs to date. Thus, this study focuses on the prevalence and genetic characterization of human NoVs and SaVs in Ethiopia.

Materials and methods

The study design for this work was cross-sectional. The study was approved by the National Health Research Ethics Committee and Institutional Review Board of Aklilu Lemma Institute of Pathobiology. The clinical samples (fecal samples, n = 213) investigated in this study were collected from self-reporting acute gastroenteritis patients attending five different government health care centers in Addis Ababa (Arada, Semen, Lideta, Kolfe and Kirkos Health Centers), between June to September 2013. Patients of all ages with diarrhea and one or more of the following symptoms – nausea, vomiting, abdominal cramps, headache, muscle pain and/or fever – were included. Addis Ababa is the capital and largest city of Ethiopia. According to the 2007 census, 98.6 % of the housing units of Addis Ababa had access to safe chlorinated drinking water, 14.9 % had flush toilets, 70.7 % had pit toilets, and 14.3 % had no toilet facilities [38].

Molecular detection and characterization

Fecal suspension preparation and RNA extraction

Fresh stool samples were collected from human acute gastroenteritis patients, using sterile containers, and stored frozen at -70 °C until processed. Ten-percent (w/v) fecal suspensions were prepared in phosphate-buffered saline (PBS, 0.01 M, pH 7.2) and clarified at 2000 × g (20 min, 4 °C) to eliminate larger debris. Viral RNA was extracted from 250 μl fecal suspensions using an RNeasy Mini Kit (QIAGEN, Valencia, CA) according to the manufacturer’s instructions. It was treated with 5 U of DNase I (Invitrogen, Carlsbad, CA, USA) to remove DNA. The extracted RNA was stored at -70 °C until used.

Reverse transcription

The cDNA was synthesized from the extracted RNA using a MAXIMA First Strand cDNA Synthesis Kit (Thermo Scientific) according to the manufacturer’s instructions. The kit contains highly sensitive primer mixture oligo (dT)18 and random hexamer. The cDNA obtained was used for the subsequent PCR assays [39].

PCR and primers

Human NoVs were detected and characterized by RT-PCR followed by sequencing using a calicivirus universal primer pair and genogroup-specific primer pairs. The universal primer pair (p290/p110) targets a conserved region in the RdRp [40, 41], which correspond to nucleotides (nt) 4568–4590 (p290) and nt 4865 to 4884 (p110) in the prototype NoV genome sequence (Hu/NoV/GI.1/Norwalk, GenBank accession no. M87661). The genogroup-specific primer pairs G1SKF/G1SKR and G2SKF/G2SKR [42], which amplified the S’ end of the capsid gene, were also used for the detection of human GI and GII NoVs, respectively. The genogroup-specific primer pair PEC66/PEC65, which is often used to detect porcine SaV GIII [43], was also employed in this study to detect potential interspecies transmission. The primer pairs used to detect NoVs and SaVs are listed in Table 1. The PCR conditions were optimized for all primers and the specificity of the primers was evaluated using plasmid DNA of laboratory reference strains of calicivirus obtained from Linda Saif’s
laboratory at the Ohio State University, which included Hu/NoV/GII.4/HS194. All RT-PCR reactions were conducted individually with each primer pair using various PCR cycling parameters. Positive controls (NoV and SaV) and negative controls (RNase-free water) were included in each run. The amount of PCR inhibitors in the samples was determined to be negligible based on results obtained by making 10-fold and 100-fold dilutions of each RNA sample and using an internal control as described previously [17].

Conditions for PCR were as follows: initial denaturation at 94°C for 3 min, followed by 40 cycles of 94°C for 30 s, 48°C for 30 s, and 72°C for 1 min, with a final extension step at 72°C for 10 min for primer pair p290/p110 and 94°C for 3 min followed by 40 cycles of 94°C for 30 s, and 55°C and 72°C for 1 min for primer pairs G1SKF/G1SKR and G2SKF/G2SKR. The amplicons obtained were analyzed in agarose gels after staining with GelRed dye (Thermo Scientific) using a UV light transilluminator.

The PCR products were separated by electrophoresis in an agarose gel, and samples were purified using a QIAquick Gel Extraction Kit (QIAGEN) before direct sequencing according to manufacturer’s instructions. Sanger sequencing of the RT-PCR products was carried out at the facilities of Segolip Unit of BecA-ILRI Hub, Nairobi, using an ABI 3730 automatic capillary sequencer (Applied Biosystem, Foster City, CA, USA). Complete sequences were obtained by assembling overlapping contigs followed by trimming off primer sequences. The nucleotide sequences obtained from the selected strains were used to search similar sequences in the NCBI genetic database using the BLAST tool (available at http://www.ncbi.nlm.nih.gov/blast/ Blast.cgi). The DNA sequences were aligned using the ClustalW method [44]. Phylogenetic trees were constructed by the neighbor-joining method using DNASTAR and MEGA6 software, with 1,000 bootstrap replicates for each gene [44]. The evolutionary history was inferred using the neighbor-joining method [45]. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Tamura-Nei method and are in units of the number of base substitutions per site. Evolutionary analyses were conducted using MEGA6 [44].

### Statistical analysis

Data were analyzed using SPSS version 15 software. Proportions were used to summarize the frequency of occurrence of target outcomes in different categories of exposure variables. The precision of the prevalence of enteric viruses across different categories was assessed by determining the 95% confidence interval around the estimates. The statistical significance of differences of proportion was evaluated using the Chi-square test. A P-value <0.05 was considered an indicator of statistical significance.

### Results

#### Profiles of the study population

During the study period, 213 individuals were enrolled and screened for NoVs and SaVs. All of the participants were residents of the city of Addis Ababa, and 54.9% were female, resulting in a female-to-male ratio of 1.22:1. The

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’–3’)</th>
<th>Polarity</th>
<th>Virus specificity</th>
<th>Target gene</th>
<th>Size (bp)</th>
<th>Location (nt)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>p290</td>
<td>GATTACTCCAAGTGGGACTCCAC</td>
<td>F</td>
<td>Calicivirus</td>
<td>RdRp</td>
<td>317 or 329</td>
<td>4568–4590a</td>
<td>Jiang et al., [40]</td>
</tr>
<tr>
<td>p110</td>
<td>DATYTCATCATCACCATACA</td>
<td>R</td>
<td>Calicivirus</td>
<td>RdRp</td>
<td>4865–4884a</td>
<td>5342–5361a</td>
<td>Le Guyader et al., [41]</td>
</tr>
<tr>
<td>G1SKF</td>
<td>CTGCCCGAATTGYGAAATGGA</td>
<td>F</td>
<td>GI NoV</td>
<td>Capsid</td>
<td>329</td>
<td>5342–5361a</td>
<td>Kojima et al., [42]</td>
</tr>
<tr>
<td>G1SKR</td>
<td>CCAACCCCARCCATATTACA</td>
<td>R</td>
<td>GI NoV</td>
<td>Capsid</td>
<td>343</td>
<td>5058–5076b</td>
<td>Kojima et al., [42]</td>
</tr>
<tr>
<td>G2SKF</td>
<td>CNTGGGAGGGCGATCGCAA</td>
<td>F</td>
<td>GII NoV</td>
<td>Capsid</td>
<td>330</td>
<td>4327–4347c</td>
<td>Guo et al. [43]</td>
</tr>
<tr>
<td>G2SKR</td>
<td>CCCGCATRCCRTTRATACAT</td>
<td>R</td>
<td>GII NoV</td>
<td>Capsid</td>
<td>4636–4656c</td>
<td>5379–5401b</td>
<td>Kojima et al., [42]</td>
</tr>
<tr>
<td>PEC66</td>
<td>GACTACAGCAAGTTGGATTCC</td>
<td>F</td>
<td>Po SaV</td>
<td>RdRp</td>
<td>4636–4656c</td>
<td>4327–4347c</td>
<td>Kojima et al., [42]</td>
</tr>
<tr>
<td>PEC65</td>
<td>ATACACACATCATCCCCGTA</td>
<td>R</td>
<td>Po SaV</td>
<td>RdRp</td>
<td>4636–4656c</td>
<td>4327–4347c</td>
<td>Kojima et al., [42]</td>
</tr>
</tbody>
</table>

Reference NoV strains with accession numbers (in parentheses) used for primer design
(a) Hu/NoV/GI.1/Norwalk (M87661)
(b) GI NoV Lordsdale strain (DDJB accession no. X86557)
(c) Po/SaV/GII/Cowden (AF182760)
age distribution of the study participants ranged from 4 months to 73 years. The proportion of the participants by age and sex are summarized in Tables 2 and 3.

**Prevalence**

**Norovirus infection**

Of the 213 diarrheic patients screened by RT-PCR, 25.3 % (54/213) were positive for NoV RNA (Table 4). Using primer pairs G1SKF/G1SKR and G2SKF/G2SKR, which were designed to detect human NoV GI and GII, respectively, 47 NoV strains were detected (explained in Table 5). For samples that were negative using the NoV genogroup-specific primer pairs, we next employed the calicivirus universal primer pair p290/p110 [40, 41]. Seven additional samples were positive using the primer pair p290/p110, and all were sequenced and identified as human NoVs (Table 5). We also analyzed the prevalence of NoVs in relation to the age of the study participants. The prevalence of NoV differed among age groups, ranging from 9.5 to 42.6 % in subjects >40 years and 1-4 years, respectively (Table 3). Statistical analysis showed that the frequency of human NoV infection was strongly correlated with the age (1-4 years) of the children ($p = 0.007$).

**Sapovirus infection**

For SaV detection, we first employed the calicivirus universal primer pair targeting the RdRp region p290/p110, which failed to detect SaV RNA in the present study. We next used the genogroup-specific primer pair PEC66/PEC65, whereby nine of the samples were positive for SaV RNA (4.2 %) (9/213) (Table 3). SaV was detected in four of the five health centers at different percentages, ranging from 1.6 to 8 % (Table 4). The percentage of the samples positive by RT-PCR in the study subjects also differed among age groups: 18.2 % in infants <1 year; 4.3 % in children 1-4 years; 7.9 % in subjects 10-19 years; 3.7 % in adults 30-39 years and 4.76 % in adults >40 years of age.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Total number (%) tested</th>
<th>Number (%) norovirus positive by RT-PCR</th>
<th>Number (%) sapovirus positive by RT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>117 (54.9 %)</td>
<td>31 (26.5 %)</td>
<td>6 (5.1 %)</td>
</tr>
<tr>
<td>Male</td>
<td>96 (45.1 %)</td>
<td>23 (24.0 %)</td>
<td>3 (3.1 %)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Total number (%) tested</th>
<th>Number (%) norovirus positive by RT-PCR</th>
<th>Number (%) sapovirus positive by RT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>11 (5.2 %)</td>
<td>3 (27.3 %)</td>
<td>2 (18.2 %)</td>
</tr>
<tr>
<td>1-4</td>
<td>47 (22.1 %)</td>
<td>20 (42.6 %)</td>
<td>2 (4.3 %)</td>
</tr>
<tr>
<td>5-9</td>
<td>18 (8.5 %)</td>
<td>2 (11.1 %)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>10-19</td>
<td>38 (17.8 %)</td>
<td>7 (18.4 %)</td>
<td>3 (7.9 %)</td>
</tr>
<tr>
<td>20-29</td>
<td>51 (23.9 %)</td>
<td>14 (27.4 %)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>30-39</td>
<td>27 (12.7 %)</td>
<td>6 (22.2 %)</td>
<td>1 (3.7 %)</td>
</tr>
<tr>
<td>&gt;40</td>
<td>21 (9.9 %)</td>
<td>2 (9.5 %)</td>
<td>1 (4.76 %)</td>
</tr>
<tr>
<td>Total</td>
<td>54 (25.3 %)</td>
<td>9 (4.2 %)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Sex distribution of study subjects and those found positive for noroviruses and sapoviruses by RT-PCR in Addis Ababa, 2013

Table 3 Age distribution of study subjects and those found positive for noroviruses and sapoviruses by RT-PCR in Addis Ababa, 2013

Table 4 Number and timing of fecal sample collections from different health centers in Addis Ababa in 2013
Genetic analysis of noroviruses

Phylogenetic analysis of a portion of the RNA polymerase region

Seven of the 54 positive samples detected using the primer pair p290/p110 targeting the RNA polymerase region were sequenced (Table 5). After performing a BLAST search, the sequenced strains were identified as human NoVs (ZS-AH21-ETH-2013, ZS-AH74-ETH-2013, ZS-KH05-ETH-2013, ZS-KH21-ETH-2013, ZS-AH12-ETH-2013, ZS-L6-ETH-2013, ZS-AH35-ETH-2013) (Fig. 1). The GenBank accession numbers of the partial NoV RdRp-specific sequences are KM589655, KM589656, KM589657, KM589658, KM589659, KM589660, and KM589661, respectively. A phylogenetic tree was generated using the partial RdRp region of seven NoV sequences that we determined, together with published sequences of NoV reference and representative strains. A neighbor-joining tree based on the partial RdRp gene and BLAST search indicated that the seven sequences we obtained were potential NoV strains of groups GII and GIV. Most of the strains (75 %, 6/7) that were identified were of genogroup GII, with nucleotide identities ranging from 67 to 100 % (Fig. 1). The GII strains were further divided into three potential genotypes (GII.4, GII.g/GII.12, GII.17). The potential recombinant GII.g/GII.12-like genotypes were dominant (4/7), with 95.0 to 100 % nucleotide sequence identity among themselves. They were highly related (95.2 to 96.5 % sequence identity) to human NoV GII.4/GII.12 isolates from Sydney, Australia (Gunnedah/ NSW895P/2010/AU: JQ613569 and Wahroonga/NSW004P/2009/AU: JQ613568) and China (F106/Beijing/2009: JQ889442) (Fig. 1). Phylogenetic analysis and BLAST search also demonstrated that one strain, identified as L6-ETH-2013 clustered with the GII.17 strain from Japan (Kawasaki/2014/JP: AB983218), with the highest nucleotide sequence similarity (95.4 % identity). The only GII.4 isolate sequenced in the RNA polymerase region, AH12-ETH-2013, shared 98.3 % nucleotide sequence identity with a strain isolated in Italy (113196-4/2014/Italy: KJ598058). The other strain identified, AH35-ETH-2013, clustered with the GII.17 strain from Japan (Kawasaki/2014/JP: AB983218), with the highest nucleotide sequence similarity (95.4 % identity). The only GII.4 isolate sequenced in the RNA polymerase region, AH12-ETH-2013, shared 98.3 % nucleotide sequence identity with a strain isolated in Italy (113196-4/2014/Italy: KJ598058). The other strain identified, AH35-ETH-2013, clustered with a separate genogroup (GIV), displaying high nucleotide sequence similarity (94.2-95.0 % identity) to several strains from China (CCDCGR 1113-59/CHN/2011:KC894731), Australia (LakeMacquarie/NSW268O/2010/AU: JQ613567) and Spain (NV2942/VLC/Spain/2012: KC119502) (Fig. 1).

Sequence and phylogenetic analysis of a portion of the capsid region

Of the 47 NoV positives amplified using primers designed based on a portion of the capsid region, we sequenced 15 samples (Table 5), and a BLAST search confirmed that the sequenced strains were all human NoV GI and GII strains (ZS-L5-ETH-2013, ZS-L512-ETH-2013, ZS-L38-ETH-2013, ZS-L587-ETH-2013, ZS-L567-ETH-2013, ZS-KH07-ETH-2013, ZS-L58-ETH-2013, ZS-L80-ETH-2013, ZS-L87-ETH-2013, ZS-AH177-ETH-2013, ZS-KH17-ETH-2013, ZS-50A-ETH-2013, ZS-5847-ETH-2013, ZS-5688-ETH-2013, ZS-5595-ETH-2013, ZS-5379-ETH-2013). We submitted these NoV sequences of the partial capsid region to the GenBank database (KM589642, KM589643, KM589644, KM589645, KM589646, KM589647, KM589648, KM589649, KM589650, KM589651, KM589652, KM589654, and F106/Beijing/2009: JQ899442) (Fig. 1). Phylogenetic analysis and BLAST search also demonstrated that one strain, identified as L6-ETH-2013 clustered with the GII.17 strain from Japan (Kawasaki/2014/JP: AB983218), with the highest nucleotide sequence similarity (95.4 % identity). The only GII.4 isolate sequenced in the RNA polymerase region, AH12-ETH-2013, shared 98.3 % nucleotide sequence identity with a strain isolated in Italy (113196-4/2014/Italy: KJ598058). The other strain identified, AH35-ETH-2013, clustered with a separate genogroup (GIV), displaying high nucleotide sequence similarity (94.2-95.0 % identity) to several strains from China (CCDCGR 1113-59/CHN/2011:KC894731), Australia (LakeMacquarie/NSW268O/2010/AU: JQ613567) and Spain (NV2942/VLC/Spain/2012: KC119502) (Fig. 1).
KU705716, KU705717, and KU705718, respectively). A neighbor-joining tree indicated that the 15 NoV strains from this study were diverse and belonged to the GI and GII types. The majority of the NoVs identified were GII (11/15: 73.3 %), which occurs in the majority of NoV infections worldwide, while four strains (26.7 %) belonged to GI. The GII strains identified were further divided into two different genotypes (GII.4 and GII.17) (Fig. 2). Of the GII isolates analyzed in this study (L5, L38, L80, L58, KH07, 5688, 5152, 5379, 5847, 5595), the majority (63.6 %, 7/11) belonged to GII.17 (L80, L58, KH07, 5688, 5152, 5687, 5847), and four belonged to GII.4 (L5, L38, 5379, 5595). Six of the seven GII.17 strains (5152, 5687, 5847, L80, L58, KH-07) identified shared 93-98 % nucleotide sequence similarity among themselves; however, strain 5688-ETH-2013 shared lower nucleotide sequence similarity (66-70 % identity) with the strains identified from this study as well as with previously published strains. The other strains that clustered as GII-17 shared high nucleotide sequence similarity (90-95 % identity) with the prototype strains from the USA (CS-E1/2002/USA: AY502009), French Guiana (C142/GF/1978: JN699043 and C142/1978/GUF: KC597139), Ethiopia (ETH176/2009/Ethiopia: JF909057) and Japan (JP/2002/Saitama/T87: KJ196286) (Fig. 2).

The phylogenetic analysis and BLAST search also demonstrated that our strains L5-ETH-2013 and L38-ETH-2013 belong to GII.4. The two GII.4 strains from this study clustered with the newly emergent variant Sydney 2012, but with different levels of nucleotide sequence similarity. Strain 5379 shared 72 % sequence identity with strains from Australia (Sydney/NSW0514/2012/AU: JX459908, AY502009).
Fig. 2 Neighbor-joining phylogenetic tree based on the 300-nucleotide capsid gene of human NoV. Only bootstrap values greater than 70% are shown. The prefix ZS and boldface type indicate samples from the current study. GenBank accession numbers are shown in parentheses.

GII.4 SYDNEY 2012

GII.17

GII.8

GII.5

GII.4

Primer pair G1SKF/G1SKR identified 13 samples, of which four were sequenced (Table 5) and grouped as GI (50A [KM589652], KH17-ETH-2013 [KM589651], L87-ETH-2013 [KM589649], and AH77-ETH-2013 [KM589650]) sharing 77 to 100 % nucleotide sequence identity among themselves. Upon analysis, of these four GI strains, two clustered together, L87-ETH-2013, AH77-ETH-2013, and shared high (99 %) nucleotide sequence identity with one another and 95-97 % identity with GL5 strains from South Africa (10-07-05-K/2010/ZAF: JN191360), Hungary (Siklos-HUN/2013: KJ402295) and Russia (2003/3168/ Moscow/ RUS: FJ383818). They also shared 91 % nucleotide sequence identity with the prototype strain, MUSGROV-GBR00 (AJ277614). The other strain, KH17-ETH-2013, shared high nucleotide sequence similarity (94.3 % identity) with a strain isolated from Russia (2006/Nizhny Novgorod/RUS: FJ383816) and 80 % and 81 % identity with the prototype GL7 (Winchester/94/UK: AJ277609) and GI.8 strains (NLV/Boxer/2001/US: AF538679), respectively. Strain 50A-2013-ETH shared very high similarity (99.7 % identity) with GI.4 strains from Taiwan (GQ401129-GI-BFDA-07/2009/TAW and GQ401127-/GI/BFDA-05/2009/TAW) and 94.2 % identity with the prototype AB042808-G1.4-CHIBA-JPN00 from Japan (Fig. 2).

Genetic and phylogenetic analysis of sapoviruses

Sapoviruses are currently divided into five genogroups, although additional potential genogroups have also been proposed [17–19]. In the present study, a total of nine SaVs were detected using the genogroup-specific primer pair PEC66/PEC65, which is often used to detect porcine SaV GIII. Of the nine positive SaV samples detected, one from a 2-year-old child that contained adequate RNA levels was sequenced (ZS-22SA-2013-ETH) (Table 5). A BLAST search confirmed that the sequenced strain was a human SaV GII.2 strain with 91 to 94 % sequence identity to strains isolated in Brazil (VIG-AM-116006/BRA: KF924392 and NSCO666/PA/BRA/1993: EU295952) and in Malawi (MW514: AB234230) (Fig. 3). We submitted this partial SaV RdRp sequence to the GenBank database (accession number: KM589661). A phylogenetic tree was generated using SaV sequences from humans and swine, based on the published sequences of SaV reference and representative strains from NCBI (Fig. 3).

Discussion

Prevalence and genogrouping/genotyping of noroviruses and sapoviruses

The results from this study, based on direct molecular detection and characterization, confirmed the presence of NoVs (25.3 %) and SaVs (4.2 %) in diarrheic patients in all age groups. Recent studies have shown that NoV plays an important etiological role in sporadic diarrhea in infants and young children [46, 47]. In our study also, we observed that the rate of NoV infection was significantly higher (p = 0.007) in children <5 years old than in older children and adults. This is in agreement with other study that reported the high prevalence of this pathogen in children <5 years of age [48]. However, unlike several studies conducted in Africa and elsewhere [49], we found NoV-associated diarrhea occurring mainly among children 1-4 years old, more frequently than among infants <1 year old. This could be related to the small number of samples from infants <1 year old in the present study. However, similar findings were reported from industrialized countries, including England and the Netherlands, with the prevalence of NoV infection peaking in children 1-4 years old, indicating acquisition of NoV at a later age or widespread protection by maternal antibodies in infants <1 year old [50–52].

Genogrouping/genotyping of human enteric viruses

Among the 54 NoV-positive samples, sufficient RNA to enable genotype determination by sequence analysis was obtained from 22 samples (22/54: 40.7 %). Low viral load or unstable RNA in the samples might account for failure in genotyping of some of the viruses. Although NoVs could only be genotyped in 40.7 % of positive samples, the observed genetic diversity was very high. Multiple genogroups (all of the genogroups classified as human NoVs, GI, GII and GIV) were detected from all age groups, with different frequencies. Norovirus GII was the dominant circulating genogroup detected (n = 17: 77.2 %). This finding is consistent with most studies worldwide, including those from Africa, in which NoV GII was reported to be predominantly responsible for gastroenteritis outbreaks and sporadic cases [3, 6, 21, 22, 29, 49, 53, 54]. Nucleotide sequence and phylogenetic analysis of the NoV GII strains
revealed that GII.17 NoVs comprise the predominant genotype (36.4 %) detected in this study. This is in contrast to most studies in Africa, as well as in Europe and the USA, where GII.4 is the dominant genotype. However, similar to this study, genotype GII.17 has been identified as the dominant genotype in water in Kenya [55]. It might not be surprising to detect similar newly detected unusual strains dominating in both countries, as the countries share borders and cross-country rivers. Moreover, some previous studies have suggested that non-GII.4 NoVs have been predominant in the past, worldwide [56]. For example, analysis of archived samples from 1974 through 1991 has shown that the frequency of GII.3 was 48 %, compared with 16 % for GII.4 and 14 % for GII.7 strains [57]. Therefore, it is essential to study increases in non-GII.4 strains to determine possible reasons for their increased transmissibility or population susceptibility. GII.17 strains were also reported in patients with acute gastroenteritis in other countries: Brazil, Paraguay, Argentina, South Korea, Thailand, Kenya, the USA and Japan [23, 58–62]. Although not dominant, GII.17 was also detected from clinical specimens in some African countries, including Djibouti [53], Cameroon [49], Morocco [63] and Burkina Faso [64]. Norovirus GII.17 is of clinical relevance, as it has been implicated in nosocomial NoV infection [65] and chronic NoV infection in a kidney transplant patient [66].

Strains homologous to GII.4 NoV strains that are associated with large epidemics of NoV worldwide were also detected in this study. Several NoV pandemics associated with GII.4 have occurred every 2 to 3 years [67–74]. The recent or new pandemic isolate GII.4 2012 Sydney, which was first identified in March 2012 in Australia [75], is currently the dominant GII.4 variant identified in NoV-associated gastroenteritis outbreaks worldwide [75]. The GII.4 Sydney 2012 variant from Australia accounted for most of the GII.4 isolates identified in this study. The sample collection for the current study was conducted during June to September 2013 (Table 4), the time when this new variant was spreading worldwide. This shows the global distribution of this GII.4-Sydney variant. While Australia is geographically isolated in the world, phylogenetic analysis revealed that the same strains are circulating in Asia, Europe, and the Americas. Modern methods of transport and population movement are facilitating the importation of global NoV strains [76]. This study suggests that the majority of the NoV GII outbreak strains isolated in Australia mirror the strains causing outbreaks during similar time frames in Africa too, which demonstrates the ability of these viruses to spread globally very rapidly. Additionally, two of the GII.4 strains identified in this study (5595-ETH, 5379-ETH) were about 75 % identical over 282 nucleotides to GII.4 variant Sydney 2012. According to Zheng et al. [12], to be classified into a genotype, strains are expected to have 80 % nucleotide sequence identity in the capsid (VP1) region. Therefore, this finding indicates that these two viruses, which were characterized as GII.4 NoV using region B primers, may represent novel NoV types, with gene sequences slightly different from the previously recognized human NoVs.

Along with mutations, recombination may also contribute to the genetic variability and evolution of NoVs. The four potential intergenotypic recombinant GII.g/GII.12-like strains from this study were highly related to human NoV recombinants from Sydney. Recombination, which has already been described for NoVs, is an important mechanism in the evolution of RNA viruses, since it can create changes in virus genomes by exchanging sequences and producing new viruses. It commonly occurs at the ORF1-ORF2 junction [77], although other recombination sites, such as the one between the ORF2 and ORF3 regions, have been reported [78, 79]. Recently there has been a significant increase in the number of NoV recombinants that have been detected. Norovirus recombinant GII.g/GII.12, emerged and was first detected in sporadic cases in
strains, while GIV.2 NoV has been designated as a canine and as a GII.17 strain in another paper [96], are human classified as Alphatron strain in most typing schemes [11, 97]. In conclusion, the present study confirms the significant role of NoV infection in diarrheic patients and provides evidence for the genetic diversity of NoVs in Ethiopia. We detected an undefined NoV genotype, potential recombinant-like strains, newly emerging variants, and rarely detected human NoV strains. This indicated that the emerging and undefined NoV genotypes have importance in Africa, including Ethiopia. Moreover, the detection of multiple genotypes of NoVs other than GI.4, which is the dominant genotype associated with NoV infection worldwide, is also important, since these strains may be circulating silently in the community, but with high clinical significance. The results presented here also suggest that recombination is not a rare phenomenon among NoVs, which thus contributes to their genetic diversity. Additionally, NoV affects more-vulnerable groups, such as the very young, which has important implications in a country like Ethiopia, where the number of children under five is high. Therefore, it is important to protect these vulnerable groups who are at the highest risk and establish a NoV surveillance reporting system that could help in understanding the public-health impact of diarrheal diseases.

Acknowledgments We thank the staffs of the different health centers, especially the laboratory technologists at Lideta Health Center, who helped us a lot in sample collection and filling out the questionnaire. We are very grateful to the patients who participated in this project. We highly acknowledge the help of Dr. Girmay Medhin in the statistical analysis. We appreciate the support of the Segolip unit at the BecA-ILRI Hub in sequencing the samples from this study. We
are also grateful for the cooperation and support of the Department of Molecular, Cellular and Biological Sciences and Aklilu Lemma Institute of Pathobiology, Addis Ababa University, in facilitating the successful accomplishment of the project. We greatly acknowledge the financial support of the African Biosciences Challenge Fund (ABCF)/ Biosciences of east and central Africa (BecA). We also thank the VPH-Biotech East Africa Consortium and the National Institutes of Health-Fogarty (Fogarty Grant D43TW008650, W Gebreyes, PI, The Ohio State University) for their financial support.

Compliance with ethical standards

Funding This study was funded by African Biosciences Challenge Fund (ABCF) Program and VPH-Biotech East Africa Consortium and the National Institutes of Health-Fogarty (Fogarty Grant D43TW008650, W Gebreyes, PI, The Ohio State University). The ABCF program is funded by the Australian Department for Foreign Affairs (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).

Conflict of interest None.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References


