

RESEARCH ARTICLE

# Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda

Akshaya Ramesh<sup>1,2</sup>\*, Sara Nakielny<sup>3</sup>, Jennifer Hsu<sup>4</sup>, Mary Kyohere<sup>5</sup>, Oswald Byaruhanga<sup>5</sup>, Charles de Bourcy<sup>6</sup>, Rebecca Egger<sup>6</sup>, Boris Dimitrov<sup>6</sup>, Yun-Fang Juan<sup>6</sup>, Jonathan Sheu<sup>6</sup>, James Wang<sup>6</sup>, Katrina Kalantar<sup>3</sup>, Charles Langelier<sup>4</sup>, Theodore Ruel<sup>7</sup>, Arthur Mpimbaza<sup>8</sup>, Michael R. Wilson<sup>1,2</sup>, Philip J. Rosenthal<sup>4</sup>, Joseph L. DeRisi<sup>3,6</sup>\*



**1** Weill Institute for Neurosciences, University of California, San Francisco, California, United States of America, **2** Department of Neurology, University of California, San Francisco, California, United States of America, **3** Department of Biochemistry and Biophysics, University of California, San Francisco, California, United States of America, **4** Division of Infectious Diseases, Department of Medicine, University of California, San Francisco, California, United States of America, **5** Infectious Diseases Research Collaboration, Kampala, Uganda, **6** Chan Zuckerberg Biohub, San Francisco, California, United States of America, **7** Division of Pediatric Infectious Diseases and Global Health, Department of Pediatrics, University of California, San Francisco, California, United States of America, **8** Child Health and Development Centre, Makerere University, Kampala, Uganda

\* These authors contributed equally to this work.

\* [Akshaya.Ramesh@ucsf.edu](mailto:Akshaya.Ramesh@ucsf.edu) (AR); [joe@derisilab.ucsf.edu](mailto:joe@derisilab.ucsf.edu) (JLD)

**OPEN ACCESS**

**Citation:** Ramesh A, Nakielny S, Hsu J, Kyohere M, Byaruhanga O, de Bourcy C, et al. (2019) Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. PLoS ONE 14(6): e0218318. <https://doi.org/10.1371/journal.pone.0218318>

**Editor:** Baochuan Lin, Defense Threat Reduction Agency, UNITED STATES

**Received:** December 13, 2018

**Accepted:** May 31, 2019

**Published:** June 20, 2019

**Copyright:** © 2019 Ramesh et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All raw data have been deposited under Bioproject ID: PRJNA483304. Assembled genomes can be accessed in GenBank: Accession numbers: MH685676-MH685701, MH685703-MH685719, MH684286-MH684293, MH684298-MH684334. All the raw data, intermediate data and IDseq reports can also be accessed at <https://idseq.net> [Project ID: Uganda]. All IDseq scripts and user instructions are available at <https://github.com/chanzuckerberg/idseq-dag> and the graphical user interface web application for sample upload is

## Abstract

Febrile illness is a major burden in African children, and non-malarial causes of fever are uncertain. In this retrospective exploratory study, we used metagenomic next-generation sequencing (mNGS) to evaluate serum, nasopharyngeal, and stool specimens from 94 children (aged 2–54 months) with febrile illness admitted to Tororo District Hospital, Uganda. The most common microbes identified were *Plasmodium falciparum* (51.1% of samples) and parvovirus B19 (4.4%) from serum; human rhinoviruses A and C (40%), respiratory syncytial virus (10%), and human herpesvirus 5 (10%) from nasopharyngeal swabs; and rotavirus A (50% of those with diarrhea) from stool. We also report the near complete genome of a highly divergent orthobunyavirus, tentatively named Nyangole virus, identified from the serum of a child diagnosed with malaria and pneumonia, a Bwamba orthobunyavirus in the nasopharynx of a child with rash and sepsis, and the genomes of two novel human rhinovirus C species. In this retrospective exploratory study, mNGS identified multiple potential pathogens, including 3 new viral species, associated with fever in Ugandan children.

## Introduction

The evaluation of children with fever is challenging, particularly in Low and Middle income countries (LMIC). A febrile child in sub-Saharan Africa may have a mild self-resolving viral infection or may be suffering from bacterial sepsis or malaria—major causes of disability and

available at <https://github.com/chanzuckerberg/idseq-web>.

**Funding:** JLD is supported by the Chan Zuckerberg Biohub. CdB, BD, YJ, JS, RE and JW are funded by the Chan Zuckerberg Initiative. SN was supported by Howard Hughes Medical Institute. PJR is funded by the National Institutes of Health, and work on this project by JH, MK, OB, and PJR was funded by the Doris Duke Charitable Foundation. MRW is supported by the Rachleff Foundation, NINDS K08NS096117 and the University of California, San Francisco Center for Next-Gen Precision Diagnostics, which is supported by the Sandler Foundation and the William K. Bowes, Jr. Foundation. AR is supported by the Rachleff Foundation. CL is supported by NHLBI K23HL138461-01A1, Nina Ireland Foundation and Marcus Program in Precision Medicine. KK is supported by the UC Berkeley UC San Francisco Joint Program in Bioengineering.

**Competing interests:** The authors have declared that no competing interests exist.

death [1, 2]. Historically, febrile illness in much of Africa has been treated empirically as malaria due to the limited availability of diagnostics and the risk of untreated malaria progressing to life-threatening illness. This strategy changed in 2010 following revised guidelines from the World Health Organization (WHO), which recommended limiting malaria therapy to those with a confirmed diagnosis [3]. However, standard recommendations for management of febrile children who do not have malaria are lacking. Increased knowledge about the prevalence of non-malarial pathogens associated with fever is needed to inform management strategies for febrile children [2], especially in low resource settings.

Advances in genome sequencing hold promise for addressing global infectious disease challenges by enabling unbiased detection of microbial pathogens that can be used to design directed diagnostics, and improve surveillance in LMIC [4–5]. The unbiased approach to detection of sequence-based diagnostics have led to the successful detection of pathogens in some rare or complex cases where traditional methods have failed [6–10]. Sequence-based diagnostics are complementary to serological assays and may contribute to a better understanding of pathogen landscapes in LMIC. Towards this aim, we conducted an exploratory retrospective mNGS analysis on samples available from a cohort of children hospitalized in rural Uganda with febrile illnesses to characterize potential pathogens associated with fever. The results, which include the detection of 3 novel viral species, suggest that mNGS will likely be a valuable tool in the arsenal of assays to understand the microbial landscape in human infections.

## Results

### Clinical characteristics of subjects with febrile illness

From October to December 2013, 94 children admitted to Tororo District Hospital were enrolled (Table 1). Their mean age was 16.4 (IQR: 8.0–21.0) months, and 66 (70.2%) were female. Chief symptoms reported in addition to fever were cough (88.3%), vomiting (56.4%), diarrhea (47.9%), and convulsions (27.7%). Top admitting diagnoses were respiratory tract infection (57.4%), gastroenteritis/diarrhea (29.8%), and septicemia (11.7%) (Table A in S1 Data). Of the 90 blood samples that were collected, thick blood smears identified *P. falciparum* in 12 samples that underwent mNGS analysis (Table B in S1 Data).

### Metagenomic sequencing findings

mNGS was performed on RNA extracted from 90 serum, 90 NP swab, and 10 stool samples following library preparation. A mean of 11.5 million (IQR 6.4–15.2 million) paired-end reads were obtained per sample; sequencing statistics are in S2 Data. For one batch of serum samples, only a single read, rather than paired-end reads, was produced. Bioinformatic analysis was

**Table 1. Overview of the patients enrolled in the study.** The other category includes: unknown (10), urinary infection (1), meningitis (2), hepatitis (1), and fever (1). Information on gender was missing for three and on age for two children.

Clinical category (number of patients in category)	Age (mean, months)	Gender	
		Male	Female
Respiratory illness (54)	14.0	25	28
Diarrhea/gastroenteritis (28)	12.1	9	19
Malaria (11)	15.5	3	7
Sepsis (11)	19.9	5	6
Malnutrition (5)	18.4	2	3
Other (15)	21.1	11	3

<https://doi.org/10.1371/journal.pone.0218318.t001>

carried out using the IDseq pipeline, a cloud-based, open-source platform designed for detection of microbes from metagenomic data (<https://github.com/chanzuckerberg/idseq-dag>) that incorporates several features of previously developed pathogen detection pipelines [11–17].

In this section, we discuss the mNGS results identified in a given sample type. Detailed findings on microbes identified in every patient, along with the total reads per million (rpm) are reported in Table B and Table C in [S1 Data](#), respectively.

### mNGS of serum

At least one microbial species was detected in 60 (66.7%) of the serum samples; more than one microbe was detected in 11 (12.2%) samples ([Fig 1A](#)). No microbial species were identified in the serum of 30 (33.3%) individuals. The most commonly identified microbes were *Plasmodium falciparum* (46, 51.1%) and parvovirus B19 (4, 4.4%). *P. falciparum* was detected in 10/12 samples from patients reported as smear-positive. mNGS detected *Plasmodium spp.* in 37 additional samples that were smear negative (36 *P. falciparum*, 1 *P. malariae*). Viruses detected in serum included human immunodeficiency 1 virus (HIV-1), hepatitis A virus, rotavirus A, human herpesvirus (HHV) type 6, HHV type 4, HHV type 7, human rhinovirus (HRV)-C, HRV-A, enteroviruses (enterovirus A71, Coxsackievirus B2 and echovirus E30), human parechovirus 2, hepatitis B virus, a novel orthobunyavirus (described in greater detail below), human cardiovirus (Saffold virus), mamastrovirus 1 and Norwalk virus ([Fig 1A](#)).

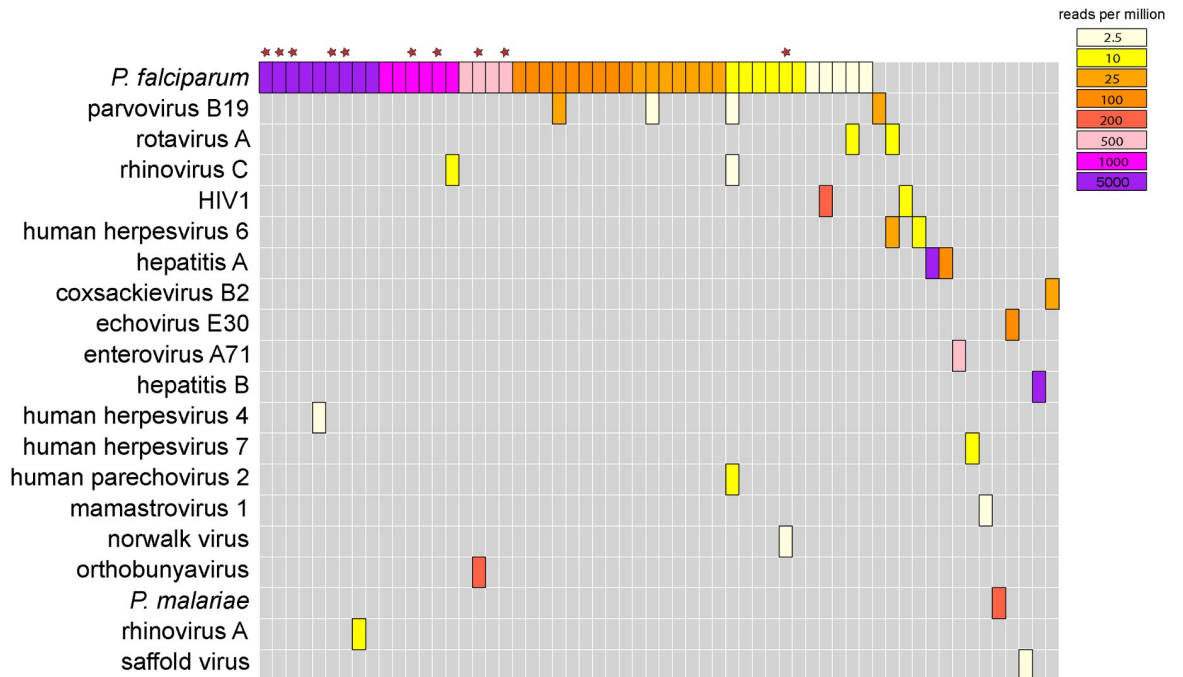
Multiple viruses were detected from serum in patients with *Plasmodium* infections (10 of 46 (21.7%) samples; [S1A Table](#)). Three of the four identified parvovirus B19 cases were associated with *P. falciparum*. Additionally, GB virus C and torque teno virus (TTV), which are of unknown clinical significance [18], [19], were identified in the serum of 25 (27.8%) and 37 (41.1%) children, respectively. There have been reports on associations between immunosuppression and TTV abundance [20, 21]. Interestingly, a prior study has reported a higher abundance of TTV in children with fevers [22].

### mNGS of NP swabs

90 NP swabs were collected and processed; 52 (57.7%) of these were from patients with admission diagnoses of pneumonia, respiratory tract infection, or bronchiolitis ([Table 1](#)). Chest imaging was not available to further assess these diagnoses. 72 NP samples (80%) contained at least 1 viral species ([Fig 1B](#)), with no microbes meeting our required cut-offs in 18 samples (20%). HRV-A and HRV-C were the most prevalent, followed by respiratory syncytial virus (RSV), cytomegalovirus (HHV-5), influenza B, and coronavirus OC43. Other respiratory viruses identified included influenza A (H1N1), HRV-B, Human mastadenovirus B (type 7), three human parainfluenza virus types (type 1, 3 and 4), metapneumovirus, coronavirus NL63, avian coronavirus, coxsackievirus A2, coxsackievirus B2, polyomaviruses (KI), HHV-6 and HHV-7. Other viruses identified that are not typically considered respiratory pathogens included hepatitis A virus, hepatitis B virus, parvovirus B19, mamastrovirus 1, Bwamba orthobunyavirus, betapapillomavirus 1, and rotavirus. Additionally, TTV was found in 49 (54.4%) NP swab samples, including one sample with both gemykrogvirus and TTV. For 26 (28.8%) patients, mNGS identified respiratory viral co-infections, most commonly with HRV-C (n = 11) and HRV-A (n = 5) ([S1B Table](#)). The same microbial species was identified in the NP swab and serum samples in six patients (on independent sequencing runs), one each with HRV-A, HRV-C, hepatitis A virus, hepatitis B virus, rotavirus A, and parvovirus B19.

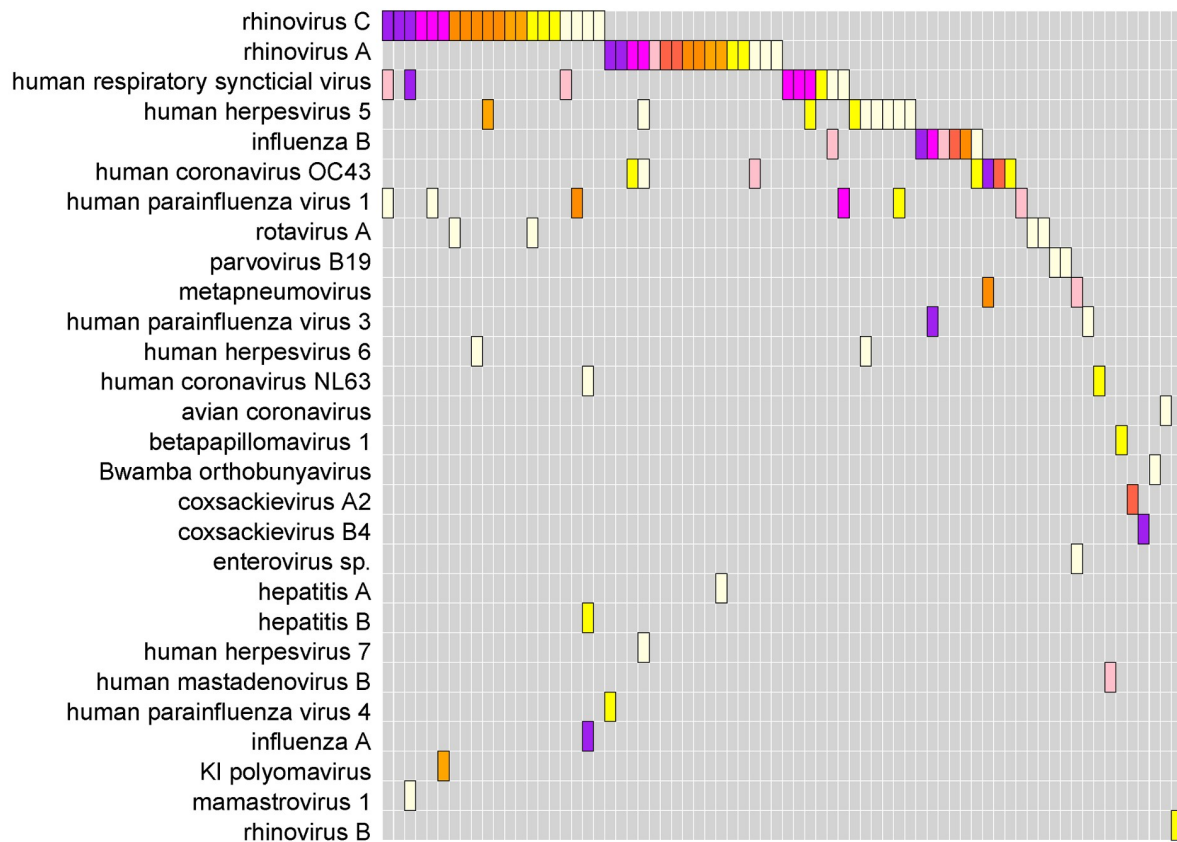
Bacteria identified in NP samples included four dominant genera, which together comprised 79% of all microbial reads—*Moraxella* (39.4%), *Haemophilus* (16.7%), *Streptococcus* (16.2%), and *Corynebacterium* (6.6%). Given that diversity loss in the microbial flora in lower

A



\* Malaria smear positive

B



**Fig 1. (A) Microbial landscape found in serum samples in Ugandan children.** Each column represents a febrile child. Results for GB virus C and torque teno virus, which are of uncertain clinical significance, are not included. **(B) Microbial landscape found in nasopharyngeal (NP) swab samples in Ugandan children.** Note that bacterial species were not considered in Fig 1B. Each column represents a febrile child, and the color bars represent the total reads per million (rpM) of a particular microbe present in the sample. Results for GB virus C and torque teno virus, which are of uncertain clinical significance, are not included.

<https://doi.org/10.1371/journal.pone.0218318.g001>

respiratory tract samples correlates with pneumonia [23, 24], we compared the Simpson Diversity Index (SDI) in patients with and without clinical diagnoses of respiratory tract infection. We found no significant difference in the SDI of upper airway samples between patients with (mean SDI = 0.51, IQR 0.37–0.65) or without (mean SDI = 0.51, IQR = 0.42–0.65;  $p = 0.86$ ) diagnoses of respiratory infection (S1 Fig). [25–27].

### mNGS of stool

Among the 10 stool samples collected and sequenced, potential non-bacterial pathogens were detected in 9/10 samples. The three most common microbes identified were rotavirus A (50%), *Cryptosporidium* (40%), and human parechovirus (40%). Our sequencing did not provide enough information to type the identified human parechovirus. Seven children had additional microbes identified: two rotavirus A and human parechovirus, and one each rotavirus A and *Cryptosporidium*, rotavirus A and enterovirus, *Cryptosporidium* and human parechovirus, *Cryptosporidium* and Norwalk virus, and *Giardia* and human parechovirus. Five samples also had *Blastocystis hominis*, a protozoan of uncertain pathogenicity.

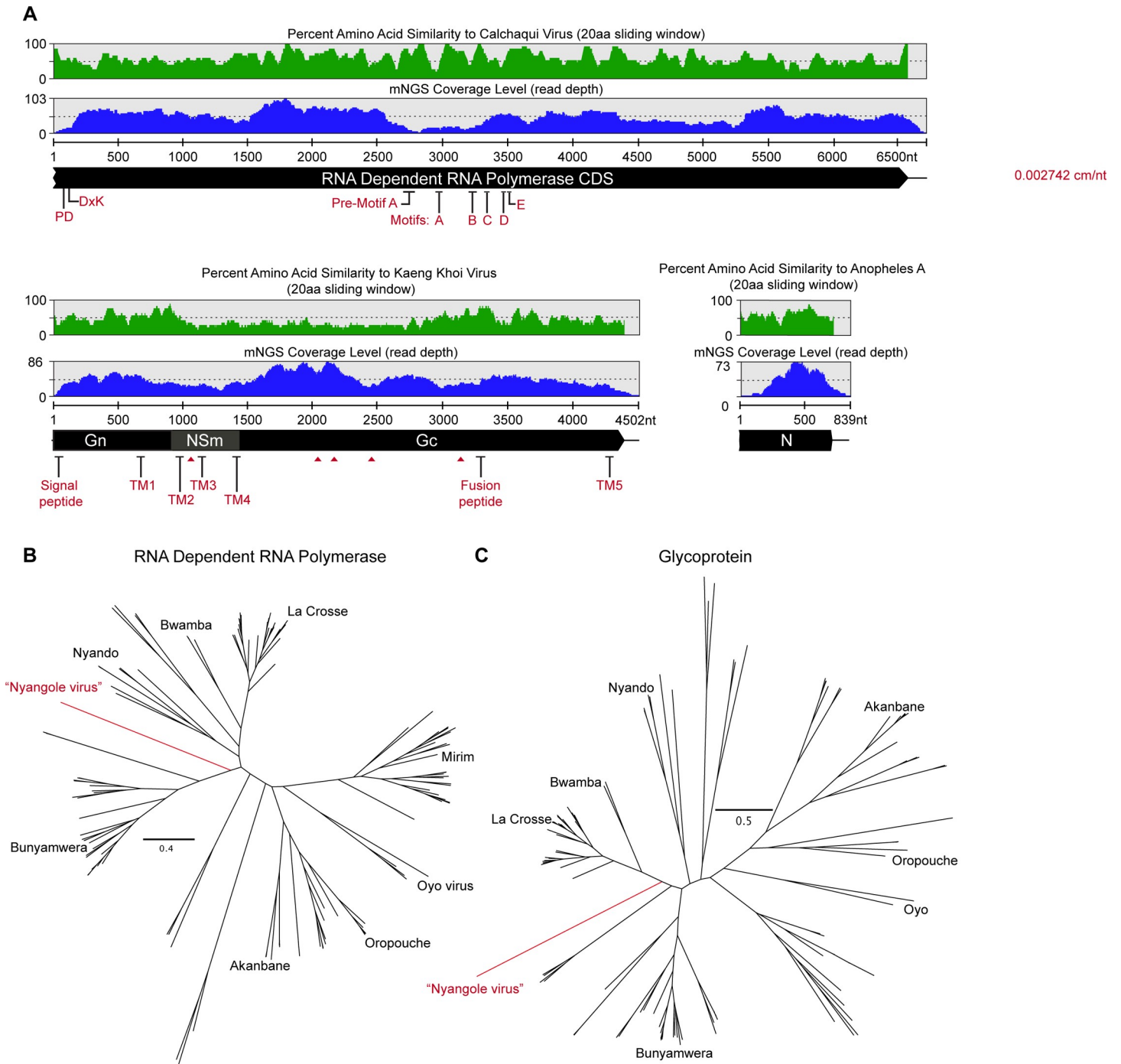
### Genomic characterization of viruses

Representative genomes for all viruses identified in this study were assembled and deposited in GenBank (Accession numbers: MH685676–MH685701, MH685703–MH685719, MH684286–MH684293, MH684298–MH684334). In this section, we describe in detail a novel orthobunyavirus identified in the serum of one individual, two novel HRV-C species identified in two individuals and diversity of influenza B viruses assembled in the nasopharynx from five individuals. We studied genomic diversity of influenza B in detail, given its potential implication to inform vaccine design. We did not include influenza A in this analysis, as we were able to assemble the complete genome sequence only from one individual.

### Orthobunyaviruses

Serum from one patient admitted with a clinical diagnosis of malaria and pneumonia contained a novel orthobunyavirus in addition to *P. falciparum*. Assembly of a near-complete genome and comparison with existing orthobunyavirus genomes indicated that this sequence includes 97.5%, 100% and 91% of the L, M, and S coding regions, respectively (Fig 2). Average read coverage across the segments was 86-fold. Phylogenetic comparison showed that the novel virus was significantly divergent from known orthobunyaviruses, sharing 44.9–55.1% amino acid identity with the closest known relatives, Calchaqui virus, Kaeng Khoi virus, and Anopheles A virus (Figs 2 and S2, S3 and S4). The virus was isolated from a patient from Nyan-gole village, Tororo District—hence, we propose the name “Nyangolo virus”, consistent with nomenclature guidelines for the family *Bunyaviridae*.

In addition, a second orthobunyavirus, Bwamba virus, was identified in the NP swab sample from a patient admitted with rash, sepsis, and diarrhea. Insufficient sample and sequencing reads precluded genome assembly of this virus.



**Fig 2. Characterization of the novel orthobunyavirus identified in a febrile child.** (A) Schematic representation of the large (L) or RNA dependent RNA polymerase, medium (M) or polyprotein of Gn, NSm and Gc proteins and small (S) segment encoding the nucleocapsid (N) protein of Nyangole virus and percentage identity with the most closely related virus. Phylogenetic tree of all complete orthobunyavirus genome sequences along with Nyangole virus are represented in (B) for the RNA dependent RNA polymerase and (C) for the glycoprotein.

<https://doi.org/10.1371/journal.pone.0218318.g002>

### Human rhinoviruses

Within the species rhinovirus, we assembled *de novo* a total of 13 HRV-C (mean coverage: 39-fold) and 13 HRV-A (mean coverage: 268-fold) genomes (> 500 bp). Of these, 10 HRV-A

and nine HRV-C genomes had complete coverage of the VP1 region, which is used to define enterovirus types [28]. Unique HRV types are defined by <73% similarity in the VP1 gene. As such, we found three HRV-A and eight HRV-C types in this cohort. One individual harbored two distinct HRV-A types (genome pairwise identity = 75.3%, VP1 pairwise identity = 67.1%). Additionally, we assembled two novel HRV-C species from two patients admitted with gastroenteritis (patient ID: EOFI-014) and with pneumonia, malaria and diarrhea (patient ID: EOFI-133), that shared 70.1% and 70.7% nucleotide sequence identity at VP1 compared to the closest known HRV-C (Accession JQ245968 and KF688606, respectively) (Fig 3). The Picornavirus Working Group has established that novel HRV-Cs should exhibit at least 13% nucleotide sequence divergence in the VP1 gene [29], qualifying these two as novel.

### Influenza B virus

We assembled influenza B genome segments (>500bp, mean-coverage: 4-fold) from six of seven samples containing influenza B virus (one sample had insufficient sequencing reads). The viruses assembled were >99% similar to each other and >99% identical to the B/Massachusetts/02/2012-like virus included in the vaccine recommended by WHO for the 2013–2014 northern hemisphere and 2014 southern hemisphere influenza seasons (Accession numbers: KC891816.1, KC891879.1 and KC892119.1). For the Ugandan viruses, three of the four major epitopes (150 loop, 160 loop and 190 helix) and their surrounding regions were 100% identical; a Ser->Thr substitution was observed at amino acid 136 (in the 120 loop) compared to the B/Massachusetts/02/2012-like virus.

### Discussion

A better understanding of the microbial agents causing fever in African children is needed to inform the development of better diagnostic algorithms, therapeutic guidelines and public health strategies. We performed an exploratory retrospective study with unbiased mNGS on various tissue types to determine whether this technology has potential to contribute to our understanding of the etiologies of fever in African children. In this limited sample set, mNGS identified a wide range of potential pathogens, including three novel viral species.

Other studies evaluating causes of febrile illness in African children have focused on a limited number of pathogens [30–33]. In a study of febrile children in Tanzania utilizing serologic, culture, and molecular assays, viruses accounted for 51% of lower respiratory infections, 78% of systemic infections, and 100% of upper respiratory infections [34]. Additionally, in the above study, 9% of the children had malaria and 4.2% had bacteremia. In febrile children in Kenya, reported pathogens were spotted fever group *Rickettsiae* (22.4%), influenza (22.4%), adenovirus (10.5%), parainfluenza virus 1–3 (10.1%), Q fever (8.9%), RSV (5.3%), malaria (5.2%), scrub typhus (3.6%), human metapneumovirus (3.2%), group A *Streptococcus* (2.3%) and typhus group *Rickettsiae* (1.0%) [35, 36]. Another study reported bacteremia in 19.1% of children admitted to a referral hospital in Uganda [37]. Additionally, in patients (across all age groups) with severe febrile illness, bacteremia was detected in 10.1% in North Africa, 10.4% in East Africa, and 12.4% in West Africa [38]. In this small study, *Plasmodium falciparum* was identified in the serum of 51.1% of the children, human rhinoviruses A and C dominated in the nasopharyngeal swab of 40% of the children and rotavirus A was identified in the stool samples of 50% of the children studied. For 20% of NP swabs and 33.3% of serum samples, no microbial species met our thresholds for detection. These proportions are consistent with previous reports [32, 34, 35, 39].

Unbiased sequencing approaches are designed to identify all potential pathogens but have also been limited by high cost and infrastructure needs. Given the exploratory nature of this

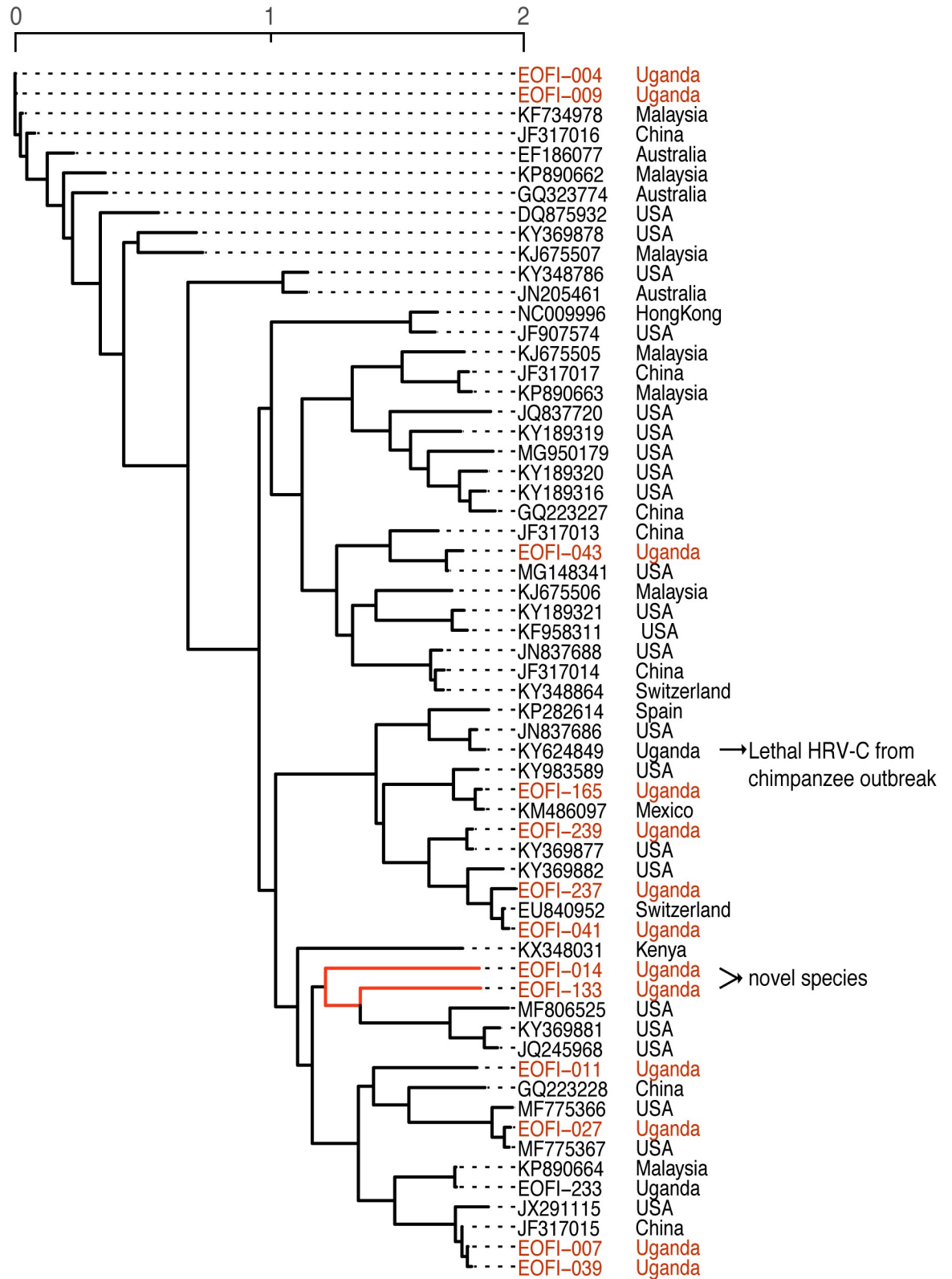


Fig 3. Phylogenetic tree of all complete HRV-C genomes from NCBI (with information on country of isolation) and HRV-C genomes assembled in this study.

<https://doi.org/10.1371/journal.pone.0218318.g003>



mNGS study, we cannot ascertain population level incidence or prevalence of particular infections. As expected, in the serum samples, *P. falciparum* was most commonly identified [40]. Some discrepancies were seen compared to blood smear readings, with false positive smears probably due to errors in slide reading, a common problem in under resourced clinics [41], and false negative smears due to the expected greater sensitivity of mNGS for identification of *P. falciparum*. In children with only sub-microscopic parasitemia, it is uncertain whether fevers can be ascribed to malaria, and in fact many children had both *P. falciparum* and additional microbes identified. Interestingly, three of the four cases of parvovirus B19 were found in association with *P. falciparum*; this co-infection has been associated with severe anemia with life-threatening consequences [42–44].

For NP and stool samples, given that the nasopharynx and intestines are normally colonized with commensal bacteria [45–48], and the lack of samples from healthy Ugandan controls, we focused on non-bacterial species. HRV was the most commonly identified virus in NP swab samples, consistent with findings previously reported in sub-Saharan Africa and developed countries [49–53]. HRV-C was most frequently encountered (54.1%), followed by HRV-A (43.2%) and HRV-B (2.7%), similar to the distribution of HRVs previously reported in Kenya [49]. We identified two novel HRV-C species which were approximately 70% identical to the most closely related previously described HRV-C species [29]. Overall, we detected at least three HRV-A and eight HRV-C types co-circulating in Tororo District. Of note, during the same collection period, a lethal HRV-C outbreak was reported in chimpanzees in Kibale National Park, in western Uganda [54]; that HRV-C was modestly related to an isolate observed in our study (74% nucleotide identity; 81% amino acid identity) (Fig 3) [54]. Our results confirm that a wide spectrum of HRVs infects Ugandan children. In addition to HRV, we detected a number of other known respiratory viruses, including RSV, human parainfluenza viruses, human coronaviruses, and adenovirus.

Diarrheal disease is one of the leading causes of death in children in Africa [55]. Approximately 48% of febrile children in our study presented with diarrhea, but due to logistical constraints stool specimens were available for only 10 cases. Rotavirus A, the leading cause of pediatric diarrhea worldwide [56], was the most commonly identified microbe in this cohort. Rotavirus vaccination, known to be highly effective, is yet to be implemented in Uganda, but the need is clear [56]. In addition to rotavirus A, we detected *Cryptosporidium*, norovirus, *Giardia*, *B. hominis* and several enteroviruses in stool specimens. Enteroviruses, HRV-C, and mamastrovirus were also identified in the serum of three children with clinical diagnoses of gastroenteritis or diarrhea.

Unbiased inspection of microbial sequences from sera revealed a novel member of the orthobunyavirus genus, tentatively named Nyangole virus, which was identified along with *P. falciparum* in a child with clinical diagnoses of malaria and pneumonia. The virus was surprisingly divergent from known viruses, with an average amino acid similarity of 51.6% to its nearest known relatives, including Calchaqui, Anopheles A and Kaeng Khoi viruses. Mosquitoes have been proposed as a vector for Calchaqui and Anopheles A viruses; Kaeng Khoi virus has been isolated from bedbugs [57–59]. Antibodies to these viruses have been detected in human sera, but their role as human pathogens is uncertain [46–49]. However, other orthobunyaviruses are responsible for severe human illnesses (e.g., Oropouche, Bunyamwera virus, California encephalitis virus, La Crosse virus, Jamestown Canyon virus, and Cache Valley virus) [60]. While the coverage depth of the assembled Nyangole virus genome in our patient suggests significant viremia, it is unknown whether the identified virus was responsible for the patient's febrile illness.

NP swab analysis identified another orthobunyavirus, Bwamba virus, in a child admitted with rash, sepsis and diarrhea. This virus has previously been described to cause fever in Uganda [61]. Our identification of two orthobunyaviruses, including one novel virus, in a small sample of

febrile Ugandan children suggests that the landscape of previously unidentified viruses that potentially infect African children and potentially cause febrile illness, is significantly under explored.

In addition to pathogen identification, the capacity of mNGS to provide viral strain resolution suggests its utility for monitoring vaccine efficacy by assessing prevalence of vaccine-targeted versus non-targeted strains. In the case of influenza B virus, the WHO recommended vaccine for 2013/2014 was highly conserved to the virus present in Uganda during that season.

Our exploratory pilot study had important limitations. First, our samples were not collected randomly, but rather were a retrospective convenience sample due to logistical constraints; as such, the results are not necessarily representative of pathogens infecting Ugandan children. In particular, the lack of identification of bacteremia in study subjects may have been due to a relative paucity of severe illness, compared to that in other studies. Second, the samples were collected only over a period of three months (October—December). Hence, we are unable to comment on seasonal trends in identified pathogens. Third, clinical evaluation of children followed the standards of a rural African hospital, so diagnostic evaluation was limited to physical examination and malaria blood smears. This study was not designed to compare mNGS to other clinical or laboratory assays. It is clear that more will be learned by linking rigorous clinical evaluation with mNGS results, and thereby more comprehensively assessing associations between clinical syndromes and specific pathogens. Fourth, healthy controls from the same population were not recruited in this study, hence we were unable to include them in the background model to filter out commensal microbial species specific to the Ugandan microbiome. Fifth, we were unable to use orthogonal techniques such as PCR to confirm the microbial species identified by mNGS due to lack of sample availability.

Given these limitations, we hesitate to integrate all the clinical specimens on a per sample basis, and rather present a portrait of all the microbes identified in febrile children. For readers interested in a breakdown of all microbial species from all samples collected per child, Table B in [S1 Data](#) contains this information. Future metagenomic studies should include rigorous clinical and microbiological phenotyping, along with samples collected from healthy individuals. This would facilitate design of an appropriate background model to identify potential pathogens, with confirmation using orthogonal techniques. Despite these limitations, our study provides an important snapshot of causes of fever in African children that could not be identified by available diagnostics, and suggests mNGS will be an important tool for future investigations. Given the yield of novel species in this small study alone, it is likely that an expanded use of this approach will continue to yield an increasingly rich portrait of microbial diversity associated with disease in this region.

## Methods

This study was approved by the Makerere University Research and Ethics Committee, the Uganda National Council of Science and Technology, and the University of California, San Francisco Committee on Human Research. Written informed consent was obtained from the parent or guardian on the child's behalf for all child participants enrolled in this study.

## Enrollment of study subjects

We studied children admitted to Tororo District Hospital, Tororo, Uganda, with febrile illnesses. Potential subjects were identified by clinic staff, who notified study personnel, who subsequently evaluated the children for study eligibility. Inclusion criteria were: 1) age 2–60 months; 2) admission to Tororo District Hospital for acute illness; 3) documentation of axillary temperature  $>38.0^{\circ}\text{C}$  on admission or within 24 hours of admission; and 4) provision of

informed consent from the parent or guardian for study procedures. The only exclusion criterion was unwillingness or inability of parents/guardians to provide consent.

### Sample collection

Serum and nasopharyngeal (NP) swab samples were collected from 90 children each; for four children, only one of the two sample types was successfully collected. Although 45 (47.9%) of the children had a presenting symptom of diarrhea, stool samples were available for only 10 due to logistical constraints. All samples that were collected were processed and included in the analysis.

### Study specimens

NP swabs and serum were collected from each enrolled subject within 24 hours of hospital admission. Approximately 5 ml of serum was collected by phlebotomy, the sample was centrifuged at room temperature, and serum was then stored at  $-80^{\circ}\text{C}$ . NP swab samples collected with FLOQSwabs swabs (COPAN) were placed into cryovials with Trizol (Invitrogen), and stored at  $-80^{\circ}\text{C}$  within  $\sim 5$  min of collection. For subjects with acute diarrhea ( $\geq$  three loose or watery stools in 24 hours), stool was collected into clean plastic containers and stored at  $-80^{\circ}\text{C}$  within  $\sim 5$  min of collection. Samples were stored at  $-80^{\circ}\text{C}$  until shipment on dry ice to UCSF for sequencing.

### Clinical data

Clinical information was obtained from interviews with parents or guardians, with specific data entered onto a standardized case record form that included admission diagnosis and physical examination as well as malaria blood smear results. For malaria diagnosis, thick Blood smears were Giemsa stained and evaluated by Tororo District Hospital laboratory personnel following routine standard-of-care practices. No efforts were made to improve on routine practice, so malaria smear readings represent routine standard-of-care rather than optimal quality controlled reads.

### Metagenomic next-generation sequencing (mNGS)

After shipment to University of California, San Francisco, RNA was extracted from clinical samples as well as positive (HeLa cells) and negative (water) controls, and unbiased cDNA libraries were generated using previously described methods (see sections “Sequencing library preparation” and “Metagenomic Library Preparation”, respectively) [62, 5]. Barcoded samples were pooled, size selected (Blue Pippin), and run on an Illumina HiSeq2500 to obtain 135 base pair (bp) paired-end reads.

### Bioinformatic analysis and pathogen identification

Microbial pathogens were identified from raw sequencing reads using the IDseq (v1.6) Portal (<https://idseq.net>), a cloud-based, open-source bioinformatics platform designed for detection of microbes from metagenomic data. IDseq is a scalable and cloud based implementation of a previously published pipeline (see Fig 1 in [11]). In brief, initial host read filtering is performed using Spliced Transcripts Alignment to a Reference (STAR) algorithm [17], followed by removal of duplicate/low quality and low complexity sequences [14], [16]. Next, reads are aligned once again to the host genome of interest using bowtie2 [12] to remove any remaining host reads. The non-human reads are then aligned to the NCBI nucleotide and protein database, using GSNAPL and RAPSearch, respectively [13], [15]. Additionally, reads that were identified as HHV-5 were assessed individually using BLAST, to verify specificity to this virus.

All IDseq scripts and user instructions are available at <https://github.com/chanzuckerberg/idseq-dag> and the graphical user interface web application for sample upload is available at <https://github.com/chanzuckerberg/idseq-web>. To distinguish potential pathogens from ubiquitous environmental agents including laboratory reagent contaminants and skin commensal flora, a Z-score was calculated for both nucleic acid and protein alignments, for each genus relative to a background of non-templated (“water only”) controls in addition to a previously published set of uninfected clinical mNGS samples [11]. CSF samples acquired through lumbar puncture from uninfected controls are included in this background model as there is extensive representation of the skin microbiome given the need to puncture the skin. Only human papillomavirus was identified in the positive (HeLa cell) controls.

For this study we report species greater than 0 reads per million (rpM) and a Z-score  $> 0$  (for both nucleic acid and protein alignments) detected in the serum, stool, and NP samples. We chose to report all species satisfying these criteria, rather than restricting to particular species, to offer an unbiased representation of microbes present in the sample. IDseq uses the CD-HIT-DUP tool to compress duplicate reads, hence final assignment of the rpM in a given sample represents coverage across the genome, rather than a single portion. Consistent with previous studies, low levels of “index bleed through” or “barcode hopping” (assignment of sequencing reads to the wrong barcode/index) was observed within the non-templated water control samples [63]. To account for barcode mis-assignment, when a microbe was found in more than one sample, it was reported only when present at levels at least four times the level of mis-assigned reads observed in the control samples. Given the extremely high levels of rotavirus found in stool samples, these samples were run in duplicate, and only microbes identified in both replicates and present at levels at least four times the number of reads mis-assigned in the control samples were reported. If the reads identified for a given microbe were not species-specific, we reported the corresponding genus. For NP and stool samples, because the nasopharynx and intestines are normally colonized with commensal bacteria [45–48], and because of a lack of healthy Ugandan NP and stool samples to serve as controls, only non-bacterial species were reported though we did analyze NP microbiome diversity (see below).

## Genome assembly, annotation and phylogenetic analysis

To more comprehensively characterize the genomes of identified microbes, the paired-read iterative contig extension (PRICE) assembler [16] and the St. Petersburg genome assembler (SPAdes) [64] were used to *de novo* assemble short read sequences into larger contiguous sequences (contigs). Assembled contigs were queried against the National Center for Biotechnology Information (NCBI) nucleotide (nt) database using the basic local alignment search tool (BLAST) to identify the closest related microbes. GenBank annotation files from genome sequence records corresponding to the highest scoring alignments were used to identify potential features within the *de novo* assembled genomes. Geneious v10.3.2 was used to annotate newly assembled genomes. Reference genomes for multiple sequence alignments and phylogenetic analyses were downloaded from NCBI. Multiple sequence (nucleotide) alignments were generated using the default settings in MUSCLEv3.8.1551 [65], and ModelTest-NGv0.1.5 was used to identify the best-fitting evolutionary model. Using the best-fitting model of for evolution, we reconstructed a maximum-likelihood phylogeny using RAXML-ng v0.6.0 using default settings [66]. Annotation of protein domains in the novel orthobunyavirus was performed using the InterPro webserver [67] as well as direct alignment against previously known orthobunyaviruses. The TOPCONS webserver [68] was used for the identification of transmembrane regions and signal peptides, and the NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) for the identification of glycosylation sites.

## Evaluation of NP microbiome diversity

We applied SDI to evaluate alpha diversity of microbes identified in NP samples. For this analysis, patients were stratified into two categories based on clinical assignment: respiratory infections (admitting diagnosis of pneumonia, respiratory tract infection, or bronchiolitis;  $n = 52$ ) and all other syndromes ( $n = 39$ ); cases with unknown admitting diagnosis were excluded. SDI was calculated in R using the Veganv2.4.4 package on genus-level reads per million values for all microbes, including bacteria. A Wilcoxon Rank Sum test was used to evaluate differences in SDI between patients in the two categories.

## Supporting information

**S1 Fig. Simpsons diversity index (SDI) for samples with pneumonia versus other etiologies.** Each triangle represents one sample.

(TIF)

**S2 Fig. Complete phylogenetic tree of Large (L) or RNA dependent RNA polymerase.**

(TIF)

**S3 Fig. Complete phylogenetic tree of Medium (M) segment polyprotein encoding Gn, NSm and Gc proteins.**

(TIF)

**S4 Fig. Complete phylogenetic tree of Small (S) segment encoding Nucleocapsid segments.**

(TIF)

**S1 Table.** (A) Co-infection table for *P. falciparum*, (B) Co-infection table for HRV.

(DOCX)

**S1 Data.** (A) Admission categories of patients enrolled in the study. (B) mNGS findings in patients enrolled in the study. (C) Total rpM identified of microbial species in the serum, NP swab and stool.

(XLSX)

**S2 Data. Total number of sequencing reads and unique non-human reads in all samples analyzed.**

(XLSX)

## Acknowledgments

We thank the children and families in Tororo District, Uganda, who participated in this study. We thank Eric Chow, Jessica Lund, and the UCSF Center for Advanced Technology for assistance with sequencing. We recognize Dr. Amy Kistler and Dr. Senjuti Saha for intellectual discussions and comments on the paper. We thank Dr. Mark Stenglein for his helpful discussions and advice.

## Author Contributions

**Conceptualization:** Sara Nakielny, Theodore Ruel, Arthur Mpimbaza, Philip J. Rosenthal, Joseph L. DeRisi.

**Data curation:** Sara Nakielny, Jennifer Hsu, Mary Kyohere, Oswald Byaruhanga, Philip J. Rosenthal, Joseph L. DeRisi.

**Formal analysis:** Akshaya Ramesh, Sara Nakielny, Katrina Kalantar, Charles Langelier, Michael R. Wilson, Philip J. Rosenthal, Joseph L. DeRisi.

**Funding acquisition:** Theodore Ruel, Arthur Mpimbaza, Philip J. Rosenthal, Joseph L. DeRisi.

**Investigation:** Philip J. Rosenthal, Joseph L. DeRisi.

**Methodology:** Akshaya Ramesh, Sara Nakielny, Philip J. Rosenthal, Joseph L. DeRisi.

**Project administration:** Jennifer Hsu, Mary Kyohere, Oswald Byaruhanga, Philip J. Rosenthal, Joseph L. DeRisi.

**Resources:** Philip J. Rosenthal, Joseph L. DeRisi.

**Software:** Charles de Bourcy, Rebecca Egger, Boris Dimitrov, Yun-Fang Juan, Jonathan Sheu, James Wang.

**Supervision:** Philip J. Rosenthal, Joseph L. DeRisi.

**Visualization:** Akshaya Ramesh, Katrina Kalantar, Charles Langelier, Joseph L. DeRisi.

**Writing – original draft:** Akshaya Ramesh, Michael R. Wilson, Philip J. Rosenthal, Joseph L. DeRisi.

**Writing – review & editing:** Akshaya Ramesh, Sara Nakielny, Katrina Kalantar, Charles Langelier, Theodore Ruel, Arthur Mpimbaza, Michael R. Wilson, Philip J. Rosenthal, Joseph L. DeRisi.

## References

1. Maze MJ, Bassat Q, Feasey NA, Mandomando I, Musicha P, Crump JA. The epidemiology of febrile illness in sub-Saharan Africa: implications for diagnosis and management. *Clin Microbiol Infect.* 2018; 23(8):808–814. <https://doi.org/10.1016/j.cmi.2018.02.011> PMID: 29454844
2. Prasad N, Sharples KJ, Murdoch DR, Crump JA. Community prevalence of fever and relationship with malaria among infants and children in low-resource areas. *Am J Trop Med Hyg.* 2015; 93(1):178–180. <https://doi.org/10.4269/ajtmh.14-0646> PMID: 25918207
3. Guidelines for the treatment of malaria, 2nd edition. World Health Organization. 2010.
4. Bibby K. Metagenomic identification of viral pathogens. *Trends Biotechnol.* 2013; 31(5):275–279. <https://doi.org/10.1016/j.tibtech.2013.01.016> PMID: 23415279
5. Yozwiak NL, Skewes-Cox P, Stenglein MD, Balmaseda A, Harris E, DeRisi JL. Virus identification in unknown tropical febrile illness cases using deep sequencing. *PLoS Negl Trop Dis.* 2012; 6(2):e1485. <https://doi.org/10.1371/journal.pntd.0001485> PMID: 22347512
6. Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, et al. Actionable Diagnosis of Neuroleptospirosis by Next-Generation Sequencing. *N Engl J Med.* 2014; 370(25):2408–2417. <https://doi.org/10.1056/NEJMoa1401268> PMID: 24896819
7. Wilson MR, Shanbhag NM, Reid MJ, Singhal NS, Gelfand JM, Sample HA, et al. Diagnosing Balamuthia mandrillaris Encephalitis with Metagenomic Deep Sequencing. *Ann Neurol.* 2015; 78(5):722–730. <https://doi.org/10.1002/ana.24499> PMID: 26290222
8. Doan T, Wilson MR, Crawford ED, Chow ED, Khan LM, Knopp KA, et al. Illuminating uveitis: metagenomic deep sequencing identifies common and rare pathogens. *Genome Med.* 2016; 8:1–2. <https://doi.org/10.1186/s13073-015-0257-9>
9. Quan PL, Wagner TA, Briese T, Torgerson TR, Hornig M, Tashmukhamedova A, et al. Astrovirus encephalitis in boy with X-linked agammaglobulinemia. *Emerg Infect Dis.* 2010; 16(6):918–925. <https://doi.org/10.3201/eid1606.091536> PMID: 20507741
10. Palacios G, Druce J, Du L, Tran T, Birch C, Briese T, et al. A New Arenavirus in a Cluster of Fatal Transplant-Associated Diseases. *N Engl J Med.* 2008; 358:991–998. <https://doi.org/10.1056/NEJMoa073785> PMID: 18256387
11. Wilson MR, O'Donovan BD, Gelfand JM, Sample HA, Chow FC, Betjemann JP, et al. Chronic Meningitis Investigated via Metagenomic Next-Generation Sequencing. *JAMA Neurol.* 2018;94:158. <https://doi.org/10.1001/jamaneurol.2018.0463> PMID: 29710329

12. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9(4):357–359. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286
13. Ye Y, Choi J-H, Tang H. RAPSearch: a fast protein similarity search tool for short reads. *BMC Bioinformatics*. 2011; 12(1):159. <https://doi.org/10.1186/1471-2105-12-159> PMID: 21575167
14. Li W, Godzik A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 2006; 22(13):1658–1659. <https://doi.org/10.1093/bioinformatics/btl158> PMID: 16731699
15. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*. 2010; 26(7):873–881. <https://doi.org/10.1093/bioinformatics/btq057> PMID: 20147302
16. Ruby JG, Bellare P, Derisi JL. PRICE: software for the targeted assembly of components of (Meta) genomic sequence data. *G3 (Bethesda)*. 2013; 3(5):865–880. <https://doi.org/10.1534/g3.113.005967> PMID: 23550143
17. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013; 29(1):15–21. <https://doi.org/10.1093/bioinformatics/bts635> PMID: 23104886
18. Naoumov N V. TT virus—highly prevalent, but still in search of a disease. *J Hepatol*. 2000; 33(13):157–159. [https://doi.org/10.1016/S0168-8278\(00\)80173-7](https://doi.org/10.1016/S0168-8278(00)80173-7)
19. Hafez MM, Shaarawy SM, Hassan AA, Salim RF, Abd El Salam FM, Ali AE. Prevalence of transfusion transmitted virus (TTV) genotypes among HCC patients in Qalubia governorate. *Virol J*. 2007; 4:1–6. <https://doi.org/10.1186/1743-422X-4-1>
20. De Vlaminc I, Khush KK, Strehl C, Kohli B, Luikart H, Neff NF, et al. Temporal response of the human virome to immunosuppression and antiviral therapy. *Cell*. 2013; 155(5):1178–87. <https://doi.org/10.1016/j.cell.2013.10.034> PMID: 24267896
21. De Vlaminc I, Martin L, Kertesz M, Patel K, Kowarsky M, Strehl C, et al. Noninvasive monitoring of infection and rejection after lung transplantation, *Proc. Natl. Acad. Sci*. 2015; 112(43):13336–41. <https://doi.org/10.1073/pnas.1517494112> PMID: 26460048
22. McElvania TeKippe E, Wylie KM, Deych E, Sodergren E, Weinstock G, and Storch GA. Increased Prevalence of Anellovirus in Pediatric Patients with Fever. *PLoS One*. 2012; 7 (11): e50937. <https://doi.org/10.1371/journal.pone.0050937> PMID: 23226428
23. Abreu NA, Nagalingam NA, Song Y, Roediger FC, Pletcher SD, Goldberg AN, et al. Sinus Microbiome Diversity Depletion and *Corynebacterium tuberculoearicum* Enrichment Mediates Rhinosinusitis. *Sci Transl Med*. 2012; 4(151):151ra124–151ra124. <https://doi.org/10.1126/scitranslmed.3003783> PMID: 22972842
24. Langelier C, Zinter M, Kalantar K, Yanik G, Christenson S, Odonovan B, et al. Metagenomic Next-Generation Sequencing Detects Pulmonary Pathogens in Hematopoietic Cellular Transplant Patients with Acute Respiratory Illnesses. *Am J Respir Crit Care Med*. 2018 Feb 15; 197(4):524–528. <https://doi.org/10.1164/rccm.201706-1097LE> PMID: 28686513
25. Park DE, Baggett HC, Howie SRC, Shi Q, Watson NL, Brooks WA, et al. Colonization density of the upper respiratory tract as a predictor of pneumonia—*Haemophilus influenzae*, *Moraxella catarrhalis*, *Staphylococcus aureus*, and *Pneumocystis jirovecii*. *Clin Infect Dis*. 2017; 64:S328–S336. <https://doi.org/10.1093/cid/cix104> PMID: 28575367
26. Loens K, Van Heirstraeten L, Malhotra-Kumar S, Goossens H, Ieven M. Optimal sampling sites and methods for detection of pathogen possibly causing community-acquired lower respiratory tract infections. *J Clin Microbiol*. 2009; 47(1):21–31. <https://doi.org/10.1128/JCM.02037-08> PMID: 19020070
27. Zar HJ, Barnett W, Stadler A, Gardner-Lubbe S, Myer L, Nicol MP. Aetiology of childhood pneumonia in a well vaccinated South African birth cohort: A nested case-control study of the Drakenstein Child Health Study. *Lancet Respir Med*. 2016; 4(6):463–472. [https://doi.org/10.1016/S2213-2600\(16\)00096-5](https://doi.org/10.1016/S2213-2600(16)00096-5) PMID: 27117547
28. Lukashev AN, Vakulenko YA. Molecular evolution of types in non-polio enteroviruses. *J Gen Virol*. 2017; 98(12):2968–2981. <https://doi.org/10.1099/jgv.0.000966> PMID: 29095688
29. McIntyre CL, Knowles NJ, Simmonds P. Proposals for the classification of human rhinovirus species A, B and C into genotypically assigned types. *J Gen Virol*. 2013; 94(PART8):1791–1806. <https://doi.org/10.1099/vir.0.053686-0>
30. Chipwaza B, Mugasa JP, Selemani M, Amuri M, Mosha F, Ngatunga SD, et al. Dengue and Chikungunya Fever among Viral Diseases in Outpatient Febrile Children in Kilosa District Hospital, Tanzania. *PLoS Negl Trop Dis*. 2014; 8(11). <https://doi.org/10.1371/journal.pntd.0003335> PMID: 25412076
31. Jacob ST, Pavlinac PB, Nakiyingi L, Banura P, Baeten JM, Morgan K, et al. *Mycobacterium tuberculosis* Bacteremia in a Cohort of HIV-Infected Patients Hospitalized with Severe Sepsis in Uganda—High

- Frequency, Low Clinical Sand Derivation of a Clinical Prediction Score. *PLoS One*. 2013; 8(8). <https://doi.org/10.1371/journal.pone.0070305> PMID: 23940557
32. Crump JA, Morrissey AB, Nicholson WL, Massung RF, Stoddard RA, Galloway RL, et al. Etiology of Severe Non-malaria Febrile Illness in Northern Tanzania: A Prospective Cohort Study. *PLoS Negl Trop Dis*. 2013; 7(7). <https://doi.org/10.1371/journal.pntd.0002324> PMID: 23875053
  33. Chipwaza B, Mhamphi GG, Ngatunga SD, Selemani M, Amuri M, Mugasa JP, et al. Prevalence of Bacterial Febrile Illnesses in Children in Kilosa District, Tanzania. *PLoS Negl Trop Dis*. 2015; 9(5). <https://doi.org/10.1371/journal.pntd.0003750> PMID: 25955522
  34. D'Acremont V, Kilowoko M, Kyungu E, Philipina S, Sangu W, Kahama-Maró J, et al. Beyond Malaria—Causes of Fever in Outpatient Tanzanian Children. *N Engl J Med*. 2014; 370(9):809–817. <https://doi.org/10.1056/NEJMoa1214482> PMID: 24571753
  35. O'Meara WP, Mott JA, Laktabai J, Wamburu K, Fields B, Armstrong J, et al. Etiology of pediatric fever in Western Kenya: A case-control study of falciparum Malaria, Respiratory Viruses, and Streptococcal Pharyngitis. *Am J Trop Med Hyg*. 2015; 92(5):1030–1037. <https://doi.org/10.4269/ajtmh.14-0560> PMID: 25758648
  36. Maina AN, Farris CM, Odhiambo A, Jiang J, Laktabai J, Armstrong J, et al. Q fever, scrub typhus, and rickettsial diseases in children, Kenya, 2011–2012. *Emerg Infect Dis*. 2016; 22(5):883–886. <https://doi.org/10.3201/eid2205.150953> PMID: 27088502
  37. Kibuuka A, Byakika-Kibwika P, Achan J, Yeka A, Nalyazi JN, Mpimbaza A, et al. Bacteremia among febrile ugandan children treated with antimalarials despite a negative malaria test. *Am J Trop Med Hyg*. 2015; 93(2):276–280. <https://doi.org/10.4269/ajtmh.14-0494> PMID: 26055736
  38. Prasad N, Murdoch DR, Reyburn H, Crump JA. Etiology of severe febrile illness in low- and middle-income countries: A systematic review. *PLoS One*. 2015; 10(6):1–25. <https://doi.org/10.1371/journal.pone.0127962.38>
  39. Baba M, Logue CH, Oderinde B, Abdulmaleek H, Williams J, Lewis J, et al. Evidence of arbovirus co-infection in suspected febrile malaria and typhoid patients in Nigeria. *J Infect Dev Ctries*. 2013 Jan 15; 7(1):51–9. <https://doi.org/10.3855/jidc.2411> PMID: 23324821
  40. Oguttu DW, Matovu JKB, Okumu DC, Ario AR, Okullo AE, Opigo J, et al. Rapid reduction of malaria following introduction of vector control interventions in Tororo District, Uganda: a descriptive study. *Malar J*. 2017; 16(1):1–8. <https://doi.org/10.1186/s12936-016-1650-6>
  41. Mekonnen SK, Aseffa A, Medhin G, Berhe N, Velavan TP. Re-evaluation of microscopy confirmed *Plasmodium falciparum* and *Plasmodium vivax* malaria by nested PCR detection in southern Ethiopia. *Malar J*. 2014; 13(48). <https://doi.org/10.1186/1475-2875-13-48> PMID: 24502664
  42. Agarwal R, Baid R, Datta R, Saha M, Sarkar N. Falciparum malaria and parvovirus B19 coinfection: A rare entity. *Trop Parasitol*. 2017; 7(1):47–48. <https://doi.org/10.4103/2229-5070.202299> PMID: 28459015
  43. Duedu KO, Sagoe KWC, Ayeh-Kumi PF, Afrim RB, Adiku T. The effects of co-infection with human parvovirus B19 and *Plasmodium falciparum* on type and degree of anaemia in Ghanaian children. *Asian Pac J Trop Biomed*. 2013; 3(2):129–139. [https://doi.org/10.1016/S2221-1691\(13\)60037-4](https://doi.org/10.1016/S2221-1691(13)60037-4) PMID: 23593592
  44. Toan NL, Sy BT, Song LH, Luong H V., Binh NT, Binh VQ, et al. Co-infection of human parvovirus B19 with *Plasmodium falciparum* contributes to malaria disease severity in Gabonese patients. *BMC Infect Dis*. 2013; 13(1). <https://doi.org/10.1186/1471-2334-13-375> PMID: 23945350
  45. Bogaert D, Keijser B, Huse S, Rossen J, Veenhoven R, van Gils E, et al. Variability and diversity of nasopharyngeal microbiota in children: A metagenomic analysis. *PLoS One*. 2011; 6(2). <https://doi.org/10.1371/journal.pone.0017035> PMID: 21386965
  46. Pérez-Losada M, Alamri L, Crandall KA, Freishtat RJ. Nasopharyngeal microbiome diversity changes over time in children with asthma. *PLoS One*. 2017; 12(1):1–13. <https://doi.org/10.1371/journal.pone.0170543> PMID: 28107528
  47. Hooper LV and Gordon JI. Commensal Host-Bacterial Relationships in the Gut. *Science*. 2001; 292(5519):1115–1118. <https://doi.org/10.1126/science.1058709> PMID: 11352068
  48. Sekirov I, Russell SL, Antunes LC and Finlay BB. Gut microbiota in health and disease. *Physiol Rev*. 2010. 90(3):859–904. <https://doi.org/10.1152/physrev.00045.2009> PMID: 20664075
  49. Onyango CO, Welch SR, Munywoki PK, Agoti CN, Bett A, Ngama M, et al. Molecular seroprevalence of human rhinovirus infections in Kilifi, coastal Kenya. *J Med Virol*. 2012; 84(5):823–831. <https://doi.org/10.1002/jmv.23251> PMID: 22431032
  50. O'Callaghan-Gordo C, Bassat Q, Morais L, Díez-Padrís N, MacHevo S, Nhampossa T, et al. Etiology and epidemiology of viral pneumonia among hospitalized children in rural mozambique: A malaria



- endemic area with high prevalence of human immunodeficiency virus. *Pediatr Infect Dis J*. 2011; 30(1):39–44. <https://doi.org/10.1097/INF.0b013e3181f232fe> PMID: 20805786
51. Niang MN, Diop OM, Sarr FD, Goudiaby D, Malou-Sompy H, Ndiaye K, et al. Viral Etiology of Respiratory Infections in Children Under 5 Years Old Living in Tropical Rural Areas of Senegal: The EVIRA Project. *J Med Virol*. 2010; 82:866–872. <https://doi.org/10.1002/jmv.21665> PMID: 20336732
  52. Smuts HE, Workman LJ, Zar HJ. Human rhinovirus infection in young African children with acute wheezing. *BMC Infect Dis*. 2011; 11(1):65. <https://doi.org/10.1186/1471-2334-11-65> PMID: 21401965
  53. Jain S, Self WH, Wunderink RG, Fakhran S, Balk R, Bramley AM, et al. Community-Acquired Pneumonia Requiring Hospitalization among U.S. Adults. *N Engl J Med*. 2015; 373(5):415–427. <https://doi.org/10.1056/NEJMoa1500245> PMID: 26172429
  54. Scully EJ, Basnet S, Wrangham RW, Muller MN, Oтали E, Hyeroba D, et al. Lethal Respiratory Disease Associated with Human Rhinovirus C in Wild Chimpanzees, Uganda, 2013. *Emerg Infect Dis*. 2018; 24(2). <https://doi.org/10.3201/eid2402.170778> PMID: 29350142
  55. Diarrhoeal Disease. World Health Organization. 2017. <http://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease>. Accessed 20 June 2018.
  56. Mwenda JM, Burke RM, Shaba K, Mihigo R, Tevi-Benissan MC, Mumba M, et al. Implementation of Rotavirus Surveillance and Vaccine Introduction—World Health Organization African Region, 2007–2016. *MMWR Morb Mortal Wkly Rep*. 2017; 66(43):1192–1196. <https://doi.org/10.15585/mmwr.mm6643a7> PMID: 29095805
  57. Calisher CH, Monath TP, Sabbatini MS, Mitchell CJ, Lazuick JS, Tesh RB, et al. A newly recognized vesiculovirus, Calchaqui virus, and subtypes of Melao and Maguari viruses from Argentina, with serologic evidence for infections of humans and horses. *Am J Trop Med Hyg*. 1987; 36(1):114–119. <https://doi.org/10.4269/ajtmh.1987.36.114> PMID: 2880522
  58. Mohamed M, McLees A, Elliott RM. Viruses in the Anopheles A, Anopheles B, and Tete Serogroups in the Orthobunyavirus Genus (Family Bunyviridae) Do Not Encode an NSs Protein. *J Virol*. 2009; 83(15):7612–7618. <https://doi.org/10.1128/JVI.02080-08> PMID: 19439468
  59. Williams JE, Imlarp S, Top FH, Cavanaugh DC, Russell PK. Kaeng Khoi virus from naturally infected bedbugs (Cimicidae) and immature free tailed bats. *Bull World Health Organ*. 1976; 53(4):365–369. PMID: 1086729
  60. Plyusnin A, Elliott RM. *Bunyaviridae: Molecular and Cellular Biology*; 2011.
  61. Lutwama JJ, Rwaguma EB, Nawanga PL, Mukuye A. Isolations of Bwamba virus from south central Uganda and north eastern Tanzania. *Afr Health Sci*. 2002; 2(1):24–28. PMID: 12789111
  62. Yozwiak NL, Skewes-Cox P, Gordon A, Saborio S, Kuan G, Balmaseda A, et al. Human Enterovirus 109: a Novel Interspecies Recombinant Enterovirus Isolated from a Case of Acute Pediatric Respiratory Illness in Nicaragua. *J. Virol*. 2010; 84(18):9047–58. <https://doi.org/10.1128/JVI.00698-10> PMID: 20592079
  63. Wilson MR, Fedewa G, Stenglein MD, Olejnik J, Rennick LJ, Nambulli S, et al. Multiplexed Metagenomic Deep Sequencing To Analyze the Composition of High-Priority Pathogen Reagents. *mSystems*. 2016; 1(4):e00058–16. <https://doi.org/10.1128/mSystems.00058-16> PMID: 27822544
  64. Bankevich A, Nurk S, Antipov D, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol*. 2012; 19(5):455–477. <https://doi.org/10.1089/cmb.2012.0021> PMID: 22506599
  65. Edgar R. C.. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004; 32(5):1792–7. <https://doi.org/10.1093/nar/gkh340> PMID: 15034147
  66. Kozlov A., Darriba D., Flouri T., Morel B., and Stamatakis A. RAxML-NG: A fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. *bioRxiv*. 2018. pp. 1–5.
  67. Mitchell A, Chang HY, Daugherty L, et al. The InterPro protein families database: The classification resource after 15 years. *Nucleic Acids Res*. 2015; 43(D1):D213–D221. <https://doi.org/10.1093/nar/gku1243> PMID: 25428371
  68. Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res*. 2015; 43(W1):W401–W407. <https://doi.org/10.1093/nar/gkv485> PMID: 25969446