1 Spatio-temporal distribution of rhinovirus types in Kenya: A retrospective 2 analysis. 2014 3 John Mwita Morobe^{1,*}, Everlyn Kamau², Martha M. Luka¹, Nickson Murunga¹, Clement 4 Lewa¹, Martin Mutunga¹, Godfrey Bigogo³, Nancy Otieno³, Bryan Nyawanda³, Clayton 5 6 Onyango⁴, D. James Nokes^{1,5}, Charles N. Agoti^{1,6}, Patrick K. Munywoki^{1,4}, 7 8 Affiliations 9 ¹ Epidemiology and Demography Department, KEMRI-Wellcome Trust Research 10 Programme, Centre for Geographic Medicine Research – Coast, Kilifi, Kenya. 11 ² Nuffield Department of Medicine, University of Oxford, Oxford, UK 12 ³ KEMRI- Centre for Global Health Research, Kisumu, Kenya. 13 ⁴ U.S Centers for Disease Control and Prevention (CDC), Centers for Global Health, Division 14 of Global Health Protection, Nairobi, Kenya. 15 ⁵ School of Life Sciences and Zeeman Institute for Systems Biology and Infectious Disease 16 Epidemiology Research (SBIDER), University of Warwick, Coventry, United Kingdom. 17 ⁶ Department of Public Health, Pwani University, Kilifi, Kenya. 18 Corresponding author: Jmwita@kemri-wellcome.org 19 20 21 Key words: rhinovirus, phylogenetics, transmission, spatio-temporal distribution, Kenva 22 23 24 25 26 NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

27 Abstract

28 Rhinoviruses (RVs) are one of the most commonly detected viruses in people with acute respiratory illness (ARI). Despite their significant disease burden, RV 29 30 epidemiology at national levels is underexplored. The circulation patterns of RV types throughout a population and the role of virus genotype in this distribution are ill-31 defined. We generated 803 VP4/VP2 gene sequences from rhinovirus-positive 32 33 samples collected from ARI patients, including both in-patient and outpatient cases, between 1st January and 31st December 2014 from eleven surveillance sites across 34 35 Kenya and used phylogenetics to characterise virus introductions and spread. RVs 36 were detected throughout the year, with the highest detection rates observed from 37 January to March and June to July. We detected a total of 114 of the 169 currently 38 classified types. Our analysis revealed numerous virus introductions into Kenya 39 characterized by local expansion and extinction, and extensive spatial mixing of types within the country due to the widespread transmission of the virus after an introduction. 40 41 This work demonstrates that in a single year, the circulation of rhinovirus in Kenya was 42 characterized by substantial genetic diversity, multiple introductions, and extensive 43 geographical spread.

44 Introduction

Rhinoviruses (RVs) are one of the leading cause of acute upper respiratory tract illnesses, commonly referred to as "common cold"^{1,2}. RV infections occur all year, although they peak in early autumn and late spring in many temperate countries, and in the rainy seasons in tropical countries^{2,3}. RVs cause 8 – 12 episodes of respiratory infections in children and 2 – 4 episodes in adults each year ^{4–6}, posing significant social and economic burden due to time lost from work or school, medical attendance, and reduced performance of regular duties⁷. Despite the high prevalence and clinical

importance of RV infections, little is known about the patterns of occurrence of RV
 types, as well as the role of virus type in the RV distribution.

RVs belong to the genus Enterovirus of the family Picornaviridae. The virus 54 genome consists of ~7200 nucleotides and encodes four structural proteins (VP4, 55 VP2, VP3 and VP1) and seven non-structural proteins (2A, 2B, 2C, 3A, 3B, 3C and 56 57 3D)². The single-stranded positive sense RNA genome is enclosed within a protein 58 capsid that is made up of sixty copies of each of the four structural proteins, of which VP1, VP2, and VP3 are exposed outside, while VP4 is completely masked in the 59 60 capsid^{2,8,9}. The three surface-exposed capsid proteins carry the antigenically important sites^{2,8,9}. Genetic variation in the VP4/VP2 and VP1 genomic regions have 61 62 been useful in molecular typing^{10,11} and epidemiological investigations^{12–14}. As of 15th 63 March 2023, a total of 169 RV types have been described and classified into three distinct species, i.e., Rhinovirus A, Rhinovirus B, and Rhinovirus C¹⁵. 64

Viral sequence data are increasingly used to track the geographic spread and 65 transmission of viral pathogens^{16,17}. Genomic data has considerably informed public 66 health interventions and outbreak management for viral pathogens, e.g., SARS-CoV-67 2, Ebolavirus, and Zika virus^{18–21}. Widespread spatial and temporal transmission 68 patterns of RV was previously described within the Kilifi Health and Demographic 69 Surveillance System (KHDSS) on the Kenyan Coast covering around 300,000 70 residents²² and marked by multiple virus introductions¹². Transmission patterns of RV 71 at a nationwide scale have not been documented in Kenya. Here we analyze 803 72 73 sequences obtained from individuals presenting with acute respiratory illness (ARIs) 74 at 11 sentinel surveillance sites across Kenya in 2014²³, to explore the temporal and spatial circulation patterns of RV types at the countrywide scale. 75

76

77 Methods

78 Ethical considerations

The study was approved by the Kenya Medical Research Institute - Scientific and Ethical Review Unit (SSC #3044) and CDC Institutional Review Board (#6806) to use pre-existent, pseudonymized specimens, and data. All individuals, parents and guardians gave written informed consent for themselves or their children to participate in the original studies²³. For older children aged 13–17 years, assent was obtained as part of the individual informed consenting process.

85

86 Surveillance sites

This study utilised samples collected throughout the surveillance period covering 1st 87 January to 31st December 2014 from three health facility-based respiratory virus 88 89 surveillance programs in Kenya (Table 1 and Figure 1). The first program included 90 eight inpatient hospitals participating in the influenza sentinel surveillance : Dadaab 91 Refugee Camp (RC), Kakuma RC, Kenyatta National Hospital (KNH), Nakuru County 92 Referral Hospital (CRH), Kakamega CRH, Nyeri CRH, Siaya CRH and Coast General Teaching and Referral Hospital^{24–26}; the second included two outpatient clinics; Lwak 93 Mission Hospital, Asembo and Tabitha Medical Clinic, Kibera under Population-Based 94 Infectious Disease Surveillance (PBIDS)²⁷, and the third program was a pediatric 95 96 inpatient sampling at the Kilifi County Hospital (KCH)²⁸. The influenza sentinel 97 surveillance was established by Kenya Medical Research Institute-Centre for Global 98 Research (KEMRI-CGHR), the US Centers for Disease Control and Prevention 99 (CDC)-Kenya, and the Ministry of Health as part of the Global Influenza Surveillance and Response System since 2006²⁴. The PBIDS platform in Asembo and Kibera is 100 run by KEMRI-CGHR with financial and technical support from CDC since 2006²⁹. 101

Surveillance at KCH is conducted by the KEMRI-Centre for Geographic Medical
Research Coast (CGMRC) under KEMRI-Wellcome Trust Research Programme
(KWTRP) in Kilifi, Kenya²⁸. Overall, the selected surveillance sites offer a good
representation of the various geographical settings in Kenya.

106

107 **Patient enrollment**

108 The surveillance sites recruited patients of various ages presenting with acute respiratory illness (ARI) with a measured fever of ≥38°C and a cough with an onset of 109 110 symptoms within the last 7 days (influenza-like illness, ILI) or acute respiratory illness 111 requiring hospitalization with a history of fever or measured fever \geq 38°C and a cough 112 with an onset of symptoms within the last 14 days (severe acute respiratory illness, 113 SARI), or acute lower respiratory tract illness (ALRTI) defined as the presence of 114 cough OR difficulty in breathing with one of the following danger signs: chest in-115 drawing, stridor, unable to breastfeed, vomit everything, convulsions, lethargy, or 116 unconsciousness or an adaptation of WHO severe/very severe pneumonia as 117 described in a **Table 1**. The different case definitions were consistently applied within 118 each platform over the course of the study period.

119 Data and specimen collection

Patient demographic data and clinical features of presenting illness were collected in real-time using custom designed databases. Nasopharyngeal (NP) and/or oropharyngeal (OP) swabs were collected from eligible patients who presented to health facilities with clinical features of an acute respiratory illness as describe in patient enrolment section. NP/OP samples were stored in viral transport media, and transported to the laboratory for long-term storage at -80°C.

126

127 **RV screening and sequencing**

Viral RNA was extracted using RNeasy Mini Kit (Qiagen, Germany) in Qiacube HT as 128 129 per the manufacturer's instructions and screened for respiratory viruses (i.e. RSV 130 (groups A and B), rhinovirus, human coronaviruses (hCoV-OC43, -NL63, -229E), influenza A virus and adenovirus) using a multiplex real-time reverse-transcription 131 PCR (rt-RT-PCR) assay^{30,31}. A sample was considered RV positive if the rt-RT-PCR 132 cycle threshold (Ct) was $\leq 35.0^{32}$. The VP4/VP2 viral genomic region (~420 nucleotides 133 long) of positive samples was amplified and sequenced as previously described³³. 134 Consensus sequences were assembled using the Sequencher software version 5.4.6 135 136 (Gene Codes Corporation, Ann Arbor, USA).

137

138 **RV** species and type assignment

We used the term 'type' to refer to RV sequences classified as distinct by genetic 139 comparisons as described previously¹¹. Sequences were assigned into the same RV 140 141 type based on >90% nucleotide similarity to rhinovirus prototype sequences (also referred to as reference sequences¹¹ and phylogenetic clustering with bootstrap 142 support value >70%¹¹. Distributions of pairwise genetic distances were assessed for 143 144 evaluation of intertype and intra-type divergence¹¹. Intra-type 'variant' was defined based on a divergence cut-off or threshold value determined as the least frequent 145 146 value between the first (same phylogenetic clade) and second (different phylogenetic clades) modes in a pairwise nucleotide difference distribution plot as previously 147 described³⁴. Viruses within the same phylogenetic clade were assumed to belong to 148 149 the same variant of an RV type.

150

151 Global dataset

152 Rhinovirus VP4/VP2 sequences from other regions around the world deposited in 153 GenBank as of 31 December 2021, whose sequenced regions overlapped the 803 154 Kenyan sequences generated in this study and derived from viruses sampled between 1st January 2005 and 31st December 2014, were collated and phylogenetically 155 156 compared with the Kenya virus sequences. The complete comparison global data set 157 comprised 4,448 VP4/VP2 sequences from 40 countries, including 653 sequences 158 from Kenya (Kilifi (n=627), Nairobi (n=10), Mombasa (n=4), Malindi (n=3), Alupe (n=3), 159 Kisumu (n=2), Isiolo (n=2), Kisii (n=1) and Kericho (n=1) (Appendix Table 1).

160

161 **Phylogenetic analysis**

Multiple sequence alignments (MSA) were generated using MAFFT v7.220³⁵ and maximum likelihood phylogenetic trees estimated using IQ-TREE v1.6.12³⁶. Branch support was assessed by 1000 bootstrap iterations. The temporal signal in the data was examined using TempEst v1.5.3³⁷. TreeTime v.0.8.1 was used to transform the ML tree topologies into dated trees. Phylogenetic trees were visualized using ggtree v2.2.4 package³⁸ in R³⁹.

168

169 Statistical analysis.

Statistical analysis was conducted using STATA version 15.1 (College Station,
Texas). Site-specific monthly percent virus positivity was computed.

- 172
- 173
- 174
- 175
- 176

177 Type diversity calculation

RV type diversity was calculated using diversity indices (Shannon (H) and Simpson 178

- (D)) 40,41 . The calculations were as follows: 179
- 180 Shannon Index (H) = - $\sum_{i=1}^{n} p_i \ln p_i$ 181 182
- 183

184 Simpson Index (D) =
$$_{s}^{1/}$$

185 $\sum_{i=1}^{186} p_{i2}$

- 185 186
- 187

188 Where p is the proportion (n/N) of individual types found in one individual species (n)

divided by the total number of types found (N), and s is the total number of species. 189

190 Higher scores of H and D indicates high diversity.

191 **Results**

Between 1st January and 31st December 2014, a total of 6398 NP and/or OP swabs 192 193 samples were collected from patients who voluntarily presented to the eleven facilities 194 with ARI. Testing was done on 5859 (91.6%) available samples, Table 2. Of the tested 195 samples, 5665 (96.7%) were linked with their respective demographic and clinical 196 data. Two hundred and fourteen samples (214) collected from Siava and Kibera 197 patients with missing data on respiratory symptoms were excluded from the study, the 198 final analytical dataset comprised of 5451 specimens. RV was detected in 17.0% 199 (924/5451) of samples collected from the 11 surveillance sites (Table 2). The percent 200 of samples that were virus positive for RV infections varied by surveillance site; Siaya 201 recorded the highest positives (24.0%) while Dadaab (10.4%) recorded the lowest 202 (Table 2). Rhinovirus was detected throughout the year in most sites, although the 203 prevalence of detections varied by month of sampling due to seasonal variation in 204 SARI (Figure 2). Different sites experienced peak occurrence in different months, but a majority saw infection peaks between January - March (7/11) and June - July (5/11) 205 206 (Figure 2). There was no RV detected in April in Kenyatta National Hospital (KNH) and 207 in May in Kakuma and Kibera sites (Figure 2). There were no samples collected in Dadaab between July and September. 208

VP4/VP2 sequences were successfully obtained from 803/924 (87.0%) RV positive samples (Table 2). The remaining samples either failed to amplify with the VP4/VP2 specific primers (n=111) or were identified as non-RV enteroviruses (n=10). Overall, 492 (61.3%) sequences were classified as *Rhinovirus A* comprising 58 types; 63 (7.8%) sequences were *Rhinovirus B* comprising 16 types, and 248 (30.9%) were *Rhinovirus C* comprising 40 types. The most detected types were A34 (n = 35), A22 (n=30), A58 (n = 27), A12 (n = 25) and A78 (n = 23) (Supplementary Table 1). The 216 number of types largely reflected the number of samples sequenced and varied 217 between sites (Supplementary Figure 1A). Comparison by care setting (i.e., outpatient 218 vs. inpatient) or ARI definition (i.e., ILI vs. SARI vs. ALRTI) across different RV types revealed no significant differences in ARI definition (p-value 0.06) and care setting (p-219 220 value 0.2) (Supplementary Table 2). Rhinovirus type diversity per site calculated using 221 the Shannon (H) and Simpson (D) indices indicated that Siava had the highest value 222 of diversity, followed by Lwak and Kilifi. Dadaab reported the lowest value of diversity 223 (Table 3). All three RV species were detected in all sites, except for Dadaab, where 224 Rhinovirus B was not detected (Supplementary Figure 1B). The proportions of RV species were similar in all the sites; *Rhinovirus A* (range, 51 - 75%) and *Rhinovirus C* 225 226 (range, 17 - 46%) were frequently detected, while *Rhinovirus B* (range, 0 - 13%) 227 infections were low or not detected (Supplementary Figure 1B).

228

229 Spatial-temporal distribution of RV types in Kenya

230 We identified up to 49 unique RV types co-circulating in a single month across the country, and/or up to 20 unique types within a single location in a single month 231 232 (Supplementary Figure 2). The duration of circulation varied by RV type; several types occurred at a high prevalence and for consecutive months, while others occurred once 233 234 or intermittently during the study period (Supplementary Figure 3). For example, RV-235 A22 was detected throughout the year, A34 was present for 11 consecutive months (February to December 2014), and RV-C7 and A21 circulated consecutively for 8 236 237 months (February to September and May to December, respectively) (Supplementary 238 Figure 3). Spatially, several RV types circulated widely; some circulated in multiple sites during the same timeframe, while others circulated in multiple sites at different 239 240 times (Figure 3). For example, RV-A22 was detected in all 11 sites, while A34 was detected in 10 sites. We observed the occurrence of localized type-specific epidemics
or outbreaks; the distribution of RV types was similar between neighbouring locations
(Figure 3). For example, types A21 and A78 circulated in Asembo, Siaya, and
Kakamega in the same period (June-July), while A49 was seen in Mombasa and Kilifi
between July and December (Figure 3). Other types were more random in occurrence
with no discernable temporal or spatial pattern.

247

248 **Phylogenetic clustering**

249 We reconstructed time-scaled phylogenies for thirteen prevalent types. In the global 250 context, Kenyan viruses formed monophyletic clusters/clades containing sequences 251 from different sampling sites in Kenya, suggesting multiple introductions and local 252 transmission chains (Figure 4). For example, the Kenyan RV-A22 viruses collected 253 between 2008 and 2014 separated into 3 major clusters, each comprising sequences 254 from different locations in Kenya. Similarly, A34 viruses collected in 2014 separated 255 into 2 major clusters, each with sequences from different sites (Figure 4). Similar 256 observations were observed for other RV types including, A12, A49, A58, A78, and 257 C10 (Figure 4). Certain clades or clusters were location-specific and genetically distinct from other Kenyan 2014 sequences. For example, an A78 variant from Nyeri 258 259 and an A58 variant in Dadaab (Figure 4), showing evidence of localised transmission 260 clusters.

261 Sequences originating from Kenya did not form a single monophyletic group on 262 the phylogenetic tree, instead, they were interspersed mostly as singleton or clusters 263 located separately or with clusters of viruses from other countries. For example, RV-264 A22 sequences from Kenya clustered closely with sequences from China and France 265 in 2014, while A12 sequences clustered closely with sequences from China, the USA, France, Uganda, and Nepal in 2014. The Kenyan A58 sequences clustered closely with those from USA and France, while the A78 sequences clustered closely with Chinese and American sequences, and the A34 resembled Nepalese sequences.

269

270 **Discussion**

This study provides a detailed description of the spatial-temporal dynamics of introductions and spread of rhinovirus in Kenya based on phylogenetic analyses. We show that the persistent circulation of RV in Kenya was mostly driven by multiple introductions of different types throughout the one-year period, leading to established local transmission.

276 Rhinovirus A was the predominant species circulating in Kenya in 2014, 277 followed by *Rhinovirus C*, and the least common was *Rhinovirus B*. This is consistent 278 with a previous surveillance study across 8 sampling sites in Kenya in 2008⁴². 279 Although differences in subject recruitment strategies may not allow one-on-one 280 comparison between this study and others, these observations are similar to studies reported in Africa, USA, Asia, Europe, the Middle East, and Australia^{11,43}. The 281 282 similarities in frequencies of RV species in diverse geographies would suggest that RV circulates unrestricted globally. 283

RV was detected all year, with the highest detection rates observed in January to March and June to July. The year-round occurrence of rhinovirus in the country was largely sustained by contemporaneous and successive mini epidemics or outbreaks caused by distinct types and variants that were probably introduced separately into the country or diversified locally after a single introduction.

289 Spatial-temporal analysis revealed occurrence of multiple types and variants 290 across space and time in agreement with previous studies^{12,14,44}. The observed 291 variations in time and duration of circulation among the rhinovirus types could be due 292 to (i) differences in the duration of type-specific immunity, (ii) frequency of introductions or, iii) level of antigenic similarity (heterologous immunity)³⁴. The concurrent and 293 294 sequential circulation of RV types seen as the occurrence of multiple peaks of the 295 same type in the same or distant locations could signify multiple introductions, antigenic variation, or infections in different population strata³⁴. The heterogeneity in 296 297 RV detection by month and location could be attributed to the seasonal variation in 298 SARI or sampling methods or to regional differences in environmental and climatic 299 factors. Climatic factors such as temperature, rainfall, and relative humidity have been hypothesized to influence RV activity in the tropics^{3,45,46}. 300

301 Even with non-uniform sampling and short duration of study (1 year), we 302 showed close genetic association among sequences from different sites as a result of 303 widespread transmission of the virus in the country. On a finer spatial scale, study 304 sites occurring within the same geographical region (as described in Table 1) had a 305 similar distribution of RV types, an indicator of a point source outbreak, in which a 306 single virus type enters a location and diffuses through the interconnected populations probably sharing social amenities¹². This agrees with our previous work that showed 307 greater similarity in RV types among the locations in proximity¹². 308

The global phylogenies showed that RV viruses circulating in Kenya were closely related to strains circulating in Europe, Asia, and North America. The Kenyan RV diversity appears to be nested within the global diversity as a result of transmission facilitated by unrestricted movement, increased connectivity, and social mixing.

Although our analyses were limited to the VP4/VP2 genomic region, we highlight the use of sequence data to trace the introduction and spread of rhinovirus at the countrywide level and show the benefit of systematic, continuous, and geographically representative surveillance to detect and monitor the occurrence of types at a larger scale. Whole genome sequencing could provide more insight into virus diversity and transmission⁴⁷, and we recommend that future studies should combine genomic data with epidemiological and anthropological data (e.g., host migration, immunity profiles, population densities, and social contact patterns) to further elucidate patterns of RV infections.

322 This study had some limitations. First, due to the retrospective approach of the study, it was not possible to recover sequences from all the samples due to sample 323 324 degradation; samples (n=111/924) that failed sequencing had considerably higher Ct 325 values (low viral load) compared to samples that were successfully sequenced with 326 low Ct values (high viral load). Second, there were varying study design by sites, for 327 instance, the disease case definition and ages enrolled. The varying designs would 328 mainly affect the virus positivity estimates observed over the one-year surveillance 329 period, in that sites that had more inclusive case definition that includes mild cases of 330 respiratory illnesses may exhibit higher rates of rhinovirus positivity compared to sites 331 using a more restrictive case definition that only includes severe cases of respiratory 332 illnesses. Third, the inclusion of fever as a criterion in most of the case definitions used may lead to the underestimation of RV prevalence since many RV infections, including 333 334 severe cases, do not necessarily manifest with fever. Consequently, there is a 335 possibility that we might have missed RV patients who do not exhibit fever symptoms, 336 potentially impacting the overall assessment of RV prevalence in our study. Fourth, 337 the short sequence fragment analysed in this study may result in spurious 338 phylogenetic connections. Lastly, the study focused on samples that were collected in 339 hospital settings, which may have missed genotypes associated with asymptomatic or 340 mildly ill infections.

341 In conclusion, our study demonstrates that sustained circulation of the virus in 342 2014 was due to frequent introductions of different types and variants into the country followed by local spread for some of these introductions. Temporal differences in 343 344 persistence of rhinovirus types over the one-year period, could be attributed to 345 differences in the frequency and number of virus introductions into the country. Spatial 346 patterns show extensive spread of the virus, and the evidence of similar distribution of 347 types in locations that are in proximity to each other may indicate local partitioning or spatial structures of virus transmission. 348

349 **Reference**

- Ouédraogo S, Traoré B, Nene Bi ZAB, et al. Viral Etiology of Respiratory Tract
 Infections in Children at the Pediatric Hospital in Ouagadougou (Burkina Faso). *PLoS One.* 2014;9(10):e110435.
- Jacobs SE, Lamson DM, St George K, Walsh TJ. Human rhinoviruses. *Clin Microbiol Rev.* 2013;26(1):135-162. doi:10.1128/CMR.00077-12
- Garcia J, Espejo V, Nelson M, et al. Human rhinoviruses and enteroviruses in influenza-like illness in Latin America. *Virol J.* 2013;10:305. doi:10.1186/1743-422X-10-305
- 358
 4. Nguyen DNT, Mai LQ, Bryant JE, et al. Epidemiology and etiology of influenza-likeillness in households in Vietnam; it's not all about the kids! *J Clin Virol*. 2016;82:126132. doi:10.1016/j.jcv.2016.07.014
- Kamau E, Onyango CO, Otieno GP, et al. An Intensive, Active Surveillance Reveals
 Continuous Invasion and High Diversity of Rhinovirus in Households. *J Infect Dis.* 2019;219(7):1049-1057. doi:10.1093/infdis/jiy621
- Arruda E, Pitkaranta A, Witek TJJ, Doyle CA, Hayden FG. Frequency and natural
 history of rhinovirus infections in adults during autumn. *J Clin Microbiol*.
 1997;35(11):2864-2868.
- Fendrick AM, Monto AS, Nightengale B, Sarnes M. The economic burden of noninfluenza-related viral respiratory tract infection in the United States. *Arch Intern Med.*2003;163(4):487-494. doi:ioi20041 [pii]
- Esneau C, Bartlett N, Bochkov YA. Chapter 1 Rhinovirus structure, replication, and
 classification. In: Bartlett N, Wark P, Knight D, eds. *Rhinovirus Infections*. Academic
 Press; 2019:1-23. doi:https://doi.org/10.1016/B978-0-12-816417-4.00001-9
- Stobart CC, Nosek JM, Moore ML. Rhinovirus Biology, Antigenic Diversity, and
 Advancements in the Design of a Human Rhinovirus Vaccine. *Front Microbiol*.
 2017;8. doi:10.3389/fmicb.2017.02412
- Simmonds P, McIntyre C, Savolainen-Kopra C, Tapparel C, Mackay IM, Hovi T.
 Proposals for the classification of human rhinovirus species C in genotypically
 assigned types. *J Gen Virol*. 2010;91. doi:10.1099/vir.0.023994-0 LB Simmonds2010
- 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(Pt 8):1791-1806. doi:10.1099/vir.0.053686-0
- Morobe JM, Nyiro JU, Brand S, et al. Human rhinovirus spatial-temporal
 epidemiology in rural coastal Kenya, 2015-2016, observed through outpatient
 surveillance [version 2; peer review: 2 approved]. *Wellcome Open Res.* 2019;3(128).
 doi:10.12688/wellcomeopenres.14836.2
- Arakawa M, Okamoto-Nakagawa R, Toda S, et al. Molecular epidemiological study of
 human rhinovirus species A, B and C from patients with acute respiratory illnesses in
 Japan. J Med Microbiol. 2012;61:410-419. doi:10.1099/jmm.0.035006-0
- 390 14. van der Linden L, Bruning AHL, Thomas X V, et al. A molecular epidemiological 391 perspective of rhinovirus types circulating in Amsterdam from 2007 to 2012. *Clin* 392 *Microbiol Infect*. 2016;22(12):1002.e9-1002.e14. doi:10.1016/j.cmi.2016.08.007
- 393 15. The Picornavirus Pages. Enteroviruses. Enteroviruses. Published 2006. Accessed
 394 March 15, 2023.
- 395 https://www.picornaviridae.com/ensavirinae/enterovirus/enterovirus.htm
- Grubaugh ND, Ladner JT, Lemey P, et al. Tracking virus outbreaks in the twenty-first century. *Nat Microbiol*. 2019;4(1):10-19. doi:10.1038/s41564-018-0296-2

398	17.	Gardy JL, Loman NJ. Towards a genomics-informed, real-time, global pathogen
399		surveillance system. Nat Rev Genet. 2018;19(1):9-20. doi:10.1038/nrg.2017.88
400	18.	Quick J, Loman NJ, Duraffour S, et al. Real-time, portable genome sequencing for
401		Ebola surveillance. Nature. 2016;530(7589):228-232. doi:10.1038/nature16996
402	19.	Eduan W, Marta G, Houriiyah T, et al. A year of genomic surveillance reveals how the
403		SARS-CoV-2 pandemic unfolded in Africa. Science (1979). 2021;374(6566):423-431.
404		doi:10.1126/science.abi4336
405	20	Gire SK, Goba A, Andersen KG, et al. Genomic surveillance elucidates Ebola virus
406	20.	origin and transmission during the 2014 outbreak <i>Science</i> (1979) 2014:345:1369-
400		1372 doi:10.1126/science.1259657
407	21	Faria NR Ouick I Claro IM at al Establishment and cryptic transmission of Zika
400	21.	virus in Brazil and the Americas Nature 2017:546(7658):406.410
409		$d_{0.101} = 1028/naturo 22401$
410	22	Cost IAC Down E Mais IC at al Drafile. The Kilifi Health and Demographic
411	22.	Scott JAG, Bauni E, Moisi JC, et al. Profile: The Kliff Health and Demographic
412		Surveillance System (KHDSS). Int J Epidemiol. 2012;41(3):650-657.
413		doi:10.1093/1je/dys062
414	23.	Murunga N, Nyawanda B, Nyiro JU, et al. Surveillance of respiratory viruses at health
415		facilities from across Kenya, 2014 [version 2; peer review: 1 approved with
416		reservations]. <i>Wellcome Open Res</i> . 2022;7(234).
417		doi:10.12688/wellcomeopenres.17908.2
418	24.	Katz MA, Muthoka P, Emukule GO, et al. Results From the First Six Years of
419		National Sentinel Surveillance for Influenza in Kenya, July 2007–June 2013. PLoS
420		One. 2014;9(6):e98615 https://doi.org/10.1371/journal.pone.0098615
421	25.	Bigogo GM, Breiman RF, Feikin DR, et al. Epidemiology of Respiratory Syncytial
422		Virus Infection in Rural and Urban Kenya. J Infect Dis. 2013;208(suppl_3):S207-
423		S216. doi:10.1093/infdis/jit489
424	26.	Emukule GO, Khagayi S, McMorrow ML, et al. The Burden of Influenza and RSV
425		among Inpatients and Outpatients in Rural Western Kenya, 2009–2012. PLoS One.
426		2014;9(8):e105543 https://doi.org/10.1371/journal.pone.0105543
427	27.	Feikin DR. Olack B. Bigogo GM. et al. The Burden of Common Infectious Disease
428		Syndromes at the Clinic and Household Level from Population-Based Surveillance in
429		Rural and Urban Kenya, <i>PLoS One</i> , 2011:6(1):e16085-
430		https://doi.org/10.1371/journal.pone.0016085
431	28	Nokes DI Ngama M Bett A et al Incidence and severity of respiratory syncytial
432	20.	virus pneumonia in rural Kenvan children identified through hospital surveillance. <i>Clin</i>
432 //33		Infact Dis 2000.40(0).1341-1340 doi:10.1086/606055
121	20	Foikin DP Nienge MK Bigage G at al Etiology and Incidence of Viral and Pacterial
434	29.	A cute Despiratory Illness among Older Children and Adults in Dural Western Kenye
435		2007 2010 DLoS One 2012.7(8):e42656
430		2007-2010. FLOS One. 2012, 7(8).e45050
43/	20	nups://doi.org/10.15/1/journal.pone.0043050
438	30.	Gunson RN, Collins TC, Carman WF. Real-time RT-PCR detection of 12 respiratory
439		viral infections in four triplex reactions. J Clin Virol. 2005;33(4):341-344.
440	21	doi:10.1016/j.jcv.2004.11.025
441	31.	Hammitt LL, Kazungu S, Welch S, et al. Added value of an oropharyngeal swab in
442		detection of viruses in children hospitalized with lower respiratory tract infection. J
443		Clin Microbiol. 2011;49(6):2318-2320. doi:10.1128/JCM.02605-10
444	32.	Onyango CO, Welch SR, Munywoki PK, et al. Molecular epidemiology of human
445		rhinovirus infections in Kilifi, coastal Kenya. J Med Virol. 2012;84(5):823-831.
446		doi:10.1002/jmv.23251

447 33. Morobe JM, Nyiro JU, Brand S, et al. Human rhinovirus spatial-temporal 448 epidemiology in rural coastal Kenya, 2015-2016, observed through outpatient surveillance [version 2; peer review: 2 approved]. Wellcome Open Res. 2019;3(128). 449 doi:10.12688/wellcomeopenres.14836.2 450 Mwita Morobe J, Kamau E, Murunga N, et al. Trends and Intensity of Rhinovirus 451 34. 452 Invasions in Kilifi, Coastal Kenya, Over a 12-Year Period, 2007–2018. Open Forum 453 Infect Dis. 2021;8(12):ofab571. doi:10.1093/ofid/ofab571 454 Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: 35. 455 improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-780. 456 doi:10.1093/molbev/mst010 457 Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective 36. 458 stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 459 2015;32(1):268-274. doi:10.1093/molbev/msu300 460 37. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. Exploring the temporal structure of 461 heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. 2016;2(1). doi:10.1093/ve/vew007 462 463 38. Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. ggtree: an r package for visualization 464 and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol. 2017;8(1):28-36. doi:https://doi.org/10.1111/2041-210X.12628 465 466 39. R Core Team. R: A language and environment for statistical computing. Published 467 online 2013. Accessed November 9, 2021. URL http://www.R-project.org/ 40. Spellerberg IF, Fedor PJ. A tribute to Claude Shannon (1916–2001) and a plea for 468 469 more rigorous use of species richness, species diversity and the 'Shannon-Wiener' 470 Index. Global Ecology and Biogeography. 2003;12(3):177-179. 471 doi:https://doi.org/10.1046/j.1466-822X.2003.00015.x 472 41. SIMPSON EH. Measurement of Diversity. Nature. 1949;163(4148):688. 473 doi:10.1038/163688a0 474 42. Milanoi S, Ongus JR, Gachara G, Coldren R, Bulimo W. Serotype and genetic 475 diversity of human rhinovirus strains that circulated in Kenya in 2008. Influenza 476 Other Respir Viruses. 2016;10(3):185-191. doi:10.1111/irv.12373 477 Briese T, Renwick N, Venter M, et al. Global distribution of novel rhinovirus 43. 478 genotype. Emerg Infect Dis. 2008;14(6):944-947. doi:10.3201/eid1406.080271 479 44. Sansone M, Andersson M, Brittain-Long R, et al. Rhinovirus infections in western 480 Sweden: a four-year molecular epidemiology study comparing local and globally 481 appearing types. Eur J Clin Microbiol Infect Dis. 2013;32(7):947-954. 482 doi:10.1007/s10096-013-1832-x 483 45. Monto AS. Epidemiology of viral respiratory infections. Am J Med. 2002;112 484 Suppl:4S-12S. doi:10.1016/s0002-9343(01)01058-0 485 46. Arden KE, McErlean P, Nissen MD, Sloots TP, Mackay IM. Frequent detection of 486 human rhinoviruses, paramyxoviruses, coronaviruses, and bocavirus during acute 487 respiratory tract infections. J Med Virol. 2006;78(9):1232-1240. 488 doi:https://doi.org/10.1002/jmv.20689 Luka MM, Kamau E, de Laurent ZR, et al. Whole genome sequencing of two human 489 47. 490 rhinovirus A types (A101 and A15) detected in Kenya, 2016-2018. Wellcome Open 491 Research 2021 6:178. 2021;6:178. doi:10.12688/wellcomeopenres.16911.1 492 48. World Health Organisation. Pocket Book of Hospital Care for Children: Guidelines for 493 the Management of Common Childhood Illnesses. 2013;(2nd edition.). 494 495 496

497 Acknowledgements

498	We thank all the study participants for their contribution of study specimens and data
499	and surveillance staff from all the study sites included. We acknowledge the laboratory
500	staff at KEMRI-Wellcome Trust Research Programme and CDC-supported laboratory
501	at KEMRI-Centre for Global Health Research for specimen handling and testing. The
502	primary data and specimen collection in Influenza surveillance and PBIDS sites were
503	supported by U.S. Centers for Disease Control and Prevention while pediatric
504	surveillance at Kilifi County Hospital was funded by the Wellcome Trust.
505	
506	Author Contributions
507	J.M.M, E.K., C.A., PKM, D.J.N., Conceptualised and designed the study
508	J.M.M.,M.M.L, C.L., M.M., Provided laboratory sequencing and support
509	J.M.M., M.M.L, E.K., C.A., D.J.N., Provided phylogenetic analysis
510	J.M.M., N.M. Provided data curation and analysis
511	J.M.M., E.K., C.A., D.J.N., Wrote the original manuscript draft
512	All authors read and approved the final manuscript.
513	
514	Funding
515	This work was supported by Wellcome [102975], and U.S. Centers for Disease Control
516	and Prevention through Cooperative Agreement [#6U01GH002143].
517	
518	Data availability

519 Sequence data generated in this study are available in GenBank under accession
520 numbers: MZ129390 - MZ130096. Additional data and analysis scripts for this

521	manuscript	are	available	at	the	VEC	Harvard	Dataverse:
522	https://doi.org	/10.791	0/DVN/CBHV	<u>TA</u>				

523

524 Conflicts of Interest

- 525 The authors declare no competing interests. The funders had no role in the design of
- 526 the study; in the collection, analyses, or interpretation of data; in the writing of the
- 527 manuscript; or in the decision to publish the results.
- 528

529 **Disclaimer**

- 530 The findings and conclusions in this report are those of the author(s) and do not
- 531 necessarily represent the official position of U.S. the Centers for Disease Control and
- 532 Prevention.

533 Tables

534 **Table 1:** Description of the 11 respiratory surveillance sites across Kenya.

Surveillance site	Region	Setting	Inclusion	Age
			criteria	included
Kenyatta National Hospital (KNH)	Central	Inpatient	SARI ¹ & ILI ²	<13
				years
Nyeri County Referral Hospital	Central	Inpatient	SARI ¹	All
Mombasa County Referral	Coast	Inpatient	SARI ¹	All
Hospital				
Kilifi County Hospital	Coast	Inpatient	Pneumonia ⁴	<5 years
Kakuma Refugee Camp	North	Inpatient	SARI ¹	All
Dadaab Refugee Camp	North	Inpatient	SARI ¹	All
Nakuru County Referral Hospital	Rift	Inpatient	SARI ¹	All
	valley			
Siaya County Hospital	Western	Inpatient	SARI ¹	All
Kakamega County Referral	Western	Inpatient	SARI ¹	All
Hospital				
Lwak Mission Hospital, Asembo	Western	Outpatient	ALRTI ³ & ILI ²	All
PBIDS*				
Tabitha Medical Clinic, Kibera	Central	Outpatient	ALRTI ³ & ILI ²	All
PBIDS				

535

Key: 1, SARI, Severe Acute Respiratory Illness is defined as an acute respiratory 536 illness requiring hospitalization with a history of fever or measured fever ≥38°C AND 537 538 a cough with an onset of symptoms within the last 14 days. 2, ILI, Influenza Like Illness is defined as measured fever of \geq 38°C OR sore throat in an outpatient of any age. 3, 539 ALRTI, Acute lower respiratory tract illness is defined as presence of cough OR 540 541 difficulty in breathing with one of the following danger signs: chest in-drawing, stridor, unable to breastfeed, vomit everything, convulsions, lethargy, or unconsciousness. 4, 542 Pneumonia is defined as modified WHO syndromic severe or very severe 543 pneumonia⁴⁸; PBIDS, Population-Based Infectious Disease Surveillance. 544

Table 2: Description of samples collected per site, cases positives for RV, and samples sequenced in the VP4/VP2 coding region.

Surveillance site	NP/OP collections	Tested	Linkage with clinical and demographic data	With respiratory symptoms	Samples positive for HRV	% Positivity	VP4/VP2 sequenced (%)
Kenyatta National Hospital (KNH)	510	357	343	343	64	18.7	63 (99)
Nyeri County Referral Hospital	482	427	425	425	68	16.0	57 (84)
Mombasa County Referral Hospital	524	455	445	445	103	23.1	87 (85)
Kilifi County Hospital	722	722	722	722	113	15.7	96 (85)
Kakuma Refugee Camp	220	175	175	175	25	14.3	23 (92)
Dadaab Refugee Camp	189	158	154	154	16	10.4	13 (81)
Nakuru County Referral Hospital	811	714	712	712	126	17.7	103 (82)
Siaya County Referral Hospital	936	922	794	642	154	24.0	128 (83)
Kakamega County Referral Hospital	464	418	402	402	72	17.9	65 (91)
Lwak Mission Hospital, Asembo PBIDS*	879	862	862	862	117	13.6	108(92)
Tabitha Medical Clinic, Kibera PBIDS	640	631	631	569	66	11.6	60 (91)
Not recorded	21	18	0	0	0	0	0
Total	6398	5859	5665	5451	924	17.0	803 (86.0)

547 Abbreviations : PBIDS, Population-Based Infectious Disease Surveillance.

Site	Number of sequences	Sequences per species			Number of types	Shannon's diversity (H)	Simpson diversity (G)
		RV-A	RV-B	RV-C			
KNH	63	32	2	29	32	3.31	23.68
Nyeri	57	31	3	23	35	3.38	24.43
Mombasa	87	54	3	30	42	3.58	30.82
Kilifi	96	54	7	35	46	3.69	34.79
Kakuma	23	14	1	8	15	2.56	10.76
Dadaab	13	7	0	6	9	2.09	7.20
Nakuru	103	61	7	35	50	3.66	28.74
Siaya	128	90	16	22	56	3.79	34.66
Kakamega	65	49	4	12	33	3.29	22.59
Asembo	108	65	14	29	52	3.74	34.23
Kibera	60	33	5	22	38	3.47	26.17

Table 3: Rhinovirus type diversity measured using Shannon (H) and Simpson (D) diversity indices.

550 Figures



551

552 **Figure 1:** Map of Kenya showing the geo-location of the 11 respiratory surveillance 553 sites, distinguishing inpatient (red markers) from outpatient (green markers) facilities.



Figure 2: Monthly distribution of samples tested and cases positives for rhinovirus from the 11 surveillance sites in Kenya in 2014.



556

557 Figure 3: Temporal distributions of frequent RV types across different sites.







558

559 Figure 4: Time-resolved lineage-specific phylogenetic trees for prevalent RV types. The Kenya genomes are indicated with filled

560 circles coloured by site.