Molecular epidemiology of respiratory syncytial virus in children in sub-Saharan Africa

Ihling, Clara Marlene; Schnitzler, Paul; Heinrich, Norbert; Mangu, Chacha; Sudi, Lwitiho; Souares, Aurélia; Gies, Sabine; Sié, Ali; Coulibaly, Boubacar; Ouédraogo, Andiyam Thierry; Mordmüller, Benjamin; Held, Jana; Adegnika, Ayola Akim; Fernandes, José F [**and 5 more**]

How to cite

IHLING, Clara Marlene et al. Molecular epidemiology of respiratory syncytial virus in children in sub-Saharan Africa. In: Tropical medicine & international health, 2021, vol. 26, n° 7, p. 810–822. doi: 10.1111/tmi.13573

This publication URL: https://archive-ouverte.unige.ch/unige:155034

Publication DOI: 10.1111/tmi.13573

Tropical Medicine and International Health doi:10.1111/tmi.13573

VOLUME 26 NO 7 PP 810–822 JULY 2021

Molecular epidemiology of respiratory syncytial virus in children in sub-Saharan Africa

Clara Marlene Ihling[1,2,3], Paul Schnitzler[1], Norbert Heinrich[2,4], Chacha Mangu[5], Lwitiho Sudi[5], Aurélia Souares[2,6], Sabine Gies[2,6], Ali Sié[7], Boubacar Coulibaly[7], Andiyam Thierry Ouédraogo[7], Benjamin Mordmüller[2,8], Jana Held[2,8], Ayola Akim Adegnika[2,8,9], José F. Fernandes[8,9], Isabella Eckerle[10,11], Juergen May[2,12], Benedikt Hogan[2,12], Daniel Eibach[2,12] and Julia Tabatabai[1,2,13]

1 *Center of Infectious Diseases, Virology, Heidelberg University Hospital, Heidelberg, Germany*
2 *German Center for Infection Research (DZIF), Heidelberg, Germany*
3 *Dr. von Haunersches Kinderspital, University Hospital of the LMU Munich, Munich, Germany*
4 *Division for Infectious Diseases, University Hospital of the LMU Munich, Munich, Germany*
5 *NIMR-Mbeya Medical Research Center, Mbeya, Tanzania*
6 *Heidelberg Institute of Global Health, Heidelberg University Hospital, Heidelberg, Germany*
7 *Centre de Recherche en Santé de Nouna, Nouna, Burkina Faso*
8 *Eberhard Karls University Tuebingen, Tübingen, Germany*
9 *Centre de Recherches Médicales de Lambaréné, Lambaréné, Gabon*
10 *Institute of Virology, University of Bonn Medical Center, Bonn, Germany*
11 *Faculty of Medicine, Geneva Center for Emerging Viral Diseases, University Hospitals of Geneva, University of Geneva, Geneva, Switzerland*
12 *Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany*
13 *Center for Pediatric and Adolescent Medicine, University Hospital Heidelberg, Heidelberg, Germany*

Abstract

OBJECTIVES This study investigated the molecular epidemiology of respiratory syncytial virus (RSV) among febrile children with acute respiratory tract infection in Ghana, Gabon, Tanzania and Burkina Faso between 2014 and 2017 as well as the evolution and diversification of RSV strains from other sub-Saharan countries.

METHODS Pharyngeal swabs were collected at four study sites (Agogo, Ghana: n = 490; Lambaréné, Gabon: n = 182; Mbeya, Tanzania: n = 293; Nouna, Burkina Faso: n = 115) and analysed for RSV and other respiratory viruses using rtPCR. For RSV-positive samples, sequence analysis of the second hypervariable region of the G gene was performed. A dataset of RSV strains from sub-Saharan Africa (2011–2017) currently available in GenBank was compiled. Phylogenetic analysis was conducted to identify the diversity of circulating RSV genotypes.

RESULTS In total, 46 samples were tested RSV positive (Ghana n = 31 (6.3%), Gabon n = 4 (2.2%), Tanzania n = 9 (3.1%) and Burkina Faso n = 2 (1.7%)). The most common RSV co-infection was with rhinovirus. All RSV A strains clustered with genotype ON1 strains with a 72-nucleotide duplication and all RSV B strains belonged to genotype BAIX. Phylogenetic analysis of amino acid sequences from sub-Saharan Africa revealed the diversification into 11 different ON1 and 22 different BAIX lineages and differentiation of ON1 and BAIX strains into potential new sub-genotypes, provisionally named ON1-NGR, BAIX-KEN1, BAIX-KEN2 and BAIX-KEN3.

CONCLUSION The study contributes to an improved understanding of the molecular epidemiology of RSV infection in sub-Saharan Africa. It provides the first phylogenetic data for RSV from Tanzania, Gabon and Burkina Faso and combines it with RSV strains from all other sub-Saharan countries currently available in GenBank.

keywords RSV, children, Africa, molecular epidemiology, genotypes

Sustainable Development Goal: Good Health and Well-Being

Introduction

Respiratory syncytial virus (RSV) is a major pathogen of acute respiratory tract infection (RTI) and the leading cause of hospitalisation with severe bronchiolitis and pneumonia among infants and young children [1, 2]. Globally, the annual incidence of RSV-related episodes of acute lower RTI amounts to over 33 million among children under 5 years, of whom around 10% require hospital care [3, 4]. Mortality estimates directly or indirectly attributable to RSV vary between 59,600 and 199,000 deaths per year, with almost all occurring in resource-limited settings [3, 4], where RSV testing is not routinely performed. In temperate regions, annual RSV epidemics peak during the winter months, whereas in tropical climate zones the peak occurs during the rainy season [5, 6]. The majority of hospital admissions due to acute lower RTI is observed among children below 2 years of age with RSV being the main pathogen [7]. Accordingly, almost all children have been infected with RSV at least once by the age of 2 years [8]. In older children and adults, reinfections are associated with milder disease indicating that RSV induces only partial immunity [9].

RSV is divided into two major subgroups: RSV A and RSV B [10]. Strain variation is thought to contribute to the ability of RSV to cause frequent reinfections [11].

The main differences between RSV A and RSV B are found in the attachment (G) glycoprotein [10]. The glycoprotein is able to accommodate drastic changes with the emergence of new variants. Diversity occurs mainly in the two hypervariable regions of the ectodomain which are separated by a highly conserved 13-aa length domain [12]. Sequencing of the second hypervariable region of the G protein has been widely used to further subdivide subgroups RSV A and RSV B into genotypes. In 1998, Peret et al. established a classification of different genotypes based on the analysis of the variability of the RSV G gene [13]. To date, 11 RSV A genotypes [13–17] and more than 30 RSV B genotypes [13, 15, 18–27] have been described based on phylogenetic analysis of nucleotide sequences.

In the past two decades, the most drastic changes and subsequent emergence of new genotypes have been attributed to a nucleotide duplication within the G glycoprotein [28]. In 1999, a new RSV B genotype (BA) emerged in Buenos Aires, Argentina, containing a 60-nt duplication in the second hypervariable region of the G gene [29]. In the following ten years, the BA genotype spread worldwide, largely replaced previously described RSV B genotypes and differentiated into fourteen sub-genotypes BAI-XIV [26, 30–32]. During the 2010/11 winter season, a novel RSV A genotype ON1 with a 72-nt duplication has been reported in Canada for the first time [17]. In later RSV epidemics, ON1 spread worldwide and is today the dominating RSV A genotype [30–33]. In line with the diversification of the BA genotype, a differentiation of ON1 into different virus strains can be observed [31,34].

Little is known about the molecular epidemiology and the emergence of novel viral strains of RSV in sub-Saharan Africa. Although in Kenya and South Africa, extensive molecular analysis of RSV strains was performed [20, 31, 35–45], only few other sub-Saharan countries conducted singular studies [35, 36, 46–49].

For the present study, respiratory samples from children presenting with fever and clinical symptoms of acute RTI were collected at study sites in Mbeya and Matema, Tanzania, Agogo, Ghana, Lambaréné, Gabon and Nouna, Burkina Faso and analysed for RSV and other respiratory viruses. We investigated the genetic diversity and patterns of co-circulating RSV genotypes in sub-Saharan Africa by analysing all RSV G gene sequences from the region available in GenBank.

Methods

Study sites and patients

In order to gain more information about the epidemiology of febrile illnesses in children in Africa, a study group of the German Center for Infection Research (DZIF) rolled out a series of 'fever without source' (FWS) studies at African partner sites in Mbeya and Matema, Tanzania, Agogo, Ghana, Lambaréné, Gabon and Nouna, Burkina Faso. Depending on the individual symptoms of the patients, a variety of clinical samples and systematic clinical data were collected. The present study focusses on a subset of patients with respiratory symptoms from whom pharyngeal swabs were collected to analyse the viral causes of acute RTI with particular focus on the molecular epidemiology of RSV infection. The diagnosis of acute RTI was made if the patient presented with symptoms or showed clinical signs of an upper (rhinitis, pharyngitis, tonsillitis and otitis) or lower (pneumonia, bronchitis and bronchiolitis) respiratory tract infection. All FWS study sites used standardised questionnaires collecting demographic and clinical information of the study patients.

In *Mbeya and Matema, Tanzania* respiratory samples from 293 (sub)febrile children were collected in cooperation with the Ludwig-Maximillian University (LMU) Munich/Germany, the NIMR-Mbeya Medical Research Center (NIMR-MMRC) Mbeya/Tanzania, and the German Army Institute for Microbiology.

In *Nouna, Burkina Faso* the FWS study was performed in cooperation with the University Hospital of

Heidelberg/Germany (UKHD) and the Centre de Recherche en Santé de Nouna/Burkina Faso (CRSN). A total number of 115 pharyngeal swabs from (sub)febrile children with acute RTI admitted to the outpatient department and paediatric ward of the district hospital of the Boucle du Mouhoun region were collected.

In *Lambaréné, Gabon* the FWS study was conducted in collaboration with the Institute for Tropical Medicine of the University of Tübingen/Germany (ITM-EKUT) and the Centre de Recherches Médicales de Lambaréné/Gabon (CERMEL) [50]. A total of 182 (sub)febrile children with acute RTI admitted to the outpatient department and paediatric ward of the Albert Schweitzer Hospital Lambaréné were included in the study.

In *Agogo, Ghana* the FWS study was performed in collaboration with the Bernhard Nocht Institute for Tropical Medicine (BNITM) Hamburg/Germany and the Kumasi Centre for Collaborative Research in Tropical Medicine (KCCR) Agogo/Ghana [51]. A total number of 490 (sub) febrile children with acute RTI admitted to the Agogo Presbyterian Hospital were recruited.

An overview of inclusion criteria per study site as well as the total number of recruited patients with acute RTI is shown in Figure S1.

Multiplex polymerase chain reaction (PCR) for respiratory pathogens

For PCR analysis of respiratory samples from Tanzania and Burkina Faso, RNA was extracted from pharyngeal swabs using the QIAamp® viral RNA mini kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Amplification and detection of viral RNA were performed with the multiplex real-time PCR kit (FTD respiratory pathogens 21, Fast-track diagnostics, Luxembourg) on a LightCycler® 480 instrument II (Roche, Mannheim, Germany). This assay identified the viral pathogens RSV A/B, influenzavirus A, B, H1N1; rhinovirus, coronavirus NL63, 229E, OC43, HKU1; parainfluenzavirus 1-4; human metapneumovirus A/B; bocavirus; adenovirus; enterovirus and parechovirus as well as mycoplasma pneumoniae. Samples from Ghana and Gabon were analysed using an in-house multiplex PCR protocol identifying RSV A/B, influenzavirus A, B, rhinovirus, coronavirus NL63, 229E, OC43, HKU1; parainfluenzavirus 1-4; human metapneumovirus A/B; adenovirus; enterovirus; and parechovirus.

Sequencing and phylogenetic analysis

For sequencing and identification of RSV genotypes, extracted RNA of RSV-positive samples from all four study sites was initially reverse transcribed and cDNA was synthesised using random hexamer primers. Subsequently, PCR targeting the second hypervariable region of the G gene was performed using primer pairs as described by Peret et al. [13]. Nucleotide sequences detected in this study were deposited in GenBank (http://www.ncbi.nlm.nih.gov) under accession numbers MT862896-MT862910. In order to investigate the genetic diversity of circulating RSV genotypes in sub-Saharan Africa, we downloaded RSV A and RSV B G gene sequences detected in sub-Saharan African countries between 2011 and 2016 (deposited in GenBank as of 22 July 2019). An overview of the retrieved datasets can be seen in Table S1. Multiple sequence alignments and phylogenetic analysis of the second hypervariable region of the G gene were conducted using the Clustal W 1.6 method of MEGA software version 7 [52]. Phylogenetic trees were generated using the maximum-likelihood method and bootstrap values with 1,000 replicates were calculated to evaluate confidence estimates. Pairwise nucleotide distances were calculated to compare the differences within genotypes of subgroup RSV A and RSV B using MEGA software version 7 [52]. Alignments of sub-Saharan African RSV A and RSV B strains were compared to references strains from GenBank to identify amino acid substitutions. At least three sequences sharing at least two signature mutations were defined as a new lineage.

N-/O-Glycosylation and positive and negative selection sites

Putative N-glycosylation sites were predicted if the encoded amino acid sequence was aspargine-Xaa-threonine/serine (N-X-T/S), where Xaa was not a proline and accepted if the glycosylation potential was ≥0.5 in NetNGlyc 1.0 server [53]. O-glycosylation was determined using the NetOGlyc 3.1 server and sites were predicted using a G-score ≥0.5 [37]. To identify codons under positive or negative selection among globally emerging ON1 and BAIX strains, positively and negatively selected sites were estimated by use of the Datamonkey Web server (http://www.datamonkey.org) [38]. A consensus of four methods [39] was used to calculate synonymous (dS) and non-synonymous (dN) rates per codon of the amino acid alignment.

Statistical analysis of epidemiological factors

Demographic data and the patients' baseline characteristics were summarised per study site using descriptive statistics (Stata/IC15.0; TX, USA). Group comparisons were performed using χ^2 or Fisher's exact test for categorical variables and by Student's *t*-test or Wilcoxon

rank-sum test for continuous variables, as appropriate. *P*-values < 0.05 were considered statistically significant.

Ethical approval and patient confidentiality

The Ethical Committee of the University of Heidelberg/Germany approved the overall study (S-267/2016). Furthermore, FWS studies were also approved by local ethical committees.

Results

Characteristics of children with acute respiratory tract infection

In Table 1, the basic demographics, admission findings and leading clinical diagnosis of children with acute RTI are summarised. The slightly varying inclusion criteria are reflected in the composition of the respective study cohorts. It is important to note that the study cohort in Tanzania comprises mainly outpatients, whereas all other study sites recruited mainly hospitalised patients. The median age of children with acute RTI was 2 years (IQR 1.0–3.8) for Burkina Faso, Gabon and Ghana, whereas children from Tanzania had a median age of 4.3 (2.8–6.8) as the study recruitment started above the age of 1 year. The most common primary diagnosis was malaria, followed by acute RTI. Most patients with primary respiratory diagnoses presented with symptoms of a lower respiratory tract infection.

Detection of respiratory viruses and clinical signs of RSV infection

A total number of 1,080 respiratory samples were collected from children who presented with acute RTI and molecular analysis for multiple viral pathogens was performed. The results are summarised in the lower part of Table 1. RSV was detected in a total of 46 patients. Common co-infections with RSV were rhinovirus (Tanzania n = 3/9; 33.3% and Ghana 6/31; 19.4%), enterovirus (Tanzania n = 2/9; 22.2% and Ghana n = 6/31; 19.4%) and adenovirus (Burkina Faso n = 1/2; 50.0% and Ghana n = 3/31; 9.7%).

In Tanzania, RSV-positive children showed a shift to younger age groups (P = 0.022) compared to RSV-negative children, with the majority of RSV-positive tested patients in the age group 2–5 years, and no child older than 5 years. The majority of all RSV-positive children (n = 8/9; 88.9%) suffered primarily from a respiratory tract infection, with 66.7% (n = 6/9) being affected by an upper respiratory tract infection.

In Burkina Faso, both RSV-positive children were hospitalised with lower respiratory tract infection without fever but with signs of respiratory insufficiency such as tachypnoea.

In Gabon, RSV-positive children were in median 1.7 years old. All RSV-positive children suffered from a lower RTI and 3 of 4 showed signs of respiratory insufficiency.

In Ghana, RSV-positive patients were in median 10 months old and thus significantly younger than RSV-negative patients (median 2.0 years; P < 0.0001). Most RSV-positive children from Ghana (80.7%) primarily suffered from an acute RTI (RSV-negative children: 33.3%; P < 0.001) and also significantly more frequently from a lower RTI than RSV-negative children (P = 0.008). RSV-positive children also showed significantly more signs of respiratory insufficiency (P = 0.01) such as tachypnoea (P = 0.038), thoracic retractions (P = 0.005) and nasal flaring (P = 0.003). All RSV-positive children were admitted to hospital. They remained in hospital for a median duration of 3 days (IQA 3–5 days).

Phylogenetic analysis of studied sequences

PCR analysis of all respiratory swabs from Burkina Faso, Tanzania, Gabon and Ghana detected 46 RSV-positive samples. Of these, the second hypervariable region of the RSV G gene was successfully sequenced for 36 samples and the RSV strains could therefore be assigned to RSV subtypes and genotypes. More than half of the RSV strains belonged to the subtype RSV A (n = 21/36), the rest belonged to the subtype RSV B (n = 15/36).

All RSV A strains clustered with strains of the ON1 genotype and formed three clusters in the phylogenetic tree, all around known international and sub-Saharan Africa ON1 reference sequences retrieved from GenBank (Figure 1a).

All RSV B strains belonged to the sub-genotype BAIX (n = 15) of genotype BA and formed a total of three clusters around known international and sub-Saharan Africa BAIX reference sequences retrieved from GenBank (Figure 1b).

Patterns of circulating RSV genotypes in sub-Saharan Africa

A total of 1,004 RSV sequences (RSV A: n = 553; RSV B: n = 451) were derived from GenBank from 5 different countries (Kenya, South Africa, Ghana, Cameroon, Nigeria). For further phylogenetic analysis of RSV strains in sub-Saharan Africa, we selected RSV A genotype ON1 and RSV B genotype BAIX sequences in order to compare these with our studied sequences.

Table 1 Basic demographic and clinical diagnosis characteristics and analysis of respiratory samples of the study cohorts in Tanzania, Burkina Faso, Gabon and Ghana

	Tanzania *n* = 293	Burkina Faso *n* = 115	Gabon *n* = 182	Ghana *n* = 490
Demographic characteristics				
Age in years, median (IQR)	4.3 (2.8–6.8)	1.7 (0.9–2.8)	2.1 (1.0–5.3)	2.0 (1.1–3.6)
Age group, *n*/*N* (%)				
0–6 months	n.a.	13/115 (11.3)	14/182 (7.7)	34/490 (6.9)
>6–24 months	45/293 (15.4)[1]	61/115 (53.0)	76/182 (41.8)	217/490 (44.3)
>2–5 years	140/293 (47.8)	32/115 (27.8)	41/182 (22.5)	175/490 (35.7)
>5 years	108/293 (36.9)	97115 (7.8)	51/182 (28.0)	64/490 (13.1)
Gender, *n*/*N* (%)				
Male	127/293 (43.3)	63/115 (54.8)	101/182 (55.5)	266/490 (54.3)
Female	166/293 (56.7)	52/115 (45.2)	81/182 (44.5)	224/490 (45.7)
Admission findings				
Weight in kg, median (IQR)	15.3 (12.6–20.8)	8.0 (6.6–10.8)	12.0 (9.2–17.4)	10.9 (8.5–13.5)[3]
Temperature in °C, median (IQR)	38.3 (37.7–39.2)[2]	38.7 (38.2–39.2)	38.9 (38.4–39.6)	39.0 (38.5–39.5)[4]
Fever ≥ 38.5 °C, *n*/*N* (%)	132/192 (45.2)	70/114 (61.4)	127/181 (70.2)	382/488 (78.3)
Clinical signs, *n*/*N* (%)				
Cough	212/293 (72.4)	78/115 (67.8)	170/182 (93.4)	366/488 (75.0)
Wheezing	8/117 (6.8)	4/115 (3.5)	5/176 (2.8)	13/483 (2.7)
Pulmonary rales	110/117 (94.0)	66/115 (57.4)	44/176 (25.0)	60/486 (12.4)
Dyspnea	11/293 (3.8)	25/115 (21.7)	53/180 (29.4)	49/488 (10.0)
Tachypnea	n.a.	90/115 (78.3)	139/174 (79.9)	292/489 (59.7)
Retractions	n.a.	9/115 (7.8)	23/182 (12.6)	73/488 (15.0)
Nasal flaring	n.a.	10/115 (8.7)	40/180 (22.2)	90/488 (18.4)
Leading clinical diagnosis, *n*/*N* (%)				
Non-respiratory	48/280 (17.1)	85/115 (73.9)	129/182 (70.9)	312/490 (63.7)
Respiratory	232/280 (82.9)	30/115 (26.1)	53/182 (29.1)	178/490 (36.3)
Upper RTI	118/280 (42.1)	1/115 (0.9)	4/182 (2.2)	48/490 (9.8)
Lower RTI	114/280 (40.7)	29/115 (25.2)	49/182 (26.9)	130/490 (16.5)
Course of disease				
Hospital stay in days, median (IQR)*	n.a.	3 (2–5)	5 (4–6)	3 (2–5)
Mortality, *n*/*N* (%)	n.a.	15/114 (13.2)	n.a.	7/484 (1.5)
Analysis of respiratory samples, *n*/*N* (%)				
At least one viral pathogen detected	159/293 (54.2)	69/115 (60.0)	25/182 (13.7)	300/490 (61.2)
RSV	9/293 (3.1)	2/115 (1.7)	4/182 (2.2)	31/490 (6.3)
Influenzavirus A	27/293 (9.2)	1/115 (0.9)	3/182 (1.7)	32/490 (6.5)
Influenzavirus B	39/293 (13.3)	6/115 (5.2)	0/182 (0.0)	13/490 (2.7)
Parainfluenzavirus	17/293 (5.8)	11/115 (9.6)	2/182 (1.1)	38/490 (7.8)
Metapneumovirus	43/293 (14.7)	9/115 (7.8)	0/182 (0.0)	16/490 (3.3)
Rhinovirus	23/293 (7.9)	35/115 (30.4)	3/182 (1.7)	100/490 (20.4)
Coronavirus	11/293 (3.8)	4/115 (3.5)	5/182 (2.7)	32/490 (6.5)
Parechovirus	2/293 (0.7)	1/115 (0.9)	0/182 (0.0)	3/490 (0.6)
Adenovirus	20/293 (6.8)	14/115 (12.1)	7/182 (3.9)	97/490 (19.8)
Enterovirus	8/293 (2.7)	12/115 (10.4)	3/182 (1.7)	76/490 (15.5)
Bocavirus	3/293 (1.0)	6/115 (5.2)	0/182 (0.0)	0/490 (0.0)

Categorical variables are reported as *n*/*N* (%) with *N* as total number of available observations per variable and continuous variables as medians with 25%-75% interquartile range.

IQR, interquartile range; n.a., not applicable/available; RSV, Respiratory Syncytial Virus; RTI, respiratory tract infection.

[1]Tanzania recruited children 12 months and older only.

[2]Calculated for 292/293 patients.

[3]Calculated for 485/490 patients.

[4]Calculated for 488/490 patients.

*Duration of hospital stay was calculated for patients who stayed at least 24h in hospital care (Burkina Faso *n* = 105, Gabon *n* = 163, Ghana *n* = 480).

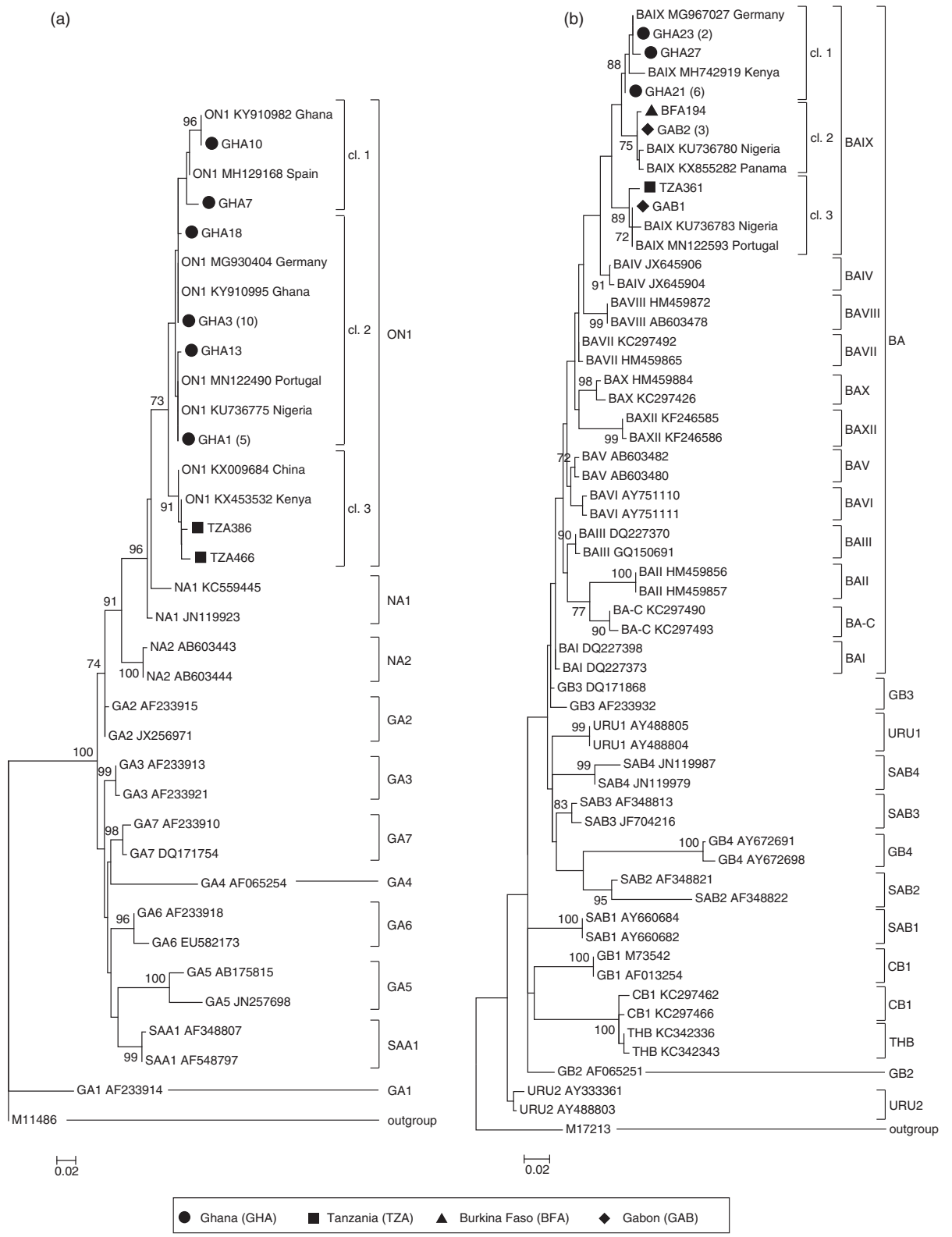
(a)
ON1 KY910982 Ghana
GHA10
ON1 MH129168 Spain
GHA7
GHA18
ON1 MG930404 Germany
ON1 KY910995 Ghana
GHA3 (10)
GHA13
ON1 MN122490 Portugal
ON1 KU736775 Nigeria
GHA1 (5)
ON1 KX009684 China
ON1 KX453532 Kenya
TZA386
TZA466
cl. 1
cl. 2
cl. 3
ON1
NA1 KC559445
NA1 JN119923
NA1
NA2 AB603443
NA2 AB603444
NA2
GA2 AF233915
GA2 JX256971
GA2
GA3 AF233913
GA3 AF233921
GA3
GA7 AF233910
GA7 DQ171754
GA7
GA4 AF065254
GA4
GA6 AF233918
GA6 EU582173
GA6
GA5 AB175815
GA5 JN257698
GA5
SAA1 AF348807
SAA1 AF548797
SAA1
GA1 AF233914
GA1
M11486
outgroup
0.02
(b)
BAIX MG967027 Germany
GHA23 (2)
GHA27
BAIX MH742919 Kenya
GHA21 (6)
BFA194
GAB2 (3)
BAIX KU736780 Nigeria
BAIX KX855282 Panama
TZA361
GAB1
BAIX KU736783 Nigeria
BAIX MN122593 Portugal
cl. 1
cl. 2
cl. 3
BAIX
BAIV JX645906
BAIV JX645904
BAIV
BAVIII HM459872
BAVIII AB603478
BAVIII
BAVII KC297492
BAVII HM459865
BAVII
BAX HM459884
BAX KC297426
BAX
BAXII KF246585
BAXII KF246586
BAXII
BAV AB603482
BAV AB603480
BAV
BAVI AY751110
BAVI AY751111
BAVI
BAIII DQ227370
BAIII GQ150691
BAIII
BAII HM459856
BAII HM459857
BAII
BA-C KC297490
BA-C KC297493
BA-C
BAI DQ227398
BAI DQ227373
BAI
BA
GB3 DQ171868
GB3 AF233932
GB3
URU1 AY488805
URU1 AY488804
URU1
SAB4 JN119987
SAB4 JN119979
SAB4
SAB3 AF348813
SAB3 JF704216
SAB3
GB4 AY672691
GB4 AY672698
GB4
SAB2 AF348821
SAB2 AF348822
SAB2
SAB1 AY660684
SAB1 AY660682
SAB1
GB1 M73542
GB1 AF013254
CB1
CB1 KC297462
CB1 KC297466
CB1
THB KC342336
THB KC342343
THB
GB2 AF065251
GB2
URU2 AY333361
URU2 AY488803
URU2
M17213
outgroup
0.02
Ghana (GHA)
Tanzania (TZA)
Burkina Faso (BFA)
Gabon (GAB)

Figure 1 Phylogenetic trees of RSV A and RSV B strains from Ghana, Gabon, Tanzania and Burkina Faso (studied sequences). Phylogenetic trees for nucleotide sequences of RSV A (a) and RSV B (b) strains were constructed with maximum-likelihood method with 1,000 bootstrap replicates using MEGA 7 software. RSV strains from Ghana, Gabon, Burkina Faso and Tanzania are indicated by their country code followed by their strain identification number. The number of identical strains is indicated in brackets after the strain identifier. Reference strains representing known genotypes were retrieved from GenBank and included in the tree (labels include accession number). The genotype assignment is shown on the right by brackets. Prototype strains (M11486 for subgroup A and M17213 for subgroup B) were used as outgroup. Bootstrap values greater than 70% are indicated at the branch nodes. The scale bar represents the number of nucleotide substitutions per site. cl. = cluster.

The phylogenetic tree based on amino acid (aa) sequences of RSV A genotype ON1 strains comprised 21 studied sequences from Ghana and Tanzania and 345 ON1 sequences retrieved from GenBank from Kenya, South Africa, Nigeria and Ghana (Figure 2a). The RSV A tree formed five clusters (a-e). Nigerian sequences from GenBank formed cluster d, here named ON1-NGR, which was supported by a bootstrap value of 87% and *P*-distances of 0.011.

The phylogenetic tree based on amino acid (aa) sequences of RSV B genotype BAIX strains comprised 15 studied sequences from Ghana, Gabon, Burkina Faso and Tanzania and 255 BAIX sequences retrieved from GenBank from Kenya, South Africa, Nigeria, Cameroon and Ghana (Figure 2b). The RSV B tree formed five clusters (a-e). Cluster d was the most heterogenous cluster with *P*-distances of 0.041 and the three sub-branches BAIX-KEN1, BAIX-KEN2 and BAIX-KEN3 supported by bootstrap values of 82%, 91% and 74% and with *P*-distances of 0.005, 0.01 and 0.006, respectively. In cluster e, studied sequences from Gabon and Tanzania formed a subbranch with GenBank sequences from Nigeria.

Evolution and diversification of RSV strains in sub-Saharan Africa

Comparison of the deduced amino acid (aa) sequences of the compiled dataset of in GenBank available ON1 strains from sub-Saharan Africa including FWS study sequences (n = 345+21) revealed different patterns of signature amino acid substitutions summarised in 11 lineages being distinct from the original strain (Figure 3). A total of four amino acid sequences were identical to the ON1 reference strain with identical nucleotide sequences or synonymous nucleotide substitutions. Further 19 strains had unique non-synonymous substitutions or did not fulfil the criteria of a separate lineage.

The analysis of the deduced amino acid (aa) sequences of the compiled dataset of in GenBank available BAIX strains from sub-Saharan Africa including FWS study sequences (n = 255+15) revealed different patterns of signature amino acid substitutions summarised in 22 lineages being distinct from the original strain (Figure 4) [20]. A total of two sequences were identical to the BAIX reference strain with identical nucleotide sequences or synonymous nucleotide substitutions. Further 87 (32.2%) strains had unique non-synonymous substitutions (n = 74; 27.4%) or did not fulfil the criteria for their own lineage (n = 13; 4.8%).

To identify positively and negatively selected sites in the ON1 and BAIX datasets from Sub-Saharan Africa, four and three different models were used, respectively. A codon site was considered as positively or negatively selected if a consensus of two or more models was reached (Tables S2 and S3). The mean non-synonymous/synonymous substitution rate ratio (dN/dS) was 0.817 for ON1 and 0.723 for BAIX strains.

Discussion

This study investigated the molecular epidemiology of RSV infection among febrile children with acute RTI in Ghana, Gabon, Tanzania and Burkina Faso between 2014 and 2017 as well as the evolution and diversification of RSV strains from other sub-Saharan countries collected between 2011 and 2016 and retrieved from GenBank. The data contributes to an improved understanding of the molecular epidemiology of RSV infection in sub-Saharan Africa. Our study provides the first phylogenetic data for RSV from Tanzania, Gabon and Burkina Faso and combines it with RSV strains from all other sub-Saharan countries currently available in GenBank.

The investigation of respiratory samples from children with acute RTI from the study cohort showed RSV to be an important representative of viral pathogens. In line with previous findings, all RSV-infected children were under 5 years old [8, 47] and in Ghana RSV-positive children were significantly younger than RSV-negative children. The most severe primary infection with RSV usually occurs during the first year of life [8]. Reinfections, on the other hand, usually lead to a milder clinical picture and thus less frequently to hospitalisation [40].

RSV-infected children also suffered more frequently from a lower RTI than RSV-negative children and were

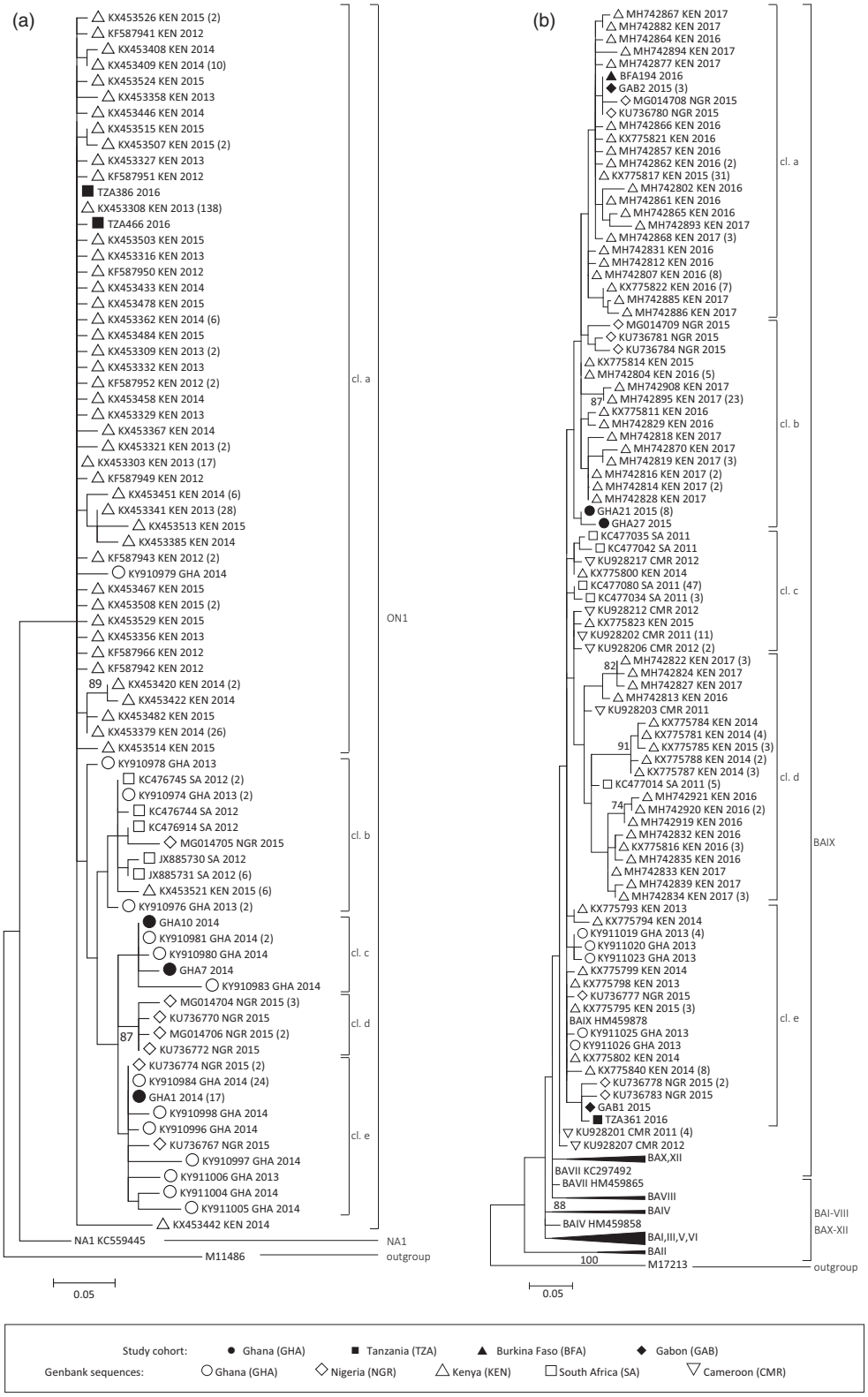
(a)
KX453526 KEN 2015 (2)
KF587941 KEN 2012
KX453408 KEN 2014
KX453409 KEN 2014 (10)
KX453524 KEN 2015
KX453358 KEN 2013
KX453446 KEN 2014
KX453515 KEN 2015
KX453507 KEN 2015 (2)
KX453327 KEN 2013
KF587951 KEN 2012
TZA386 2016
KX453308 KEN 2013 (138)
TZA466 2016
KX453503 KEN 2015
KX453316 KEN 2013
KF587950 KEN 2012
KX453433 KEN 2014
KX453478 KEN 2015
KX453362 KEN 2014 (6)
KX453484 KEN 2015
KX453309 KEN 2013 (2)
KX453332 KEN 2013
KF587952 KEN 2012 (2)
KX453458 KEN 2014
KX453329 KEN 2013
KX453367 KEN 2014
KX453321 KEN 2013 (2)
KX453303 KEN 2013 (17)
KF587949 KEN 2012
KX453451 KEN 2014 (6)
KX453341 KEN 2013 (28)
KX453513 KEN 2015
KX453385 KEN 2014
KF587943 KEN 2012 (2)
KY910979 GHA 2014
KX453467 KEN 2015
KX453508 KEN 2015 (2)
KX453529 KEN 2015
KX453356 KEN 2013
KF587966 KEN 2012
KF587942 KEN 2012
89
KX453420 KEN 2014 (2)
KX453422 KEN 2014
KX453482 KEN 2015
KX453379 KEN 2014 (26)
KX453514 KEN 2015
KY910978 GHA 2013
KC476745 SA 2012 (2)
KY910974 GHA 2013 (2)
KC476744 SA 2012
KC476914 SA 2012
MG014705 NGR 2015
JX885730 SA 2012
JX885731 SA 2012 (6)
KX453521 KEN 2015 (6)
KY910976 GHA 2013 (2)
GHA10 2014
KY910981 GHA 2014 (2)
KY910980 GHA 2014
GHA7 2014
KY910983 GHA 2014
MG014704 NGR 2015 (3)
KU736770 NGR 2015
87
MG014706 NGR 2015 (2)
KU736772 NGR 2015
KU736774 NGR 2015 (2)
KY910984 GHA 2014 (24)
GHA1 2014 (17)
KY910998 GHA 2014
KY910996 GHA 2014
KU736767 NGR 2015
KY910997 GHA 2014
KY911006 GHA 2013
KY911004 GHA 2014
KY911005 GHA 2014
KX453442 KEN 2014
NA1 KC559445
M11486
cl. a
cl. b
cl. c
cl. d
cl. e
ON1
NA1
outgroup
0.05
(b)
MH742867 KEN 2017
MH742882 KEN 2017
MH742864 KEN 2016
MH742894 KEN 2017
MH742877 KEN 2017
BFA194 2016
GAB2 2015 (3)
MG014708 NGR 2015
KU736780 NGR 2015
MH742866 KEN 2016
KX775821 KEN 2016
MH742857 KEN 2016
MH742862 KEN 2016 (2)
KX775817 KEN 2015 (31)
MH742802 KEN 2016
MH742861 KEN 2016
MH742865 KEN 2016
MH742893 KEN 2017
MH742868 KEN 2017 (3)
MH742831 KEN 2016
MH742812 KEN 2016
MH742807 KEN 2016 (8)
KX775822 KEN 2016 (7)
MH742885 KEN 2017
MH742886 KEN 2017
MG014709 NGR 2015
KU736781 NGR 2015
KU736784 NGR 2015
KX775814 KEN 2015
MH742804 KEN 2016 (5)
MH742908 KEN 2017
87
MH742895 KEN 2017 (23)
KX775811 KEN 2016
MH742829 KEN 2016
MH742818 KEN 2017
MH742870 KEN 2017
MH742819 KEN 2017 (3)
MH742816 KEN 2017 (2)
MH742814 KEN 2017 (2)
MH742828 KEN 2017
GHA21 2015 (8)
GHA27 2015
KC477035 SA 2011
KC477042 SA 2011
KU928217 CMR 2012
KX775800 KEN 2014
KC477080 SA 2011 (47)
KC477034 SA 2011 (3)
KU928212 CMR 2012
KX775823 KEN 2015
KU928202 CMR 2011 (11)
KU928206 CMR 2012 (2)
82
MH742822 KEN 2017 (3)
MH742824 KEN 2017
MH742827 KEN 2017
MH742813 KEN 2016
KU928203 CMR 2011
KX775784 KEN 2014
KX775781 KEN 2014 (4)
91
KX775785 KEN 2015 (3)
KX775788 KEN 2014 (2)
KX775787 KEN 2014 (3)
KC477014 SA 2011 (5)
MH742921 KEN 2016
74
MH742920 KEN 2016 (2)
MH742919 KEN 2016
MH742832 KEN 2016
KX775816 KEN 2016 (3)
MH742835 KEN 2016
MH742833 KEN 2017
MH742839 KEN 2017
MH742834 KEN 2017 (3)
KX775793 KEN 2013
KX775794 KEN 2014
KY911019 GHA 2013 (4)
KY911020 GHA 2013
KY911023 GHA 2013
KX775799 KEN 2014
KX775798 KEN 2013
KU736777 NGR 2015
KX775795 KEN 2015 (3)
BAIX HM459878
KY911025 GHA 2013
KY911026 GHA 2013
KX775802 KEN 2014
KX775840 KEN 2014 (8)
KU736778 NGR 2015 (2)
KU736783 NGR 2015
GAB1 2015
TZA361 2016
KU928201 CMR 2011 (4)
KU928207 CMR 2012
BAX,XII
BAVII KC297492
BAVII HM459865
BAVIII
88
BAIV
BAIV HM459858
BAI,III,V,VI
BAII
100
M17213
cl. a
cl. b
cl. c
cl. d
cl. e
BAIX
BAI-VIII
BAX-XII
outgroup
0.05
Study cohort: Ghana (GHA) Tanzania (TZA) Burkina Faso (BFA) Gabon (GAB)
Genbank sequences: Ghana (GHA) Nigeria (NGR) Kenya (KEN) South Africa (SA) Cameroon (CMR)

Figure 2 Phylogenetic trees of RSV A and RSV B strains from sub-Saharan Africa (including GenBank data set). Phylogenetic trees for amino acid sequences of RSV A (a) and RSV B (b) strains were constructed with maximum-likelihood method with 1,000 bootstrap replicates using MEGA 7 software. RSV strains from Ghana, Gabon, Burkina Faso and Tanzania are indicated by their country code followed by their strain identification number and the year of isolation. GenBank strains from sub-Saharan Africa are indicated by their country code followed by their accession number and the year of isolation. The number of identical strains is indicated in brackets after the strain identifier. Reference strains representing known genotypes were retrieved from GenBank and included in the tree (labels include accession number). The genotype assignment is shown on the right by brackets. Prototype strains (M11486 for subgroup A and M17213 for subgroup B) were used as outgroup. Bootstrap values greater than 70% are indicated at the branch nodes. The scale bar represents the number of amino acid substitutions per site. cl. = cluster.

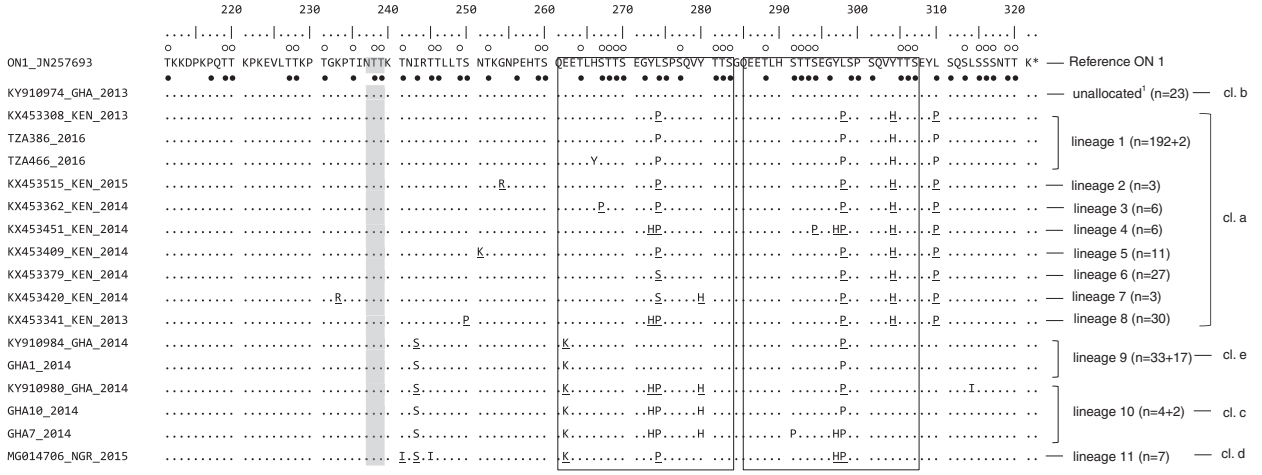

Figure 3 Alignment of deduced amino acid sequences of the second hypervariable region of the G gene of RSV A ON1 strains from sub-Saharan Africa. Alignments are shown relatively to the sequence of the ON1 strain first described from Canada (GenBank accession number JN257693). Strain identifier of studied sequences are shown in bold with the number of identical amino acid sequences in brackets. GenBank sequences are summarised by one exemplary strain per lineage. The amino acid positions correspond to positions 211 to 322 of the G protein of the prototype strain A2. Identical residues are indicated by dots, asterisks indicate the position of stop codons. Boxes frame the 23 amino acid duplication. Signature coding mutations are underlined. Grey shading highlights predicted N-glycosylation sites. Unfilled circles indicate predicted O-glycosylation sites of the Canadian reference ON1 strain; potential O-glycosylation sites of sub-Saharan strains are indicated by black dots. Lineages and corresponding clusters from the phylogenetic tree are shown on the right by brackets. cl. = cluster.

usually hospitalised. RSV infections are the most common cause of hospitalisation of infants and young children with respiratory infections [1–4, 41–43]. Hospitalisation rates and duration differ depending on the study population, region and time of study and correlate with the children's age and with risk factors for a severe course of the disease [7, 40, 44].

Besides malaria, acute RTI is the most frequent cause of febrile illness in Africa [45, 54]. All patients included in this study suffered from clinical symptoms of respiratory infections, whereas only in Tanzania respiratory infections were the most common primary diagnosis. In Burkina Faso, Ghana and Gabon, malaria was the most common primary reason for consultation, and the affected children only suffered of secondary respiratory symptoms. In contrast, the majority (50% in Gabon) of RSV-positive patients at all study sites were primarily affected by respiratory infections.

More than half of the studied RSV sequences from sub-Saharan Africa belonged to the subtype RSV A (n = 21/36; 58.3%). The novel RSV genotypes ON1 and BAIX were the only circulating genotypes in our study population.

Similarly, more than half of the RSV sequences from GenBank belonged to the subtype RSV A (n = 553, 55.1%). The majority of the GenBank RSV B sequences could also be assigned to the BAIX sub-genotype of the BA genotype (n = 270/451; 59.9%). These findings are in line with recent studies from other parts of the world that confirm ON1 and BAIX to be the globally predominating RSV genotypes [28, 31, 55]. The phylogenetic analysis of RSV A (ON1) and RSV B (BAIX) amino acid sequences

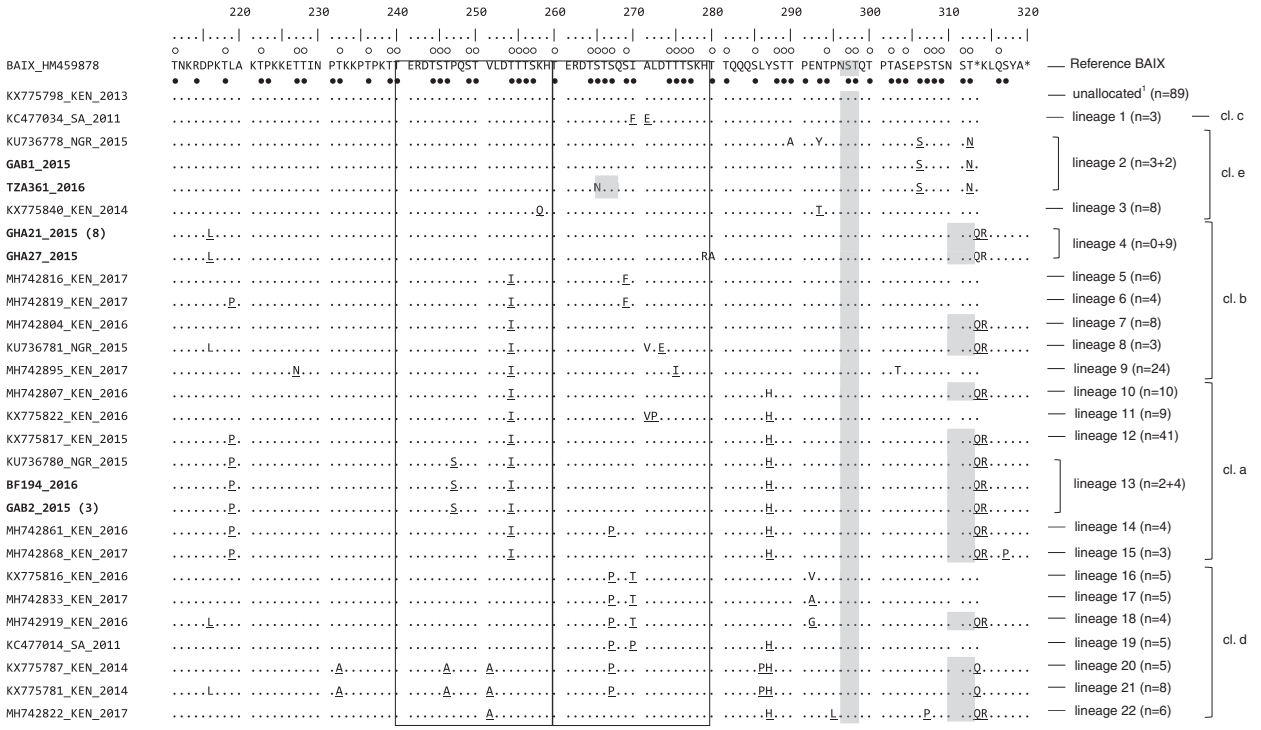

Figure 4 Alignment of deduced amino acid sequences of the second hypervariable region of the G gene of RSV B BAIX strains from sub-Saharan Africa. Alignments are shown relatively to the sequence of the BAIX strain first described from Japan (GenBank accession number HM459878). Strain identifier of studied sequences are shown in bold with the number of identical amino acid sequences in brackets. GenBank sequences are summarised by one exemplary strain per lineage. The amino acid positions correspond to positions 211 to 320 of the G protein of the prototype BA strain. Identical residues are indicated by dots, asterisks indicate the position of stop codons. Boxes frame the 20 amino acid duplication. Signature coding mutations are underlined. Grey shading highlights predicted N-glycosylation sites. Unfilled circles indicate predicted O-glycosylation sites of the Japanese reference BAIX strain; potential O-glycosylation sites of sub-Saharan strains are indicated by black dots. Lineages and corresponding clusters from the phylogenetic tree are shown on the right by brackets. cl. = cluster.

from sub-Saharan Africa revealed 11 different ON1 and 22 different BAIX lineages with characteristic signature amino acid substitutions.

With bootstrap values >70% and pairwise nucleotide distances <0.07, cluster ON1-NGR (cluster d) of ON1 strains and the three sub-clusters BAIX-KEN1, BAIX-KEN2 and BAIX-KEN3 of BAIX cluster d strains from sub-Saharan Africa formally fulfilled the criteria for the definition of new RSV genotypes [15]. As these clusters were only supported by a total number of seven (ON1) and four to thirteen (BAIX) strains, respectively, this observation has to be confirmed in further studies. However, these findings could be a first trend of differentiation of RSV genotypes ON1 and BAIX into further subgenotypes.

The majority of identified ON1 and BAIX lineages were confined to one area of isolation, whereas some lineages were also found in several countries. These data suggest that different lineages of RSV evolve and circulate both locally as well as globally by introduction of new lineages into distinct communities. While ON1 strains had a stable G protein length of 321 amino acids, the length of the BAIX G protein was more variable with either 312 or 319 amino acids. Genetic alterations such as amino acid substitutions can cause antigen variation and might therefore provide a survival advantage to viral strains by evasion of host immune response [56–60]. Glycosylation of the C-terminal end of the G protein affects the expression of epitopes, either hiding or unmasking them from antibody recognition [6, 61]. ON1 strains from sub-Saharan Africa showed a total of three additional putative O-glycosylation sites in the duplicated region when compared to the ON1 reference strain. BAIX strains from sub-Saharan Africa also showed an increased number of predicted O-glycosylation sites when compared to the BAIX reference strain, including one

O-glycosylation site located in the duplicated region (aa position 270). Positively selected codons can represent an evolutionary advantage and amino acid substitutions of these codon sites may predict which emerging strain variant will predominate in future epidemics.

A total of three positively and eight negatively selected sites were identified for genotype ON1 and five positively and eleven negatively selected sites were found for genotype BAIX strains. Only one of the positively selected sites of ON1 (codon site 250) that were found in this study have been described previously [31, 62–64], whereas four of the BAIX positively selected sites have already been identified in previous studies (amino acid positions 219, 267, 270, 287) [64–66]. The overall mean non-synonymous/synonymous substitution rate (dN/dS) was 0.817 for ON1 and 0.723 for BAIX strains, indicating a rather purifying selection rather than a positive selection [36, 67].

Conclusions

Patterns of circulating RSV genotypes provide valuable information about the geographical spread, transmission dynamics and the pathogenicity of RSV strains. Especially in resource-poor settings, further surveillance of circulating and newly emerging RSV genotypes in combination with corresponding clinical data is needed to understand their full implications.

Acknowledgements

The authors wish to acknowledge and thank the following institutions and their teams for contributing to this study: Kumasi Center for Collaborative Research in Tropical Medicine, Kumasi/Ghana; Bernhard Nocht Institute for Tropical Medicine, Hamburg/Germany; NIMR-Mbeya Medical Research Center, Mbeya/Tanzania; Division for Infectious Diseases, Ludwig-Maximilian University Munich/Germany; Centre de Recherche en Santé de Nouna (CRSN), Nouna/Burkina Faso; Heidelberg Institute of Global Health, Heidelberg University Hospital, Heidelberg/Germany; Centre de Recherches Médicales de Lambaréné, Lambarné/Gabon; Eberhard Karls University, Tuebingen/Germany; Institute of Virology, University of Bonn Medical Center, Bonn/Germany. The authors would also like to thank the technicians in the virology diagnostic laboratory at the University Hospital Heidelberg for excellent technical assistance. The study was supported by a grant of the German Center for Infection Research (DZIF). CMI and JT were also financially supported by DZIF. Open Access funding enabled and organized by Projekt DEAL. WOA Institution: RUPRECHT KARLS UNIVERSITAET HEIDELBERGBlended DEAL: Projekt DEAL.

References

1. Szabo SM, Gooch KL, Bibby MM *et al.* The risk of mortality among young children hospitalized for severe respiratory syncytial virus infection. *Paediatr Respir Rev.* 2013: **13** (Suppl 2): S1–S8.
2. Glezen WP. Viral pneumonia as a cause and result of hospitalization. *J Infect Dis.* 1983: **147**: 765–770.
3. Shi T, McAllister DA, O'Brien KL *et al.* Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *Lancet* 2017: **390**: 946–958.
4. Nair H, Nokes DJ, Gessner BD *et al.* Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *Lancet* 2010: **375**: 1545–1555.
5. Weber MW, Mulholland EK, Greenwood BM. Respiratory syncytial virus infection in tropical and developing countries. *Trop Med Int Health* 1998: **3**: 268–280.
6. Cane PA. Molecular epidemiology of respiratory syncytial virus. *Rev Med Virol.* 2001: **11**: 103–116.
7. Weigl JA, Puppe W, Rockahr S, Schmitt HJ. Burden of disease in hospitalized RSV-positive children in Germany. *Klin Padiatr.* 2002: **214**: 334–342.
8. Glezen WP, Taber LH, Frank AL, Kasel JA. Risk of primary infection and reinfection with respiratory syncytial virus. *Am J Dis Child.* 1986: **140**: 543–546.
9. Henderson FW, Collier AM, Clyde WA, Denny FW. Respiratory-syncytial-virus infections, reinfections and immunity. A prospective, longitudinal study in young children. *N Engl J Med.* 1979: **300**: 530–534.
10. Mufson MA, Orvell C, Rafnar B, Norrby E. Two distinct subtypes of human respiratory syncytial virus. *J Gen Virol.* 1985: **66**: 2111–2124.
11. García O, Martín M, Dopazo J *et al.* Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *J Virol* 1994: **68**: 5448–5459.
12. Johnson PR, Spriggs MK, Olmsted RA, Collins PL. The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc Natl Acad Sci USA* 1987: **84**: 5625–5629.
13. Peret TC, Hall CB, Schnabel KC, Golub JA, Anderson LJ. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. *J Gen Virol.* 1998: **79**: 2221–2229.
14. Peret TCT, Hall CB, Hammond GW *et al.* Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. *J Infect Dis.* 2000: **181**: 1891–1896.
15. Venter M, Madhi SA, Tiemessen CT, Schoub BD. Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. *J Gen Virol.* 2001: **82**(Pt 9): 2117–2124.

16. Shobugawa Y, Saito R, Sano Y *et al.* Emerging genotypes of human respiratory syncytial virus subgroup A among patients in Japan. *J Clin Microbiol.* 2009: **47**: 2475–2482.
17. Eshaghi A, Duvvuri VR, Lai R *et al.* Genetic variability of human respiratory syncytial virus A strains circulating in Ontario: a novel genotype with a 72 nucleotide G gene duplication. *PLoS One* 2012: **7**: e32807.
18. Arnott A, Vong S, Mardy S *et al.* A study of the genetic variability of human respiratory syncytial virus (HRSV) in Cambodia reveals the existence of a new HRSV group B genotype. *J Clin Microbiol.* 2011: **49**: 3504–3513.
19. Blanc A, Delfraro A, Frabasile S, Arbiza J. Genotypes of respiratory syncytial virus group B identified in Uruguay. *Arch Virol.* 2005: **150**: 603–609.
20. Dapat IC, Dapat IC, Dapat C *et al.* New genotypes within respiratory syncytial virus group B genotype BA in Niigata, Japan. *J Clin Microbiol.* 2010: **48**: 3423–3427.
21. Baek YH, Choi EH, Song M-S *et al.* Prevalence and genetic characterization of respiratory syncytial virus (RSV) in hospitalized children in Korea. *Arch Virol.* 2012: **157**: 1039–1050.
22. Khor C-S, Sam I-C, Hooi P-S, Chan Y-F. Displacement of predominant respiratory syncytial virus genotypes in Malaysia between 1989 and 2011. *Infect Genet Evol.* 2013: **14**: 357–360.
23. Auksornkitti V, Kamprasert N, Thongkomplew S *et al.* Molecular characterization of human respiratory syncytial virus, 2010–2011: identification of genotype ON1 and a new subgroup B genotype in Thailand. *Arch Virol.* 2014: **159**: 499–507.
24. Cui G, Zhu R, Qian Y *et al.* Genetic variation in attachment glycoprotein genes of human respiratory syncytial virus subgroups a and B in children in recent five consecutive years. *PLoS One* 2013: **8**: e75020.
25. Gimferrer L, Campins M, Codina MG *et al.* Molecular epidemiology and molecular characterization of respiratory syncytial viruses at a tertiary care university hospital in Catalonia (Spain) during the 2013–2014 season. *J Clin Virol.* 2015: **66**: 27–32.
26. Ábrego LE, Delfraro A, Franco D *et al.* Genetic variability of human respiratory syncytial virus group B in Panama reveals a novel genotype BA14. *J Med Virol.* 2017: **89**: 1734–1742.
27. Trento A, Viegas M, Galiano M *et al.* Natural history of human respiratory syncytial virus inferred from phylogenetic analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication. *J Virol.* 2006: **80**: 975–984.
28. Agoti CN, Otieno JR, Gitahi CW, Cane PA, Nokes DJ. Rapid spread and diversification of respiratory syncytial virus genotype ON1. *Kenya. Emerging Infect Dis.* 2014: **20**: 950–959.
29. Trento A, Galiano M, Videla C *et al.* Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. *J Gen Virol.* 2003: **84**(Pt 11): 3115–3120.
30. Calderón A, Pozo F, Calvo C *et al.* Genetic variability of respiratory syncytial virus A in hospitalized children in the last five consecutive winter seasons in Central Spain. *J Med Virol.* 2016: **89**: 767–774.
31. Duvvuri VR, Granados A, Rosenfeld P, Bahl J, Eshaghi A, Gubbay JB. Genetic diversity and evolutionary insights of respiratory syncytial virus A ON1 genotype: global and local transmission dynamics. *Sci Rep.* 2015: **5**: 14268.
32. Pierangeli A, Trotta D, Scagnolari C *et al.* Rapid spread of the novel respiratory syncytial virus A ON1 genotype, central Italy, 2011 to 2013. *Eurosurveillance* 2014: **19**).
33. Panayiotou C, Richter J, Koliou M, Kalogirou N, Georgiou E, Christodoulou C. Epidemiology of respiratory syncytial virus in children in Cyprus during three consecutive winter seasons (2010–2013): age distribution, seasonality and association between prevalent genotypes and disease severity. *Epidemiol Infect.* 2014: **142**: 1–6.
34. Tabatabai J, Prifert C, Pfeil J, Grulich-Henn J, Schnitzler P. Novel respiratory syncytial virus (RSV) genotype ON1 predominates in Germany during winter season 2012–13. *PLoS One* 2014: **9**: e109191.
35. Ogunsemowo O, Olaleye DO, Odaibo GN. Genetic diversity of human respiratory syncytial virus circulating among children in Ibadan, Nigeria. *PLoS One* 2018: **13**: e0191494.
36. Obodai E, Odoom JK, Adiku T *et al.* The significance of human respiratory syncytial virus (HRSV) in children from Ghana with acute lower respiratory tract infection: A molecular epidemiological analysis, 2006 and 2013–2014. *PLoS One* 2018: **13**: e0203788.
37. Julenius K, Mølgaard A, Gupta R, Brunak S. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 2005: **15**: 153–164.
38. Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 2010: **26**: 2455–2457.
39. Haider MSH, Khan WH, Deeba F *et al.* BA9 lineage of respiratory syncytial virus from across the globe and its evolutionary dynamics. *PLoS One* 2018: **13**: e0193525.
40. Hall CB, Weinberg GA, Iwane MK *et al.* The burden of respiratory syncytial virus infection in young children. *N Engl J Med.* 2009: **360**: 588–598.
41. Stockman LJ, Curns AT, Anderson LJ, Fischer-Langley G. Respiratory syncytial virus-associated hospitalizations among infants and young children in the United States, 1997–2006. *Pediatr Infect Dis J.* 2012: **31**: 5–9.
42. Stein RT, Bont LJ, Zar H *et al.* Respiratory syncytial virus hospitalization and mortality: Systematic review and meta-analysis. *Pediatr Pulmonol.* 2017: **52**: 556–569.
43. Shay DK, Holman RC, Newman RD, Liu LL, Stout JW, Anderson LJ. Bronchiolitis-associated hospitalizations among US children, 1980–1996. *JAMA* 1999: **282**: 1440–1446.
44. Boyce TG, Mellen BG, Mitchel EF, Wright PF, Griffin MR. Rates of hospitalization for respiratory syncytial virus infection among children in medicaid. *J Pediatr.* 2000: **137**: 865–870.
45. Muro F, Reyburn R, Reyburn H. Acute respiratory infection and bacteraemia as causes of non-malarial febrile illness in African children: a narrative review. *Pneumonia* 2015: **6**: 6–17.
46. Kenmoe S, Vernet M-A, Miszczak F *et al.* Genetic diversity of human respiratory syncytial virus isolated among children

with acute respiratory infections in Southern Cameroon during three consecutive epidemic seasons, 2011–2013. *Trop Med Health*. 2018: **46**: 7.
47. Fall A, Dia N, Cisse EHAK *et al*. Epidemiology and molecular characterization of human respiratory syncytial virus in Senegal after Four consecutive years of surveillance, 2012–2015. *PLoS One* 2016: **11**: e0157163.
48. Khalil SO, Enan KA, Salim B, Elkhidir IM. Detection and molecular characterization of respiratory syncytial virus (RSV) in children with respiratory signs in Khartoum State, Sudan 2011–2012. *Am J Infect Dis Microbiol*. 2015: **3**: 6–13.
49. Cane PA, Weber M, Sanneh M, Dackour R, Pringle CR, Whittle H. Molecular epidemiology of respiratory syncytial virus in The Gambia. *Epidemiol Infect* 1999: **122**: 155–160.
50. Fernandes JF, Held J, Dorn M *et al*. Causes of fever in Gabonese children: a cross-sectional hospital-based study. *Sci Rep*. 2020: **10**: 2080–2111.
51. Hogan B, Ammer L, Zimmermann M *et al*. Burden of influenza among hospitalized febrile children in Ghana. *Influenza Other Respir Viruses* 2017: **11**: 497–501.
52. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016: **33**: 1870–1874.
53. Gupta R, Jung E, Brunak S. Prediction of N-glycosylation sites in human proteins. [Internet]. NetNGlyc. Server, 2004. (Available from: http://www.cbs.dtu.dk/services/NetNGlyc/) [cited 2014 Apr 6]
54. Mahende C, Ngasala B, Lusingu J *et al*. Aetiology of acute febrile episodes in children attending Korogwe District Hospital in north-eastern Tanzania. *PLoS One* 2014: **9**: e104197.
55. Otieno JR, Kamau EM, Oketch JW *et al*. Whole genome analysis of local Kenyan and global sequences unravels the epidemiological and molecular evolutionary dynamics of RSV genotype ON1 strains. *Virus Evol*. 2018: **4**: vey027.
56. Roca A, Loscertales MP, Quintó L *et al*. Genetic variability amo ng group A and B respiratory syncytial viruses in Mozambique: identification of a new cluster of group B isolates. *J Gen Virol*. 2001: **82**(Pt 1): 103–111.
57. Rueda P, Delgado T, Portela A, Melero JA, García-Barreno B. Premature stop codons in the G glycoprotein of human respiratory syncytial viruses resistant to neutralization by monoclonal antibodies. *J Virol*. 1991: **65**: 3374–3378.
58. Martinez X, Li X, Kovarik J, Klein M, Lambert PH, Siegrist CA. Combining DNA and protein vaccines for early life immunization against respiratory syncytial virus in mice. *Eur J Immunol*. 1999: **29**: 3390–3400.
59. Zheng X, Song Z, Li Y, Zhang J, Wang X-L. Possible interference between seasonal epidemics of influenza and other respiratory viruses in Hong Kong, 2014–2017. *BMC Infect Dis*. 2017: **17**: 772–777.
60. Melero JA, García-Barreno B, Martínez I, Pringle CR, Cane PA. Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. *J Gen Virol*. 1997: **78**: 2411–2418.
61. Palomo C, García-Barreno B, Peñas C, Melero JA. The G protein of human respiratory syncytial virus: significance of carbohydrate side-chains and the C-terminal end to its antigenicity. *J Gen Virol*. 1991: **72**(Pt 3): 669–675.
62. Martinelli M, Frati ER, Zappa A *et al*. Phylogeny and population dynamics of respiratory syncytial virus (Rsv) A and B. *Virus Res*. 2014: **189**: 293–302.
63. Tan L, Lemey P, Houspie L *et al*. Genetic variability among complete human respiratory syncytial virus subgroup A genomes: bridging molecular evolutionary dynamics and epidemiology. *PLoS One* 2012: **7**: e51439.
64. Gaunt ER, Harvala H, McIntyre C, Templeton KE, Simmonds P. Disease burden of the most commonly detected respiratory viruses in hospitalized patients calculated using the disability adjusted life year (DALY) model. *J Clin Virol*. 2011: **52**: 215–221.
65. Haider MSH, Deeba F, Khan WH *et al*. Global distribution of NA1 genotype of respiratory syncytial virus and its evolutionary dynamics assessed from the past 11 years. *Infect Genet Evol*. 2018: **60**: 140–150.
66. Botosso VF, de Zanotto PMA, Ueda M *et al*. Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus. *PLoS Pathog*. 2009: **5**: e1000254.
67. Zlateva KT, Lemey P, Vandamme A-M, Van Ranst M. Molecular evolution and circulation patterns of human respiratory syncytial virus subgroup A: positively selected sites in the attachment G glycoprotein. *J Virol*. 2004: **78**: 4675–4683.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Study population, inclusion criteria and laboratory analyses per study site.

Table S1. Overview of sequences of the hypervariable region of the RSV G-gene from Sub-Saharan countries retrieved from GenBank between 2011–2016.

Table S2. Positive selective sites of RSV A – ON1 strains in sub-Sahara Africa.

Table S3. Positive selective sites of RSV B – BAIX strains in sub-Sahara Africa.

Corresponding Author Julia Tabatabai, Center of Infectious Diseases, Virology, Im Neuenheimer Feld 324, 69120 Heidelberg, Germany. Tel.: +49-6221-56 39149; Fax +49-6221-56 50 03; E-mail: Julia.tabatabai@med.uni-heidelberg.de