


ORIGINAL ARTICLE OPEN ACCESS

# Molecular Analysis of Human Respiratory Syncytial Virus Group B Strains Isolated in Kenya Before and During the Emergence of Pandemic Influenza A/H1N1

Julia Wangui<sup>1,2</sup>  | George Gachara<sup>3</sup> | Victor Mobegi<sup>4</sup> | Charles Agoti<sup>5</sup> | James Otieno<sup>6</sup> | Silvanos Opanda<sup>1</sup> | Benjamin Opot<sup>1</sup> | Joseph N. Ngeranwa<sup>2</sup> | Regina Njeru<sup>7</sup> | Wallace Bulimo<sup>1</sup>

<sup>1</sup>Centre For Virus Research, Kenya Medical Research Institute (KEMRI), Nairobi, Kenya | <sup>2</sup>Department of Biochemistry, Kenyatta University, Nairobi, Kenya | <sup>3</sup>Department of Medical Laboratory Sciences, Kenyatta University, Nairobi, Kenya | <sup>4</sup>Department of Biochemistry, University of Nairobi, Nairobi, Kenya | <sup>5</sup>Department of Epidemiology and Demography, Kenya Medical Research Institute (KEMRI) - Wellcome Trust Program, Nairobi, Kenya | <sup>6</sup>Theiagen Genomics, Highlands Ranch, Colorado, USA | <sup>7</sup>International Livestock Research Institute, Nairobi, Kenya

**Correspondence:** Julia Wangui ([julia.wanguimwangi@gmail.com](mailto:julia.wanguimwangi@gmail.com))

**Received:** 10 September 2024 | **Revised:** 18 January 2025 | **Accepted:** 4 February 2025

**Funding:** The study was funded in part through AFHSC-GEIS funds at US Army Medical Research Directorate–Kenya through the vote of the DEID director as the sponsor and by the Wellcome Trust (grant # 102975).

**Keywords:** COVID-19 pandemic | evolution | genetic variability | HRSV-B genotypes | pandemic influenza | recombination | SARS-CoV-2 | spatial-temporal trends

## ABSTRACT

**Background:** We conducted a retrospective study to explore molecular insights into human respiratory syncytial virus (HRSV) group B strains among patients attending outpatient clinics at government medical facilities both prior and during the onset of Influenza A/H1N1/2009 pandemic outbreak.

**Methods:** We screened 2300 nasopharyngeal swabs using multiplex real time reverse transcriptase polymerase chain reaction. We amplified a segment of the first and second hypervariable regions, as well as the conserved portion of the third domain of the G-gene using HRSV-B specific primers, sequenced by Sanger di-deoxy chain termination method and thereafter analyzed the sequences.

**Results:** We characterized the circulating strains into three known genotypes: SAB4 (1.4%), BA7 (1.4%), and multiple variants of BA9 (97.2%). The majority of BA9 viruses were uniquely Kenyan with only 4% aligning with BA9 lineages found elsewhere. The mean evolutionary rate of the HRSV-B was estimated to be  $3.08 \times 10^{-3}$  substitutions per site per year.

**Conclusion:** Our findings indicate that the circulating HRSV-B viruses in Kenya underwent a slower evolution during the period of 2007–2010. Additionally, our findings reveal the existence of a unique lineage as well as new variants that have not been reported elsewhere to date.

## 1 | Introduction

Human respiratory syncytial virus (HRSV) is a leading viral pathogen that causes acute lower respiratory tract infections (ALRI) in children resulting in hospitalization and severe illness in the elderly and immuno-compromised populations globally

[1, 2]. The virus is estimated to cause over 30,000 million ALRI in children under the age of 5, resulting in up to 199,000 deaths annually [3].

HRSV is enveloped with a negative-sense, single-stranded RNA genome of approximately 15.2 kb. It belongs to

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Influenza and Other Respiratory Viruses* published by John Wiley & Sons Ltd.

*Orthopneumovirus* genus of the *Pneumoviridae* family [4, 5]. It is classified into two HRSV groups (A and B) and further categorized into various genotypes [6, 7]. The nucleotide sequence variability in the C-terminus of the G protein has been widely utilized to classify HRSV into genotypes until recently with efforts to standardize the classification HRSV genotypes [8, 9]. The allocation of genotypes in HRSV has been a topic of ongoing debate within the scientific community. Despite the absence of a unanimous consensus regarding the criteria for genotype classification, a myriad of HRSV-B genotypes persistently emerges in research findings. Notable among these are GB 1-5, SAB 1-4, BA1-14, BA-C, BA-CCA, BA-CCB, CB1, JAB1, NZB1, and NZB2 [9–12]. A new proposed classification system reduces the number of genotypes into seven (GB1–GB7) [9].

HRSV-B with a 60-nucleotide (20-amino acid) duplication was first reported in Buenos Aires, Argentina, in 1999 [13]. As the decade drew to a close, a noteworthy phenomenon in the virological landscape unfolded with the global dissemination of genotypes in this subgroup (BA subgroup). It was first detected in Kenya in 2003, and within a decade, it had replaced the existing genotypes in the country with exception of the sporadic SAB4 also reported elsewhere [14–16]. Over the past decade, the BA9 genotype has exhibited a growing prevalence, gradually asserting its dominance on a global scale and consequently displacing other genotypes [17–19].

The primary objective of this study was to genetically profile and deduce the evolutionary dynamics of human respiratory syncytial virus subgroup B (HRSV-B) among individuals exhibiting influenza-like illness (ILI). These were a proportion of patients seeking treatment at eight public hospitals in the country during the period from 2007 to 2010. It is noteworthy that this time-frame corresponds to the time just before and during the emergence of the influenza pandemic attributed to the novel A/H1N1 pandemic strain.

## 2 | Materials and Methods

### 2.1 | Study Population

This study was carried out within the department of emerging infectious diseases' (DEID) surveillance network of the US Army Medical Research Directorate–Kenya. The network included eight outpatient hospitals spread across the country. The study used 2300 samples from the Kenyan Influenza Surveillance Protocol, which were randomly selected from an estimated 11,000 samples collected between March 2007 and February 2010 [20]. The surveillance network comprised of the following public hospitals: Port Reitz and Malindi (Coast region), Mbagathi (Nairobi region), Isiolo (Eastern region), Kisii and Kericho (Highlands region), and Alupe and New Nyanza Children's hospitals (Western region) (Figure 1) [20].

### 2.2 | Ethics Statement

This study was approved by the Kenya Medical Research Institute's Scientific and Ethical Review Unit (SERU) (SSC 2051)

and the Walter Reed Army Institute of Research's Institutional Review Board (IRB) (WRAIR 1267 subprotocol 1).

### 2.3 | RNA Extraction, Screening, Amplification, and Nucleotide Sequencing

A previously published approach was utilized to screen for HRSV using modified multiplex real-time reverse transcription polymerase chain reaction (MPX RT-PCR). The assay used HRSV group-specific Taqman primers (HRSV-A and HRSV-B) and a primer specific to adenoviruses [21, 22].

The cDNA was obtained from 5  $\mu$ L of the MPX RT-PCR positive RNAs in a 25- $\mu$ L reaction assay utilizing the QIAGEN one-step Reverse Transcriptase Polymerase Chain Reaction (RT-PCT) kit. To amplify the larger PCR fragment containing a portion of the second domain of the G-gene, the third domain of the G-gene, and a portion of the F-gene, the reaction used two flanking primers consisting of a forward primer, AG20 (5'-GGGGCAAATGCAAACATGTCC-3'), and reverse primer, F164 (5'-GTTATGACACTGGTATACCAACC-3'). The reaction involved the following steps: 1 cycle of 30 min at 50°C, 1 cycle of 15 min at 95°C, 40 cycles of (30 s at 94°C followed by 54°C for 30 s and 1 min at 72°C) and 10 min at 72°C.

Thereafter, 2  $\mu$ L of the One-Step RT-PCR product was utilized in a nested PCR reaction to amplify the third domain of the G-gene containing the first hypervariable region, conserved region, and the second hypervariable region, using Taqman PCR master mix kit (Qiagen). The reaction utilized a forward primer, BG10 (5'-GCAATGATAATCTCAACCTC-3'), and F1 reverse primer (5'-CAACTCCATTGTTATTTGCC-3') [21].

The amplicons were sequenced utilizing the Sanger ABI dideoxy chain termination technology. Four primers were utilized; two of the primers used in the nested PCR (BG10 and F1) and two group-specific primers, G533F (5'-TG TAGTATATGTGGCAACAA-3') and G533R (5'-TTGTTGCCACATATACTACA-3') [21]. The resulting nucleotide sequences were edited, assembled, and aligned using multiple tools including DNABaser version 3.2 [23] for sequence assembly, BioEdit program version 7.7 [24] and Muscle v3.8.31 software [25] for multiple sequence alignment.

### 2.4 | Phylogenetic and Evolution Analysis

These analyses utilized the complete portion of the G-gene sequence recovered, with length varying from 534 to 594 base pairs depending on the virus. We aligned 73 local sequences with 31 global sequences of known genotypes retrieved from GenBank for up to 2017. The criteria used to curate these sequences include (i) highest percentage nucleotide identity and (ii) length that is equals to the local sequences.

The phylogenetic trees were inferred using MrBayes version 3.2 utilizing the Metropolis-coupled Markov chain Monte Carlo (MC3) algorithm [26, 27]. Further selection was conducted to exclude multiple global sequences in a single cluster and those clusters devoid of local sequences in the final presentation of the phylogenetic tree.



**FIGURE 1** | Kenyan map showing the locations of health facilities in the previously known provinces. The five regions examined in this study comprise of one or two public hospitals. The Port Reitz and Malindi subcounty hospitals serve the Coast region, while the Mbagathi county hospital serves the Nairobi region and the Isiolo county hospital represents the Eastern region. The Highlands region encompasses the Kericho county hospital in the middle belt of the rift valley and the Kisii county hospital in the southeast of the larger Nyanza region. The Western region is represented by the Alupe subcounty hospital and the New Nyanza Children's Hospital, which is located in the northeast of Nyanza province.

MEGA 11 was utilized to compute the genetic diversity within the population, evaluate the variability of the virus, and to identify the best fit model for subsequent analysis. Models with the lowest Bayesian information criterion (BIC) scores were considered for evolutionary analysis [28].

The evolutionary rate was estimated using the Bayesian Markov chain Monte Carlo (MCMC)-based approach implemented in BEAST v1.8.2 [29]. The evolutionary analysis assumed an uncorrelated lognormal relaxed molecular clock model and implemented the Hasegawa-Kishino-Yano and gamma (G) plus invariant (I) nucleotide (HKY + G + I) substitution model. BEAST run convergence confirmation (ESS > 200) and analysis of parameter estimates was accomplished using Tracer v1.6 [30]. The BEAST trees were summarized using TreeAnnotator v1.8.2, and the maximum clade

credibility (MCC) tree was visualized using Figtree v1.4.3 software [31].

## 2.5 | Selection Pressure and Recombination Analysis

Selection pressure was estimated using four methods implemented in the web based DATAMONKEY application: single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), fixed unconstrained Bayesian approximation (FUBAR), and mixed effect model of evolution (MEME) [32, 33]. A *p*-value of 0.01 (SLAC, FEL, and MEME) and a posterior probability of 0.9 (FUBAR) were applied in determining the positively ( $dN/dS > 1$ ) and negatively ( $dN/dS < 1$ ) selected sites, respectively. The genetic algorithm for recombination detection (GARD)

implemented within the DATAMONKEY application [34] was utilized to detect potential recombination events and thereafter verified using the RDP, BootScan, 3SEQ, Chimaera, and GENECONV methods implemented the Recombination Detection Program version 4.39 (RDP4) [35].

### 3 | Results

#### 3.1 | Detection, Spatial, and Temporal Distribution of HRSV-B

Of the 2300 clinical samples tested, 145 (6.3%) were positive for HRSV-B. The virus was observed in varying proportions with distinct peaks each year. HRSV-B cases were detected in all regions during Q2 and Q1 of 2008 and 2010, respectively (Figure 2a).

HRSV-B cases were highest in the Coast and Nairobi regions during the first quarter with a variation seen in 2009 (Figure 2c,d). The incident peaks varied in highlands region from 2007 to 2009 (Figure 2e). Quarters 1 of 2007 and 2010 exhibited the highest incidence in the Western region (Figure 2b). The lowest incidence was observed in 2009, with no cases reported from the Eastern region (Figure 2f). This analysis revealed an upswing in HRSV-B activity during the initial quarter of 2010 in four of the regions, except for the Highlands region, where there was a declining trend.

#### 3.2 | Characterization of HRSV-B and Genotype Distribution in Kenya, 2007–2010

The sequences were deposited in the GenBank (accession numbers PP101687–PP101759). The phylogenetic analysis showed that 71 of these sequences (97.2%) formed the BA9 cluster, followed by two sequences, (2.8%) clustering with the BA7 and SAB4 genotypes each (Figure 3). In the Nairobi region, SAB4 and BA9 genotypes were identified, while BA7 and BA9 genotypes were detected in the Highlands. Conversely, only the BA9 genotype was observed in the Coast, Eastern, and Western regions with the lowest cases seen in the Eastern (10%) and Western (9%) regions.

The phylogenetic tree displays BA9 clade, (clade credibility value of 70%), delineated into seven clusters which we designated as BA9.1–BA9.7. When compared to the prototype (JX198143), these clusters presented distinct characteristics at the nucleotide level and amino acid level. The BA9.1 cluster was made up of four sequences that remained on the BA9 clade as well as two sequences from an adjacent cluster. The 9.1 and 9.2 clusters comprised sequences collected during the first two years of the study, whereas 9.3 cluster included sequences collected throughout the study. This was the most extensive of the BA9 clusters, encompassing 33 sequences, distinguished by a singular nucleotide substitution, C843T.

The sequences in the remaining four clusters were collected over 2–4 years of the study. Markedly, three of the clusters (BA9.4, BA9.5, and BA9.7) showed features marked by one to three nucleotide substitutions with an amino acid change. The

BA9.4 cluster exhibited three nucleotide substitutions leading to one amino acid change: A345T, C489T, and A539G (resulting in Q180R). On the other hand, BA9.7 cluster showed a single nucleotide change, T752C, (resulting in V251A) whereas BA9.5 cluster exhibited two nucleotide changes and one amino acid change: C620T (resulting in T207L) and C645T.

Remarkably, only three of the local sequences in the BA9 genotype clustered with global sequences collected between 2009 and 2017, as seen in Figure 3, cluster BA9.6. These sequences showed a single nucleotide mutation (C648T) and were collected in 2008 and 2009.

#### 3.3 | Genetic Analysis and Amino Acids Deduction

In the realm of genetic analysis and amino acid deduction, the mean distance within the BA subgroup (BA7 and BA9 genotypes shown in Figure 3) extended up 0.008 with an overall distance mean of 0.013. Conversely, the mean distance between the BA9 clusters exhibited a range of 0.005 to 0.021 (Table 1).

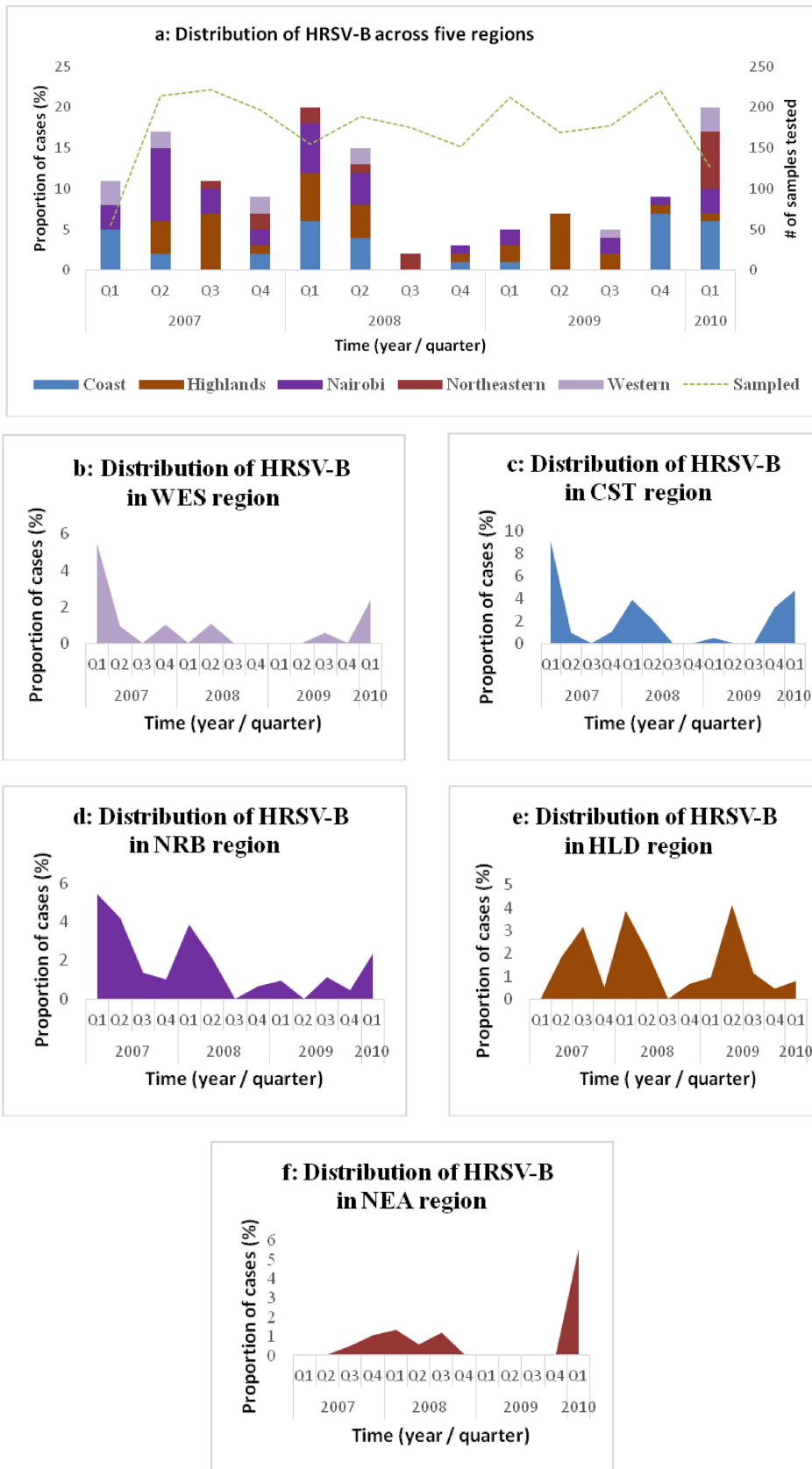
The Kenyan BA subgroup exhibited amino acid variability in comparison to the BA prototype (AY333364), with variations observed in 30 out of the 196 amino acids. All the local sequences had a 3-amino acid deletion at positions 157–159 and 4 substitutions at the following positions: Y112H, K218T, L223P, and S247P (Figure 5a,b). These sequences displayed mutations within the region of 20-amino acid duplication (comprising a 60-nucleotide duplication), amounting to a total of seven mutations. Among the Kenyan sequences, significant mutations were identified at positions 11 (T to I) and 12 (V to A), observed in 62 (86%) and 69 (95%) of the sequences, respectively (Figure 5b).

#### 3.4 | Evolution, Selection Pressure, and Recombination Analyses

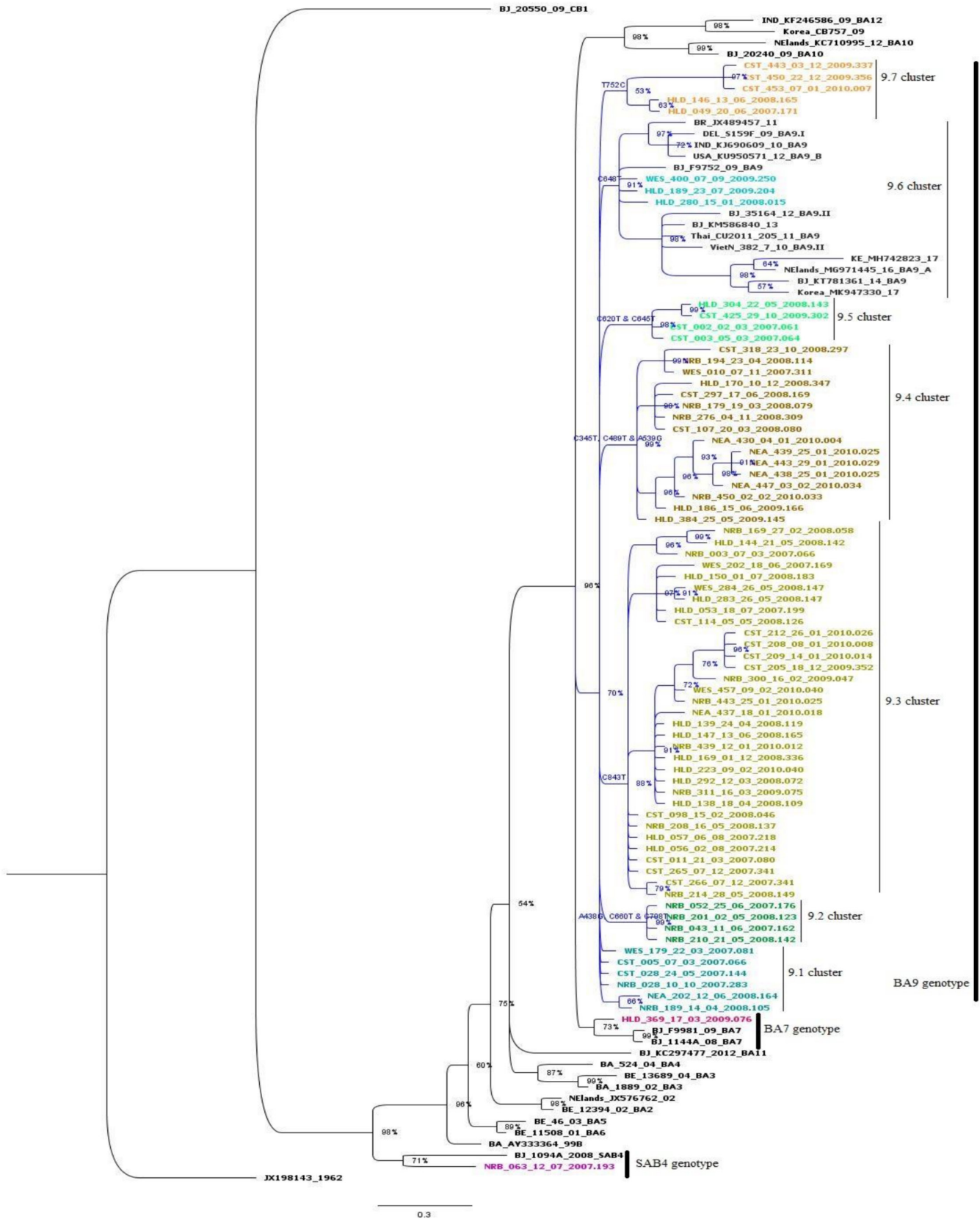
The mean evolutionary rate for the local HRSV-B was estimated at  $3.08 \times 10^{-3}$  substitutions per site per year (95% high posterior density [HPD],  $1.81 \times 10^{-3}$  to  $6.48 \times 10^{-4}$ ) during the study period. We further estimated the mean evolutionary rate for the BA subgroup at  $2.78 \times 10^{-3}$  substitutions per site per year (95% high posterior density [HPD],  $1.65 \times 10^{-3}$  to  $6.77 \times 10^{-4}$ ).

The Bayesian MCMC analysis showed that the local SAB4 genotype first emerged around 1990. A major divergence took place around 2001 leading to the emergence of two distinct lineages within the BA subgroup (Figures 4 and S1). In 2004, each of the two lineages diverged further, resulting in two distinct sublineages. Following this split, further diversification occurred, leading to the appearance of the BA7 genotype in 2006 among other sublineages.

The FEL, MEME, and FUBAR models concurred in pinpointing two specific sites (270 and 287) subjected to episodic and pervasive diversification (Table 2). The SLAC model identified three sites subjected to purifying selection and none under positive selection. In contrast, the FEL model revealed the potential influence of selection on up to 11 codons. The three selection models



**FIGURE 2** | (a–f) Temporal spatial distribution of HRSV-B in Kenya. (a) Quarterly distribution of HRSV-B from 2007 to 2010. It is displaying the proportion of cases in the five regions and the total number of samples tested per quarter. The quarters followed the calendar year starting January to December. (b–f) Temporal distribution of HRSV-B in five regions. (b) Western (WES), (c) Coast (CST), (d) Nairobi (NRB), (e) Highlands (HLD), and (f) Eastern (NEA).



**FIGURE 3** | Phylogenetic tree based on the G-gene segment of the HRSV-B recovered from the ILI cases across Kenya. It is made up of 104 sequences rooted on the prototype isolated in 1962, accession number JX198143. The local sequences (73) are represented by the colored tips. Each individual cluster or subcluster has a distinct color for the tips. We only assigned genotype names to clusters comprising local sequences. The BA9 genotype comprised of seven clusters labeled as BA9.1–BA9.7. The nucleotide differences among these sequences were annotated based on the prototype sequence and therefore does not display the changes in the duplicated segment among these sequences.

**TABLE 1** | The group means between the clusters within the BA9 genotype.

	<b>BA9.7</b>	<b>BA9.5</b>	<b>BA9.4</b>	<b>BA9.3</b>	<b>BA9.2</b>	<b>BA9.6</b>	<b>BA9.1</b>
BA9.7							
BA9.5	0.01816						
BA9.4	0.02102	0.02009					
BA9.3	0.01673	0.01688	0.01779				
BA9.2	0.01905	0.01750	0.01716	0.01500			
BA9.6	0.01603	0.01451	0.01405	0.01202	0.01105		
BA9.1	0.01265	0.01151	0.01103	0.00905	0.00807	0.00458	

Note: The analysis utilized MEGA 11 for evolutionary divergence using 71 sequences. The 7 clusters within the BA9 genotype were considered as groups in this analysis.

were further used to analyze sites under selection pressure for the population by year of collection (Table 2). Codon 270 was seen to be under diversification by at least two models in 2008 and 2010. This residue was under both pervasive diversification and episodic diversification in 2008.

The genetic algorithm for recombination detection (GARD) program detected a potential breakpoint while no recombination events were detected by the RDP, BootScan, 3SEQ, Chimaera, and GENECONV methods in the Recombination Detection Program version 4.39.

#### 4 | Discussion

The current research offers a comprehensive analysis of the spatiotemporal distribution patterns, genetic diversity and molecular evolution of human respiratory syncytial virus B (HRSV-B) among individuals exhibiting acute respiratory illness with influenza-like symptoms in Kenya. This investigation is particularly pertinent, as it took place during a pivotal historical period marked by the global spread of a new pandemic influenza variant. The study period aligns with the Influenza A/H1N1/2009 outbreak and precedes the widespread emergence of various HRSV-B genotypes. Spanning from March 2007 to February 2010, the timeframe captures key developments in the evolution and distribution of HRSV-B during a period of significant global health concern.

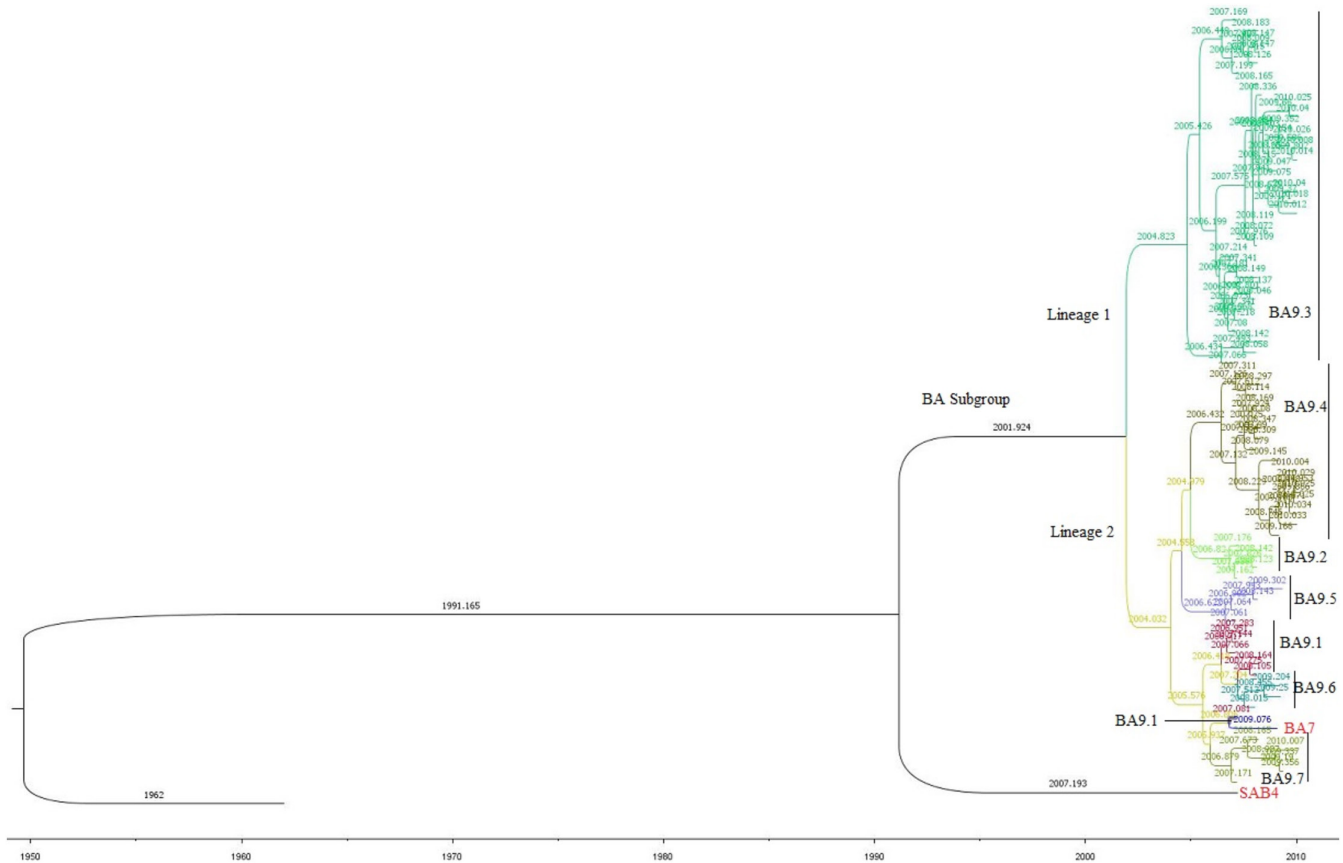
Our data were grouped in quarters following the calendar year due to the low monthly number of HRSV-B cases. The geographic temporal distribution pattern varied by location; however, certain regions showed comparable patterns. This analysis reveals variable seasonal patterns in the country with three to five HRSV-B peaks from March 2007 to February 2010. First, in the Nairobi region, the highest HRSV-B activity was in the first quarter (January–March), with the lowest in the fourth quarter (October–December). Second, the highest activity in the Coast region was in the first quarter with the lowest occurrences seen during the third quarter (July–September). Finally, the highest activity in the Western and Highlands regions demonstrated a dynamic pattern of fluctuation ranging from the first to third quarters. The virus was detected in low proportions in the Eastern region with a notable increase in transmission at the end of the study period.

There was reduced or no HRSV-B activity in 2009 across the regions. While the reasons are not clear, they may be attributed to (i) a biennial HRSV cycle with significant decline of HRSV-B during the year [36] and (ii) low HRSV incidences during the high Influenza activity as reported in the Western Kenya [37]. A study utilizing same sample population recorded the highest influenza activity during the third quarter with Influenza/H1N1/2009 recording the highest incidence in the third and fourth quarters within this demographic [38]. This study highlights the molecular epidemiology of HRSV-B during the Influenza/H1N1/2009 pandemic, offering insights into the ecological dynamics of respiratory viruses. The reduced HRSV-B activity during that period suggests possible viral interference or host immune interactions.

Similar patterns emerged during the COVID-19 pandemic, where SARS-CoV-2 significantly altered RSV transmission, leading to delayed or off-season outbreaks due to non-pharmaceutical interventions (NPIs), shifts in population immunity, and viral interactions [39–41]. While our study predates SARS-CoV-2 by a decade, the parallels underscore how novel respiratory pathogens can disrupt endemic virus epidemiology [42]. These findings emphasize the need to investigate interactions between emerging and seasonal viruses to better inform surveillance, vaccination, and outbreak preparedness strategies.

HRSV-B with 60 nucleotide duplication in the second hyper-variable region of G gene was first reported in 1999 [43]. This strain, BA subgroup, was reported to have gradually replaced the existing genotypes within the decade in most of the continents [44]. The current study has identified one sample belonging to the SAB4 genotype, a non-BA group collected in 2007. SAB4 genotype has been sporadically detected elsewhere in the past decade [15, 17, 45]. These findings indicate that the BA subgroup had not fully supplanted the pre-existing genotypes. The HRSV-B, BA subgroup has evolved to over 12 genotypes (BA1–BA12) to date [10, 19, 46]. During the last decade, the BA9 genotype has been reported to be gradually increasing its global dominance [17, 46, 47]. It was the predominant genotype during the period of this study (2007–2010) accounting for over 95% of the sequences, consistent with other studies worldwide [48, 49]. It is worth noting that it belongs to the newly proposed GB5.0.5a genotype [9].

We further evaluated the BA9 genotype using a broader definition approach of variants, which are distinct strains with



**FIGURE 4** | Bayesian Markov chain Monte Carlo (MCMC) tree based on the G-gene segment of the HRSV-B recovered from the ILI cases across Kenya. The tree comprised of 74 sequences. It is rooted on the prototype, accession number JX198143. The two major lineages are represented by distinct colored clades. Lineage 1 comprises of 33 sequences found in cluster 9.3 of Figure 2 while lineage 2 comprises of the remaining 38 BA9 sequences and 1 BA7 sequence represented in different shades. The BA9 clusters shown in Figure 2 are labeled in black font while the BA7 and SAB4 genotypes are labeled in a colored font.

specific mutations in comparison to the prototype strain. We therefore recognize the previously identified 9.1–9.7 BA9 clusters in Figure 3 as variants. These variants differed substantially, forming seven variants (BA9.1–BA9.7) with a genetic distance (P-distance) of  $<0.3$ . Notably, only the BA9.6 variants clustered with sequences from elsewhere in four subvariants: (i) BA9.A from Netherlands, (ii) BA9.B and BA9.I from the United States and India respectively, (iii) BA9.II from China and Vietnam, and (iv) local variant alongside with a variant from China. The current classification criterion is based on monophyletic clusters with  $>70\%$  bootstrap value and  $<0.07$  inter genotypic distances of based on full genome analysis [9, 35, 46, 50].

The genetic changes in a variant can alter viral features, potentially increasing prevalence in certain populations thus classifying high risk variants as variant of concern [8, 12, 51]. While our study did not evaluate the viral characteristics of the local variants, we utilized the Bayesian MCMC analysis to illustrate the ancestor-offspring relationship (lineages) among the BA subgroup (proposed GB5.0.1 genotype) [9]. The BA9.3 variant represented the earliest divergence of the Bayesian MCMC tree. It appeared in 2005 thus representing the first lineage among the local BA subgroup (shown in Figure 4). The remaining BA9 variants (9.1, 9.2, 9.4, 9.5, 9.6, and 9.7) and the BA7 genotype made up the second lineage. The BA9.2 and BA9.5 variants

appeared to be the oldest in the second lineage, followed by the BA7 genotype and the youngest being the BA9.4 variant.

Viral evolution rates vary due to regional factors and population immunity, as evidenced by the disparity in HRSV-B evolution rates across the globe over time [45, 52]. The local BA9 genotype had a slower mean evolution rate of  $2.56 \times 10^{-3}$  substitutions/site/year in comparison with those from elsewhere. For example, during the 2006–2016 time frame, China and India reported a mean evolutionary rate of  $4.93 \times 10^{-3}$  and  $4.58 \times 10^{-3}$  substitutions/site/year, respectively [14, 17].

A number of amino acid variations were observed across the entire segment in comparison with both the prototype sequence (JX198143) and a 1999 BA sequence (AY333364). A mutation at position 180 (Q180R) was observed within the local BA9.4 variant. This mutation was also reported sporadically in other areas [47]. The local BA sequences lacked some of the mutations exhibited by the BA reference strain in positions 112 (H-Y), 218 (T-K), and 223 (P-L), hence similar to the prototype. They further exhibited parallel mutation in position 247 (S247P), which was reported elsewhere [17, 49].

We further observed T270I, V271A, and H287Y substitutions in over 75% of the local sequences, which was consistent with

(a)

	* 10	* 20	* 30	* 40	* 50	* 60	* 70	* 80	* 90	* 100	
	H										
AY333364_99B	:	TTTPPIYINSATISPNKSETHHTTAQTKGR	TTTTPTQNNKPKSTKFFPKNPKKPKK	DDYHFEVFN	FVPC	SICGNNQLCK	SICKTIP	SNKPKK	KFTIKPTNK	:	100
CST_453_07_01_2010.0	:	H							T	:	97
CST_450_22_12_2009.3	:	H							T	:	97
CST_443_03_12_2009.3	:	H							T	:	97
HLD_304_22_05_2008.1	:	H		P						:	97
CST_425_29_10_2009.3	:	H		P						:	97
CST_003_05_03_2007.0	:	H								:	97
CST_002_02_03_2007.0	:	H								:	97
HLD_280_15_01_2008.0	:	TH								:	97
HLD_189_23_07_2009.2	:	H								:	97
WES_400_07_09_2009.2	:	H								:	97
HLD_056_02_08_2007.2	:	H								:	97
HLD_057_06_08_2007.2	:	H								:	97
CST_212_26_01_2010.0	:	H								:	97
CST_098_15_02_2008.0	:	H								:	97
CST_114_05_05_2008.1	:	H								:	97
HLD_053_18_07_2007.1	:	H								:	97
NRB_311_16_03_2009.0	:	H								:	97
NRB_439_12_01_2010.0	:	H								:	97
NRB_003_07_03_2007.0	:	H								:	97
CST_265_07_12_2007.3	:	H								:	97
CST_266_07_12_2007.3	:	H								:	97
NRB_300_16_02_2009.0	:	H								:	97
CST_011_21_03_2007.0	:	H								:	97
HLD_223_09_02_2010.0	:	H								:	97
HLD_283_26_05_2008.1	:	H								:	97
HLD_292_12_03_2008.0	:	H								:	97
HLD_138_18_04_2008.1	:	H								:	97
HLD_139_24_04_2008.1	:	H								:	97
HLD_144_21_05_2008.1	:	H								:	97
CST_205_18_12_2009.3	:	H								:	97
CST_208_08_01_2010.0	:	H								:	97
CST_209_14_01_2010.0	:	H								:	97
HLD_147_13_06_2008.1	:	H								:	97
HLD_150_01_07_2008.1	:	H								:	97
HLD_169_01_12_2008.3	:	H								:	97
NRB_169_27_02_2008.0	:	H								:	97
WES_284_26_05_2008.1	:	H								:	97
WES_457_09_02_2010.0	:	H								:	97
NEA_437_18_01_2010.0	:	H								:	97
WES_202_18_06_2007.1	:	H								:	97
NRB_208_16_05_2008.1	:	H								:	97
NRB_214_28_05_2008.1	:	H								:	97
NRB_443_25_01_2010.0	:	H								:	97
HLD_170_10_12_2008.3	:	H							R	:	97
NEA_430_04_01_2010.0	:	H							R	:	97
NEA_443_29_01_2010.0	:	H							R	:	97
NEA_447_03_02_2010.0	:	H							R	:	97
NEA_438_25_01_2010.0	:	H							R	:	97
NRB_179_19_03_2008.0	:	H							R	:	97
HLD_384_25_05_2009.1	:	H							R	:	97
NEA_439_25_01_2010.0	:	H							R	:	97
NRB_450_02_02_2010.0	:	H							R	:	97
CST_318_23_10_2008.2	:	H		P					R	:	97
CST_107_20_03_2008.0	:	H							R	:	97
NRB_194_23_04_2008.1	:	H		P					R	:	97
CST_297_17_06_2008.1	:	H							R	:	97
HLD_186_15_06_2009.1	:	H							R	:	97
WES_010_07_11_2007.3	:	H		P					R	:	97
NRB_276_04_11_2008.3	:	H							R	:	97
HLD_146_13_06_2008.1	:	H								:	97
NRB_043_11_06_2007.1	:	H								:	97
NRB_049_20_06_2007.1	:	H								:	97
NRB_210_21_05_2008.1	:	H								:	97
NRB_052_25_06_2007.1	:	H								:	97
NRB_201_02_05_2008.1	:	H								:	97
NRB_028_10_10_2007.2	:	H								:	97
CST_028_24_05_2007.1	:	H								:	97
WES_179_22_03_2007.0	:	H							T	:	97
CST_005_07_03_2007.0	:	H								:	97
NEA_202_12_06_2008.1	:	H								:	97
NRB_189_14_04_2008.1	:	H								:	97
HLD_369_17_03_2009.0	:	H							T	:	97

FIGURE 5 | Legend on next page.

other reports across the globe [17, 49, 50, 53]. Worth noting, two of the above mutations were found in the 20 amino acid duplication segment (which starts from 260 up to 279). The V271A and H287Y mutation were also present in the local BA7 sequence and thus not restricted to the BA9 sequences [54].

While the G gene is known to be the most variable gene in the HRSV genome, the current study revealed only two sites under diversifying selection using FEL, FUBAR, and MEME models. While none of the sites were under diversification in 2007, site 270 (threonine to isoleucine) was seen to be under both



**TABLE 2** | The residues of HRSV-BA subgroup under positive selection across Kenya from 2007 to 2010.

Methods	Positively selected sites per year				Amino acid substitution
	2007	2008	2009	2010	
SLAC	None	None	None	None	NA
FEL	270 <sup>a</sup>	270 <sup>a</sup>	None	270	T270I
FUBAR	None	270 <sup>a</sup>	270, 287 <sup>a</sup>	270, 287 <sup>a</sup>	T270I, H287Y
MEME	None	270 <sup>a</sup>	None	None	T270I

Note: The residue position is in reference to a reference strain from Argentina, AY333364. The *p* value of 0.01 was used for SLAC, MEME, and FEL, while the FUBAR model had a posterior probability of 0.9.

Abbreviation: NA: not applicable because there were no positively selected sites during the year under evaluation.

<sup>a</sup>The codon is located within the duplicated region of the G gene.

pervasive and episodic diversification in 2008 by the three models. A second site, 287 (histidine to tyrosine), was identified to be under diversification in 2009 and 2010 using the MEME model expressing episodic diversification. This site was reported elsewhere to be undergoing diversification during the same time frame [53]. Diversifying selection displays a survival advantage under the selective constraints that confront the viral population. Additionally, there was no evidence of recombination in this population. It is worthwhile noting that recombination events are known to be a rare in HRSV [12, 55].

The limitation of the study was mainly on the low numbers of cases and also the sampling period, as samples collected in 2010 were limited to the first 2 months as indicated in the methods and therefore lacking a holistic picture of the HRSV activity at the peak of the Influenza A/H1N1/2009 pandemic.

## 5 | Conclusions

The genetic profiling of HRSV-B reveals the existence of a non-BA subgroup at a time when it was almost extinct, alongside dominance of the BA9 genotype. It sheds more light on the local strain's significant genetic diversity. It emphasizes the uniqueness of the local BA9 genotype, as well as variants that are uniquely Kenyan which calls into question the possibility of unreported genotypes in the country.

### Author Contributions

**Julia Wangui:** conceptualization, investigation, writing – original draft, methodology, validation, visualization, writing – review and editing, formal analysis, project administration, data curation. **George Gachara:** formal analysis, data curation, software, validation, visualization, writing – review and editing, methodology. **Victor Mobegi:** writing – review and editing, data curation, formal analysis. **Charles Agoti:** conceptualization, methodology, data curation. **James Otieno:** validation, data curation, writing – review and editing. **Silvanos Opanda:** writing – review and editing, data curation, formal analysis, methodology. **Benjamin Opot:** data curation, formal analysis, software. **Joseph N. Ngeranwa:** supervision, conceptualization, writing – review and editing. **Regina Njeru:** methodology, validation. **Wallace Bulimo:** conceptualization, investigation, funding acquisition, writing – review and editing, supervision, resources, data curation, formal analysis.

### Acknowledgments

We thank the Virus Epidemiology and Control research group, KEMRI-Wellcome Trust, Kilifi, and the Respiratory Laboratory, KEMRI, Nairobi, staff for facilitating this work to completion. We are grateful to the Director, Kenya Medical Research Institute, for the organization's support in scientific and ethical review.

### Disclosure

The views or insertions expressed herein are private views of the authors and are not to be construed to represent those of the US Department of Defense or Army.

### Conflicts of Interest

The authors declare no conflicts of interest.

### Data Availability Statement

The data presented is available in the National Center for Biotechnology Information using the following GenBank accession numbers: PP101687 - PP101759.

### Peer Review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/irv.70082>.

### References

1. A. R. Falsey and E. E. Walsh, "Respiratory Syncytial Virus Infection in Adults," *Clinical Microbiology Reviews* 13 (2000): 371–384.
2. T. C. Peret, C. B. Hall, G. W. Hammond, et al., "Circulation Patterns of Group A and B Human Respiratory Syncytial Virus Genotypes in 5 Communities in North America," *Journal of Infectious Diseases* 181 (2000): 1891–1896.
3. H. Nair, D. J. Nokes, B. D. Gessner, et al., "Global Burden of Acute Lower Respiratory Infections due to Respiratory Syncytial Virus in Young Children: A Systematic Review and Meta-Analysis," *Lancet (London, England)* 375 (2010): 1545–1555.
4. C. Y. Lee, Y. P. Fang, L. C. Wang, T. Y. Chou, and H. F. Liu, "Genetic Diversity and Molecular Epidemiology of Circulating Respiratory Syncytial Virus in Central Taiwan, 2008–2017," *Viruses* 14 (2021): 32.
5. B. Rima, P. Collins, A. Easton, et al., "ICTV Virus Taxonomy Profile: Pneumoviridae," *Journal of General Virology* 98 (2017): 2912–2913.
6. W. M. Sullender, M. A. Mufson, L. J. Anderson, and G. W. Wertz, "Genetic Diversity of the Attachment Protein of Subgroup B Respiratory Syncytial Viruses," *Journal of Virology* 65 (1991): 5425–5434.

7. S. Goya, C. Ruis, R. A. Neher, et al., "Standardized Phylogenetic Classification of Human Respiratory Syncytial Virus Below the Subgroup Level," *Emerging Infectious Diseases* 30 (2024): 1631–1641.
8. W. M. Sullender, "Respiratory Syncytial Virus Genetic and Antigenic Diversity," *Clinical Microbiology Reviews* 13 (2000): 1–15 table of contents.
9. S. Goya, M. Galiano, I. Nauwelaers, et al., "Toward Unified Molecular Surveillance of RSV: A Proposal for Genotype Definition," *Influenza and Other Respiratory Viruses* 14 (2020): 274–285.
10. L. Houspie, P. Lemey, E. Keyaerts, et al., "Circulation of HRSV in Belgium: From Multiple Genotype Circulation to Prolonged Circulation of Predominant Genotypes," *PLoS ONE* 8 (2013): e60416.
11. P. Hu, T. Zheng, J. Chen, et al., "Alternate Circulation and Genetic Variation of Human Respiratory Syncytial Virus Genotypes in Chengdu, West China, 2009–2014," *Journal of Medical Virology* 89 (2017): 32–40.
12. J. M. Yu, Y. H. Fu, X. L. Peng, Y. P. Zheng, and J. S. He, "Genetic Diversity and Molecular Evolution of Human Respiratory Syncytial Virus A and B," *Scientific Reports* 11 (2021): 12941.
13. A. Trento, M. Galiano, C. Videla, et al., "Major Changes in the G Protein of Human Respiratory Syncytial Virus Isolates Introduced by a Duplication of 60 Nucleotides," *Journal of General Virology* 84 (2003): 3115–3120.
14. J. Song, H. Wang, J. Shi, et al., "Emergence of BA9 Genotype of Human Respiratory Syncytial Virus Subgroup B in China From 2006 to 2014," *Scientific Reports* 7 (2017): 16765.
15. C. N. Agoti, J. R. Otieno, M. Ngama, et al., "Successive Respiratory Syncytial Virus Epidemics in Local Populations Arise From Multiple Variant Introductions, Providing Insights Into Virus Persistence," *Journal of Virology* 89 (2015): 11630–11642.
16. X. Chen, Y. Zhu, W. Wang, et al., "A Multi-Center Study on Molecular Epidemiology of Human Respiratory Syncytial Virus From Children With Acute Lower Respiratory Tract Infections in the Mainland of China Between 2015 and 2019," *Virologica Sinica* 36 (2021): 1475–1483.
17. M. S. H. Haider, W. H. Khan, F. Deeba, et al., "BA9 Lineage of Respiratory Syncytial Virus From Across the Globe and Its Evolutionary Dynamics," *PLoS ONE* 13 (2018): e0193525.
18. R. Fan, C. Fan, J. Zhang, et al., "Respiratory Syncytial Virus Subtype ON1/NA1/BA9 Predominates in Hospitalized Children With Lower Respiratory Tract Infections," *Journal of Medical Virology* 89 (2017): 213–221.
19. N. Korsun, S. Angelova, I. Trifonova, et al., "Predominance of ON1 and BA9 Genotypes of Respiratory Syncytial Virus (RSV) in Bulgaria, 2016–2018," *Journal of Medical Virology* 93 (2021): 3401–3411.
20. W. D. Bulimo, J. L. Garner, D. C. Schnabel, et al., "Genetic Analysis of H3N2 Influenza A Viruses Isolated in 2006–2007 in Nairobi, Kenya," *Influenza and Other Respiratory Viruses* 2 (2008): 107–113.
21. J. Wangui, D. J. Nokes, V. A. Mobegi, et al., "Spatial-Temporal Distribution and Sequence Diversity of Group A Human Respiratory Syncytial Viruses in Kenya Preceding the Emergence of ON1 Genotype," *Influenza and Other Respiratory Viruses* 16 (2022): 501–510.
22. L. L. Hammit, S. Kazungu, S. Welch, et al., "Added Value of an Oropharyngeal Swab in Detection of Viruses in Children Hospitalized With Lower Respiratory Tract Infection," *Journal of Clinical Microbiology* 49 (2011): 2318–2320.
23. S. M. Opanda, F. Wamunyokoli, S. Khamadi, R. Coldren, and W. D. Bulimo, "Genetic Diversity of Human Enterovirus 68 Strains Isolated in Kenya Using the Hypervariable 3'- End of VP1 Gene," *PLoS ONE* 9 (2014): e102866.
24. T. Hall, I. Biosciences, and C. Carlsbad, "BioEdit: An Important Software for Molecular Biology," *GERF Bulletin of Biosciences* 2 (2011): 60–61.
25. R. C. Edgar, "MUSCLE: Multiple Sequence Alignment With High Accuracy and High Throughput," *Nucleic Acids Research* 32 (2004): 1792–1797.
26. F. Ronquist, M. Teslenko, P. van der Mark, et al., "MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space," *Systematic Biology* 61 (2012): 539–542.
27. J. P. Huelsenbeck and F. Ronquist, "MRBAYES: Bayesian Inference of Phylogenetic Trees," *Bioinformatics (Oxford, England)* 17 (2001): 754–755.
28. K. Tamura, G. Stecher, and S. Kumar, "MEGA11: Molecular Evolutionary Genetics Analysis Version 11," *Molecular Biology and Evolution* 38 (2021): 3022–3027.
29. A. J. Drummond, M. A. Suchard, D. Xie, and A. Rambaut, "Bayesian Phylogenetics With BEAUti and the BEAST 1.7," *Molecular Biology and Evolution* 29 (2012): 1969–1973.
30. A. Rambaut, A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard, "Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7," *Systematic Biology* 67 (2018): 901–904.
31. O. Robinson, D. Dylus, and C. Dessimoz, "Phylo.io: Interactive Viewing and Comparison of Large Phylogenetic Trees on the Web," *Molecular Biology and Evolution* 33 (2016): 2163–2166.
32. S. L. Kosakovsky Pond and S. D. Frost, "Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection," *Molecular Biology and Evolution* 22 (2005): 1208–1222.
33. B. Murrell, S. Moola, A. Mabona, et al., "FUBAR: A Fast, Unconstrained Bayesian Approximation for Inferring Selection," *Molecular Biology and Evolution* 30 (2013): 1196–1205.
34. S. L. Kosakovsky Pond, D. Posada, M. B. Gravenor, C. H. Woelk, and S. D. W. Frost, "Automated Phylogenetic Detection of Recombination Using a Genetic Algorithm," *Molecular Biology and Evolution* 23 (2006): 1891–1901.
35. D. P. Martin, B. Murrell, M. Golden, A. Khoosal, and B. Muhire, "RDP4: Detection and Analysis of Recombination Patterns in Virus Genomes," *Virus Evolution* 1 (2015): vev003.
36. C. N. Agoti, L. M. Mayieka, J. R. Otieno, et al., "Examining Strain Diversity and Phylogeography in Relation to an Unusual Epidemic Pattern of Respiratory Syncytial Virus (RSV) in a Long-Term Refugee Camp in Kenya," *BMC Infectious Diseases* 14 (2014): 178.
37. G. O. Emukule, S. Khagayi, M. L. McMorrow, et al., "The Burden of Influenza and RSV Among Inpatients and Outpatients in Rural Western Kenya, 2009–2012," *PLoS ONE* 9 (2014): e105543.
38. K. K. Wong, W. D. Bulimo, J. Magana, et al., "Epidemiology of 2009 Pandemic Influenza A Virus Subtype H1N1 Among Kenyans Aged 2 Months to 18 Years, 2009–2010," *Journal of Infectious Diseases* 206 (2012): S68–S73.
39. J. S. Eden, C. Sikazwe, R. Xie, et al., "Off-Season RSV Epidemics in Australia After Easing of COVID-19 Restrictions," *Nature Communications* 13 (2022): 2884.
40. R. Agha and J. R. Avner, "Delayed Seasonal RSV Surge Observed During the COVID-19 Pandemic," *Pediatrics* 148 (2021): e2021052089.
41. P. Simusika, M. Okamoto, C. Dapat, et al., "Characterization of Human Respiratory Syncytial Virus in Children With Severe Acute Respiratory Infection Before and During the COVID-19 Pandemic," *IJID Regions* 11 (2024): 100354.
42. E. Quintero-Salgado, J. Briseno-Ramírez, G. Vega-Cornejo, R. Damian-Negrete, G. Rosales-Chavez, and J. C. De Arcos-Jiménez, "Seasonal Shifts in Influenza, Respiratory Syncytial Virus, and Other Respiratory Viruses After the COVID-19 Pandemic: An Eight-Year Retrospective Study in Jalisco, Mexico," *Viruses* 16 (2024): 1892.
43. A. Trento, I. Casas, A. Calderón, et al., "Ten Years of Global Evolution of the Human Respiratory Syncytial Virus BA Genotype With a

60-Nucleotide Duplication in the G Protein Gene,” *Journal of Virology* 84 (2010): 7500–7512.

44. M. A. Pretorius, S. van Niekerk, S. Tempia, et al., “Replacement and Positive Evolution of Subtype A and B Respiratory Syncytial Virus G-Protein Genotypes From 1997–2012 in South Africa,” *Journal of Infectious Diseases* 208, no. Suppl 3 (2013): S227–S237.

45. A. Fall, F. Elawar, E. B. Hodcroft, et al., “Genetic Diversity and Evolutionary Dynamics of Respiratory Syncytial Virus Over Eleven Consecutive Years of Surveillance in Senegal,” *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 91 (2021): 104864.

46. J. C. Muñoz-Escalante, A. Comas-García, S. Bernal-Silva, and D. E. Noyola, “Respiratory Syncytial Virus B Sequence Analysis Reveals a Novel Early Genotype,” *Scientific Reports* 11 (2021): 3452.

47. H. A Al-Sharif, S. A. El-Kafrawy, J. M. Yousef, et al., “Dominance of the ON1 Genotype of RSV-A and BA9 Genotype of RSV-B in Respiratory Cases from Jeddah, Saudi Arabia,” *Genes* 11 (2020): 1323.

48. R. Malasao, W. Chaiut, W. Tantipetcharawan, et al., “Predominance of ON1 and BA9 Genotypes of Human Respiratory Syncytial Virus in Children With Acute Respiratory Infection in Chiang Mai, Thailand, 2020–2021,” *Journal of Infection and Public Health* 16 (2023): 1418–1426.

49. N. H. Razanajatovo Rahombanjanahary, K. Rybkina, T. H. Randriambolamanantsoa, H. Razafimanjato, and J. M. Heraud, “Genetic Diversity and Molecular Epidemiology of Respiratory Syncytial Virus Circulated in Antananarivo, Madagascar, From 2011 to 2017: Predominance of ON1 and BA9 Genotypes,” *Journal of Clinical Virology : the Official Publication of the Pan American Society for Clinical Virology* 129 (2020): 104506.

50. H. M. Kang, K. C. Park, J. Park, H. R. Kil, and E. A. Yang, “Circulating Respiratory Syncytial Virus Genotypes and Genetic Variability of the G Gene During 2017 and 2018/2019 Seasonal Epidemics Isolated From Children With Lower Respiratory Tract Infections in Daejeon, Korea,” *Journal of Korean Medical Science* 35 (2020): e422.

51. P. V. Markov, M. Ghafari, M. Beer, et al., “The Evolution of SARS-CoV-2,” *Nature Reviews Microbiology* 21 (2023): 361–379.

52. K. T. Zlateva, P. Lemey, E. Moës, A. M. Vandamme, and M. Van Ranst, “Genetic Variability and Molecular Evolution of the Human Respiratory Syncytial Virus Subgroup B Attachment G Protein,” *Journal of Virology* 79 (2005): 9157–9167.

53. I. Kushibuchi, M. Kobayashi, T. Kusaka, et al., “Molecular Evolution of Attachment Glycoprotein (G) Gene in Human Respiratory Syncytial Virus Detected in Japan 2008–2011,” *Infection, Genetics and Evolution* 18 (2013): 168–173.

54. A. Arnott, S. Vong, S. Mardy, 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,” *Journal of Clinical Microbiology* 49 (2011): 3504–3513.

55. K. M. Spann, P. L. Collins, and M. N. Teng, “Genetic Recombination During Coinfection of Two Mutants of Human Respiratory Syncytial Virus,” *Journal of Virology* 77 (2003): 11201–11211.

### Supporting Information

Additional supporting information can be found online in the Supporting Information section.