Contributory role of SARS-CoV-2 genomic variations and life expectancy in COVID-19 transmission and low fatality rate in Africa

Olabode E. Omotoso

Abstract

Background: The novel coronavirus disease (COVID-19) has claimed lots of lives, posing a dire threat to global health. It was predicted that the coronavirus outbreak in the African population would be very lethal and result to economic devastation owing to the prevalence of immune-compromised population, poverty, low lifespan, fragile health care systems, poor economy, and lifestyle factors. Accumulation of mutations gives virus selective advantage for host invasion and adaptation, higher transmissibility of more virulent strains, and drug resistance. The present study determined the severe acute respiratory syndrome-2 (SARS-CoV-2) genomic variability and the contributory factors to the low COVID-19 fatality in Africa. To assess the SARS-CoV-2 mutational landscape, 924 viral sequences from the Africa region with their sociobiological characteristics mined from the Global Initiative on Sharing All Influenza Data (GISAID) database were analyzed.

Results: Mutational analysis of the SARS-CoV-2 sequences revealed highly recurrent mutations in the SARS-CoV-2 spike glycoprotein D614G (97.2%), concurrent R203K, and G204R (65.2%) in the nucleocapsid phosphoprotein, and P4715L (97.2%) in the RNA-dependent RNA polymerase flagging these regions as SARS-CoV-2 mutational hotspots in the African population. COVID-19 is more severe in older people (> 65 years); Africa has a low percentage of people within this age group (4.36%). The average age of the infected patients observed in this study is 46 years with only 47 infected patients (5.1%) above 65 years in Africa in comparison to 13.12% in countries in other continents with the highest prevalence of COVID-19.

Conclusions: Africa’s young generation, the late incidence of the disease, and adherence to public health guidelines are important indicators that may have contributed to the observed low COVID-19 deaths in Africa. However, with the easing of lockdown and regulatory policies, daily increasing incidence in most countries, and low testing and sequencing rate, the epidemiology and the true impact of the pandemic in Africa remain to be unraveled.

Keywords: Coronavirus, Mutations, Low fatality, Age, SARS-CoV-2, Africa

Background

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative pathogen of the novel coronavirus disease 2019 (COVID-19), was first reported in December 2019 in Wuhan, China [1, 2]. The highly infectious SARS-CoV-2 has spread globally within 3-months of its first outbreak and has posed dire stress on public health and global economy [1, 3]. Globally, as of October 16, 2020, there have been 39,153,270 confirmed cases and 1,102,440 reported deaths due to COVID-19 [4]. It was predicted that the coronavirus outbreak in the African population would be very lethal and result to economic devastation owing to the prevalence of immune-compromised population (due to HIV/AIDS, tuberculosis, hepatitis virus, malaria, etc.), poverty, low lifespan, fragile health care systems, poor economic decisions, and lifestyle factors [3, 5]. Africa was the last continent
to be hit, and surprisingly boast of about 80% recovery rate at this stage of the pandemic [6]. Africa accounts for 3.5% of the fatality rate recorded globally with the highest death toll reported in Americas (55.1%) and Europe (25.7%) which are developed regions with high standard of living and healthcare structure [6, 7]. In line with public health and the World Health Organization (WHO) guidelines, the African governments and policy makers were quick to institute policies (restrictions such as partial/full-scale lockdowns, physical distancing, accelerated contact tracing, self-isolation, and quarantine) and adopt public health measures in light of the index case in the continent [7–9]. Africa’s index case of COVID-19 was on the 14th of February, 2020, in Egypt, North Africa [10] and on the 27th of February, Nigeria; Sub-Saharan Africa reported its index case from a European traveler [11]. As of October 16, 2020, there have been 1,623,181 confirmed cases and 39,150 deaths in Africa, out of which South Africa has the highest toll with 698,184 confirmed cases and 18,309 deaths [4].

Coronaviruses are single-positive-stranded RNA viruses with the largest genome (ranging from 26 to 32 kb) among all RNA viruses consisting of ORF1ab (21,289 nucleotides), ORF3a (827 nucleotides), ORF3b (455 nucleotides), ORF6 (185 nucleotides), ORF7a (365 nucleotides), ORF7b (131 nucleotides), ORF8 (365 nucleotides), ORF10 (116 nucleotides), and four genes encoding structural proteins: the nucleocapsid (1259 nucleotides), membrane (668 nucleotides), envelope (227 nucleotides), and spike (3,821 nucleotides) [1, 12–14]. Accumulation of mutation is central to viral evolution, transmission, and virulence. This gives virus selective advantage for host invasion and adaptation, higher transmissibility of more virulent strains, and drug resistance [9, 14, 15]. Identifying the mutational pattern of coronaviruses and monitoring their spread can be important in guiding drug/vaccine development and in making decisions to curtail the transmission [16]. The present study aims to gain insight into the mutational landscape of the SARS-CoV-2 genome in African population which may serve as eventual targets for drug design and/or vaccine development. This study also assesses the contributory factors to the low fatality rate due to COVID-19 in the African continent.

**Methods**

**Data acquisition**

A slightly modified protocol was used [17]. A total of 924 SARS-CoV-2 genomic sequences from the Global Initiative on Sharing All Influenza Data (GISAID) database (https://www.epicov.org/epi3/frontend) filtered as “high coverage only, Homo sapiens, complete, all clades and low coverage excl”, with patient’s status, “Africa” were mined. Patient’s age of all the sequences were also obtained to determine the age distribution of the infected patients. The accession numbers and laboratories in Africa that sequenced and deposited complete SARS-CoV-2 genomes on the GISAID database used in this study are provided as Supplementary file S1.

Country data of number of tests done per country with each nation’s population, confirmed cases, recoveries, and reported deaths due to COVID-19 were obtained from Worldometer (https://www.worldometers.info/coronavirus/) and the WHO database (covid19.who.int). The age distribution of countries with highest prevalence of COVID-19 cases was obtained from the World Factbook (www.cia.gov).

**Sequence and mutational analysis**

In this study, the mined 924 SARS-CoV-2 viral sequences were used to analyze the genomic variability since the index case of COVID-19 pandemic in Africa in February 2020 to identify the frequency and spread of mutations in the African population. The evolution of COVID-19 outbreak with respect to the transmission in the mutational hotspots was assessed and evaluated on the GISAID web interface (https://www.epicov.org/epi3/frontend). Recurrent mutations observed were focused on as they are likely to confer viral-host structure-function relationship promoting higher transmission rate.

**Determination of testing, fatality, and recovery rate**

The testing rate was determined for each African country as percentage of total test done from the country’s population.

\[
\text{Testing rate} (%) = \frac{\text{Total COVID-19 test done}}{\text{Country population}} \times 100
\]

The fatality rate was determined as percentage of total reported deaths due to COVID-19 from each country’s number of confirmed cases.

\[
\text{Fatality rate} (%) = \frac{\text{Total number of COVID-19 reported deaths}}{\text{Number of confirmed cases}} \times 100
\]

The recovery rate was determined as percentage of number of infectious patients who recovered from all reported confirmed cases in each country.

\[
\text{Recovery rate} (%) = \frac{\text{Number of COVID-19 infected patients who recovered}}{\text{Number of confirmed cases}} \times 100
\]

**Results**

Genomic variabilities dispersed at various sites in the SARS-CoV-2 sequence were observed; few mutations occurred more frequently (presented in Table 1). Mutational analysis of the 924 SARS-CoV-2 sequences revealed highly recurrent mutations in D614G (898 viral sequences) which falls in the spike glycoprotein (S)
region and was observed concurrently with P4715L variant (898 viral sequences) in the ORF1ab polyprotein region. More so, R203K and G204R mutations (602 viral sequences respectively) in the nucleocapsid (N) phosphoprotein region, Q57H (60 viral sequences) in the ORF3 protein, L84S (16 viral sequences) in the ORF8 protein, and T265I (26 viral sequences) in the ORF1ab polyprotein region were observed. It is worthy of note that 97.3% of viral sequences with the nsp5 G3278S mutation were from South Africa. The distribution of the mutations flags the Spike S1 domain (D614G), N protein R203K and G204R and nsp12 P4715L as SARS-CoV-2 mutational hotspots in the African population. According to mutational pattern, the viral sequences were characterized into the GISAID six (6) clades: GR (598), G (240), GH (57), S (16), O (7), V (4), and L (2).

For the age distribution, the average age of the infected patients is 46 years with only 47 infected patients (5.1%) above 65 years. According to demographic data from the World Factbook, the percentage of population of Africans ≥ 65 years is 4.36% in comparison to 13.12% in countries in other continents with highest prevalence of COVID-19 (as shown in Fig. 1a–e). At this stage of the pandemic, Africa still faces the challenge of low testing as none of the African countries meet up with the 15% testing rate standard, as South Africa has the highest testing rate (6.59%) in Africa. Nevertheless, there is a high recovery rate and low fatality rate observed across the African continent. The trend and distribution of confirmed cases, discharged cases, fatality, and recovery rate are as shown in Table 2.

### Discussion

As of September 5, 2020, over 705 million people in the world live in extreme poverty compared to about 651 million people in 2019 [5]. In 2020, 102,125,917 people are reported to be living in extreme poverty from a total population of 205,323,504 people in Nigeria, the most populous African nation compared to 200,369,248 people, among whom 92,441,550 people lived in extreme poverty in 2019 [5]. The COVID-19 has significantly ravaged economies and posed serious threat to public health and social interaction, spreading across the globe within 3 months of its outbreak [3]. It was hypothesized that low- and middle-income countries, especially African nations, would be seriously hit by the pandemic due to her vulnerability to infectious diseases and inadequate health structure. Surprisingly, there have been very low mortality rate due to COVID-19 in African nations when compared to developed nations in the Americas and Europe where the pandemic has had its highest toll of infection and fatality [7]. This study assessed the possible factors that could have accounted for the low mortality rate due to COVID-19 in Africa.

The dataset from infected patients helps to understand the COVID-19 infection rate, risks, epidemiology, and spread of the disease. The WHO suggested 10–30 tests per confirmed case as an indicator of adequate testing [7]. South Africa has the highest test rate of 3,918,478 people in Africa, representing about 7% of its 59,461,240 population, which is still a way long off from the adequate testing indicator and also when compared with countries with highest prevalence of COVID-19 in other regions: USA (37%), Russia (36%), and Australia (32%) [4]. In same vein, Nigeria, the most populous Black African nation, has only tested 440,248 people representing 0.21% of its over 207 million population. This trend holds true for many other African countries. Hence, caution should be taken while interpreting these data as the true number of COVID-19 cases in Africa might still remain undetected. However, with the current data available, though with daily increasing incidence, the mortality rate due to COVID-19 in Africa has been minimal. Other factor such as late incidence of the disease [10, 18], which gives an advantage for early preparation before the outbreak by enforcing regulatory guidelines.
(active surveillance, isolation, quarantine, contact tracing, and social distancing among others) in line with the WHO guidelines could have played a key role in managing the pandemic in Africa [7, 8]. The American Public Media (APM) Research laboratory in August 2020 reported 88.4 deaths per 100,000 Black Americans due to COVID-19 compared to 54.4 deaths per 100,000 Latino Americans and 36.4 deaths per 100,000 Asian Americans. The APM findings suggests higher mortality due to COVID-19 in Black Americans in the USA [19]. This raises a concern to understand why Blacks are dying more due to COVID-19 overseas than in Africa.

The present study also looked into the mutational landscape of SARS-CoV-2 genome in Africa. Highly recurrent mutation (D614G) was observed in the S1 domain of S glycoprotein which facilitates viral entry into host by binding to the human angiotensin-converting enzyme 2 (ACE2) receptor. The S2 domain of the S glycoprotein which plays a crucial role in viral-host cell membrane fusion was well conserved; this could serve as an important target for antiviral drug design. The T265I mutation that is prevalent in the USA which aids SARS-CoV-2 survival in the host cell [20] was only observed in 26 viral sequences, predominantly in Senegal (69.23%). P4715L mutation hotspot present in the RNA-dependent RNA polymerase (RdRp) region was first observed in Italy during the sporadic increase in incidence and fatality of COVID-19 in Europe [15]. This supports the evidence that most of Africa’s index case originates from European and American travelers [3, 18], which could have resulted to the vast distribution of these mutations across the infected patients.

The concurrent R203K and G204R mutation are present in the N phosphoprotein region, which plays an important role in viral interaction with the M protein during virion assemblage [21]. Gene variants in the nucleocapsid region alter miRNAs binding, which might contribute to the pathogenesis and progression of infection in the patient [22]. The viral ORF3a and ORF10 proteins can synergistically attack heme on the host’s hemoglobin 1-β chain, thereby disintegrating iron to form porphyrin. This results to reduced hemoglobin carrying oxygen and carbon dioxide causing extreme poisoning and inflammation of the hepatocytes [23].
Q57H mutation in the ORF3a region was observed to coincide with D614G and P4715L variants. The RdRp P4715L mutation coincides with S protein D614G mutation alongside the concurrent N protein R203K and G204R variants, and in their absence ORF3a Q57H variant. The SARS-CoV-2 membrane, envelope, ORF6, ORF7, and ORF10 proteins were well conserved. These mutational hotspots and conserved domains must be well-considered during drug or vaccine design to avoid vaccine evasion and drug resistance.

Lastly, this study looked at the age distribution of the infected patients as a contributory factor to the low fatality rate. The average life expectancy in Nigeria is 54 years, a reflection of Sub-Saharan Africa (61 years) compared to European union (81 years), China (77 years), and USA (79 years) [24]. Earlier report [3] suggests that COVID-19 has high mortality and severity in older people (> 65 years) than the younger population, which have high chances of recovery from the infection. Only 5.1% of the 924 infected Africans in this present study are over 65 years. This could have played a major role in the relatively high recovery rate (81.21%) and low mortality (2.4%) due to COVID-19 observed in the African population.

**Conclusion**

As of now, the highly infectious COVID-19 continues to spread globally, with increasing incidence and mortality. The African continent continues to enjoy the benefit of the swift actions and policies imposed to curtail the pandemic. With the gradual easing of these measures, increasing poverty, poor healthcare structure, and a large percentage of immune-compromised population, it is important that the populace adhere to regulatory guidelines. Low life expectancy, low testing rate, late incidence of the disease, and adherence to public health guidelines could be important factors that have contributed to the observed low COVID-19 fatality in Africa. However, sufficient data is still unavailable to ascertain the epidemiology, transmission, genomic variation, and the true impact of the pandemic in Africa. Genomic sequences are important in the design and development of antiviral drugs and vaccine. From the over 200,000 complete SARS-CoV-2 sequences available in public repositories,
Africa has barely contributed 4000 viral sequences (2%). Collaboration with scientists and research institutes in African nations is highly recommended so as to enhance their delivery capacity.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s43042-020-00116-x.

Additional file 1: Supplementary file S1R1.
Additional file 2: Supplementary file S2R1.

Abbreviations
ACE2: Angiotensin-converting enzyme 2; APM: American Public Media; COVID-19: Coronavirus disease 2019; E: Envelope protein; GISAID: Global Initiative on Sharing All Influenza Data; M: Membrane protein; N: Nucleocapsid phosphoprotein; RdRp: RNA-dependent RNA polymerase; S: Spike glycoprotein; SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2; WHO: World Health Organization

Acknowledgements
The author would like to appreciate the heroic effort of health workers in the frontline of tackling COVID-19 and all researchers who sequenced SARS-CoV-2 genomes and deposited them in public repositories.

Author's contributions
OO: conceptualization, data mining, methodology, analysis, writing. The author read and approved the final manuscript.

Funding
No funding was received for this study.

Availability of data and materials
All data used in this study are available in recognized public repositories as indicated.

Ethics approval and consent to participate
All dataset used are made freely available in public repositories.

Consent for publication
Not applicable

Competing interests
The author declares no conflict of interest.

Received: 17 October 2020 Accepted: 25 November 2020
Published online: 09 December 2020

References
1. Wang H, Li X, Li T, Zhang S, Wang L, Wu X et al (2020) The genetic sequence, origin, and diagnosis of SARS-CoV-2. Eur J Clin Microbiol Infect Dis China CDC 39(9):1629–1635. https://doi.org/10.1007/s10096-020-03899-4
2. WHO (2020) Coronavirus disease (COVID-19) pandemic [Internet]. [cited 2020 Oct 28]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019
3. Lome SA, Ahmad A (2020) COVID-19 pandemic – an African perspective. Emerg Microbes Infect 9:1300–1308
4. Coronavirus update (Live): 39,153,270 cases and 1,102,440 deaths from COVID-19 virus pandemic - Worldometer [Internet]. [cited 2020 Oct 16]. Available from: https://www.worldometers.info/coronavirus/
5. World Poverty Clock [Internet]. [cited 2020 Sep 5]. Available from: https://worldpoverty.io/headline
6. Coronavirus disease 2019 (COVID-19) - Africa CDC [Internet]. [cited 2020 Oct 21]. Available from: https://africacdc.org/covid-19/
7. WHO Coronavirus disease (COVID-19) dashboard | WHO Coronavirus Disease (COVID-19) Dashboard [Internet]. [cited 2020 Sep 3]. Available from: https://covid19.who.int/
8. COVID-19 in Africa: marking six months of response | WHO | Regional Office for Africa [Internet]. [cited 2020 Sep 4]. Available from: https://www.afro.who.int/covid-19-africa-marking-six-months-response
9. Omotoso O, Babalola D, Matarere A (2020) Mutational analysis of SARS-CoV-2 in African population bioRxiv
10. COVID-19 cases top 10 000 in Africa | WHO | Regional Office for Africa [Internet]. [cited 2020 Sep 3]. Available from: https://www.who.int/news/covid-19-cases-top-10-000-africa
11. Nigeria Centre for Disease Control [Internet]. [cited 2020 Sep 3]. Available from: https://ncdc.gov.ng/news/227/first-case-of-corona-virus-disease-confirmed-in-nigeria
12. Muhammad TQ, Safar AM, Alamri L-LC (2020) Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. J Pharm Anal 10(4):313–319. https://doi.org/10.1016/j.jpha.2020.03.0
13. GenBank (2020) Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome. NCBI GenBank, pp 1–28 Available from: https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seq/
14. Koyama T, Platt D, Pandla L (2020) Variant analysis of COVID-19 genomes. Bull World Health Organ 98(7):495–504. https://doi.org/10.2471/BLT.20.253591
15. Pachetti M, Marini B, Benedetti F, Giudici F, Mauro E, Storici P et al (2020) Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. J Transl Med 18:179 Available from: https://doi.org/10.1186/s12967-020-02344-6
16. Li F (2017) Structure, function and evolution of coronavirus spike proteins. Annu Rev Virol 3(1):237–261
17. Omotoso O (2020) SARS-CoV-2 genomic variation-African perspective. protocols.io [Internet]. 2020 [cited 2020 Oct 11]; Available from: https://doi.org/10.17504/protocols.io.bmpk5jn
18. Paul O (2020) First African SARS-CoV-2 genome sequence from Nigerian COVID-19 case. Genome reports, pp 1–5 Available from virological.org/t/first-african-sars-cov-2-genome-sequence-from-nigerian-covid-19-case/421
19. The color of coronavirus: COVID-19 deaths by race and ethnicity in the US [Internet]. APM Research Lab. 2020 [cited 2020 Sep 3]. Available from: https://www.apmresearchlab.org/covid-deaths-by-race
20. Shuvam B, Sohan S, Riju D, Kousik M, Pritha B. Mutational spectra of SARS-CoV-2 ORF1ab polyprotein and signature mutations in the United States of America. bioRxiv. 2020;
21. Malik YA (2020) Properties of coronavirus and SARS-CoV-2. Malays J Pathol 42(1):3–11
22. Maitra A, Sarkar MC, Raheja H, Biswas N, Chakraborti S, Singh AK et al (2020) Mutations in SARS-CoV-2 viral RNA identified in eastern India: possible implications for the ongoing outbreak in India and impact on viral structure and host susceptibility. J Biosci 45:7621–18
23. Li C, Wang L, Ren L (2020) Antiviral mechanisms of candidate chemical medicines and traditional Chinese medicines for SARS-CoV-2 infection. Virus Res 286(May):198073 Available from. https://doi.org/10.1016/j.virusres.2020.198073
24. Life Expectancy [Internet]. World Bank. [cited 2020 Sep 5]. Available from: https://data.worldbank.org/indicator/SP.DYN.Le00.IN?locations