



# Whole genome sequencing and phylogenetic classification of Tunisian SARS-CoV-2 strains from patients of the Military Hospital in Tunis

Susann Handrick<sup>1</sup> · Malena Bestehorn-Willmann<sup>1</sup> · Simone Eckstein<sup>1</sup> · Mathias C. Walter<sup>1</sup> · Markus H. Antwerpen<sup>1</sup> · Habiba Naija<sup>2</sup> · Kilian Stoecker<sup>1</sup> · Roman Wölfel<sup>1</sup> · Mohamed Ben Moussa<sup>2</sup>

Received: 2 July 2020 / Accepted: 12 September 2020 / Published online: 9 October 2020  
© The Author(s) 2020

## Abstract

In the present work, two complete genome sequences of SARS-CoV-2 were obtained from nasal swab samples of Tunisian SARS-CoV-2 PCR-positive patients using nanopore sequencing. The virus genomes of two of the patients examined, a Tunisian soldier returning from a mission in Morocco and a member of another Tunisian family, showed significant differences in analyses of the total genome and single nucleotide polymorphisms (SNPs). Phylogenetic relationships with known SARS-CoV-2 genomes in the African region, some European and Middle Eastern countries and initial epidemiological conclusions indicate that the introduction of SARS-CoV-2 into Tunisia from two independent sources was travel-related.

**Keywords** SARS-CoV-2 · COVID-19 · RT-qPCR · Sequencing · Tunisia · Morocco

The current global spread of the newly emerged severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of coronavirus disease 2019 (COVID-19), has led to worldwide social and economic disruption. The virus was first described in December 2019 in Wuhan, China, and has since spread into a pandemic. As a result, healthcare systems in almost all countries of the world are facing unprecedented challenges. A few clinically approved drugs have been shown to exhibit anti-SARS-CoV-2 activity, but none have yet been proven to be sufficiently effective in treating COVID-19 patients, nor are vaccines available [1]. By August 2020, the World Health Organization (WHO) had reported more than 21.99 million infected individuals and 775,893 deaths worldwide. In the African Region, 966,352 confirmed cases have been reported [2], but many cases

remain undetected, probably due to insufficient diagnostic capacity, limited contact tracing and the oftentimes unspecific symptoms of affected patients.

Diagnostic tests for SARS-CoV-2 have been carried out in Tunisia since early February 2020. On 2 March 2020, the first case of COVID-19, a Tunisian patient from Italy, was reported [3, 4]. Since then, the number of positively-tested SARS-CoV-2 patients has been increasing. By 18 August 2020, a total of 117,086 tests were performed and 2427 SARS-CoV-2 positive cases were reported by the Tunisian Ministry of Health. Of these positive cases, 60 patients died [5]. So far, only a few whole genome sequences from Tunisia are available online. However, these few data already indicate various independent, travel-related introductions of SARS-CoV-2 into the country [6]. Proper surveillance and monitoring of the SARS-CoV-2 epidemic in Tunisia and on the African continent, in order to guarantee faultless diagnostics, urgently requires further epidemiological and bioforensic data in addition to those currently available.

In the present work, two SARS-CoV-2 strains from Tunisian nationals were sequenced and phylogenetically compared with available SARS-CoV-2 genomes from Tunisia and other countries in Africa, Europe and the Middle East.

Four nasopharyngeal swab samples from Tunisian citizens were tested positive for SARS-CoV-2 RNA (Table 1) after viral RNA extraction with the QIAmp viral RNA Mini Kit (Qiagen, Hilden, Germany) using RT-qPCR according

---

Edited by Hartmut Hengel.

---

Susann Handrick and Malena Bestehorn-Willmann have contributed equally.

---

Roman Wölfel and Mohamed Ben Moussa are shared last authors.

---

✉ Susann Handrick  
SusannHandrick@bundeswehr.org

<sup>1</sup> Bundeswehr Institute of Microbiology, Munich, Germany

<sup>2</sup> Department of Virology, Military Hospital of Instruction of Tunis, Tunis, Tunisia

**Table 1** Patient information and sequencing results of MHT\_1 and MHT\_2

Sample	ID	PCR analyses	AccNo.	SNPs (AA exchanges)
31-year-old Tunisian soldier	MHT_1	E Ct=29.48; RdRP Ct=33.14	MT559037	C3037T C14408T (ORF1b:P314L) C18877T A23403G (S:D614G) G25563T (ORF3a:Q57H) G29474T (N:D401Y)
59-year-old mother	MHT_2	E Ct=27.18; RdRP Ct=31.17	MT559038	A361G C8782T G22468T T28144C (ORF8:L84S) G28878A (N:S202N)
62-year-old father	MHT_4	E Ct=34.34; RdRP Ct=40.33	n.d.	n.d.
23-year-old daughter	MHT_3	E Ct=34.03; RdRP no Ct detectable	n.d.	n.d.
24-year-old daughter	none	negative	n.a.	n.a.
22-year-old son	none	negative	n.a.	n.a.

*n.d.* not done, *n.a.* not applicable

to the protocol of Corman et al. [7] at the Military Hospital in Tunis (MHT). The results of this investigation were confirmed at the Bundeswehr Institute of Microbiology in Munich (IMB) using an RT-qPCR protocol targeting the viral *N* gene [8]. An asymptomatic Tunisian soldier (31-year-old male, MHT\_1) returning from a mission in Agadir, Morocco, tested positive immediately after his arrival at Tunis-Carthage Airport on 21 March 2020. In addition, three of five members of a Tunisian family from Ezzahra, 15 km south of Tunis, tested positive on 24 March 2020 (59-year-old mother, MHT\_2; 62-year-old father, MHT\_4) and 28 March 2020 (23-year-old daughter, MHT\_3; Table 1). One of the patients, the father, showed symptoms such as fever, cough, arthralgia and headache one week before the first RT-qPCR test. The other two COVID-19 infections were without development of symptoms. All patients were tested negative for SARS-CoV-2 RNA in nasopharyngeal swabs after 14 days of home quarantine.

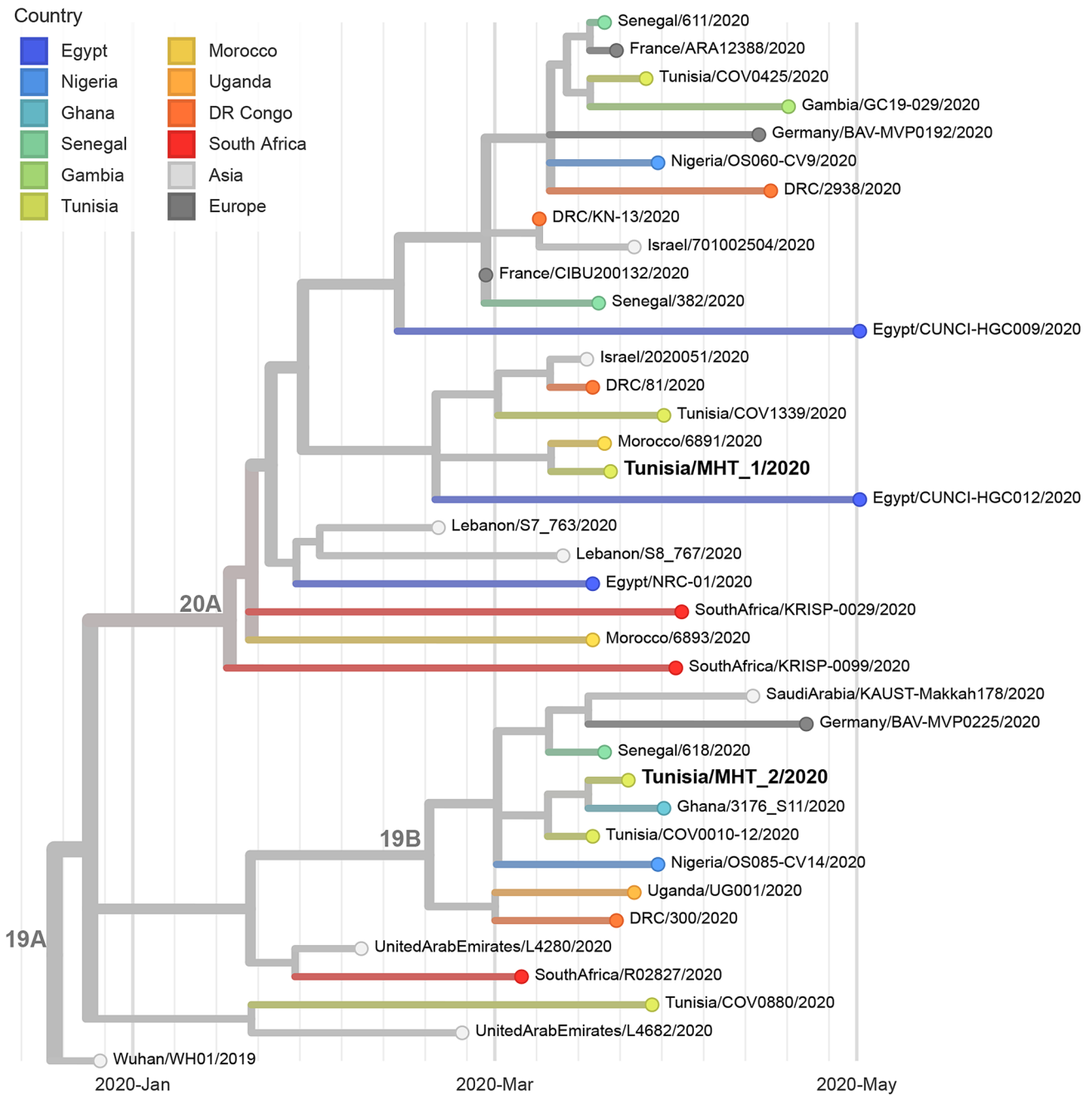
The four positively-tested RNA samples were further analysed at the IMB. For two of the four samples, having the highest virus loads (MHT\_1/MHT\_2; Table 1), whole genome sequencing of SARS-CoV-2 according to the nCoV-2019 sequencing protocol [9] on a GridION instrument was successful. Briefly, after passing quality control, demultiplexing, adapter trimming and consensus sequences generation based on the reference strain Wuhan-Hu-1 [10] using

the ARTIC pipeline [11], closed genomes lacking only the outermost nucleotides of the 5'UTR or 3'UTR regions were obtained. Both genomes were annotated by and submitted to GenBank (accession MT559037 and MT559038).

SNP and phylogenetic analyses were performed using a local installation of the nextstrain.org pipeline [12]. Strains listed in GISAID belonging to the African subgroup ( $N=1203$  of 22 June 2020) and with available whole genome sequences were included in the initial analysis and further filtered phylogenetically to minimise redundancies and selected for relevance to possible travel and/or trade routes ( $N=38$ ).

In comparison to the reference strain Wuhan-Hu-1, MHT\_1 shows six SNPs, four of which lead to amino acid changes in two different open reading frames (ORF) as well as the *S* and *N* genes respectively. Compared to Wuhan-Hu-1, MHT\_2 contains five SNPs, two of which lead to amino acid changes in the ORF8 and the *N* gene. When comparing the two Tunisian SARS-CoV-2 sequences, no insertions or deletions and no mutual SNPs were found (Table 1).

Phylogenetic analyses showed that the genomes in our study are grouped in two different branches (Fig. 1). Sample MHT\_1 belongs to clade 20A (Nextstrain.org nomenclature [13] matches Pangolin lineage B.1 [14]) with the characteristic mutation D614G (SNP A23403G) of the SARS-CoV-2 spike protein assumed to enhance viral infectivity [15];



**Fig. 1** Phylogenetic tree showing the genomic epidemiology of a Tunisia-focused subsampling of 38 SARS-CoV-2 genomes sampled between Dec 2019 and May 2020. Unpublished data are included

with the permission of the data generators and does not affect their right to publish. A complete list of sequence authors is available at [nextstrain.org](https://nextstrain.org)

sample MHT\_2 is clustering within clade 19B (Pangolin lineage A). While genomes of clade 19 were dominated by Asian samples, especially during the first weeks of the incipient pandemic outbreak, clade 20A was the group rapidly spreading worldwide in February/March/April 2020. Observation of the genomes of both clones, which are present in North Africa, suggests multiple introductions at different points in time. When comparing whole genomes from Africa

with a focus on the northern and central countries of this continent, the existence of geographically-associated genetic clusters can be observed.

Since the amount of genome sequences from the African continent is still small compared to European, Asian or American sequencing data, the interpretation of the results should be treated with caution. If more metadata are taken

into account, possible epidemiological links become obvious and could explain genetic similarities. In pandemic situations, however, not all genetic similarities can be explained, even if data is available, e.g. the Bav-MVP0225/2020 sample from Germany. The situation is similar in the case of sample MHT\_1, where the genetically closest neighbour is the isolate Morocco/6891/2020. This is conclusive, as this isolate originates from a Tunisian soldier who had returned from a stay in Morocco. The family isolate, on the other hand, is related to other virus strains, mainly from Central African countries such as Ghana, Uganda and the Democratic Republic of Congo (DRC). A direct link from Tunisia to these regions, e.g. on the basis of travel or trade, could not be established.

Therefore, correct and sufficient epidemiological data remain the greatest need for case tracing and successful infection control. In multidisciplinary approaches that take into account both epidemiological and genetic data, phylogeographic predictions can be verified on the basis of genetic analyses. While in epidemic or localised outbreaks with rapidly evolving pathogens, such as the Ebola virus, whole genomes can greatly assist in contact tracing, multiple introductions are difficult to monitor and track by phylogenetic analysis during prolonged outbreaks and pandemics. Medical and scientific cooperation is, therefore, one of the main triggers for combating pandemics like COVID-19.

**Acknowledgements** We gratefully acknowledge the Tunisian patients described in these case reports for participating in the study and for giving written consents for publication of the obtained data. We acknowledge the authors from originating and submitting laboratories of sequence data provided by GISAID on which the analysis is based.

**Author contributions** All authors contributed to the conception and design of the study. Material preparation, data collection and analysis were performed by MBW, MCW, SH, SE, HN and MBM. The first draft of the manuscript was written by SH, and all authors commented on earlier versions of the manuscript. All authors have read and approved the final manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Enable and Enhance Initiative of the German Government (Ertüchtigungsinitiative der Deutschen Bundesregierung, OR12-370.43 ERT TUN IMB).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** The use of the biological material described in the underlying study was approved by the Ethics Committee of the Military Hospital in Tunis and written consent forms were signed by all included participants.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Wu R, Wang L, Kuo H-CD, Shannar A, Peter R, Chou PJ, Li S, Hudlikar R, Liu X, Liu Z et al (2020) An update on current therapeutic drugs treating COVID-19. *Curr Pharmacol Rep*. <https://doi.org/10.1007/s40495-20-00216-7>
2. WHO Coronavirus Disease (COVID-19) Dashboard Available online: <https://covid19.who.int>. Accessed on Aug 20, 2020
3. WHO Coronavirus disease (2019) (COVID-19) Situation Report—43. WHO, Geneva, p 2020
4. Louhaichi S, Allouche A, Baili H, Jrad S, Radhouani A, Greb D, Akrouf I, Ammar J, Hamdi B, Added F et al (2020) Features of patients with 2019 novel coronavirus admitted in a pneumology department: the first retrospective Tunisian case series. *Tunis Med* 98:261–265
5. Tunesienexplorer Covid-19 Zahlen Tunesien von Dienstag, 15 Aug 2020. *Tunesienexplorer* 2020.
6. Fares W, Chouikha A, Ghedira K, Gdoura M, Rezig D, Boubaker SH, Dhifallah IB, Touzi H, Hammami W, Zi Meddeb et al (2020) First whole genome sequences and phylogenetic analysis of SARS-CoV-2 virus isolates during COVID-19 outbreak in Tunisia, North Africa. <https://doi.org/10.22541/au.159137642.26983355>
7. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, Bleicker T, Brünink S, Schneider J, Schmidt ML et al (2020) Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull* 25(3):2000045. <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>
8. Nao N, Shirato K, Katano H, Matsuyama S, Takeda M (2020) Detection of second case of 2019-nCoV infection in Japan (corrected version). National Institute of Infectious Diseases, Tokyo
9. Quick J (2020) nCoV-2019 sequencing protocol. <https://doi.org/10.17504/protocols.io.bbmuik6w>
10. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, Hu Y, Tao Z-W, Tian J-H, Pei Y-Y et al (2020) A new coronavirus associated with human respiratory disease in China. *Nature* 579:265–269. <https://doi.org/10.1038/s41586-020-2008-3>
11. Lohman N, Rowe W, Rambaut A (2020) ARTIC nCoV-2019 novel coronavirus bioinformatics protocol v1.1 Available online: <https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html>. Accessed on May 16, 2020
12. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA (2018) Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34:4121–4123. <https://doi.org/10.1093/bioinformatics/bty407>
13. Year-letter Genetic Clade Naming for SARS-CoV-2 on Nextstrain.org Available online: <https://nextstrain.org/blog/2020-06-02-SARSCoV2-clade-naming>. Accessed on Jun 25, 2020
14. Rambaut A, Holmes EC, Hill V, O'Toole Á, McCrone JT, Ruis C, du Plessis L, Pybus OG (2020) A dynamic nomenclature proposal

for SARS-CoV-2 to assist genomic epidemiology. bioRxiv. <https://doi.org/10.1101/2020.04.17.046086>

15. Hu J, He C-L, Gao Q-Z, Zhang G-J, Cao X-X, Long Q-X, Deng H-J, Huang L-Y, Chen J, Wang K et al (2020) The D614G mutation of SARS-CoV-2 spike protein enhances viral infectivity and decreases neutralization sensitivity to individual convalescent sera. bioRxiv. <https://doi.org/10.1101/2020.06.20.161323>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.