Version 2

Title: No association between the SARS-CoV-2 variants and mortality rates in the Eastern Mediterranean Region

Authors: Saad Omais[1], Samer Kharroubi[2], Hassan Zaraket[3,4,#]

Affiliations:

1, Department of Biology, Faculty of Arts and Sciences, American University of Beirut, Beirut, Lebanon.

2, Department of Nutrition and Food Sciences, Faculty of Agricultural and Food Sciences, American University of Beirut, Beirut, Lebanon.

3, Faculty of Medicine, Center for Infectious Diseases Research, American University of Beirut, 1107 2020 Beirut, Lebanon.

4, Department of Experimental Pathology, Faculty of Medicine, Immunology & Microbiology, American University of Beirut, 1107 2020 Beirut, Lebanon.

 # Corresponding author: hz34@aub.edu.lb

 # Present/permanent address: Department of Experimental Pathology, Faculty of Medicine, Immunology & Microbiology, American University of Beirut, 1107 2020 Beirut, Lebanon

Keywords: SARS-CoV-2, whole genome, variants, case fatality rate, East Mediterranean Region

Abstract

As the novel coronavirus SARS-CoV-2 continues to spread in all countries, there is a growing interest in monitoring and understanding the impact of emerging strains on virus transmission and disease severity. Here, we analyzed SARS-CoV-2 genomic sequences reported in the Eastern Mediterranean Region (EMR) countries, as of 1 January 2021. The majority (~75%) of these sequences originated from three out of 22 EMR countries, and 65.8% of all sequences belonged to GISAID clades GR, GH, G and GV. A delay ranging between 30-150 days from sample collection to sequence submission was observed across all countries, limiting the utility of such data in informing public health policies. We identified ten common non-synonymous mutations represented among SARS-CoV-2 in the EMR and several country-specific ones. Two substitutions, spike_D614G and NSP12_P323L, were predominantly concurrent in most countries. While the single incidence of NSP12_P323L was positively correlated with higher case fatality rates in EMR, no such association was established for the double (spike_D614G and NSP12_P323L) concurrent variant across the region. Our study identified critical data gaps in EMR highlighting the importance of enhancing surveillance and sequencing capacities in the region.

Introduction

In early October 2020, the World Health Organization (WHO) estimated that 10% of the global population has been infected with the novel coronavirus, SARS-CoV-2 (Keaten, 2020). Since its declaration as a pandemic seven months earlier, COVID-19 has inflicted substantial health and economic burden around the world (Fineberg, 2020; Ioannidis, 2020a). As of 22 January 2021, there have been over 96 million confirmed cases and 2 million deaths attributed to COVID-19 (WHO). Specifically, 5,461,398 confirmed cases (of whom 4,815,552 have recovered) and 130,079 deaths were reported in the Eastern Mediterranean Region (EMR) (WHO).

Prior to SARS-CoV-2 emergence, six human coronavirus (HCoV) strains were known to infect humans: two of which caused fatal respiratory diseases (SARS-CoV-1 and MERS-CoV), while the remaining four circulate annually and cause 15-29% of all common colds (HCoV-229E, -OC43, -NL63 and -HKU1) (Su et al., 2016). SARS-CoV-2 gains access to the host cell by binding to the human angiotensin-converting enzyme 2 (hACE2) through its spike (S) protein. SARS-CoV-2 S receptor-binding domain (RBD) has a 10- to 20-fold higher hACE2 binding affinity than SARS-CoV-1 (Wrapp et al., 2020). Yet, SARS-CoV-2 RBD is less accessible such that the overall S-hACE2 interaction is similar or weaker than that of SARS-CoV-1. Nonetheless, the SARS-CoV-2 spike protein possesses a polybasic cleavage site making it accessible to furin protease which enhances cell entry (Shang et al., 2020). Similar to other beta-coronaviruses, SARS-CoV-2 has a positive-sense RNA genome that is around 30-kb in length with six functional open reading frames (ORFs). ORFa and -b constitute nearly two-thirds of the genome and produce 16 non-structural proteins (NSPs), including the RNA-dependent RNA polymerase (RdRp or NSP12) and a helicase (NSP13). The structural nucleocapsid (N),

membrane (M), envelope (E), and S proteins are encoded by the remaining stretch of the RNA genome along with other accessory proteins (Hartenian et al., 2020).

There has been a tremendous global surveillance effort to closely monitor SARS-CoV-2 circulating worldwide that enabled timely detection of emerging variants. SARS-CoV-2 evolves at an estimated rate of around 6 x 10^{-4} nucleotides/genome/year (van Dorp et al., 2020a). Numerous variants have emerged but only a handful of them have been fixated such as the S_D614G that has prevailed globally. S_D614G was shown to alter infectivity and virulence of SARS-CoV-2 and was associated with an increase in mortality (Eaaswarkhanth et al., 2020; Plante et al., 2020; van Dorp et al., 2020a; Vankadari, 2020; Wang et al., 2020). Specifically, Sallam et al. reported an increase in the incidence of S_D614G variant between February and June 2020 in the Middle East and North Africa (Sallam et al., 2021). However, a detailed genomic characterization of SARS-CoV-2 in EMR is still lacking. It also remains not clear whether the case fatality rates (CFRs) in EMR are associated with any of the prevailing SARS-CoV-2 variants. In this study, we analyzed SARS-CoV-2 whole-genome sequences collected in EMR and examined the temporal and country-level associations between the predominant variants and CFRs.

Methods

Sequence Retrieval and Analysis

SARS-CoV-2 sequences were downloaded from GISAID EpiCoV™ database (Shu and McCauley, 2017) which were submitted by 1 January 2021 from EMR countries. These are Afghanistan, Bahrain, Djibouti, Egypt, Iran, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Pakistan, Palestine, Qatar, Saudi Arabia, Somalia, Sudan, Syrian Arab Republic, Tunisia,

United Arab Emirates (UAE), and Yemen (http://www.emro.who.int/countries.html). Only 'high coverage' sequences were selected, i.e. "entries with <1% Ns and <0.05% unique amino acid mutations (not seen in other sequences in the database) and no insertion/deletion unless verified by submitter". These were then analyzed by CoVsurver enabled by GISAID (https://www.gisaid.org/epiflu-applications/covsurver-mutations-app/) and compared with reference strain hCoV-19/Wuhan/WIV04/2019 (accession number: EPI_ISL_402124). Sequence metadata corresponding to age and gender were extracted from EpiCoV™ 'nextmeta' file for sequences submitted as of 12 November 2020. For clade analysis, we used the Nextclade tool from the Nextstrain project (https://clades.nextstrain.org/) (Hadfield et al., 2018) and GISAID clade designation provided by CoVsurver analysis.

Country Data and Mortality Rate

Data on population size, total COVID-19 cases, and total deaths were retrieved from the Worldometers website as of 22 January 2021 (https://www.worldometers.info/coronavirus/). Data on daily cases and deaths per country were extracted from WHO Coronavirus Disease (COVID-19) Dashboard (https://covid19.who.int/) (WHO). CFR was calculated and adjusted for a 13-day lag time from reporting to death, such that the denominator was designated as the cumulative cases from 13 days earlier, while the numerator totaled the cumulative deaths till 13 days earlier plus half the additional deaths recorded during the lag period (Wilson et al., 2020).

Statistical Analysis

The data were checked for completeness and entered into the Statistical Package for the Social Sciences (SPSS) software version 20 for Windows, which was later used for statistical analyses [IBM: Statistical Package for the Social Sciences. SPSS Statistics 2013, 20.]. Eight of the 15

EMR countries that have sequence data available were excluded from cross-country fatality correlations for the following reasons: Egypt had a relatively high CFR (5.9%) probably as a result of the low number of reported cases (1073 cases per million) likely due to limited testing capacity; Iraq, Pakistan, Tunisia, Lebanon, Kuwait, and Qatar had a relatively low sequence count each (<50, 1.5% of all sequences) (**Figure 1A**), Iran had only 18 complete sequences (>25 kbp) out of 104 sequences and was thus less comparable to other sequences (**Supplementary Figure 1)**. For the summary of the data, descriptive statistics were presented to summarize the study variables of interest as counts and percentages across countries. Pie charts and bar plots were used to chart comparison in count and percentages of variants across all countries. An independent t-test was used to chart the comparison of variants via gender (male vs female), while one-way ANOVA was carried out for age distribution (0-9, 10-18, 19-49, 50-59, and >60). Pearson's correlation coefficients were used to examine the association between variant percentages and CFR for the seven countries considered in this study (Toyoshima et al., 2020). For all analyses done, a *p-value* of less than 0.05 was used to detect statistical significance.

Results

We identified 3345 high-coverage SARS-CoV-2 genome sequences from EMR countries – no sequences were available for Afghanistan, Djibouti, Libya, Somalia, Sudan, Syria, and Yemen. Three countries accounted for ~75% of the sequence count overall – UAE (n=1233), Saudi Arabia (n=762) and Jordan (n=512) (**Figure 1A**). All but 5 EMR countries – Iran, Tunisia, Iraq, Pakistan and Kuwait – had near-full length sequences of SARS-CoV-2 (**Supplementary Figure 1**). UAE had the highest sequence to population ratio (123.8 in 1 million), but also the highest number of sequences per reported COVID-19 cases (45.5 per 10,000 cases) (**Figure 1B, 1C**). When dates of collection were considered, none of the countries showed routine and consistent

sequencing; most sequences were collected before July (e.g. from UAE, Saudi Arabia, Oman) while only few countries (mainly Jordan, Bahrain and Morocco) reported sequences from the last three months (**Figure 1D**). Moreover, a significant delay between sample collection and sequence submission was observed across all countries (ranging 30-150 days), with no noticeable progress between first and last 100 days for the top three represented countries (**Figure 1E)**.

Using CoVsurver from GISAID, amino acid substitutions were identified in each viral protein sequence and the top ten prevalent variants for each country were then plotted (excluding single-incidence variants) (**Figure 2**). In total, 2155 non-synonymous mutations were detected in EMR countries, out of which ten common variants were found to be relatively most common (i.e. in ≥ 8 EMR countries and > 2.5% of all sequences). These are: NSP12_P323L, S_D614G, NS3_Q57H, N_R203K, N_G204R, N_S194L, NSP6_L37F, NSP2_V198L, NS8_L84S and N_S202N. Besides, many country-specific substitutions were detected (here defined as found in 3 other countries or less, and not listed in other countries' most common mutations). For instance, NSP3_V1209I was found in 18% of UAE's sequences, Spike_Q957L in 80% of Jordan's sequences and NSP12_M666I in 76% of Palestine's sequences (**Figure 2**). NSP12_P323L and S_D614G were predominantly detected in 10 out of 15 countries. S_D614G was also dominant in Bahrain. In Kuwait and Iran, NSP6_L37F and NSP2_V198L were the most frequently reported variants.

Next, we used the Nextclade tool from Nextstrain to assign clades (GISAID nomenclature) to the sequences (**Figure 3A)** (GISAID, 2020). Of all sequences from EMR countries, 65.8% belonged to the G (20A per Nextstrain clades), GR (20B), GH (20C) and GV (20A.EU1) clades, which emerged following the introduction of the S_D614G variant (**Figure 3B**, (Alm et al., 2020)).

Moreover, these clades were almost exclusively represented as early as February and March in Saudi Arabia. In UAE, Jordan and Oman, clade distribution was heterogeneous with multiple clades cocirculating between February through November (**Figure 3C**).

For the sequence cohort submitted by 12 November 2020, more specimens were sequenced from males in the EMR (2.05:1 male-to-female in top four represented countries, independent t-test $p < 0.004$): Saudi Arabia (4.44:1), UAE (2.25:1), Oman (1.42:1) and Egypt (1.49:1) (**Figure 4A**). No statistically significant difference in gender distribution of the top represented SARS-CoV-2 variants were observed in these countries (independent t-test for all variants, $p > 0.270$) (**Figure 4B**). By analyzing the age distribution of the top variants in these countries, we observed a statistically significant underrepresentation in the number of sequenced viruses from children (0-9 years) (3.43% ± 1.21 S.E.M.) and young adolescents (10-18 years) (2.26% ± 1.16), while adults (19-49) were more represented (62.5% ± 8.84) (ANOVA; $F = 73.4$, $p = 1.15E-9$) (**Figure 4C**). For the three remaining age categories (19-49, 50-59 and >60 years), we report no statistically significant difference in the occurrence of the Wuhan-like or any of the identified variants (multivariate ANOVA for all variants, $F < 0.906$, $p > 0.487$) (**Figure 4D**).

We next sought to determine whether a correlation exists between the incidence of the ten most common variants and countries' CFRs. Only seven countries were included in this analysis – UAE, Saudi Arabia, Jordan, Oman, Bahrain, Palestine and Morocco – where only NSP12_P323L out of the ten variants was significantly correlated with an increase in CFR ($R=0.883$, $P=0.008$) (**Figure 5**). However, consistent with previous studies (Ilmjärv et al., 2020; Kannan et al., 2020), we confirm that NSP12_P323L seems to have quickly co-evolved following S_D614G introduction, as they are more often co-detected among SARS-CoV-2 genome sequences (**Figure 6A**). This double variant 'D614G & P323L' is found to be evenly represented among

different age groups in UAE, Saudi Arabia, Oman and Egypt (for sequences submitted by 12 November) (**Figure 6B**). After repeating the correlation analysis while accounting for the single versus double occurrence of D614G and P323L, we found that neither single nor double variants were significantly correlated with CFR. The P323L-single variant was mainly detected in Saudi Arabia (19% of it sequences) and among 0.5% sequences reported from UAE, while in the other countries it was exclusively co-detected with D614G. As such, we suggest that the earlier correlation of NSP12 P323L is likely the result of considering single- and double-variants as one group, specifically in Saudi Arabia. However, when such distinction was made, no correlation with higher CFR was reported in either group (**Figure 6C**).

Since sequences of UAE, Saudi Arabia and Oman were dominantly collected before September and those of Jordan and Bahrain were collected afterwards (**Figure 1D**), we repeated our analysis but this time accounting for the cumulative incidence of the 'D614G & P323L' double variant and the CFR at the end of each month. We observed no significant association between these two variables from March through September (**Supplementary Figure 2**). To further investigate this correlation, we considered the cumulative monthly incidence of 'D614G & P323L' and the monthly CFRs (**Figure 7A**): while UAE showed a significantly negative correlation between 'D614G & P323L' incidence and the monthly CFR (R=-0.830, P=0.021), Saudi Arabia showed a significantly positive correlation (R=0.909, P=0.012). Meanwhile, Jordan and Oman had no significant associations (R=-0.101, P=0.849; R=0.921, P=0.079 respectively) (**Figure 7B**). This suggests that by accounting for the change in CFR for each country with time the correlation between 'D614G & P323L' incidence and higher fatality does not necessarily hold true contrary to previous findings (Becerra-Flores and Cardozo, 2020; Eaaswarkhanth et al., 2020; Toyoshima et al., 2020).

Discussion

In this study, we analyzed whole-genome sequences of SARS-CoV-2 collected in the EMR countries. Significant data gaps are notable, whereby sequences were lacking for seven EMR countries – the majority of which are in war conflict zones (e.g., Syria, Libya, and Yemen (Daw, 2020)). In general, large disparities in the availability of sequences are observed across EMR. Only five countries reported ≥150 sequences spanning the first year of the pandemic, with UAE, Saudi Arabia and Jordan exceeding 500 sequence submissions. These data gaps highlight the challenges for monitoring the pandemic in developing countries particularly in conflict zones. It also emphasizes the importance of building regional sequencing capacities for timely and continuous monitoring of viral evolution during the pandemic (Al Saidi et al., 2020; Salameh, 2020).

Despite the diversity in the substitutions detected across countries, only a handful of variants predominated. For instance, a mere 9 out of 945 non-synonymous mutations (0.95%) in UAE and 6 out of 421 mutations (1.4%) in Saudi Arabia were detected in more than 10% of the respective total of sequences (**Figure 2**). This suggests that consistent with other studies, most SARS-CoV-2 mutations are neutral or deleterious (van Dorp et al., 2020a). Although a random founder's effect cannot be ruled out, the increased frequency of certain substitutions might be attributable to a fitness advantage (Callaway, 2020; Grubaugh et al., 2020; Kannan et al., 2020; Plante et al., 2020; Yurkovetskiy et al., 2020; Zhang et al., 2020). For instance, the S_D614G was found to predominate all viral strains globally (78% of sequences by May 2020) and was associated with increased viral loads in the upper respiratory tract of COVID-19 patients (Korber et al., 2020). Along the same line, Plante et al. recently demonstrated that S_D614G enhanced viral replication and infectivity of SARS-CoV-2 in human lung epithelial cells and primary

human airway tissues and increased infectious virus loads in hamsters' nasal cavities (Plante et al., 2020). Hou et al. further showed that the S_D614G variant displayed faster transmission and increased competitive fitness *in vivo* compared to wild-type virus in hamsters (Hou et al., 2020). These findings were also confirmed in hACE2 knock-in mice as well as in ferrets (Zhou et al., 2021). On a molecular level, S_D614G allows for more efficient infectivity by shifting the spike protein to a more 'open' conformation that is more favorable of ACE2 binding (Yurkovetskiy et al., 2020). Interestingly, however, a recent analysis by van Dorp et al. could not identify any single recurrent mutation, including the S_D614G, to be convincingly associated with increased viral transmission in humans (van Dorp et al., 2020b).

Several studies reported a positive correlation between S_D614G incidence and mortality (CFR) (Becerra-Flores and Cardozo, 2020; Eaaswarkhanth et al., 2020; Toyoshima et al., 2020). However, a more recent study of 25,000 SARS-CoV-2 sequences in the United Kingdom did not associate the incidence of this mutation with higher COVID-19 mortality or with clinical severity in patients (Volz et al., 2021). Similarly, Korber et al. did no find an association between S_D614G and increased disease severity or hospitalization (Korber et al., 2020). In our sequence set, we saw no association between S_D614G with CFRs of select EMR countries. Instead, the increasing incidence of NSP12_P323L (of the ten most common substiutions), when considered alone, was correlated with a higher mortality rate. However, our data as well as others' (Ilmjärv et al., 2020; Kannan et al., 2020) show that the NSP12_P323L and S_D614G are frequently coupled in most sequences, suggesting co-evolution of these variants (**Figure 6A**). Therefore, we assessed the combined effect of the 'D614G & P323L' double variant on mortality. We found no correlation between the cumulative incidence of the 'D614G & P323L' variant and CFRs in the EMR countries, consistent with earlier observations that found no association between 'D614G

& P323L' and deaths per million, CFR and patient status (Ilmjärv et al., 2020). In contrast to the well-studied S_D614G, the effect of NSP12_P323L on viral fitness is much less understood (Hillen et al., 2020; Pachetti et al., 2020). One study speculated that it negatively impacts NSP8's association with RdRp leading to an attenuated polymerase activity (Ilmjärv et al., 2020), while another suggested that it enhances hydrophobic NSP8-RdRp interactions leading to improved processivity of RdRp and enhanced viral replication overall (Kannan et al., 2020). In all cases, it is speculated that the 'epidemiological success' of the G, GH, GR and GV lineages could be attributed to the combined viral fitness advantage conferred by both NSP12_P323L and S_D614G (Ilmjärv et al., 2020).

Considering the limitations due to the discrepant number of sequences available (as mentioned above) or to country-specific variations in managing the COVID-19 pandemic (such as testing capacity, age stratification, enforcing/relaxing lockdown policies, etc.), we assessed the association of 'D614G & P323L' incidence with CFR within each of the top four represented countries over time. Interestingly, we found a significantly negative correlation between 'D614G & P323L' incidence and COVID-19 CFR in UAE, but an otherwise positive correlation in Saudi Arabia (**Figure 7B**). This suggests that the reported correlation between these variants and increased CFR might be a mere coincidence rather than true causation. Noteworthy, our method of calculating daily/monthly CFR is credited for accounting for a 13-day lag between cumulative cases and deaths (Wilson et al., 2020). Yet, it overestimates the case fatality risk early on in the pandemic (Ge and Sun, 2020), a limitation that would apply to our final analysis. Then again, considering the overwhelming discrepancy in fatality correlations between countries, it is unlikely that early bias in CFR estimation would have altered our conclusion. Additionally, a more accurate method to estimate COVID-19 mortality would be the infection fatality rate (IFR),

inferred from seroprevalence studies. However, such data is only available for Iran, Pakistan, and Qatar out of all EMR countries according to a recent WHO review (Ioannidis, 2020b).

Other genomic mutations were associated with milder disease, such as a major deletion (Δ382; nucleotide positions: 27,848 to 28,229), which results in truncated ORF7b and arrested transcription of ORF8 (Su et al., 2020). Patients infected with the SARS-CoV-2 Δ382 variant had less systemic release of proinflammatory cytokines and better clinical outcomes (i.e. lower frequency of hypoxia requiring supplemental oxygen) than wild-type (Young et al., 2020). This variant, however, was not present in any of our sequence cohort. Moreover, an S_ N439K variant was reported to enhance spike binding to its protein hACE2, resulting in immune escape from a panel of neutralizing monoclonal antibodies (Thomson et al., 2021). We detected only one sequence carrying this mutation from Morocco (EPI_ISL_728353).

On a final note, the COVID-19 pandemic cannot be viewed in the absence of other respiratory viruses, especially in view of the increased concerns regarding a 'Twindemic' (Uyeki et al., 2020; Adlhoch et al., 2021). In the case of influenza, few cases were reported in EMR countries in 2020-21 season (<20 positive specimens per week; GISRS report), a pattern that was established worldwide (Jones, 2020). Previous studies from the United States and Australia reported a co-infection rate ranging from 2% to 6.9% between SARS-CoV-2 and other respiratory viruses including rhinovirus/enterovirus, RSV, other coronaviruses, and influenza virus (Kim et al., 2020; Kim et al., 2021). Elghoudi et al. reported an incidence of 3.1% (9/288) respiratory co-infections with SARS-CoV-2 in children and young adolescents in a hospital in UAE (Elghoudi et al., 2020). In addition to co-infections with respiratory viruses a study from Pakistan reported a relatively high incidence of co-infections between SARS-CoV-2 and the endemic dengue virus – 25% (5/20) of studied samples (Saddique et al., 2020). A meta-analysis

study estimated that while 3% of COVID-19 patients (out of 1014 patients from 16 studies) had a confirmed viral co-infection, there was no significant increase in hospitalization or ICU admission in the co-infected cohort (Lansbury et al., 2020). Overall, the incidence of respiratory viral co-infections is low and may not be considered a key factor associated with mortality.

Our analysis reveals the diversity of SARS-CoV-2 in EMR and emphasizes the importance of considering temporal incidence and concurrent mutations in disease correlations. Our study is limited by the relatively low number of specimens sequenced per country spanning ten months of COVID-19 pandemic. There is a critical need for continuous and consistent surveillance of SARS-CoV-2 genomic variations in the EMR region.

Acknowledgments

S.O. is the recipient of a joint doctoral fellowship from the American University of Beirut (AUB) and the National Council for Scientific Research of Lebanon (CNRS-L). We gratefully acknowledge the authors from the originating laboratories responsible for obtaining the specimens, as well as the submitting laboratories where the genome data were generated and shared via GISAID, on which this research is based (a detailed list of authors and laboratories can be found in the supplementary material). We also thank Dr. Colin A. Smith and Dr. Heinrich zu Dohnna (Biology department, American University of Beirut) for their valuable comments.

Competing Interests: The authors declare no conflicts of interest.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Adlhoch, C., Mook, P., Lamb, F., Ferland, L., Melidou, A., Amato-Gauci, A.J., Pebody, R. and European Influenza Surveillance, N., 2021. Very little influenza in the WHO European Region during the 2020/21 season, weeks 40 2020 to 8 2021. Euro Surveill 26.

Al Saidi, A.M.O., Nur, F.A., Al-Mandhari, A.S., El Rabbat, M., Hafeez, A. and Abubakar, A., 2020. Decisive leadership is a necessity in the COVID-19 response. The Lancet 396, 295-298.

Alm, E., Broberg, E.K., Connor, T., Hodcroft, E.B., Komissarov, A.B., Maurer-Stroh, S., Melidou, A., Neher, R.A., O'Toole, A., Pereyaslov, D., laboratories, W.H.O.E.R.s., group, G.E., laboratories, W.H.O.E.R.s. and group*, G.E., 2020. Geographical and temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020. Euro Surveill 25.

Becerra-Flores, M. and Cardozo, T., 2020. SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate. Int J Clin Pract 74, e13525.

Callaway, E., 2020. The coronavirus is mutating - does it matter? Nature 585, 174-177.

Daw, M.A., 2020. Corona virus infection in Syria, Libya and Yemen; an alarming devastating threat. Travel Med Infect Dis 37, 101652.

Eaaswarkhanth, M., Al Madhoun, A. and Al-Mulla, F., 2020. Could the D614 G substitution in the SARS-CoV-2 spike (S) protein be associated with higher COVID-19 mortality? International Journal of Infectious Diseases.

Elghoudi, A., Aldhanhani, H., Ghatasheh, G., Sharif, E. and Narchi, H., 2020. Covid-19 in Children and Young Adolescents in Al Ain, United Arab Emirates- a Retrospective Cross-Sectional Study. Front Pediatr 8, 603741.

Fineberg, H.V., 2020. The Toll of COVID-19. JAMA 324, 1502-1503.

Ge, Y. and Sun, S., 2020. Estimation of Coronavirus Disease Case-Fatality Risk in Real Time. Emerg Infect Dis 26, 1922-1923.

GISAID, 2020. Clade and lineage nomenclature aids in genomic epidemiology studies of active hCoV-19 viruses. GISAID, Munich.

Grubaugh, N.D., Hanage, W.P. and Rasmussen, A.L., 2020. Making Sense of Mutation: What D614G Means for the COVID-19 Pandemic Remains Unclear. Cell 182, 794-795.

Hadfield, J., Megill, C., Bell, S.M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T. and Neher, R.A., 2018. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics 34, 4121-4123.

Hartenian, E., Nandakumar, D., Lari, A., Ly, M., Tucker, J.M. and Glaunsinger, B.A., 2020. The molecular virology of coronaviruses. J Biol Chem 295, 12910-12934.

Hillen, H.S., Kokic, G., Farnung, L., Dienemann, C., Tegunov, D. and Cramer, P., 2020. Structure of replicating SARS-CoV-2 polymerase. Nature 584, 154-156.

Hou, Y.J., Chiba, S., Halfmann, P., Ehre, C., Kuroda, M., Dinnon, K.H., 3rd, Leist, S.R., Schafer, A., Nakajima, N., Takahashi, K., Lee, R.E., Mascenik, T.M., Graham, R., Edwards, C.E., Tse, L.V., Okuda, K., Markmann, A.J., Bartelt, L., de Silva, A., Margolis, D.M., Boucher, R.C., Randell, S.H., Suzuki, T., Gralinski, L.E., Kawaoka, Y. and Baric, R.S., 2020. SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. Science.

Ilmjärv, S., Abdul, F., Acosta-Gutiérrez, S., Estarellas, C., Galdadas, I., Casimir, M., Alessandrini, M., Gervasio, F.L. and Krause, K.-H., 2020. Epidemiologically most

successful SARS-CoV-2 variant: concurrent mutations in RNA-dependent RNA polymerase and spike protein. medRxiv.
Ioannidis, J.P.A., 2020a. Global perspective of COVID-19 epidemiology for a full-cycle pandemic. Eur J Clin Invest, e13423.
Ioannidis, J.P.A., 2020b. Infection fatality rate of COVID-19 inferred from seroprevalence data. Bulletin of the World Health Organization.
Jones, N., 2020. How COVID-19 is changing the cold and flu season. Nature 588, 388-390.
Kannan, S.R., Spratt, A.N., Quinn, T.P., Heng, X., Lorson, C.L., Sonnerborg, A., Byrareddy, S.N. and Singh, K., 2020. Infectivity of SARS-CoV-2: there Is Something More than D614G? J Neuroimmune Pharmacol.
Keaten, J., 2020. WHO: 10% of world's people might have had COVID-19, Telegraph - Herald. Dubuque, Iowa.
Kim, D., Quinn, J., Pinsky, B., Shah, N.H. and Brown, I., 2020. Rates of Co-infection Between SARS-CoV-2 and Other Respiratory Pathogens. JAMA 323, 2085-2086.
Kim, K.W., Deveson, I.W., Pang, C.N.I., Yeang, M., Naing, Z., Adikari, T., Hammond, J.M., Stevanovski, I., Beukers, A.G., Verich, A., Yin, S., McFarlane, D., Wilkins, M.R., Stelzer-Braid, S., Bull, R.A., Craig, M.E., van Hal, S.J. and Rawlinson, W.D., 2021. Respiratory viral co-infections among SARS-CoV-2 cases confirmed by virome capture sequencing. Sci Rep 11, 3934.
Korber, B., Fischer, W.M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E.E., Bhattacharya, T., Foley, B., Hastie, K.M., Parker, M.D., Partridge, D.G., Evans, C.M., Freeman, T.M., de Silva, T.I., Sheffield, C.-G.G., McDanal, C., Perez, L.G., Tang, H., Moon-Walker, A., Whelan, S.P., LaBranche, C.C., Saphire, E.O. and Montefiori, D.C., 2020. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell 182, 812-827 e19.
Lansbury, L., Lim, B., Baskaran, V. and Lim, W.S., 2020. Co-infections in people with COVID-19: a systematic review and meta-analysis. J Infect 81, 266-275.
Pachetti, M., Marini, B., Benedetti, F., Giudici, F., Mauro, E., Storici, P., Masciovecchio, C., Angeletti, S., Ciccozzi, M., Gallo, R.C., Zella, D. and Ippodrino, R., 2020. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. J Transl Med 18, 179.
Plante, J.A., Liu, Y., Liu, J., Xia, H., Johnson, B.A., Lokugamage, K.G., Zhang, X., Muruato, A.E., Zou, J., Fontes-Garfias, C.R., Mirchandani, D., Scharton, D., Bilello, J.P., Ku, Z., An, Z., Kalveram, B., Freiberg, A.N., Menachery, V.D., Xie, X., Plante, K.S., Weaver, S.C. and Shi, P.Y., 2020. Spike mutation D614G alters SARS-CoV-2 fitness. Nature.
Saddique, A., Rana, M.S., Alam, M.M., Ikram, A., Usman, M., Salman, M., Faryal, R., Massab, U., Bokhari, H., Mian, M.S., Israr, A. and Safiullah, 2020. Emergence of co-infection of COVID-19 and dengue: A serious public health threat. J Infect 81, e16-e18.
Salameh, P., 2020. COVID-19 in the Eastern Mediterranean Region: testing frequency, cumulative cases and mortality analysis. East Mediterr Health J 26, 1005-1010.
Sallam, M., Ababneh, N.A., Dababseh, D., Bakri, F.G. and Mahafzah, A., 2021. Temporal increase in D614G mutation of SARS-CoV-2 in the Middle East and North Africa. Heliyon 7, e06035.
Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A. and Li, F., 2020. Cell entry mechanisms of SARS-CoV-2. Proc Natl Acad Sci U S A 117, 11727-11734.

Shu, Y. and McCauley, J., 2017. GISAID: Global initiative on sharing all influenza data - from vision to reality. Euro Surveill 22.

Su, S., Wong, G., Shi, W., Liu, J., Lai, A.C.K., Zhou, J., Liu, W., Bi, Y. and Gao, G.F., 2016. Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. Trends Microbiol 24, 490-502.

Su, Y.C.F., Anderson, D.E., Young, B.E., Linster, M., Zhu, F., Jayakumar, J., Zhuang, Y., Kalimuddin, S., Low, J.G.H., Tan, C.W., Chia, W.N., Mak, T.M., Octavia, S., Chavatte, J.M., Lee, R.T.C., Pada, S., Tan, S.Y., Sun, L., Yan, G.Z., Maurer-Stroh, S., Mendenhall, I.H., Leo, Y.S., Lye, D.C., Wang, L.F. and Smith, G.J.D., 2020. Discovery and Genomic Characterization of a 382-Nucleotide Deletion in ORF7b and ORF8 during the Early Evolution of SARS-CoV-2. mBio 11.

Thomson, E.C., Rosen, L.E., Shepherd, J.G., Spreafico, R., da Silva Filipe, A., Wojcechowskyj, J.A., Davis, C., Piccoli, L., Pascall, D.J., Dillen, J., Lytras, S., Czudnochowski, N., Shah, R., Meury, M., Jesudason, N., De Marco, A., Li, K., Bassi, J., O'Toole, A., Pinto, D., Colquhoun, R.M., Culap, K., Jackson, B., Zatta, F., Rambaut, A., Jaconi, S., Sreenu, V.B., Nix, J., Zhang, I., Jarrett, R.F., Glass, W.G., Beltramello, M., Nomikou, K., Pizzuto, M., Tong, L., Cameroni, E., Croll, T.I., Johnson, N., Di Iulio, J., Wickenhagen, A., Ceschi, A., Harbison, A.M., Mair, D., Ferrari, P., Smollett, K., Sallusto, F., Carmichael, S., Garzoni, C., Nichols, J., Galli, M., Hughes, J., Riva, A., Ho, A., Schiuma, M., Semple, M.G., Openshaw, P.J.M., Fadda, E., Baillie, J.K., Chodera, J.D., Investigators, I.C., Consortium, C.-G.U., Rihn, S.J., Lycett, S.J., Virgin, H.W., Telenti, A., Corti, D., Robertson, D.L. and Snell, G., 2021. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. Cell 184, 1171-1187 e20.

Toyoshima, Y., Nemoto, K., Matsumoto, S., Nakamura, Y. and Kiyotani, K., 2020. SARS-CoV-2 genomic variations associated with mortality rate of COVID-19. J Hum Genet.

Uyeki, T.M., Santoli, J. and Jernigan, D.B., 2020. Preparing for the 2020-2021 Influenza Season. JAMA 324, 2318-2319.

van Dorp, L., Acman, M., Richard, D., Shaw, L.P., Ford, C.E., Ormond, L., Owen, C.J., Pang, J., Tan, C.C.S., Boshier, F.A.T., Ortiz, A.T. and Balloux, F., 2020a. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. Infect Genet Evol 83, 104351.

van Dorp, L., Richard, D., Tan, C.C.S., Shaw, L.P., Acman, M. and Balloux, F., 2020b. No evidence for increased transmissibility from recurrent mutations in SARS-CoV-2. Nat Commun 11, 5986.

Vankadari, N., 2020. Overwhelming mutations or SNPs of SARS-CoV-2: A point of caution. Gene 752, 144792.

Volz, E., Hill, V., McCrone, J.T., Price, A., Jorgensen, D., O'Toole, A., Southgate, J., Johnson, R., Jackson, B., Nascimento, F.F., Rey, S.M., Nicholls, S.M., Colquhoun, R.M., da Silva Filipe, A., Shepherd, J., Pascall, D.J., Shah, R., Jesudason, N., Li, K., Jarrett, R., Pacchiarini, N., Bull, M., Geidelberg, L., Siveroni, I., Consortium, C.-U., Goodfellow, I., Loman, N.J., Pybus, O.G., Robertson, D.L., Thomson, E.C., Rambaut, A. and Connor, T.R., 2021. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. Cell 184, 64-75 e11.

Wang, C., Liu, Z., Chen, Z., Huang, X., Xu, M., He, T. and Zhang, Z., 2020. The establishment of reference sequence for SARS-CoV-2 and variation analysis. J Med Virol 92, 667-674.

WHO. Eastern Mediterranean Regional Office COVID-19 Dashboard (2021). http://www.emro.who.int/index.html. Last Accessed: Jan 22 2021.

WHO. WHO Coronavirus Disease (COVID-19) Dashboard (2021). https://covid19.who.int/. Last Accessed: Jan 22 2021.

WHO. WHO Global Influenza Surveillance and Response System (GISRS). Number of Specimens Positive for Influenza by Subtype, Eastern Mediterranean Region of WHO (2021) https://apps.who.int/flumart/Default?ReportNo=3&WHORegion= Last Accessed: Mar 21 2021.

Wilson, N., Kvalsvig, A., Barnard, L.T. and Baker, M.G., 2020. Case-Fatality Risk Estimates for COVID-19 Calculated by Using a Lag Time for Fatality. Emerg Infect Dis 26, 1339-1441.

Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.L., Abiona, O., Graham, B.S. and McLellan, J.S., 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367, 1260-1263.

Young, B.E., Fong, S.W., Chan, Y.H., Mak, T.M., Ang, L.W., Anderson, D.E., Lee, C.Y., Amrun, S.N., Lee, B., Goh, Y.S., Su, Y.C.F., Wei, W.E., Kalimuddin, S., Chai, L.Y.A., Pada, S., Tan, S.Y., Sun, L., Parthasarathy, P., Chen, Y.Y.C., Barkham, T., Lin, R.T.P., Maurer-Stroh, S., Leo, Y.S., Wang, L.F., Renia, L., Lee, V.J., Smith, G.J.D., Lye, D.C. and Ng, L.F.P., 2020. Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. Lancet 396, 603-611.

Yurkovetskiy, L., Wang, X., Pascal, K.E., Tomkins-Tinch, C., Nyalile, T.P., Wang, Y., Baum, A., Diehl, W.E., Dauphin, A., Carbone, C., Veinotte, K., Egri, S.B., Schaffner, S.F., Lemieux, J.E., Munro, J.B., Rafique, A., Barve, A., Sabeti, P.C., Kyratsous, C.A., Dudkina, N.V., Shen, K. and Luban, J., 2020. Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. Cell 183, 739-751 e8.

Zhang, L., Jackson, C.B., Mou, H., Ojha, A., Peng, H., Quinlan, B.D., Rangarajan, E.S., Pan, A., Vanderheiden, A., Suthar, M.S., Li, W., Izard, T., Rader, C., Farzan, M. and Choe, H., 2020. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. Nat Commun 11, 6013.

Zhou, B., Thi Nhu Thao, T., Hoffmann, D., Taddeo, A., Ebert, N., Labroussaa, F., Pohlmann, A., King, J., Steiner, S., Kelly, J.N., Portmann, J., Halwe, N.J., Ulrich, L., Trueb, B.S., Fan, X., Hoffmann, B., Wang, L., Thomann, L., Lin, X., Stalder, H., Pozzi, B., de Brot, S., Jiang, N., Cui, D., Hossain, J., Wilson, M., Keller, M., Stark, T.J., Barnes, J.R., Dijkman, R., Jores, J., Benarafa, C., Wentworth, D.E., Thiel, V. and Beer, M., 2021. SARS-CoV-2 spike D614G change enhances replication and transmission. Nature.

Figure Legends

Figure 1. Distribution of SARS-CoV-2 sequences in EMR countries. (A) Pie chart showing number of available sequences as of 1 January 2021. (B) Histogram showing ratio of sequences per million population. (C) Histogram showing ratio of sequences per 10,000 confirmed cases of SARS-CoV-2. (D) Temporal distribution of sequence collection in EMR countries from January through December 2020. (E) Averages of delay in sequence submission (in days) of available EMR countries. For UAE, Saudi Arabia and Jordan, only the first and last 100 collected sequences were considered. Error bars represent standard deviation.

Figure 1. Common SARS-CoV-2 substitutions detected in EMR countries. Bar graphs showing percentages of most prevalent SARS-CoV-2 variants per country, excluding single-incidence variants. Colored are the top ten most represented substitutions; in black are country-specific ones.

Figure 2. EMR sequences are dominantly distributed in G, GR and GH GISAID clades. (A) Phylogenetic tree of GISAID clades plotted over time. Red arrow points to the introduction of S_D614G, marking the emergence of G, GR and GH clades (reproduced from Nextstrain [Ref. 17]. (B) Clade distribution of all 3345 sequences. (C) Temporal distribution of sequences as clades in the top four represented countries.

Figure 4. Gender and age distribution of SARS-CoV-2 variants in four EMR countries. Gender distribution of (A) sequences of Saudi Arabia, UAE, Oman and Egypt and (B) select top most frequent variants of these countries. (C) Age groups distribution of sequences from four EMR countries and (D) of most frequent variants. Sequences in this figure are as of 12 November 2020 due to the inability to extract demographics from metafiles after this date.

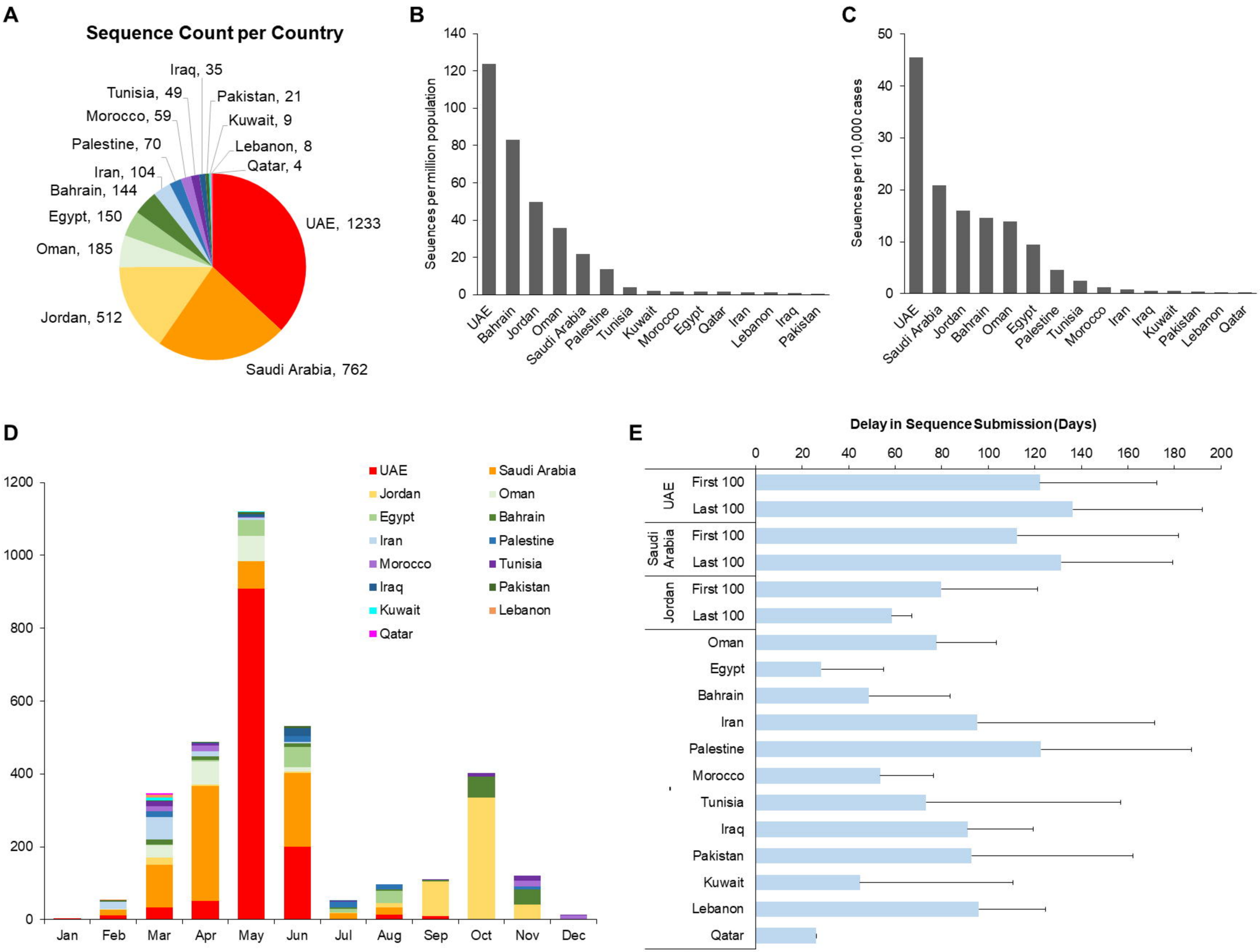
Figure 5. Only NSP12_P323L of the ten most common SARS-CoV-2 single variants correlates with higher case fatality rates in seven EMR countries. Scatter plots showing percentages of each of the top ten represented substitutions with cumulative CFR as of 1 January 2021. R, Pearson's correlation coefficient.

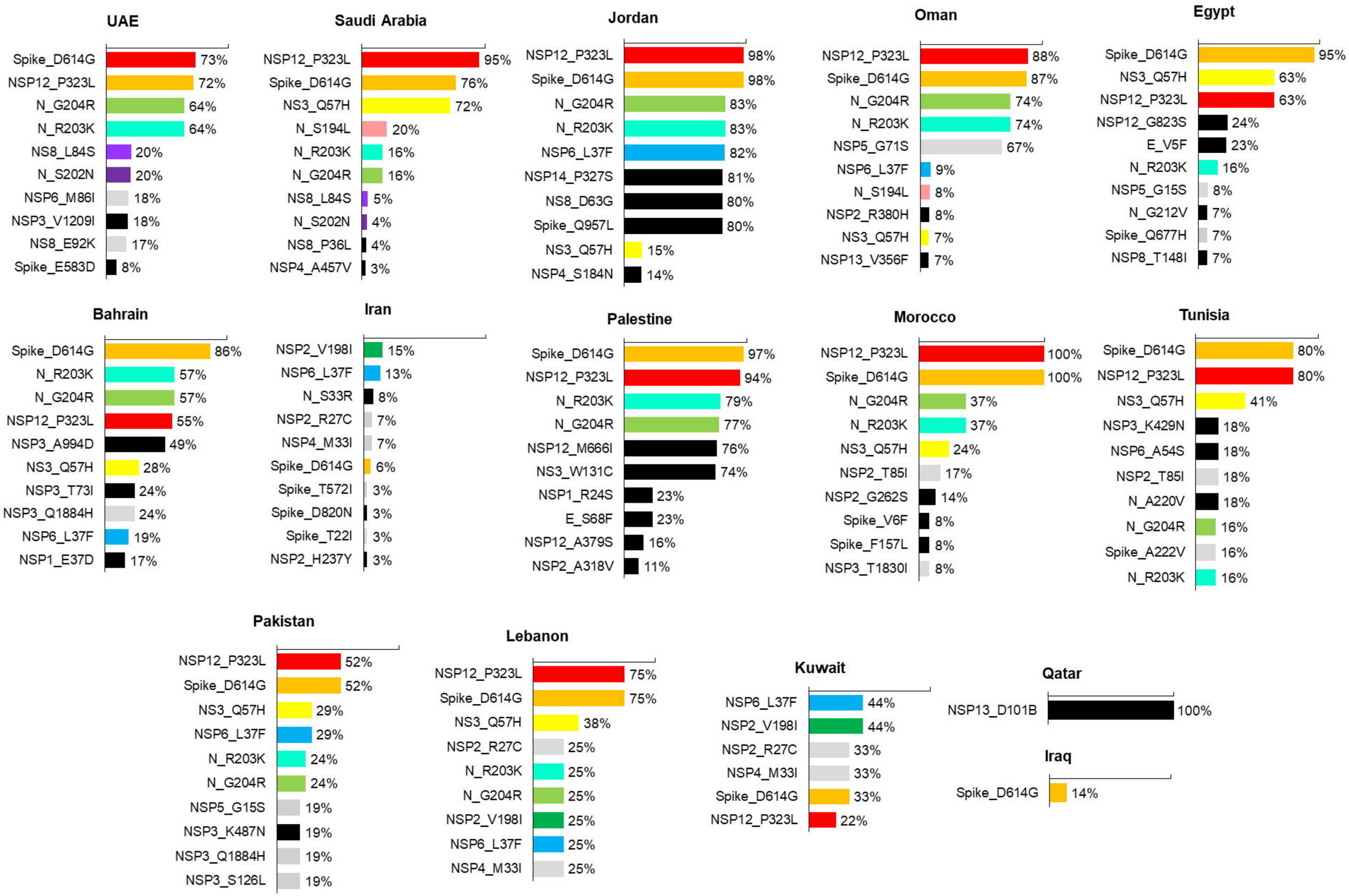
Figure 6. The 'D614G & P323L' double variant is predominant in select EMR countries, but does not correlate with increased fatality. (A) Percentages of Wuhan-like, S_D614G only, NSP_P323L only and 'D614G & P323L' double variants in eight EMR countries. (B) Distribution of the 'D614 & P323L' variant based on age groups in top four represented EMR countries. (C) Scatter plots showing correlation of three variants with CFR in seven EMR countries. R, Pearson's correlation coefficient.

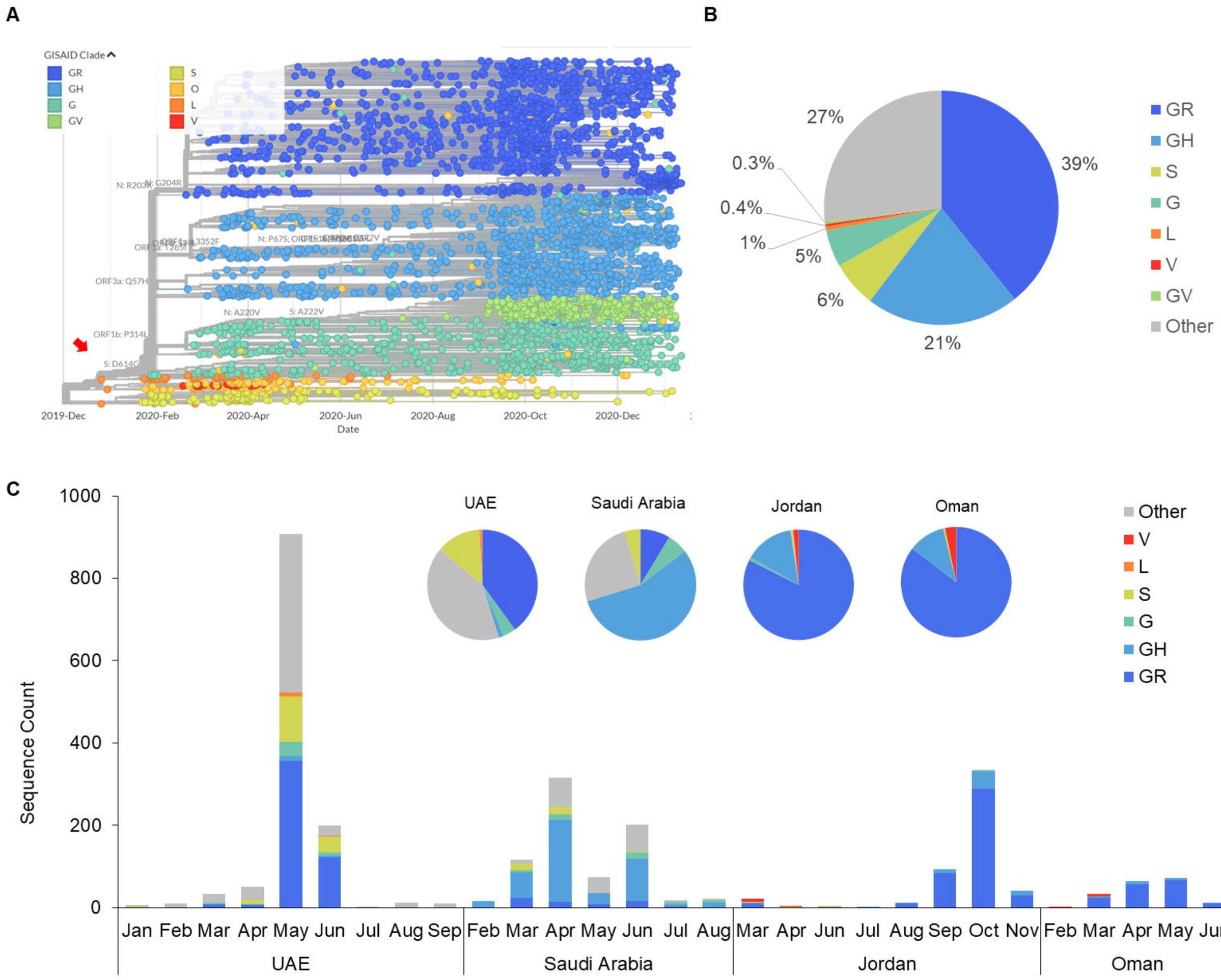
Figure 7. Temporal distribution and fatality correlation of the 'D614G & P323L' double variant with CFR in each of the top four represented EMR Countries. (A) Time-based evolution of double variant percentages and CFR in Saudi Arabia, UAE, Oman and Egypt. (B) Scatter plots showing correlations of the 'D614G & P323L' double variant with CFR over select months of pandemic life in the top four represented EMR countries. R, Pearson's correlation coefficient.

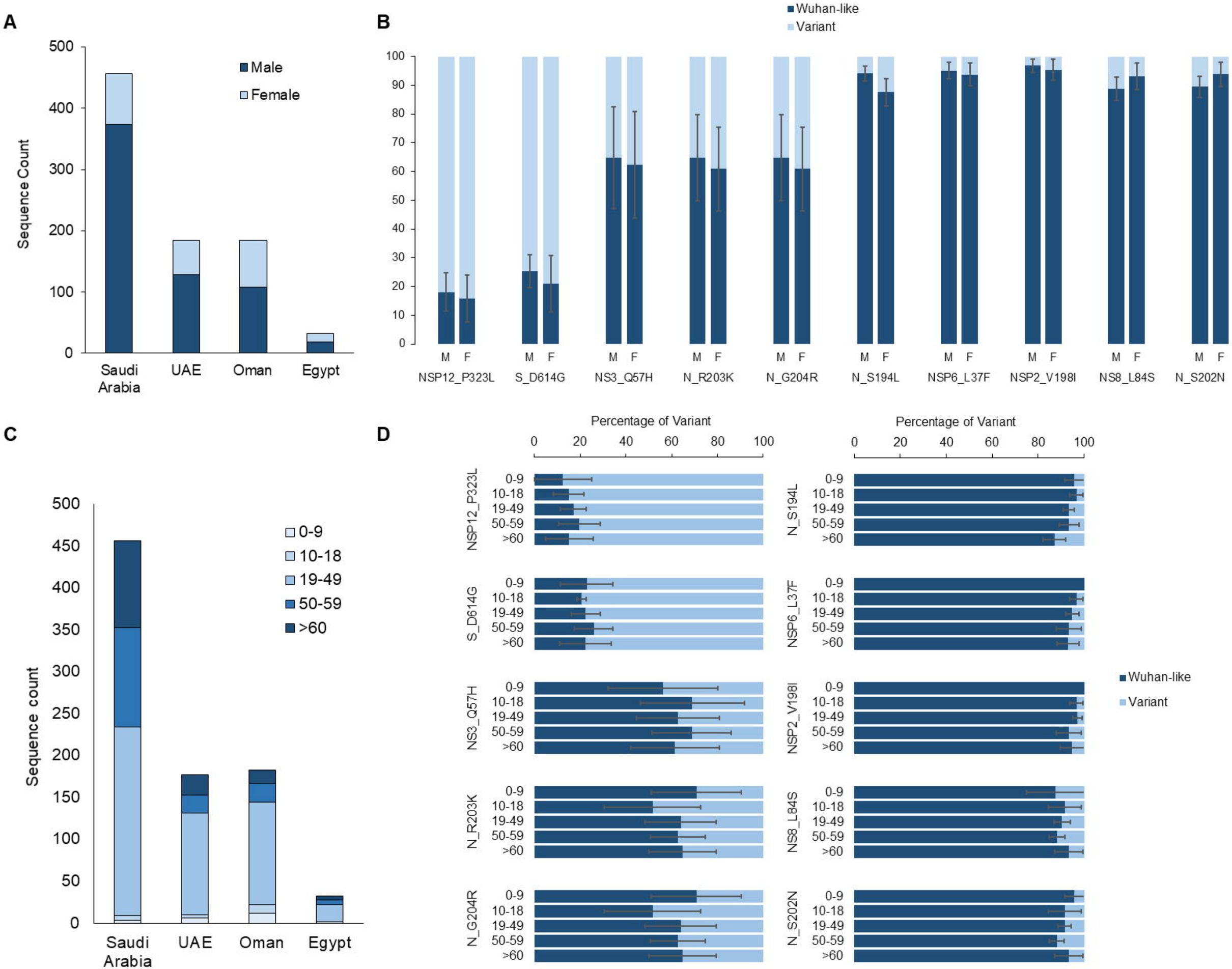
Supplementary Figure 1. Genome length of available EMR sequences. Sequence size distribution (%) of available EMR sequences in two categories: <5000 bp and >25000 bp.

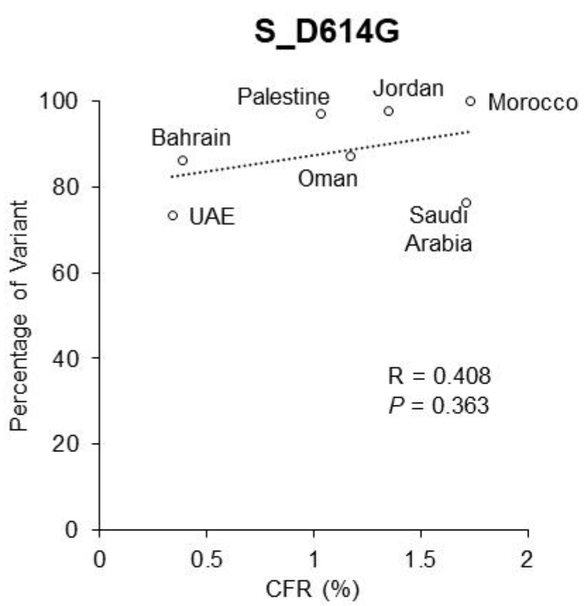
Supplementary Figure 2. Correlation between the 'D614G & P323L' double variant with CFR over eight months of the pandemic. Scatter plots showing cumulative percentages of the double variant in seven EMR countries versus respective CFR as of the end of each month, March through September and November. R, Pearson's correlation coefficient.

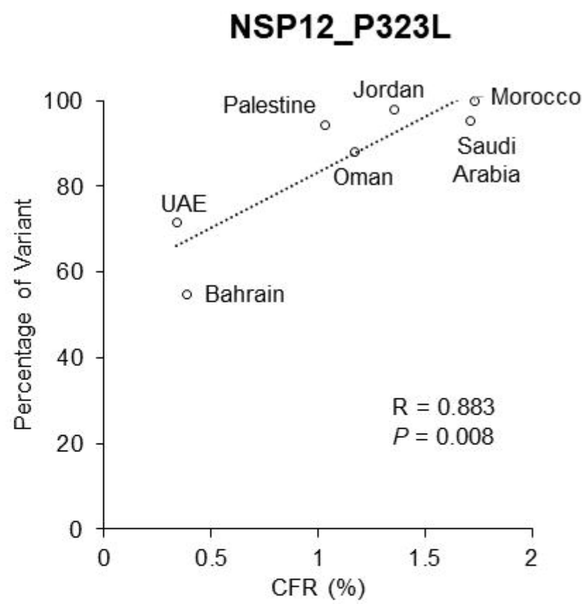

A
Sequence Count per Country
Iraq, 35
Tunisia, 49
Pakistan, 21
Morocco, 59
Kuwait, 9
Palestine, 70
Lebanon, 8
Qatar, 4
Iran, 104
Bahrain, 144
Egypt, 150
UAE, 1233
Oman, 185
Jordan, 512
Saudi Arabia, 762
B
Seuences per million population
0
20
40
60
80
100
120
140
UAE
Bahrain
Jordan
Oman
Saudi Arabia
Palestine
Tunisia
Kuwait
Morocco
Egypt
Qatar
Iran
Lebanon
Iraq
Pakistan
C
Seuences per 10,000 cases
0
10
20
30
40
50
UAE
Saudi Arabia
Jordan
Bahrain
Oman
Egypt
Palestine
Tunisia
Morocco
Iran
Iraq
Kuwait
Pakistan
Lebanon
Qatar
D
Sequence Count
0
200
400
600
800
1000
1200
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
UAE
Saudi Arabia
Jordan
Oman
Egypt
Bahrain
Iran
Palestine
Morocco
Tunisia
Iraq
Pakistan
Kuwait
Lebanon
Qatar
E
Delay in Sequence Submission (Days)
0
20
40
60
80
100
120
140
160
180
200
UAE
First 100
Last 100
Saudi Arabia
First 100
Last 100
Jordan
First 100
Last 100
Oman
Egypt
Bahrain
Iran
Palestine
Morocco
Tunisia
Iraq
Pakistan
Kuwait
Lebanon
Qatar

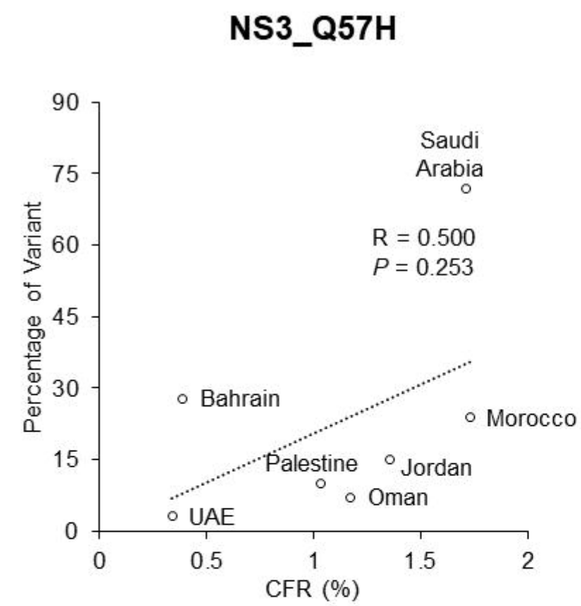

UAE
Spike_D614G 73%
NSP12_P323L 72%
N_G204R 64%
N_R203K 64%
NS8_L84S 20%
N_S202N 20%
NSP6_M86I 18%
NSP3_V1209I 18%
NS8_E92K 17%
Spike_E583D 8%
Saudi Arabia
NSP12_P323L 95%
Spike_D614G 76%
NS3_Q57H 72%
N_S194L 20%
N_R203K 16%
N_G204R 16%
NS8_L84S 5%
N_S202N 4%
NS8_P36L 4%
NSP4_A457V 3%
Jordan
NSP12_P323L 98%
Spike_D614G 98%
N_G204R 83%
N_R203K 83%
NSP6_L37F 82%
NSP14_P327S 81%
NS8_D63G 80%
Spike_Q957L 80%
NS3_Q57H 15%
NSP4_S184N 14%
Oman
NSP12_P323L 88%
Spike_D614G 87%
N_G204R 74%
N_R203K 74%
NSP5_G71S 67%
NSP6_L37F 9%
N_S194L 8%
NSP2_R380H 8%
NS3_Q57H 7%
NSP13_V356F 7%
Egypt
Spike_D614G 95%
NS3_Q57H 63%
NSP12_P323L 63%
NSP12_G823S 24%
E_V5F 23%
N_R203K 16%
NSP5_G15S 8%
N_G212V 7%
Spike_Q677H 7%
NSP8_T148I 7%
Bahrain
Spike_D614G 86%
N_R203K 57%
N_G204R 57%
NSP12_P323L 55%
NSP3_A994D 49%
NS3_Q57H 28%
NSP3_T73I 24%
NSP3_Q1884H 24%
NSP6_L37F 19%
NSP1_E37D 17%
Iran
NSP2_V198I 15%
NSP6_L37F 13%
N_S33R 8%
NSP2_R27C 7%
NSP4_M33I 7%
Spike_D614G 6%
Spike_T572I 3%
Spike_D820N 3%
Spike_T22I 3%
NSP2_H237Y 3%
Palestine
Spike_D614G 97%
NSP12_P323L 94%
N_R203K 79%
N_G204R 77%
NSP12_M666I 76%
NS3_W131C 74%
NSP1_R24S 23%
E_S68F 23%
NSP12_A379S 16%
NSP2_A318V 11%
Morocco
NSP12_P323L 100%
Spike_D614G 100%
N_G204R 37%
N_R203K 37%
NS3_Q57H 24%
NSP2_T85I 17%
NSP2_G262S 14%
Spike_V6F 8%
Spike_F157L 8%
NSP3_T1830I 8%
Tunisia
Spike_D614G 80%
NSP12_P323L 80%
NS3_Q57H 41%
NSP3_K429N 18%
NSP6_A54S 18%
NSP2_T85I 18%
N_A220V 18%
N_G204R 16%
Spike_A222V 16%
N_R203K 16%
Pakistan
NSP12_P323L 52%
Spike_D614G 52%
NS3_Q57H 29%
NSP6_L37F 29%
N_R203K 24%
N_G204R 24%
NSP5_G15S 19%
NSP3_K487N 19%
NSP3_Q1884H 19%
NSP3_S126L 19%
Lebanon
NSP12_P323L 75%
Spike_D614G 75%
NS3_Q57H 38%
NSP2_R27C 25%
N_R203K 25%
N_G204R 25%
NSP2_V198I 25%
NSP6_L37F 25%
NSP4_M33I 25%
Kuwait
NSP6_L37F 44%
NSP2_V198I 44%
NSP2_R27C 33%
NSP4_M33I 33%
Spike_D614G 33%
NSP12_P323L 22%
Qatar
NSP13_D101B 100%
Iraq
Spike_D614G 14%

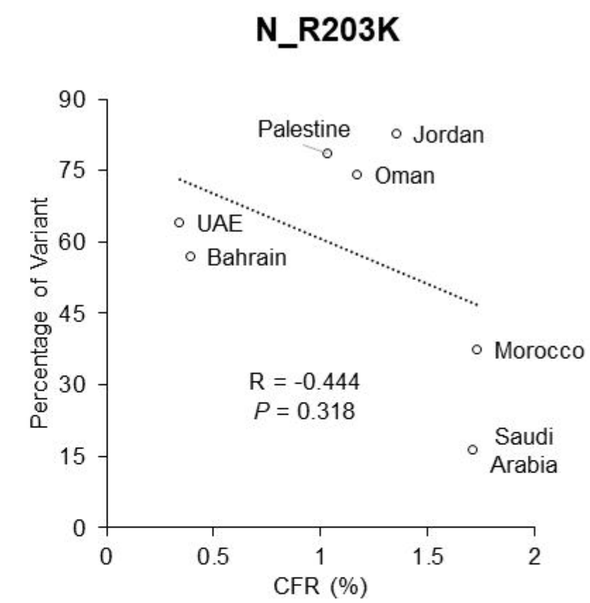

A
GISAID Clade
GR
GH
G
GV
S
O
L
V
N: R203K
N: G204R
ORF3a: Q57H
N: A220V
S: A222V
ORF1b: P314L
S: D614G
2019-Dec
2020-Feb
2020-Apr
2020-Jun
2020-Aug
2020-Oct
2020-Dec
Date
B
27%
39%
0.3%
0.4%
1%
5%
6%
21%
GR
GH
S
G
L
V
GV
Other
C
Sequence Count
1000
800
600
400
200
0
UAE
Saudi Arabia
Jordan
Oman
Other
V
L
S
G
GH
GR
Jan Feb Mar Apr May Jun Jul Aug Sep
UAE
Feb Mar Apr May Jun Jul Aug
Saudi Arabia
Mar Apr Jun Jul Aug Sep Oct Nov
Jordan
Feb Mar Apr May Jun
Oman

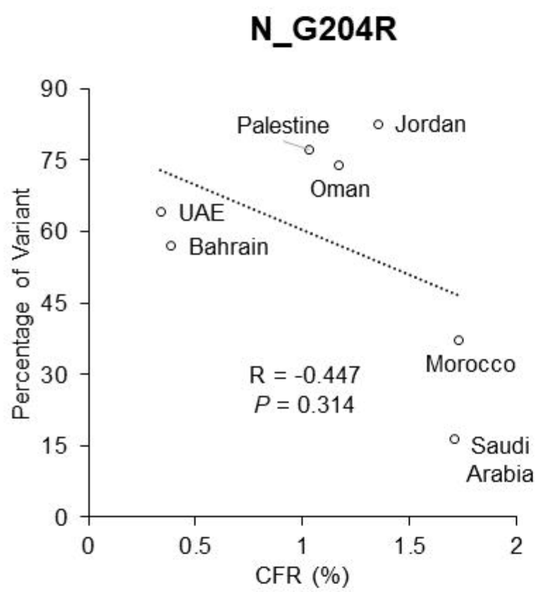

A
Male
Female
Sequence Count
500
400
300
200
100
0
Saudi Arabia
UAE
Oman
Egypt
B
Wuhan-like
Variant
100
90
80
70
60
50
40
30
20
10
0
M F
NSP12_P323L
S_D614G
NS3_Q57H
N_R203K
N_G204R
N_S194L
NSP6_L37F
NSP2_V198I
NS8_L84S
N_S202N
C
0-9
10-18
19-49
50-59
>60
Sequence count
500
450
400
350
300
250
200
150
100
50
0
Saudi Arabia
UAE
Oman
Egypt
D
Percentage of Variant
0 20 40 60 80 100
NSP12_P323L
S_D614G
NS3_Q57H
N_R203K
N_G204R
N_S194L
NSP6_L37F
NSP2_V198I
NS8_L84S
N_S202N
0-9
10-18
19-49
50-59
>60
Wuhan-like
Variant

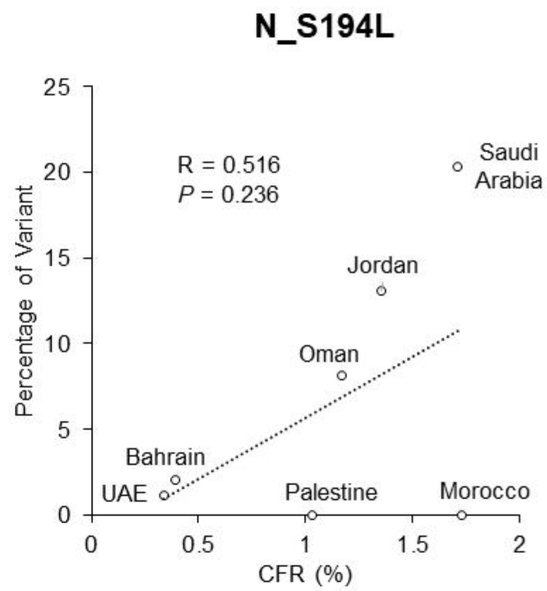

S_D614G
Percentage of Variant
CFR (%)
Palestine
Jordan
Morocco
Bahrain
Oman
UAE
Saudi Arabia
R = 0.408
P = 0.363

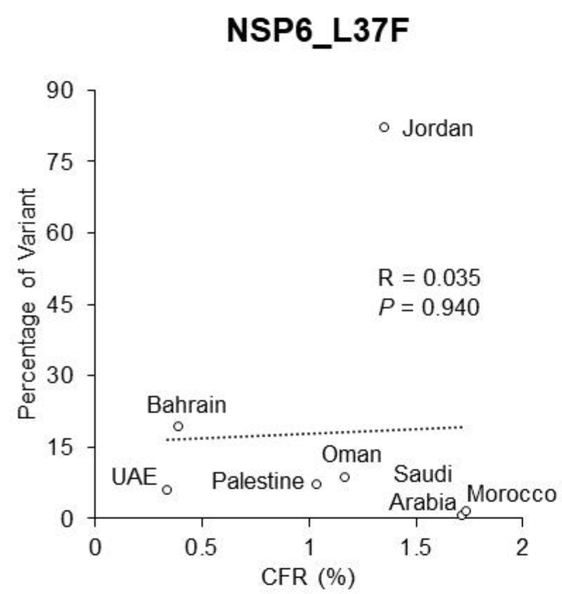

NSP12_P323L
Percentage of Variant
CFR (%)
Palestine
Jordan
Morocco
Saudi Arabia
Oman
UAE
Bahrain
R = 0.883
P = 0.008

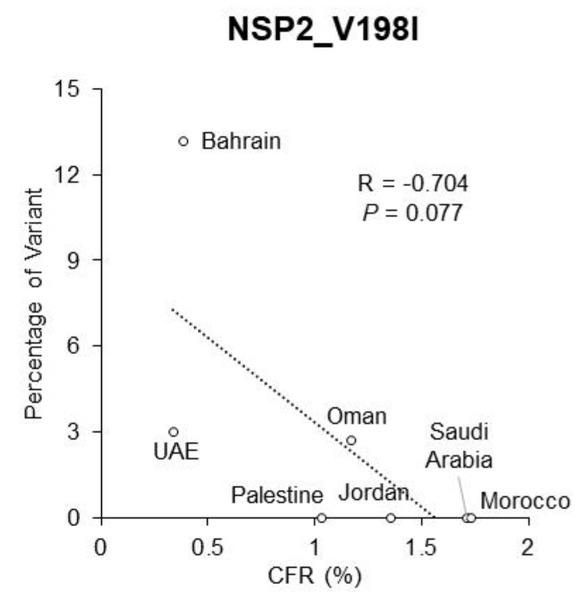

NS3_Q57H
Percentage of Variant
CFR (%)
Saudi Arabia
R = 0.500
P = 0.253
Bahrain
Morocco
Palestine
Jordan
Oman
UAE

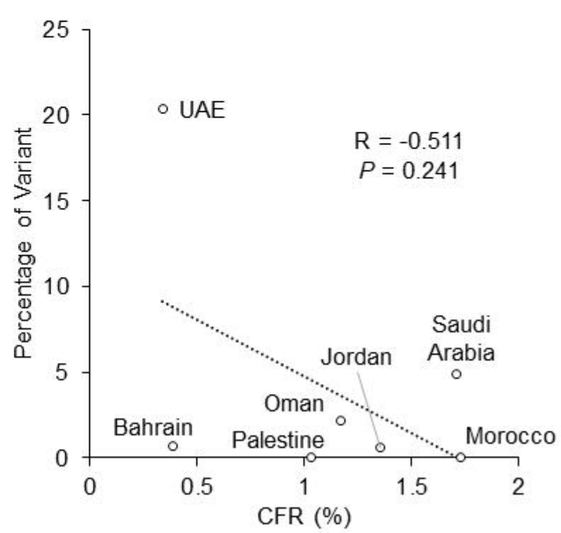

N_R203K
Percentage of Variant
CFR (%)
Palestine
Jordan
Oman
UAE
Bahrain
Morocco
R = -0.444
P = 0.318
Saudi Arabia

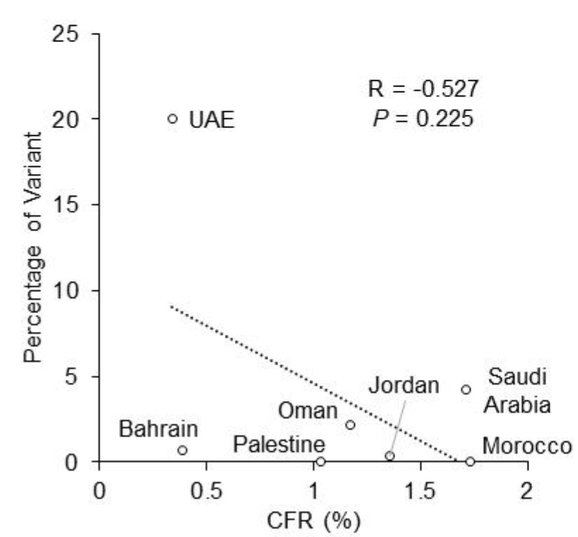

N_G204R
Percentage of Variant
CFR (%)
Palestine
Jordan
Oman
UAE
Bahrain
Morocco
R = -0.447
P = 0.314
Saudi Arabia

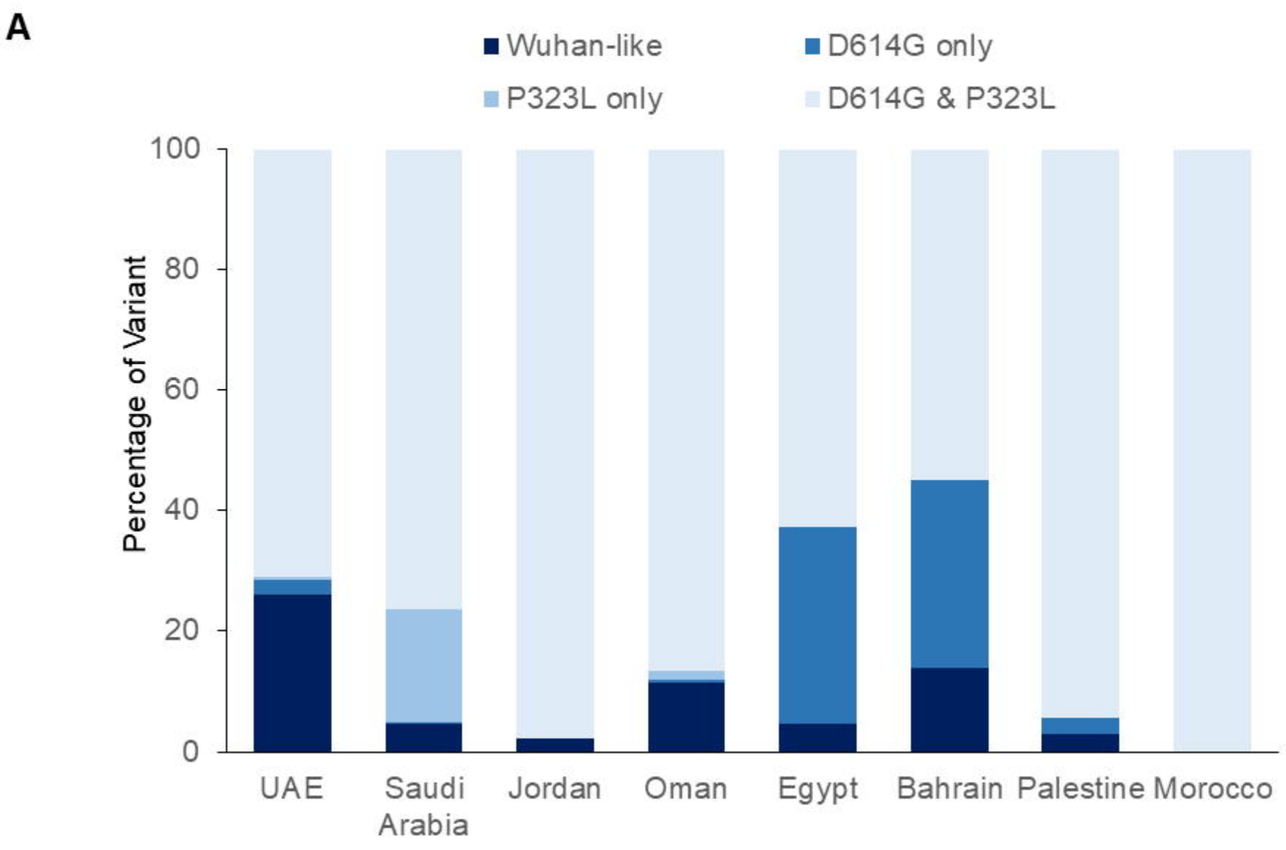

N_S194L
Percentage of Variant
CFR (%)
R = 0.516
P = 0.236
Saudi Arabia
Jordan
Oman
Bahrain
UAE
Palestine
Morocco

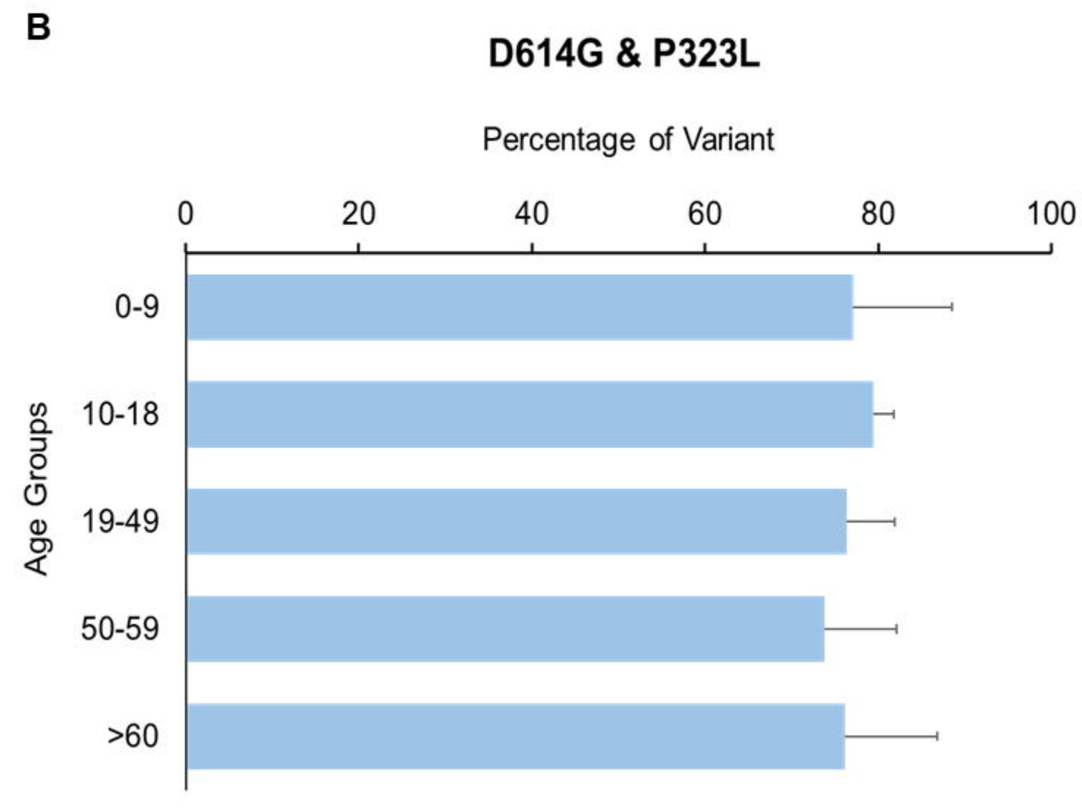

NSP6_L37F
Percentage of Variant
CFR (%)
Jordan
R = 0.035
P = 0.940
Bahrain
UAE
Palestine
Oman
Saudi Arabia
Morocco

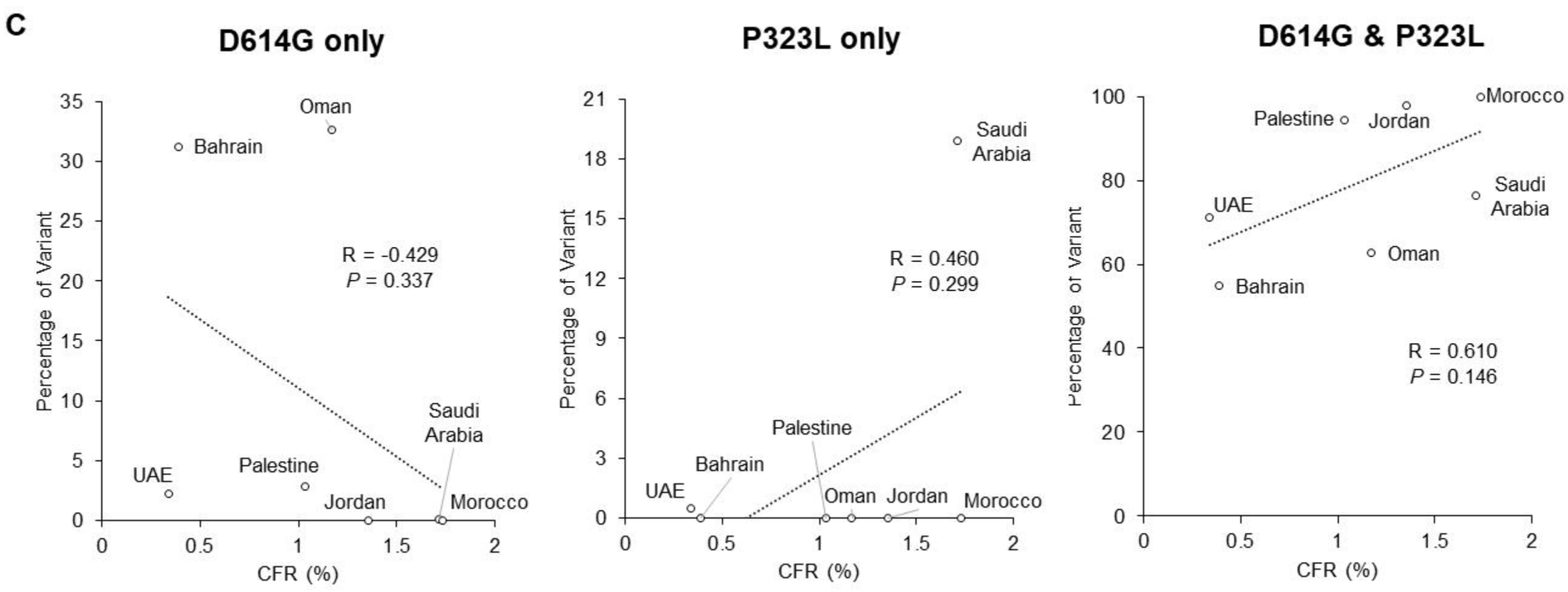

NSP2_V198I
Percentage of Variant
CFR (%)
Bahrain
R = -0.704
P = 0.077
Oman
UAE
Saudi Arabia
Palestine
Jordan
Morocco


NS8_L84S
Percentage of Variant
CFR (%)
UAE
R = -0.511
P = 0.241
Saudi Arabia
Jordan
Oman
Bahrain
Palestine
Morocco

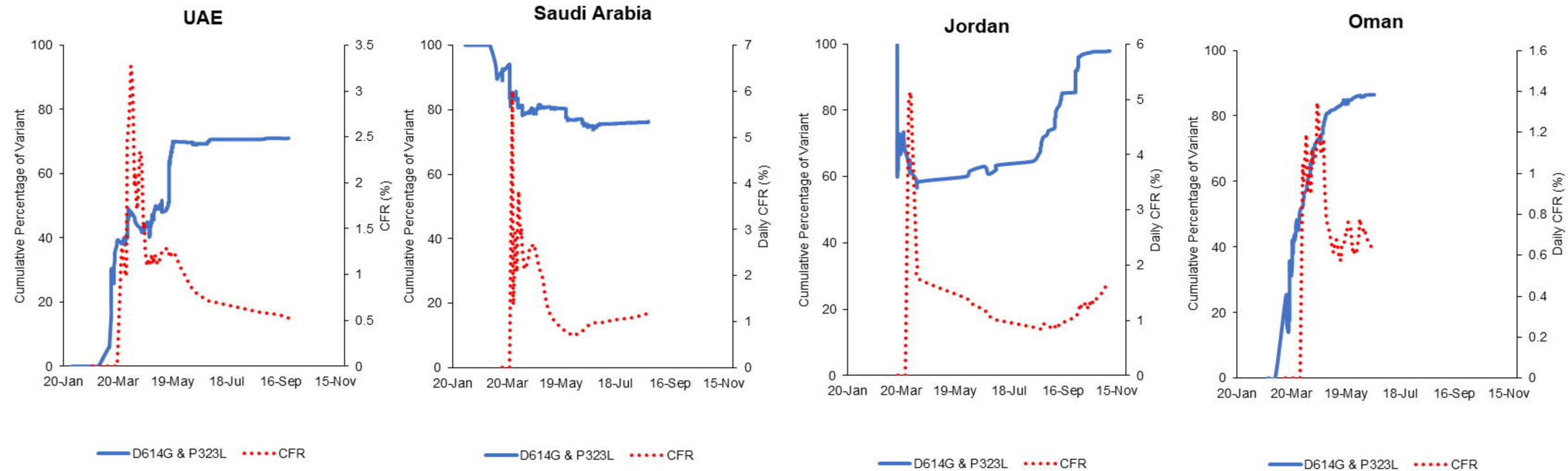

N_S202N
Percentage of Variant
CFR (%)
UAE
R = -0.527
P = 0.225
Jordan
Saudi Arabia
Oman
Bahrain
Palestine
Morocco


A
Wuhan-like
D614G only
P323L only
D614G & P323L
Percentage of Variant
100
80
60
40
20
0
UAE
Saudi Arabia
Jordan
Oman
Egypt
Bahrain
Palestine
Morocco

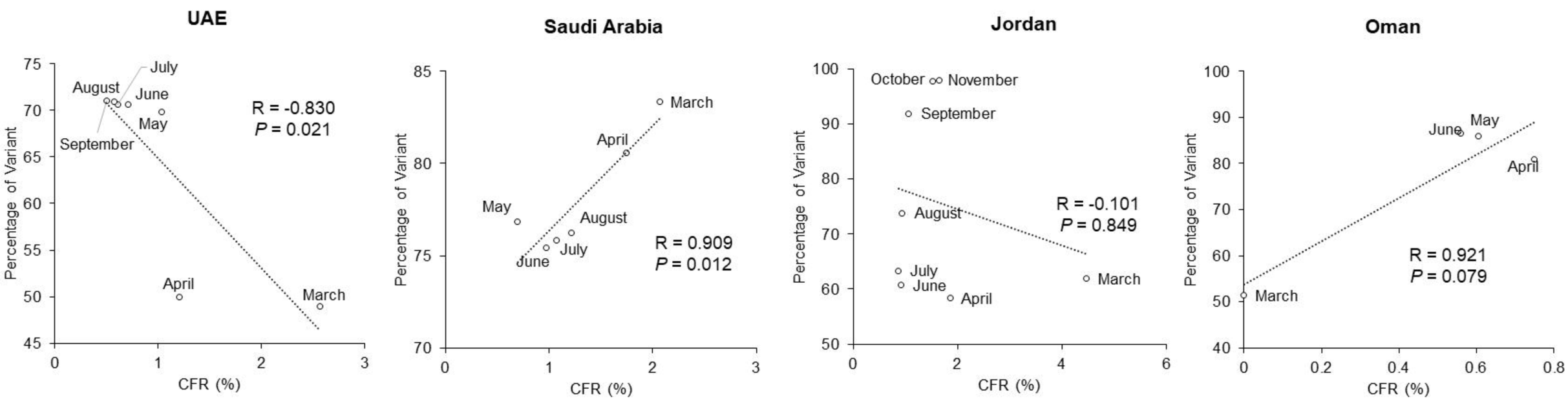

B
D614G & P323L
Percentage of Variant
0
20
40
60
80
100
Age Groups
0-9
10-18
19-49
50-59
>60


C
D614G only
Percentage of Variant
CFR (%)
R = -0.429
P = 0.337
Oman
Bahrain
Saudi Arabia
UAE
Palestine
Jordan
Morocco
P323L only
R = 0.460
P = 0.299
Saudi Arabia
Palestine
Bahrain
UAE
Oman
Jordan
Morocco
D614G & P323L
R = 0.610
P = 0.146
Morocco
Palestine
Jordan
Saudi Arabia
UAE
Oman
Bahrain

A


UAE
Saudi Arabia
Jordan
Oman
Cumulative Percentage of Variant
CFR (%)
Daily CFR (%)
20-Jan
20-Mar
19-May
18-Jul
16-Sep
15-Nov
D614G & P323L
CFR

B


UAE
R = -0.830
P = 0.021
July
August
June
May
September
April
March
Saudi Arabia
R = 0.909
P = 0.012
March
April
May
August
July
June
Jordan
R = -0.101
P = 0.849
October
November
September
August
July
June
April
March
Oman
R = 0.921
P = 0.079
June
May
April
March
Percentage of Variant
CFR (%)