Genetic drift and environmental spreading dynamics of COVID-19

R. Di Pietro, doctor [1], M. Basile, PhD [1], L. Antolini, PhD [2] and S. Alberti, doctor [3 *]

[1] Department of Medicine and Aging Sciences, Section of Biomorphology, G. d'Annunzio University of Chieti-Pescara, Italy.

[2] Center for Biostatistics, Department of Clinical Medicine, Prevention and Biotechnology, University of Milano-Bicocca, Monza, Italy.

[3] Unit of Medical Genetics, Department of Biomedical Sciences - BIOMORF, University of Messina, via Consolare Valeria, Messina, Italy.

[*] Address reprint requests to Dr. Alberti at the Unit of Medical Genetics, Department of Biomedical Sciences - BIOMORF, University of Messina, via Consolare Valeria, Messina, Italy, or at salberti@unime.it.

Drs. Di Pietro and Basile contributed equally to this article.

Abstract

Objective

To delineate the genetic and environmental determinants of COVID-19 spreading.

Design

Retrospective case series.

Setting

Spain, Italy, Sweden, Finland, Norway.

Participants

All laboratory-confirmed infection cases (n=168,089) collected from February 21st to April 14th 2020.

Main outcome measures

Infection spreading velocity according to viral mutation load and to climate region.

Results

The mean doubling time of COVID-19 was 6.63 days in northern Italy, 5.87 days in central areas, and 5.38 days in southern Italy, with shorter COVID-19 doubling time in warmer regions. Spain extended this trend, with a mean COVID-19 doubling time of 4.2 days. At the other end of the spectrum, slower diffusion across progressively colder regions was observed in Scandinavia, with

9.4 days COVID-19 doubling time in Sweden, 10.8 days in Finland and 12.95 days in Norway. Mutations and mutation rates of SARS-CoV-2 versus COVID-19 spreading were analyzed worldwide. Models of increased aggressiveness of SARS-CoV-2 upon progressive acquisition of genetic changes were not supported by regional mutation data.

Conclusion

Current propagation models suggest dependence of COVID-19 pandemic spreading on wintertime conditions, with expected waning over the summer. Our findings indicate association of COVID-19 to a sharp North/South climate gradient, with faster spreading in southern regions. Thus, warmer climate conditions may not limit SARS-CoV-2 diffusion. Very cold regions may be better spared by recurrent courses of infection.

Introduction

A first study in China on 425 cases identified initial transmission dynamics of Severe Acute Respiratory Syndrome (SARS) in 2019 (COVID-19) [1]. In its early stages, the epidemic doubled in size every 6.4 [2] to 7.4 [1] days, with a reproductive number (R_0) of infectious cases from 2.2 [1] to 2.7 [2]. Later studies described how the disease spread to Singapore [3], then to Germany [4], France and Finland (www.ecdc.europa.eu/en/covid-19-pandemic) [5-7].

One of the biggest unknowns about the epidemic caused by the newly-discovered coronavirus (SARS-CoV-2) was how it would spread around the world, and how environmental factors may shape the pandemic [8] (papers.ssrn.com/sol3/papers.cfm?abstract_id=3550308; ssrn.com/abstract=3556998; www.medrxiv.org/content/10.1101/2020.02.22.20025791v1). If SARS-CoV-2 is sensitive to environmental factors (papers.ssrn.com/sol3/papers.cfm?abstract_id=3550308; ssrn.com/abstract=3556998; www.medrxiv.org/content/10.1101/2020.02.22.20025791v1), this may provide means for optimizing COVID-19 mitigation strategies.

Current propagation models predicted a limited impact of COVID-19 in the Southern hemisphere during seasons that were infection-prone in the Northern hemisphere (papers.ssrn.com/sol3/papers.cfm?abstract_id=3550308; ssrn.com/abstract=3556998). However, early foci of infection were detected in Australia and New Zealand (Figures 1, 2). Outbreaks were also revealed in South America and extended to Central America and Mexico. Further infection foci were revealed in Saudi Arabia and Africa, and extended to sub-Saharan countries (Table S1), questioning simple models of climate-dependent COVID-19 spreading.

Coronaviruses spread to some extent similarly to the influenza virus [8], through small droplets suspended in the air, suggesting sensitivity to environmental humidity and temperature conditions. A recent meta-analysis [9], though, indicated resilience of coronaviruses to the environment. In a comparison to SARS-CoV-1, SARS-CoV-2 remained viable in aerosols for hours, and persisted over solid surfaces, 72 hours on plastic, 48 hours on stainless steel and 24 hours on cardboard [10].

SARS-CoV-2 possesses a single-strand RNA genome [11, 12] and was soon found to acquire and maintain frequent genomic mutations (nextstrain.org/ncov/europe) (Figures 2, S1). Selective pressure may apply to SARS-CoV-2 genomic drifting, leading to higher-efficiency infective particles along propagation trajectories. This, together with stochastic initial seeding at distinct geographic areas, may mimic climate-dependent diffusion.

This led us to challenge a genetic versus environmental coronavirus infection model. A robust analysis of SARS-CoV-2 spreading determinants required high-information density [13, 14] per

area. Pure case incidence models depend on a variety of factors (global traveling, founder effect versus time from initial infection [2], population clustering in big cities, social dynamics, governmental policies, infectious ability of the virus [15, 16], lead-time bias on case confirmation [2], COVID-19 containment procedures). Among them, a major confounding factor in COVID-19 is the time of initial infection at any given place, which, everything else being equal, leads to vastly different absolute numbers of derived cases [2]. Hence, we focused our attention on the velocity of infection spreading. Infected cases doubling time depends on the R_0 of the virus and on patient viral load/disease stage/severity [1]. Hence, it provided a hard composite index of the epidemic spreading force.

Vastly diverse climatic regions around the CET longitude (15ºE), were severely exposed to infection. Spain and Italy were the countries with the highest initial incidence of COVID-19 in Europe (Figure 3). The heaviest initial casualties in Italy were suffered by Lombardy and Veneto, i.e. cold and humid areas during wintertime. Markedly warmer and drier climate conditions prevail in southern regions of the country (Figure 4). A further shift toward warmer/drier conditions occurs in Spain. Scandinavian countries appeared initially spared by the infection (Table S1) and provided a reference for cold winter temperatures, over a Sweden-Finland-Norway axis. We went on to challenge a climate-dependent coronavirus infection model, through the analysis of 86,498 infection cases in Italy, 64,095 in Spain and 17,496 in Scandinavia (github.com/pcm-dpc/COVID-19) (Supplemental Appendix). This model was then challenged against a coronavirus genetic drift-driven diffusion model, according to mutation trajectories in the analyzed areas.

Methods

Incidence data

Incidence data of laboratory-confirmed infection cases over time were collected in Spain (n=64,095), Italy (n=86,498) and Scandinavia (n=17,496) (github.com/pcm-dpc/COVID-19). Incidence data were collapsed into a global database, to explore case incidence over time, and health outcome measures across countries and country provinces, Disease severity was classified as (a) hospitalized cases, (b) intensive-care unit patients, (c) recovered cases, (d) deaths. Incidence scatter plots by region were linked to Köppen–Geiger climate classification (average 1980-2016)[17] and average temperatures at the times analyzed.

Data analysis strategy

As indicated, a major confounding factor in COVID-19 is the time of initial infection at any given place, which, everything else being equal, leads to corresponding different absolute numbers of

infectious cases [2]. Hence, we focused our analysis on quantification of velocity of infection spreading at any given location.

Incidence of infection cases over time was determined for:

Spain; data from RTVE - Ministry of Health; www.rtve.es/noticias/20200415/mapa-del-coronavirus-espana/2004681.shtml.

Italy; data on 86,498 infection cases from the Ministry of Health (github.com/pcm-dpc/COVID-19).

Sweden; data by the Public Health Agency of Sweden; https://www.folkhalsomyndigheten.se/smittskydd-beredskap/utbrott/aktuella-utbrott/covid-19/; counties presenting more than 50 infection cases.

Finland; data from the National Institute for Health and Welfare THL; https://thl.fi/en/web/thlfi-en.

Norway; data from the Norwegian Institute of Public Health; https://www.fhi.no/sv/smittsomme-sykdommer/corona/dags--og-ukerapporter/dags--og-ukerapporter-om-koronavirus/; counties presenting more than 50 infection cases.

Cumulative incidence was contrasted to calendar time, to provide scatter plots, for quantitative analysis of SARS-CoV-2 spreading dynamics. The fractional approach we followed, for detailed analysis of infection growth rates in over a hundred sites, was allowed by the global nature of the collected data.

SARS-CoV-2 mutation analysis

SARS-CoV-2 genomic RNA sequences and country-correlated data were obtained from https://nextstrain.org/ncov/global. Scatter plots were generated, by strings of acquired mutations over time and overall number of mutations per genome per chosen area. Phylogeny trees for compiled mutations strings were then obtained, according to mutant branch descriptors (https://nextstrain.org/ncov/europe?branchLabel=aa).

Statistical analysis

Graphs were obtained by plotting the incidence of COVID-19 patients at weekly time points. The cumulative incidence of COVID-19 diagnoses was contrasted to calendar time for each province in a scatter plot [13, 14]. These plots acted as a smoother, for insight into the trajectory of infection records.

A landmark date for total numbers of diagnoses was set according to case incidence shape in each dataset. From this date, the doubling time for cumulative number of diagnoses was calculated backward for each province as follows. Two dates were identified: the maximum date, at which the cumulative number of diagnoses were lower than a half of the cumulative number of diagnoses at

the landmark time, and the minimum date, with a cumulative number of diagnoses greater than a half of the cumulative number of diagnoses at the landmark date. The fraction of days from the minimum date to achieve half of the cumulative number of diagnoses at the landmark date were obtained by a linear assumption for the cumulative incidence between the two dates. The doubling time was obtained by summing up the fraction of days to the number of days elapsed from minimum to landmark dates.

Software

Stata software version 16 was used for data importing, manipulation and graphics (StataCorp. 2019. *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC).

Results

Our attention was first drawn to the Southern hemisphere. Simple propagation models predicted essential absence of COVID-19 diffusion, during seasons that were infection-prone in the Northern hemisphere. However, early foci of infection were detected in Australia and New Zealand (Figure 1*A*). South and Central America appeared initially spared. Assessment at later time points indicated, though, large-scale (≥30 infection cases) outbreaks in Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Perù, Uruguay, Venezuela. Parallel outbreaks were revealed in Costa Rica, Dominican Republic, Panama and Mexico (Figure 1*C*, Table S1).

Africa, Middle-East and the Arabian peninsula also appeared spared during the initial course of COVID-19 (Figure 1*A*). However, infection foci appeared soon in Saudi Arabia, a non-high-risk country by most standards. This was soon recognized as a risk for COVID-19 spreading [18], and Saudi Arabia suspended the Umrah pilgrimage to Mecca and Medina on March 4th. Additional cases were reported in United Arab Emirates, Bahrain, Kuwait, Oman. Infectious foci were revealed in other countries facing the Persian Gulf and the Gulf of Oman, such as Iraq, Iran, Afghanistan, Pakistan. Further outbreaks were recorded in continental Africa, i.e. in Algeria, Egypt, Burkina Faso, Senegal, Democratic Republic of the Congo, Cameroon, Co□te d'Ivoire, Ghana, Nigeria, South Africa (Table S1). This led us to challenge a climate-dependent COVID-19 diffusion model.

Genetic-drift driven diffusion

The SARS-CoV-2 genomic RNA was shown to progressively mutate over time (nextstrain.org/ncov/europe). To determine whether mutation strings correlated with diffusion at distinct geographic areas, SARS-CoV-2 genomic RNA sequences from different countries were obtained at sequential times. Scatter plots were then generated, for strings of acquired mutations

versus overall number of accumulated mutations versus time. Phylogeny trees for compiled mutations were then obtained, according to mutant clade descriptors (https://nextstrain.org/ncov/europe?branchLabel=aa).

Mutation numbers, frequencies and mutation strings were compared. Sequence mutation analysis revealed up to seven major branches of mutation rows, at sites of major diffusion after spreading from China (Figure S1, Australia). This analysis was extended to Europe (Figure S2), with focus on Spain (Figure S3), Italy (Figure S4) and Sweden (Figure S5). Accumulated mutation numbers were plotted over the major mutation strings branches. The Netherlands, Belgium and France were utilized as benchmarks (Figures S6-8).

A mutation string-driven aggressiveness of SARS-CoV-2 spreading was expected to lead to (a) correlation of specific strings with highest-hit countries, (b) a late predominance of one/few dominant strings over the course of COVID-19 and (c) increase in disease severity over time. However, top-scores of accumulated mutations were revealed in SARS-CoV-2 in Wales and Senegal isolates (Figure 2). Further, the lowest number of accumulated mutations was recorded in Italy, the country with the strongest disease severity in Europe. The largest mutation load that was observed in Spain (n=14), the second hardest-hit country in Europe, appeared very close to that of Sweden (n=13), a country that was largely spared by the COVID-19 course in Europe. Large mutation loads in late-disease-insurgence countries, such as France and Belgium (n=16) suggested closer relationship with length of disease course. Four major mutation strings branches were revealed in all examined European countries, suggesting limited, if any, selective advantage of individual mutation strings versus the others. Early, high disease severity in Italy, versus a scanty SARS-CoV-2 mutation spectrum, further limited the support to the tested model.

COVID-19 doubling time by geographic area

COVID-19 doubling times by Countries, Regions and Provinces were computed as indicated. Landmark dates were set for each analyzed geographic area. From this date the time for doubling the cumulative number of diagnoses was calculated backward for each province as described.

Incidence of infection cases over time was determined for:

Italy: data on 86,498 infection cases were collected from March 3rd to March 27th (Supplemental Appendix). The global nature of disease data collection allowed to separately determine COVID-19 disease trajectories for all 105 Italian provinces. Bergamo (n=8,060), Milano (n=7,469) and Brescia (n=7,305) showed the highest incidence. Together with other high-incidence provinces, e.g. Cremona, Lodi, Parma, Pavia, Piacenza, Rimini, they showed an essentially linear increase of cases over time. An exponential growth component was detected at early time-points in other provinces,

such as Bari, Bologna, Catania, Como, Firenze, Forlì-Cesena, Novara, Padova, Perugia, Pescara, Rome, Torino. Grouping of province disease cases and disease parameters was then performed by Regions (n=22) (Figures 4, S9-10). Prevention of mass traveling after the 10th of March, and a total ban of non-commercial vehicle travel since March 22nd (Supplemental Appendix) were expected to reduce spread-estimates, with projected impact at ≈2 weeks afterwards. The latter dates land upon the final days of the period under study; some slow-down of COVID-19 was indeed revealed in multiple provinces after the 25th of March (Supplemental Appendix).

Spain: global data were obtained from February 25th to March 27th 2020 (n=64,095) (Figure S11).

Norway: data (>50 infection cases) were obtained from February 21st to April 14th 2020 (n=6,676) Troms og Finnmark, Oslo, Viken, Vestland, Rogaland, Agder, Innlandet, Møre og Romsdal, Vestfold og Telemark, Trøndelag and Nordland (Figure S12).

Finland: global data were obtained from March 1st to April 7th 2020 (n=2,646) (Figure S13).

Sweden: data (>50 infection cases) were obtained from February 26th to April 9th 2020 (n=8,995) from Dalarna, Gävleborg, Halland, Jönköping, Skåne, Stockholm, Södermanland, Uppsala, Västerbotten, Västmanland, Västra Götaland, Örebro, Östergötland (Figure S14).

Summary doubling times as grouped by geographical regions are shown in Figure 4. The average doubling time for northern Italy was 6.63 (SD=1.94) days; 5.87 (SD=1.08) days in central regions; 5.38 (SD=2.31) days in southern areas, with significantly shorter doubling time in southern regions versus northern areas (P=0.02) (Table S2, Figures S9-10). The mean COVID-19 doubling-time for the whole country was 6.06 (SD=1.95) days (Table S2).

With a doubling time of 4.2-days, Spain extended such a direction (Figure S11). At the opposite end of the climate spectrum, Scandinavia showed longer COVID-19 doubling times, over a Sweden-Finland-Norway axis. Sweden showed a doubling time of 9.4 days (SD=1.2) (P<0.0001 versus northern Italy), Finland one of 10.8 days, while Norway reached 12.95 days (SD=0.52) (P<0.0001 versus northern Italy) (Table S3, Figures S12-14), for a North-South gradient of COVID-19 spreading (Anova P<0.0001) (Figure 5).

Discussion

In this article, the spreading dynamics of COVID-19 versus environmental determinants and genetic drift were explored. A mutation-string-driven aggressiveness of SARS-CoV-2 spreading was expected to associate with overall disease impact at specific sites. It was also expected to associate with increased disease severity over time. A first clash with this model was provided by the most numerous accumulated mutations in SARS-CoV-2 isolates in Wales and Senegal, which were not listed among the hardest-hit areas. A second inconsistency was highlighted by evidence that Italy,

the country with maximum disease severity in Europe, had the lowest number of accumulated mutations in viral isolates. The largest mutation load of SARS-CoV-2 recorded in Spain, the second hardest-hit country in Europe, was very close to that of Sweden, a country with a least severe COVID-19 course in Europe. Large mutational loads in late disease-occurrence countries, such as France and Belgium, suggested relationship of increased mutation complexity with extended disease course. Virus mutant isolates falling in four major mutation-strings branches were revealed in all the countries that were examined, suggesting limited, if any, selective advantage of any one of the mutation strings versus the others. Early disease severity in Italy contrasted with a scanty SARS-CoV-2 mutation spectrum, further diminishing the strength of the proposed model.

We then went on to challenge a climate-dependent coronavirus infection model on 168,089 laboratory-confirmed infection cases, as detected from February 21st to April 14th 2020 in continental Europe. Data on 86,498 infection cases were collected from March 3rd to March 27th infection cases, over 105 provinces in Italy. In performing our analysis, we took advantage of the Italian National Health System (WHO, www.who.int/whr/en/; Bloomberg Global Health Index, 2018, www.bloomberg.com/), which operated as the sole central source of data. Province findings were then consolidated into 22 Regions. The region/province-level analysis tamed the confounding factor of non-homogeneous population density/infection clustering at urban versus countryside areas. As a consequence, infection case curves were expected to distinctly depend on (a) virus infective capacity (R_0) [1, 19] and (b) infection clinical severity, patient viral load, disease stage 15, 16. Either one or both of these factors were expected to be modulated by environmental conditions, had this been the case.

The average COVID-19 doubling time in the North of Italy was 6.63 days, versus 5.87 days in central regions, and 5.38 days in the South. Spain extended such a trend, with a 4.2-day doubling time. Scandinavia showed longer doubling times of 9.4 days in Sweden, 10.8 days in Finland, and 12.95 days in Norway. Taken together, our findings indicate that COVID-19 spreading associates to a sharp North-South gradient.

The global outbreak of the novel coronavirus SARS-CoV-2 has drawn comparison to the 1918 influenza pandemic, leading people to wonder what the Summer 2020 might bring. It has been suggested that SARS-CoV-2, just like flu viruses, is not easily transmitted in hot and dry conditions (www.scientificamerican.com/article/what-could-warming-mean-for-pathogens-like-coronavirus/). However, coronaviruses show resilience to environmental conditions [9], and SARS-CoV-2 was shown to remain viable over solid surfaces for up to 72 hours [10]. Our findings suggest that high environmental temperatures may not efficiently tame SARS-CoV-2 infectiousness. Consistent, the Middle East Respiratory Syndrome (MERS) was first reported in Saudi Arabia in

2012 (www.cdc.gov/coronavirus/mers). MERS is caused by the MERS-CoV, which is structurally and genetically related to SARS-CoV. MERS is endemic in the Arabic Peninsula, indicating resilience of the coronavirus to desert climate conditions (https://www.cdc.gov/coronavirus/mers/risk.html). Experimental testing of SARS-CoV-2 resistance to hot and dry conditions appears needed. This may have implications for public-health policies worldwide. This may also bear to models of long-term state of infection in temperate countries [20]. Very cold regions may be better spared by recurrent courses of infection.

Acknowledgments

We are much indebted to all the information curators we cite, and to the website providers the article data and graphic primers have been downloaded from. No funding was provided to this work.

Role in the article

All authors contributed to literature search, figures, study design, data collection, data analysis, data interpretation. S.A and R.DiP. wrote the manuscript draft. All authors contributed to discussing and writing the final text. R.DiP. and M.B. contributed equally to this work.

References

1. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. New England Journal of Medicine. 2020;382(13):1199-207. doi: 10.1056/NEJMoa2001316.
2. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet. 2020;395(10225):689-97. doi: 10.1016/S0140-6736(20)30260-9.
3. Pung R, Chiew CJ, Young BE, Chin S, Chen MIC, Clapham HE, et al. Investigation of three clusters of COVID-19 in Singapore: implications for surveillance and response measures. The Lancet. 2020;395(10229):1039-46. doi: 10.1016/S0140-6736(20)30528-6.
4. Hoehl S, Rabenau H, Berger A, Kortenbusch M, Cinatl J, Bojkova D, et al. Evidence of SARS-CoV-2 Infection in Returning Travelers from Wuhan, China. New England Journal of Medicine. 2020;382(13):1278-80. doi: 10.1056/NEJMc2001899.
5. Sun J, He W-T, Wang L, Lai A, Ji X, Zhai X, et al. COVID-19: Epidemiology, Evolution, and Cross-Disciplinary Perspectives. Trends in Molecular Medicine. 2020. doi: https://doi.org/10.1016/j.molmed.2020.02.008.
6. McMichael TM, Currie DW, Clark S, Pogosjans S, Kay M, Schwartz NG, et al. Epidemiology of Covid-19 in a Long-Term Care Facility in King County, Washington. New England Journal of Medicine. 2020. doi: 10.1056/NEJMoa2005412.
7. Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 — Studies Needed. New England Journal of Medicine. 2020;382(13):1194-6. doi: 10.1056/NEJMp2002125.
8. Sundell N, Andersson L-M, Brittain-Long R, Lindh M, Westin J. A four year seasonal survey of the relationship between outdoor climate and epidemiology of viral respiratory tract infections in a temperate climate. Journal of Clinical Virology. 2016;84:59-63. doi: https://doi.org/10.1016/j.jcv.2016.10.005.
9. Kampf G, Todt D, Pfaender S, Steinmann E. Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. Journal of Hospital Infection. 2020;104(3):246-51. doi: https://doi.org/10.1016/j.jhin.2020.01.022.
10. van Doremalen N, Bushmaker T, Morris DH, Holbrook MG, Gamble A, Williamson BN, et al. Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1. New England Journal of Medicine. 2020. doi: 10.1056/NEJMc2004973.

11. Egloff M-P, Ferron F, Campanacci V, Longhi S, Rancurel C, Dutartre H, et al. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. P Natl Acad Sci USA. 2004;101(11):3792-6. Epub 2004/03/08. doi: 10.1073/pnas.0307877101. PubMed PMID: 15007178.
12. Mousavizadeh L, Ghasemi S. Genotype and phenotype of COVID-19: Their roles in pathogenesis. Journal of Microbiology, Immunology and Infection. 2020. doi: https://doi.org/10.1016/j.jmii.2020.03.022.
13. Ambrogi F, Biganzoli E, Querzoli P, Ferretti S, Boracchi P, Alberti S, et al. Molecular subtyping of breast cancer from traditional tumor marker profiles using parallel clustering methods. Clin Cancer Res. 2006;12(3 Pt 1):781-90. PubMed PMID: 16467089.
14. Cimoli G, Malacarne D, Ponassi R, Valenti M, Alberti S, Parodi S. Meta-analysis of the role of p53 status in isogenic systems tested for sensitivity to cytotoxic antineoplastic drugs. Biochim Biophys Acta. 2004;1705(2):103-20. Epub 2004/12/14. doi: S0304-419X(04)00064-2 pii 10.1016/j.bbcan.2004.10.001. PubMed PMID: 15588765.
15. To KK-W, Tsang OT-Y, Leung W-S, Tam AR, Wu T-C, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. The Lancet Infectious Diseases. 2020. doi: https://doi.org/10.1016/S1473-3099(20)30196-1.
16. Chen Y, Li L. SARS-CoV-2: virus dynamics and host response. The Lancet Infectious Diseases. 2020. doi: 10.1016/S1473-3099(20)30235-8.
17. Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific data. 2018;5:180214-. doi: 10.1038/sdata.2018.214. PubMed PMID: 30375988.
18. Ebrahim SH, Memish ZA. COVID-19: preparing for superspreader potential among Umrah pilgrims to Saudi Arabia. The Lancet. 2020;395(10227):e48. doi: 10.1016/S0140-6736(20)30466-9.
19. Liu Y, Eggo RM, Kucharski AJ. Secondary attack rate and superspreading events for SARS-CoV-2. The Lancet. 2020;395(10227):e47. doi: 10.1016/S0140-6736(20)30462-1.
20. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science. 2020:eabb5793. doi: 10.1126/science.abb5793.

Figure legends

Figure 1. Worldwide progression of COVID-19.

(***A***) COVID-19 case incidence, as of 21st 2020; numbers are indicated by color code and circle diameter (www.healthmap.org/covid-19/).

(***B***) COVID-19 case incidence, as in (***A***), zoomed over Central Europe.

(***C***) COVID-19 incidence of active cases, as of March 31st, 2020; numbers are indicated by circle diameter (Johns Hopkins University, JHU; coronavirus.jhu.edu/map.html).

(***D***) Coronavirus spreading around the world as of April 4th. Confirmed cases by country since February 24th (JHU, public.flourish.studio/visualisation/1694807/).

Figure 2. COVID-19 diffusion and SARS-CoV-2 mutations.

(*A*) SARS-CoV-2 virus spread according to mutation load.

(*B*) Radial diagram of SARS-CoV-2 mutations worldwide. Concentric circles correspond to the number of genomic mutations detected in individual virus isolates.

Figure 3. Geographic areas versus climate zones.

(*top*) Worldwide temperature averages (pseudocolor) versus geographic areas.

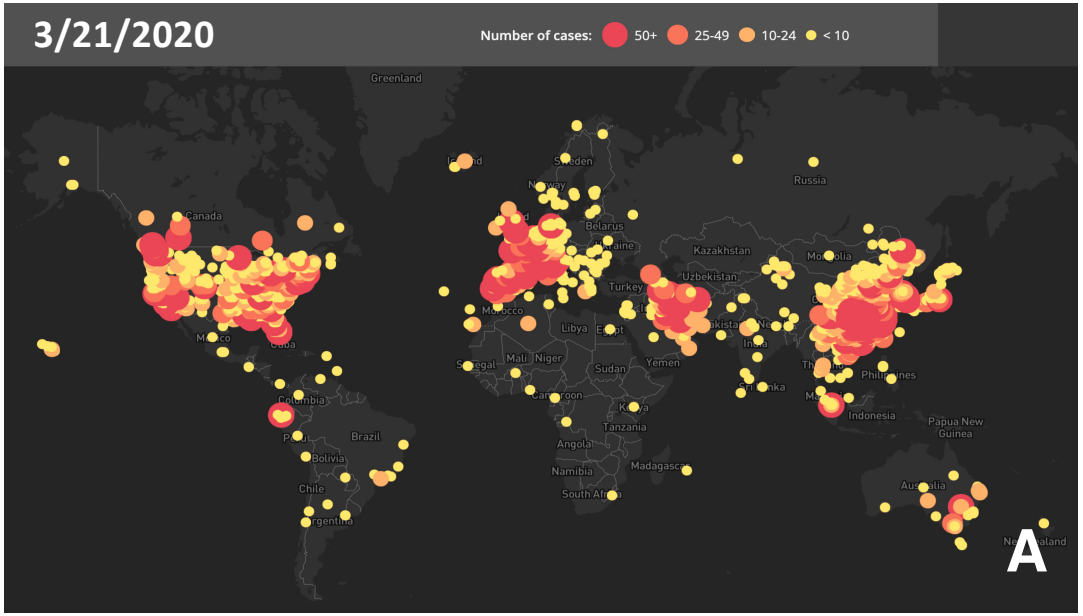
(*bottom*) Köppen–Geiger climate classification map for Europe (average 1980-2016).

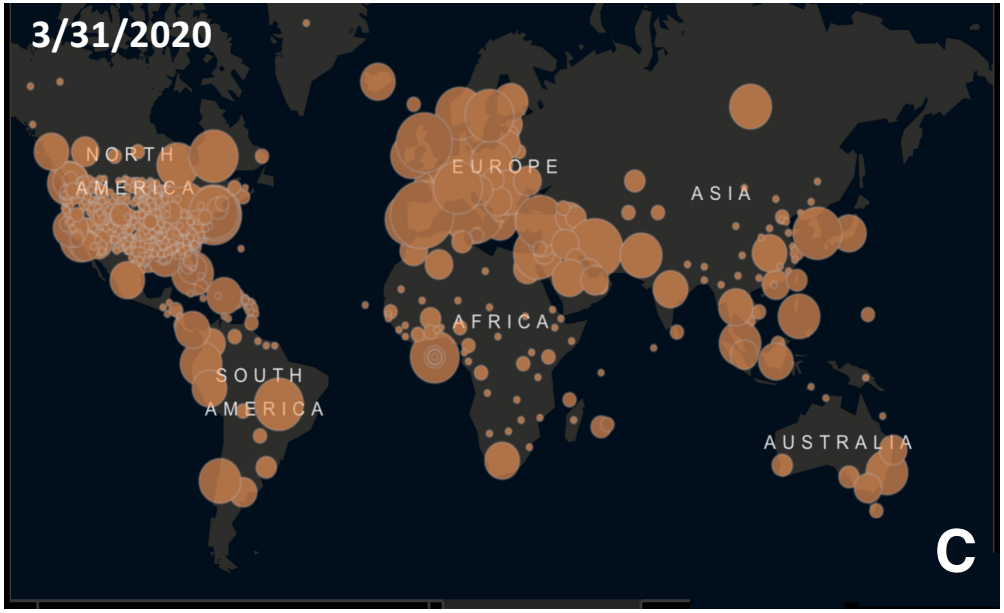
Figure 4. Epidemiology of COVID-19 versus North-South climate gradient.

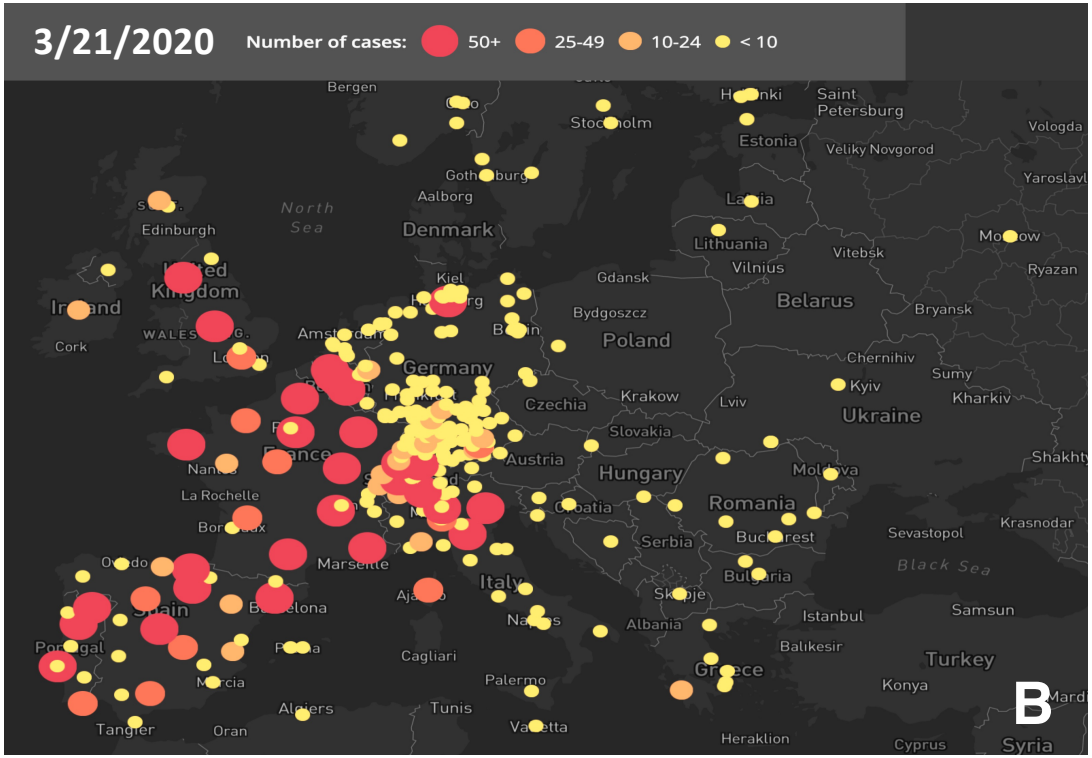
Köppen–Geiger climate classification map for Italy (average 1980-2016). Compounded data of COVID-19 spreading in individual Italy's regions are overlaid over the country's climate areas. (*insets*) Dark gray dots: SARS-CoV-2-positive cases; brown dots: hospitalized cases; green dots: intensive-care unit cases; orange dots: recovered cases. (*bottom*) color codes for areas classification.

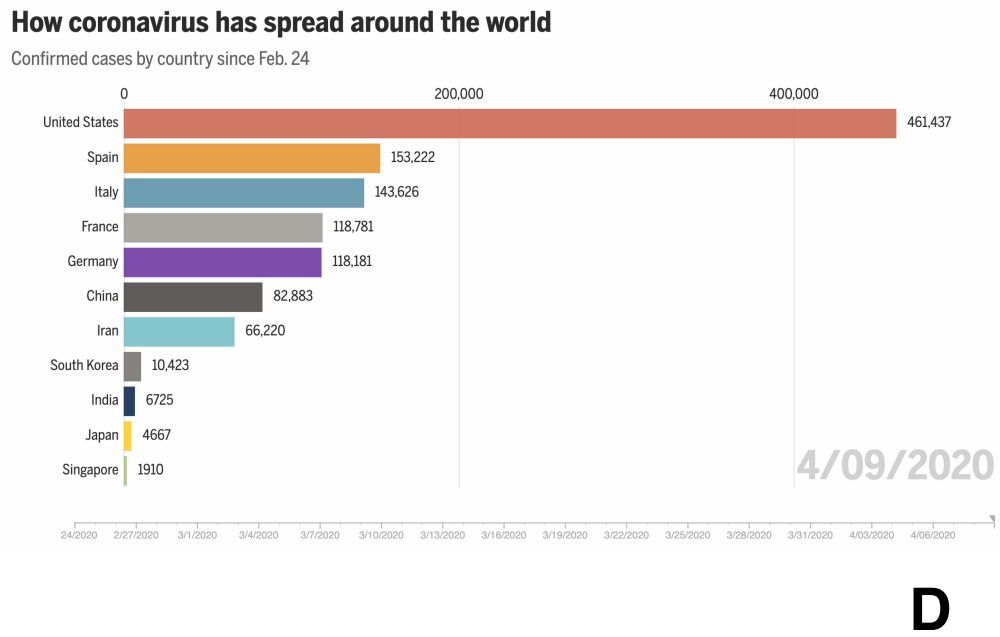
Figure 5. The COVID-19 spreading gradient across countries by climate zone.

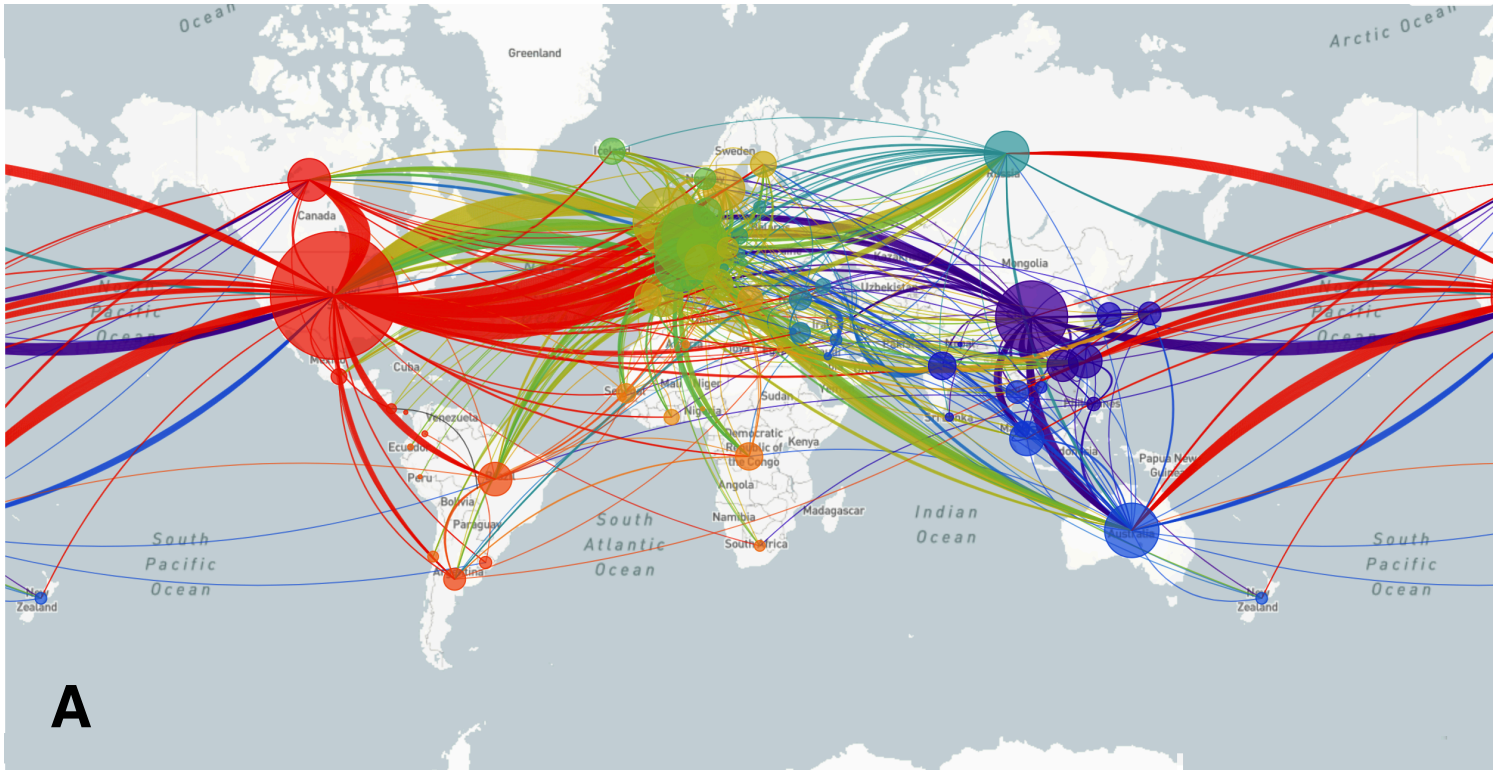
The COVID-19 North-South doubling-time gradient across countries by climate zone is depicted. The Anova P-value for trend of the curve is shown.

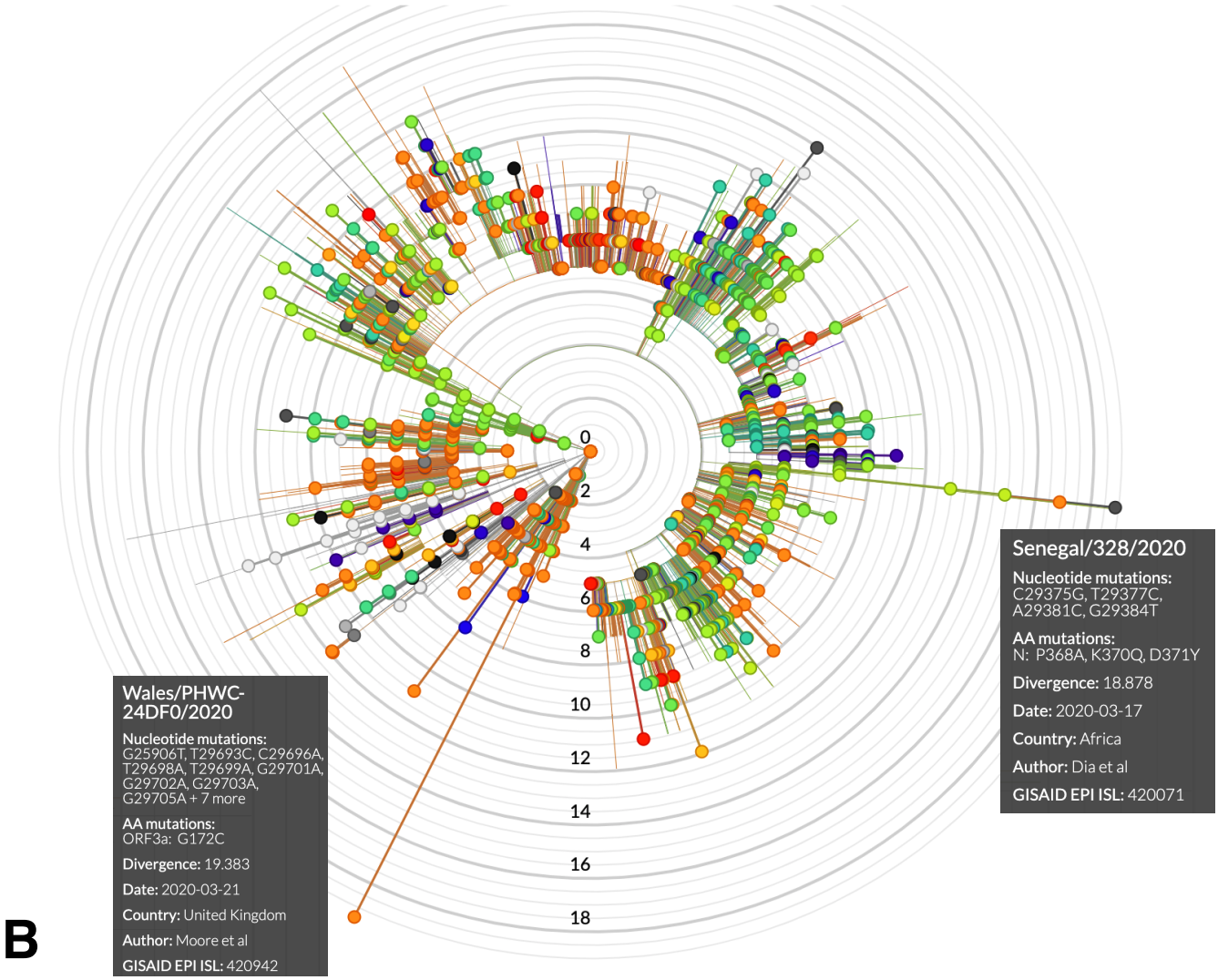

3/21/2020
Number of cases: 50+ 25-49 10-24 < 10
A

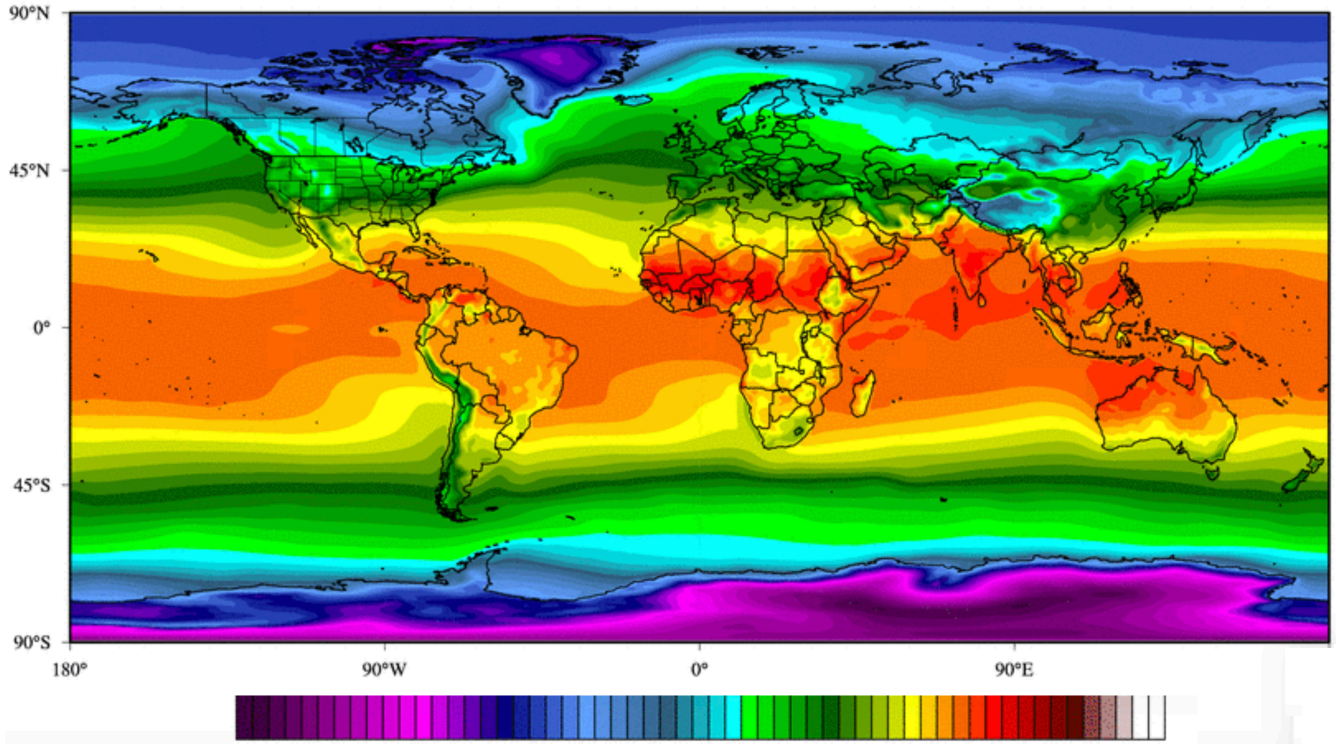

3/31/2020
NORTH AMERICA
EUROPE
ASIA
AFRICA
SOUTH AMERICA
AUSTRALIA
C


3/21/2020
Number of cases: 50+ 25-49 10-24 < 10
B

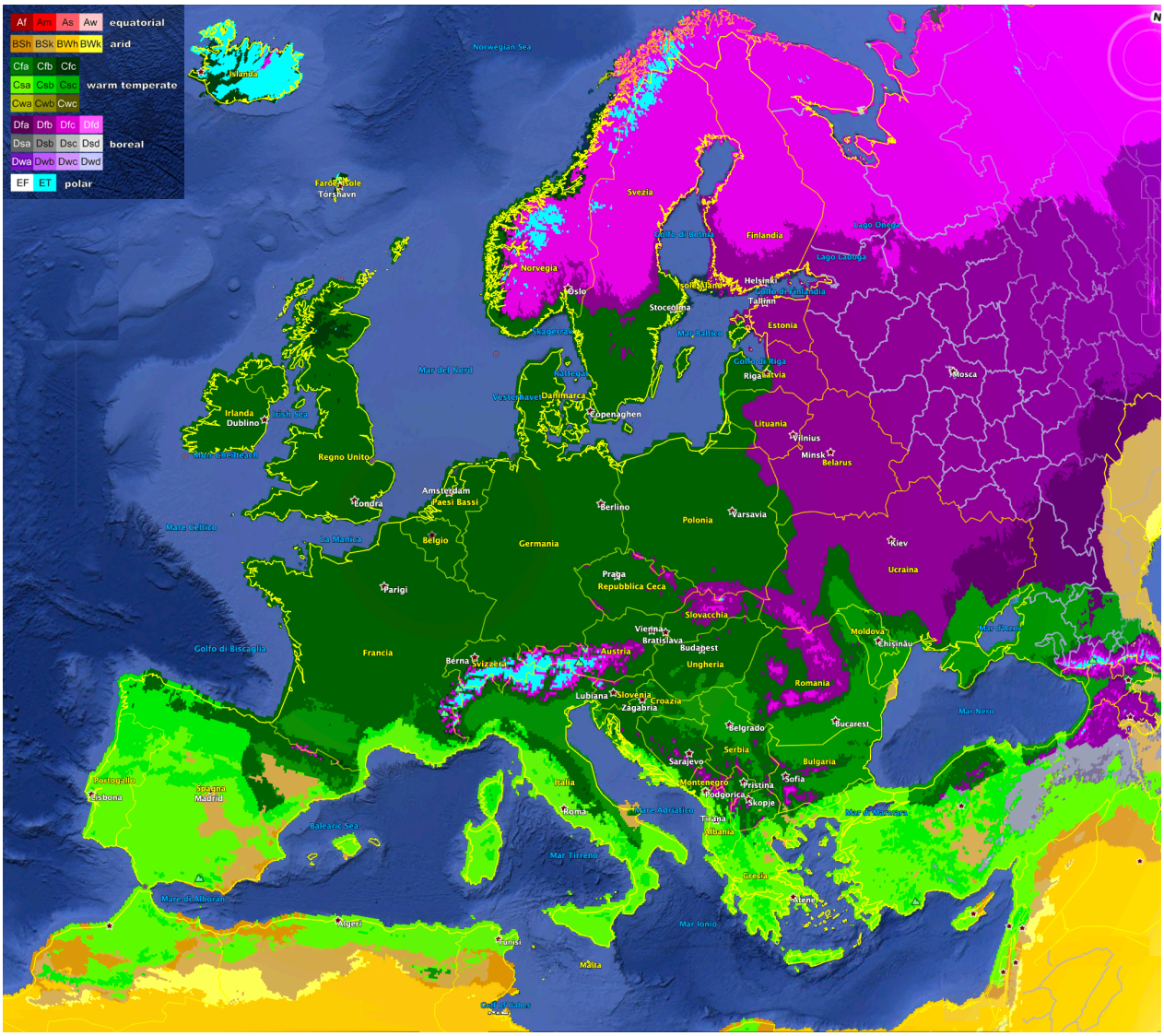

How coronavirus has spread around the world
Confirmed cases by country since Feb. 24
0
200,000
400,000
United States 461,437
Spain 153,222
Italy 143,626
France 118,781
Germany 118,181
China 82,883
Iran 66,220
South Korea 10,423
India 6725
Japan 4667
Singapore 1910
4/09/2020
24/2020
2/27/2020
3/1/2020
3/4/2020
3/7/2020
3/10/2020
3/13/2020
3/16/2020
3/19/2020
3/22/2020
3/25/2020
3/28/2020
3/31/2020
4/03/2020
4/06/2020
D


Greenland
Canada
Sweden
Mongolia
Cuba
Venezuela
Ecuador
Peru
Bolivia
Paraguay
Sudan
Kenya
Angola
Namibia
Madagascar
South Africa
Papua New Guinea
Australia
Indian Ocean
South Atlantic Ocean
South Pacific Ocean
Arctic Ocean
New Zealand
A

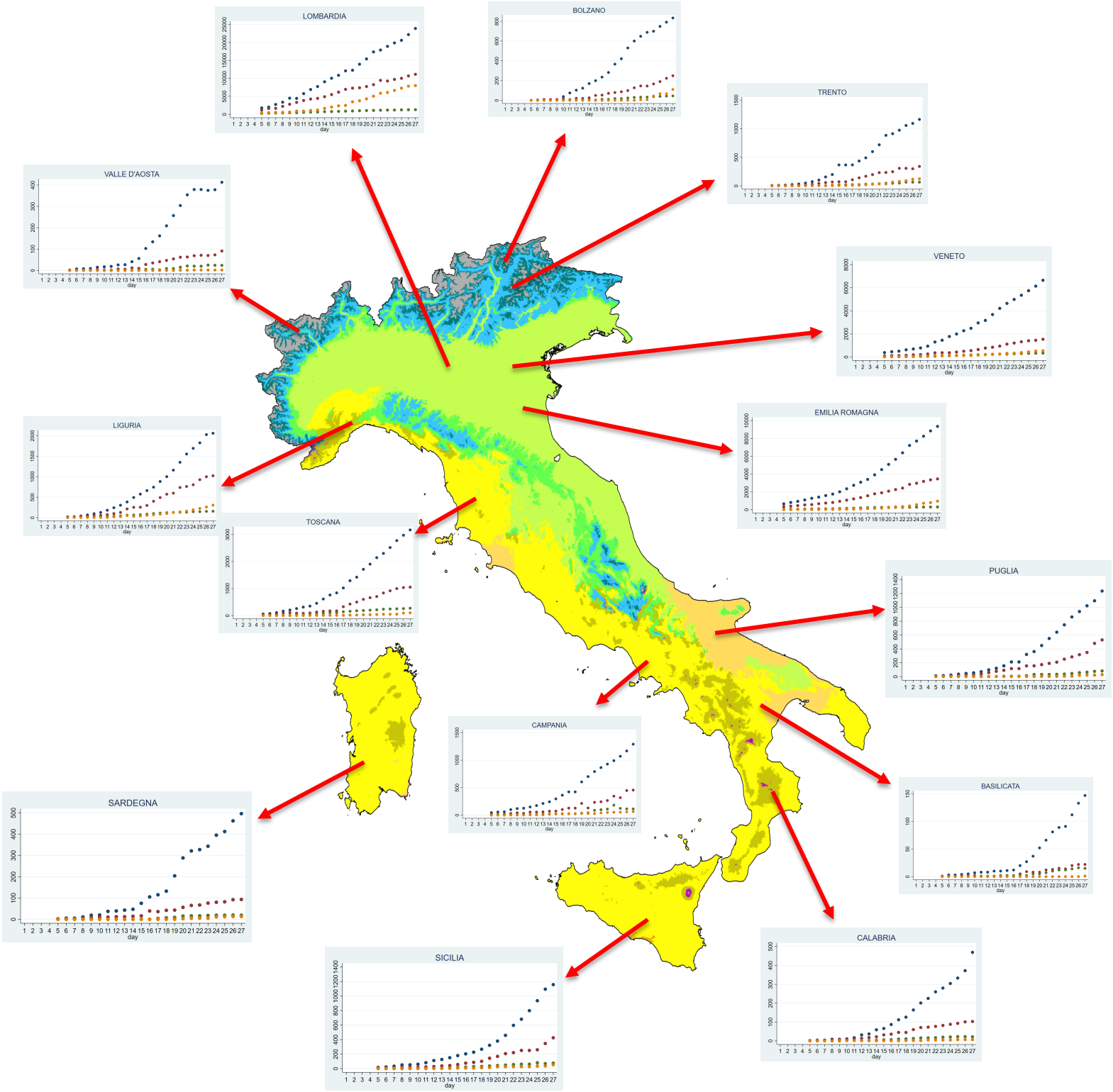

0
2
4
6
8
10
12
14
16
18
Wales/PHWC-24DF0/2020
Nucleotide mutations: G25906T, T29693C, C29696A, T29698A, T29699A, G29701A, G29702A, G29703A, G29705A + 7 more
AA mutations: ORF3a: G172C
Divergence: 19.383
Date: 2020-03-21
Country: United Kingdom
Author: Moore et al
GISAID EPI ISL: 420942
Senegal/328/2020
Nucleotide mutations: C29375G, T29377C, A29381C, G29384T
AA mutations: N: P368A, K370Q, D371Y
Divergence: 18.878
Date: 2020-03-17
Country: Africa
Author: Dia et al
GISAID EPI ISL: 420071
B


90°N
45°N
0°
45°S
90°S
180°
90°W
0°
90°E

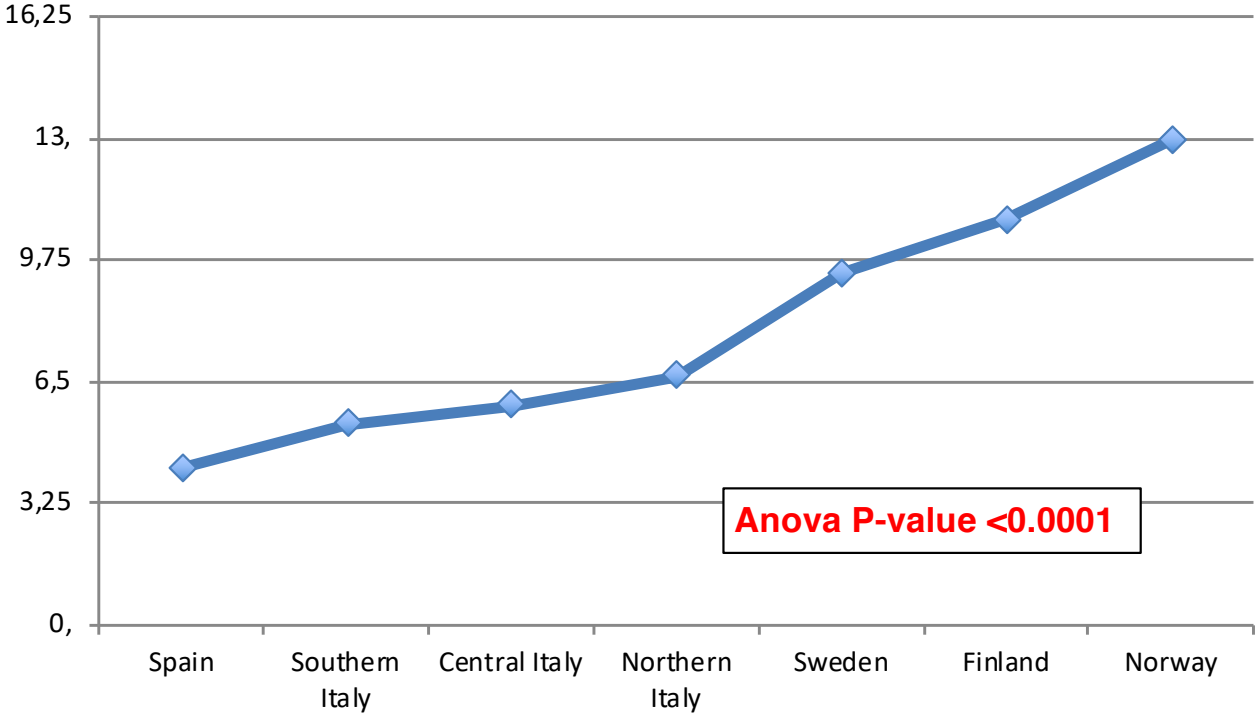
A


Af Am As Aw equatorial
BSh BSk BWh BWk arid
Cfa Cfb Cfc
Csa Csb Csc warm temperate
Cwa Cwb Cwc
Dfa Dfb Dfc Dfd
Dsa Dsb Dsc Dsd boreal
Dwa Dwb Dwc Dwd
EF ET polar
Islandia
Norwegian Sea
Torshavn
Svezia
Finlandia
Norvegia
Oslo
Helsinki
Tallinn
Estonia
Stoccolma
Riga
Lituania
Vilnius
Minsk
Belarus
Mosca
Mar del Nord
Danimarca
Copenaghen
Irlanda
Dublino
Regno Unito
Londra
Amsterdam
Paesi Bassi
Berlino
Polonia
Varsavia
Germania
Belgio
Parigi
Francia
Praga
Repubblica Ceca
Slovacchia
Vienna
Bratislava
Budapest
Ungheria
Austria
Berna
Svizzera
Lubiana
Slovenia
Zagabria
Croazia
Belgrado
Serbia
Sarajevo
Bucarest
Romania
Moldova
Chisinau
Kiev
Ucraina
Bulgaria
Sofia
Pristina
Skopje
Tirana
Albania
Italia
Roma
Lisbona
Portogallo
Madrid
Spagna
Algeri
Tunisi
Malta
Atene
Grecia

B



- Arid, desert, hot (BWh)
- Arid, desert, cold (BWk)
- Arid, steppe, hot (BSh)
- Arid, steppe, cold (BSk)
- Temperate, dry summer, hot summer (Csa)
- Temperate, dry summer, warm summer (Csb)
- Temperate, no dry season, hot summer (Cfa)
- Temperate, no dry season, warm summer (Cfb)
- Cold, dry summer, warm summer (Dsb)
- Cold, dry summer, cold summer (Dsc)
- Cold, no dry season, warm summer (Dfb)
- Cold, no dry season, cold summer (Dfc)
- Polar, tundra (ET)
- Polar, frost (EF)


16,25
13,
9,75
6,5
3,25
0,
Anova P-value <0.0001
Spain
Southern Italy
Central Italy
Northern Italy
Sweden
Finland
Norway