Evolutionary Studies Combined with Species Distribution Models are Useful to Understand COVID-19 Spread and Aid Vaccine Design

Prakash Chandra Arya[1], Priyanka Bajaj[2*]

[1]Centre for Earth Science, Indian Institute of Science, Bangalore, Karnataka, India; [2]Molecular Biophysics Unit, Indian Institute of Science, Bangalore, Karnataka, India

ABSTRACT

Different opinions on the use of Species Distribution Models (SDMs) to understand COVID-19 spread resulted in an ongoing debate in the scientific community. We agree SDMs alone cannot suffice the purpose of understanding COVID-19 spread. However, evolutionary studies combined with SDMs can overcome these limitations and, mutational insights gained can aid design of an ideal vaccine against SARS-CoV-2.

Keywords: Species Distribution Models (SDM); Phylogenetic studies; Vaccine design

DESCRIPTION

Nearly after a century of the Spanish flu pandemic, SARS-CoV-2 pathogen responsible for COVID-19 pandemic has appeared as a global threat as well as a challenge in front of the scientific community. Worldwide research is being conducted to prepare a vaccine against this deadly disease. The initial cases of COVID-19 were identified in Wuhan (China), at the end of 2019. Since then the disease has up surged and managed to reach in every part of the world. A well-known fact about COVID-19 spread, which is identified and discussed by several authors [1,2] is that COVID-19 has not spread simultaneously in all regions of the world. Chronologically the spread of the disease can be divided into four stages.

Stage 1: Initially in Wuhan and south east China.

Stage 2: In the regions with a similar climate to that of Wuhan (e.g., Europe, Washington, Florida, New South Whales, and Sau-Paulo)

Stage 3: Regions colder than Wuhan's climate (Russia and Canada) and

Stage 4: Regions warmer to Wuhan's climate (Brazil and India) [1].

This time lag is still not understood why such a pattern of spread was followed by SARS-CoV-2. COVID-19 spread controlled by climate has been suggested by several authors [1,2]. However, there is an ongoing debate on the use of Species Distribution Models (SDMs) to comprehend SARS-CoV-2 spread [1-3]. Some authors suggests the use of SDMs are inappropriate for COVID-19 [3]. While few authors believed the use of SDMs are useful for COVID-19 [2]. Although SDMs are very strong tools for modeling the distribution of the species across space and time, the question arises 'Can SDMs be used in case of COVID-19?' Insights about the functioning of the SDMs will provide its strengths and limitation.

Broadly, there are two types of SDMs, one "Correlative SDMs", in these, the observed species distribution is modelled as a function of environmental conditions by projecting distribution of species on various climate maps to understand where they can sustain [4].

Another type of SDMs are "Mechanistic SDMs" which are processed based models, that uses physiological information of species in controlled environments (laboratories) and determines the range of environmental conditions suitable for species [5]. Traditional correlative SDMs use only geolocation and climatic data that to some extent are insufficient to capture different transmission pathways of the pathogen [3]. Even the most advanced SDMs might be unable to distinguish between different pathways of the spread of a vector borne disease. For example in case of Zika virus, the disease can spread through sexual transmission, blood transfusion, or during delivery which can occur far beyond the associated niche of Aedes mosquito [4]. This is certainly true and is a fundamental challenge for the users of SDMs, to demarcate the different pathways of the spread, but such cases will be far lower than the actual vector transmission and these cases can be delimited if medical history and travel history are well recorded. Even in the absence of medical/travel history it is most likely that an SDM will statistically inclined towards the niche of Aedes mosquito but such uncertainty in the data cannot be ignored. These inherited characteristics of transmission limits the use of SDMs for a vector borne disease to some extent, but if the results of SDMs are integrated with other techniques such as phylogenetic analysis. The combined results will overcome the limitations of SDMs and provide a better picture to the user. For example, if one checks whether the genomic sequence of the virus varies in two different places, and if, there is a significant change in the genome sequence supported with enough statistical data, would be sufficient to say at least there might be a possibility that the virus is responding to

the external environment. Another limitation of SDMs is they can produce misleading risk maps with a very high accuracy.

For example anyone can access HIV database and can produce a risk map with a very high accuracy with a known methodology, showing Africa has a more suitable climate for HIV transmission and Arctic/Antarctica are too cold for the risk [4]. This is another major limitation of SDMs that produce highly misleading risk maps. Although settlement maps and population density maps could be of great help in discriminating the validity of the risk map. This problem can also be addressed by integrating phylogenetic analysis with SDMs. For example, if one checks whether the genome sequence of any virus shows similarities across similar and differences across different climate types all over the world, if there is a significant correlation or a consistent pattern with enough supporting statistical data, this would be sufficient to say at least there is a high possibility that virus is responding to its external environment. This kind of approach is far more superior to any of the presently available SDMs and overcomes the limitations of SDMs. Such highly infectious pathogens pose a global threat to civilizations as they manage to escape prior immunity and therapeutic interventions. Such a problem can be addressed by combining genomic information obtained from studying their evolution and subsequently integrating it with SDMs. This combination will enhance the results of SDMs, minimize its limitations/complications, and will provide a lucid understanding of the abiotic factors affecting the virus evolution and natural spreading pathways adopted by such infectious pathogens, thus displaying the role of natural selection on their evolution.

Certainly, the different transmission pathways of a vector borne disease to some extent outreaches the ability of SDMs to generate a true model. This hints the user to apply SDMs in a judicious way by integrating SDMs results with other studies to generate the results that enhance our understanding of the spread of the disease. At present it is unknown that geographical clustering of immune response exists, if it exists, it could be possibly due to climate [2]. In such a case, judicious use of SDMs could be useful for understanding COVID-19 spread which is identified by few authors and In such a case, judicious use of SDMs could be useful for understanding COVID-19 spread which is also identified by few authors [2]. Combing genomic studies with SDMs will enhance such clustering and can be easily modeled with higher accuracy than that of SDM alone. Studies involving only SDMs in some cases may be misleading in some cases where medical/travel histories are absent, and the number of patient cases are not correctly reported. Studies involving genomic sequencing of infectious pathogens in different climate types, tracing their evolution, and detecting signatures as they move into new regions are required. This can be uncovered by molecular detection assays and phylogenetic studies that complement metagenomic sequencing and help understand mechanistic relationships between climate variability and pathogen transmission and evolution. The virus isolates of SARS-CoV-2 pertaining to different climates are genomically diverse [1]. Vaccine development for rapidly mutating RNA viruses has always been an erroneous task due to which vaccines for majority of RNA viruses such as HIV, Zika, Nepa, Ebola, SARS-CoV, MERS-CoV are unavailable. This hints towards the need of a vaccine which will work effectively against all variants of SARS-CoV-2. Since, SDMs combined with phylogenetic studies provides information on tolerance to abiotic factors (temperature, humidity etc.) of evolved variants, such mutational insights can aid design of an efficacious vaccine against all strains of SARS-CoV-2 making it functional in the long run. Such a study can also be helpful in understanding the expected time for the second wave of COVID-19. Our understanding of, and responses of viruses to, such interactions and their impacts can be enhanced through an amalgam of both traditional and novel strategies. Finer-grained analysis of the evolutionary dynamics of the poorly understood SARS-CoV-2 genomes in different climates will aid manufacturers to design thermotolerant vaccines, enabling the efficient deployment of these vaccines at all temperatures prevailing on Earth. From animal and clinical studies carried out till now, since it seems 2 doses per person may be required, it will account for 16 billion doses throughout the globe [6]. Simply manufacturing an effective vaccine is not enough to ensure the long-term efficacy of the product, as vaccines spend considerable time at the country level storage facility, as well as being transported between warehouses, and all service points. Thus, countries require robust storage and transport system which can be a limiting factor for many countries. Despite expedited SARS-CoV-2 vaccines being made at various platforms, we still have no information about the temperature sensitivity, storage, and distribution conditions of various vaccine candidates. Limited research is being conducted in evaluation of the thermo tolerant properties of a vaccine candidate against SARS-CoV-2 [7]. Although supply infrastructure exists in many countries, the cool capacity (2°C-8°C) for storage and transportation is unlikely to be adequate for many countries. Depending on the temperature requirement, special boxes and temperature monitoring devices would be required to be able to transport vaccines internationally as well as within the country, and the existing systems in countries will not be able to accommodate this demand. Several researchers have separately studied the evolutionary aspect of SARS-CoV-2 and SDMs related to SARS-CoV-2. Integrating these studies will provide significant information related to virus response towards the external environment which will aid the design of a vaccine that can overcome these limitations, be transported at room temperature and reach the entire world population in a short time. Such interdisciplinary studies are not only useful for COVID-19 pandemic but will be useful for other future outbreaks as well.

ACKNOWLEDGMENT

REFERENCES

1. Bajaj P, Arya PC. Climatic-niche evolution of SARS CoV-2. bioRxiv. 2020.

2. Araújo MB, Mestre F, Naimi B. Ecological and epidemiological models are both useful for SARS-CoV-2. Nat Ecol Evol. 2020;4(9):1153-1154.

3. Carlson CJ, Chipperfield JD, Benito BM, Telford RJ, O'Hara RB. Species distribution models are inappropriate for COVID-19. Nat Ecol Evol. 2020;4(6):770-771.

4. Elith J, Leathwick JR. Species distribution models: ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst. 2009;40:677-697.

5. Kearney M, Porter W. Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. Ecol Lett. 2009;12(4):334-350.

6. World Population Data Sheet Shows Older Populations Growing, Total Fertility Rates Declining: Population Reference Bureau. 2020.

7. Malladi SK, Singh R, Pandey S, Gayathri S, Kanjo K, Ahmed S, et al. Design of a highly thermotolerant, immunogenic SARS-CoV-2 spike fragment. J Biol Chem. 2020.