Understanding COVID-19 spread in the National Capital Region, Philippines using Genomic Sequences: A Phylodynamic Investigation

IMARI JOY C. BORDA[1,2], **SHERYL GRACE BUENAVENTURA[1,2]**, ZYTHON PAUL LACHICA[1,2], IVY GRACE PANOGALINOG[2],HANNA LYKA BONTILAO[1,2], MAY ANNE MATA[1,2], RICARDO DEL ROSARIO[3]

[1] Center for Applied Modeling, Data Analytics, and Bioinformatics for Decision Support Systems, University of the Philippines Mindanao,Mintal, Tugbok District, Davao City, Philippines
[2] Department of Mathematics, Physics, and Computer Science, University of the Philippines Mindanao, Mintal, Tugbok District, Davao City, Philippines
[3] Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, United States of America

AMDABiDSS-Health

Introduction

- Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) evolves rapidly like all other disease-causing viruses. Therefore, to understand its real dynamics, it is ideal to account the virus' evolutionary history in estimating epidemiological parameters such as the reproductive number (R_0) and transmission or birth rate (λ).
- Phylodynamic models such as Birth-Death Susceptible-Infected - Removed (BDSIR) can simultaneously estimate evolutionary and epidemiological parameters using Bayesian inference (Figure 1)[1,2].
- BDSIR model was used by neighboring countries (e.g. Japan, Taiwan and South Korea) for analyzing the spread of COVID-19 (Figure 2) [3].
- The Philippines' first and second reported cases were confirmed on January 30, 2020 and February 2, 2020, respectively (Figure 3) [4].
- The earliest available SARS-CoV-2 whole genome sequence data uploaded in Global Initiative on Sharing All Influenza Data (GISAID) and National Center for Biotechnology Information (NCBI) were sampled in the National Capital Region (NCR), Philippines [5,6].
- Our study aimed to understand the spread of COVID-19 using phylodynamic analysis focused on NCR as a first case study in the Philippines using BDSIR model.

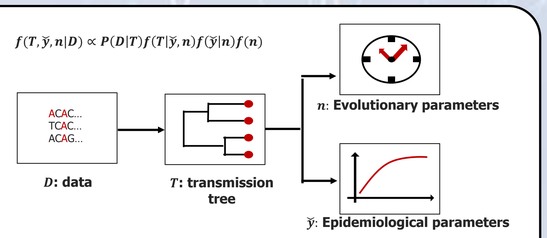

Figure 1. Schematic overview of phylodynamic model inference. The function f is the probability density while P is the probability mass.

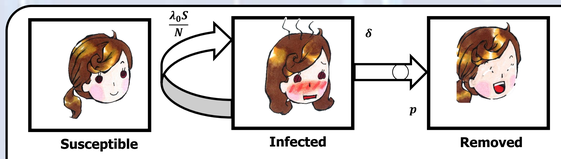

Figure 2. BDSIR model for COVID-19 transmission.

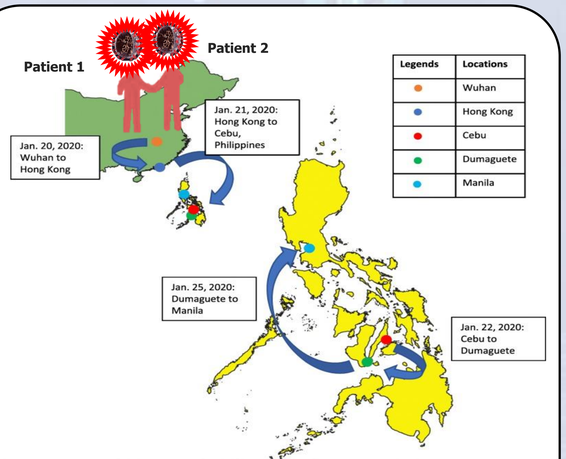

Figure 3. Travel history of Philippine's first recorded COVID-19 cases. Lifted from Edrada et al. (2020).

Methodology

DATA COLLECTION

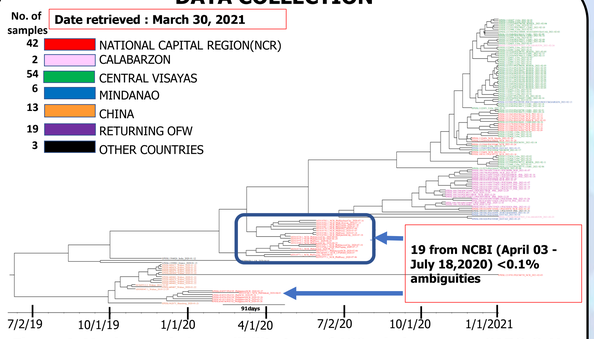

Figure 4. Maximum clade credibility tree of 139 whole genome SARS-CoV-2 sequences from the Philippines and other neighboring countries reconstructed using BEAST v1.10.4 (Sequence source: GISAID and NCBI).

BDSIR MODEL PARAMETERIZATION

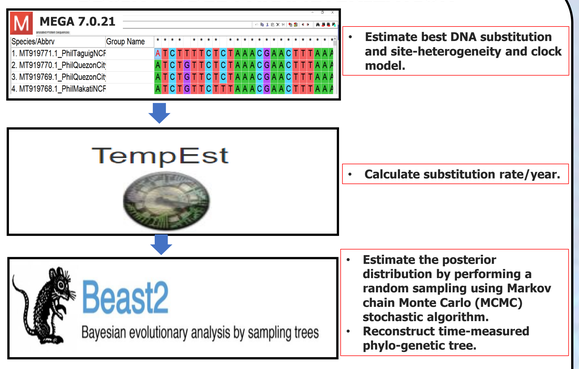

Figure 5. Workflow for phylodynamic analysis using BDSIR model.

BEAST 2 OUTPUT ANALYSIS USING TRACER

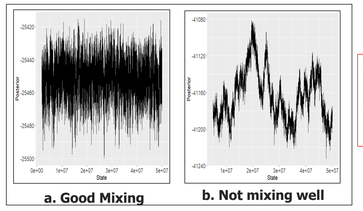

Figure 6. Comparison of posterior trace plots. Each plot shows the sampled values of posterior over time.

EXTRACTING PARAMETER ESTIMATES

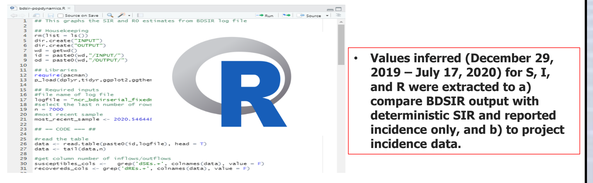

Figure 7. Extracting parameter estimates with R codes.

Results and Discussion

Parameter Estimates from BDSIR model and comparison with Deterministic SIR model

- The BDSIR estimated origin of the epidemic is between November 2, 2019 to February 24, 2020, with a median estimate of **29 December 2019**.
- The BDSIR model estimated the higher R_0 compared to the reported and projected R_0 (Table 1) indicating that the spread of COVID-19 in NCR was actually faster and cases were expected to exponentially grow through time.
- The estimated birth rate from BDSIR (2.42E-6) was also higher than the deterministic SIR estimate (3.15E-8) which means that the rate per unit time that the host passed on the disease to another host was underestimated using deterministic SIR approach.
- All the marginal density distributions of the key epidemiological parameters estimated in BDSIR model are right-skewed (Figure 8).

Table 1. Reproductive number estimates

BDSIR (using genomic data)	Reported	Deterministic SIR (using incidence data)
3.04	2.5 (maximum value reported)	1.22

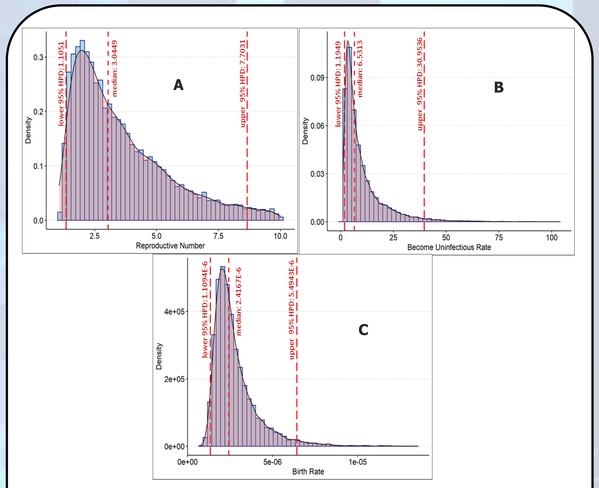

Figure 8. Marginal density distributions of (A) Reproductive Number (B) Become Uninfectious Rate and (C) Birth Rate using the BDSIR model. HPD: Highest Posterior Density

- The BDSIR estimated the highest number of COVID-19 infected individuals in NCR across time which suggests that there was underreporting of incidence data (Figure 9).

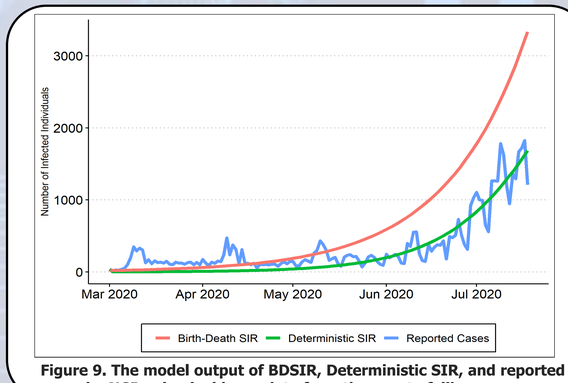

Figure 9. The model output of BDSIR, Deterministic SIR, and reported cases in NCR using incidence data from the onset of illness.

Results and Discussion

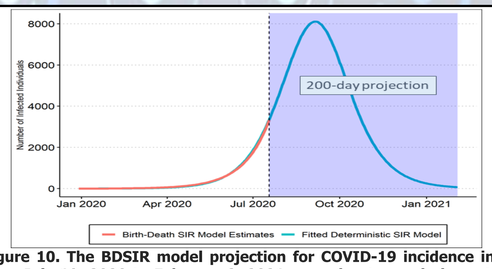

Figure 10. The BDSIR model projection for COVID-19 incidence in NCR from July 19, 2020 to February 2, 2021 assuming transmission rate and become uninfectious stays the same.

- COVID-19 cases in NCR from 19 July 2020 to 2 February 2021, 200 days after the latest sampling date of the genomic sequences, was projected using the incidence estimates from BDSIR fed to the deterministic SIR model (Figure 10).
- The peak of the incidences appears to occur by mid-September 2020 and died out by January 2021.

Conclusions and Recommendations

- Phylodynamic analysis using BDSIR model reconstructed realistic estimates of epidemiological parameters, such as R_0 and transmission rate, describing the actual COVID-19 dynamics in NCR using the 19 whole genome sequences. Genomic sequences provided information on the temporal dynamics of the virus capturing its evolution.
- The BDSIR model estimates suggest that most likely there could be an underreporting of epidemiological estimates using deterministic SIR and it also implied that COVID-19 at that point in time was spreading rapidly jumping from one host to another giving more opportunities for the virus to mutate hence the emergence of new variants of concern.
- The BDSIR projection of COVID-19 incidence data in NCR from July 19, 2020 to February 2, 2021 appears to have the same trend with the reported cases.
- This study suggests to integrate phylodynamic analysis in disease surveillance programs of the country to accurately monitor viral spread
- Future works: Explore other phylodynamic models such as coalescent SIR and SEIR.

Acknowledgement

Department of Science and Technology
Philippine Council for Health Research and Development
PGC PHILIPPINE GENOME CENTER

References

1. Kühnert, D., Wu, C. H., & Drummond, A. J. (2011). Phylogenetic and epidemic modeling of rapidly evolving infectious diseases. Infection, genetics and evolution, 11(8), 1825-1841.
2. Kühnert, D., Stadler, T., Vaughan, T. G., & Drummond, A. J. (2014). Simultaneous reconstruction of evolutionary history and epidemiological dynamics from viral sequences with the birth–death SIR model. Journal of the Royal Society Interface, 11(94), 20131106.
3. Edrada, E.M., Lopez, E.B., Villarama, J.B. et al.(2020). First COVID-19 infections in the Philippines: a case report. Trop Med Health 48, 21.
4. Choi SC. A phylodynamic analysis of epidemiological situation of East Asia due to the coronavirus disease of 2019Korean J. Microbiol. 2020;56:241-253.
5. Elbe, S., and Buckland-Merrett, G. (2017) Data, disease and diplomacy: GISAID's innovative contribution to global health. Global Challenges, 1:33-46.
6. Velasco, John Mark et al. "Coding-Complete Genome Sequences of 23 SARS-CoV-2 Samples from the Philippines." Microbiology resource announcements vol. 9,43 e01031-20.

Contact Information

Corresponding authors :
Imari Joy Borda icborda@up.edu.ph
Sheryl Grace Buenaventura scbuenaventura@up.edu.ph
Zython Paul Lachica ztlachica@up.edu.ph
Ricardo del Rosario rcdelros@broadinstiture.org
May Anne Mata memata@up.edu.ph
Phone: 0917 129 3986
Email: nicer.upmin@up.edu.ph
https://sites.google.com/up.edu.ph/nicer-upmin/home