BioNORAD: Fast Scalable Pandemic Risk Assessment of Influenza A Strains Circulating In Non-human Hosts

Animal influenza viruses emerging into humans have triggered devastating pandemics in the past. Yet, our ability to evaluate the pandemic potential of individual strains that do not yet circulate in humans, remains limited. In thsi study, our goal is to develop an experimentally validated platform called the Emergenet (Enet), to predict in near-real-time where and when new variants of concern would emerge, using only observed sequences of key viral proteins, procured in ongoing global surveillance of Influenza A viruses. We bring together new machine learning algorithms customized to the problem at hand, key insights from information theory, evolutionary theory, epidemiology and precise statistical uncertainty quantification to develop a rigorous framework, to track evolutionary trajectories of pathogens through a complex, poorly characterized, and dynamically changing fitness landscape. Our deliverable is best described as the foundations for creating a platform akin to bio-NORAD, *identifying when and where an imminent zoonotic emergence event is likely, and if such novel strains are likely to achieve human-to-human transmission capability.*

Influenza viruses constantly evolve[1], sufficiently altering surface protein structures to evade the prevailing host immunity, and cause the recurring seasonal epidemic. These periodic infection peaks claim a quarter to half a million lives[2] globally. Additionally, Influenza A, partly on account of its segmented genome and its wide prevalence in animal hosts, can easily incorporate genes from multiple strains and (re)emerge as novel human pathogens[3], thus harboring a high pandemic potential. Strains spilling over into humans from animal reservoirs is thought to have triggered pandemics at least four times (1918 Spanish flu/H1N1, 1957 Asian flu/H2N2, 1968 Hong Kong flu/H3N2, 2009 swine flu/H1N1) in the past 100 years[4]. One approach to mitigating such risk is to identify animal strains that do not yet circulate in humans, but is likely to spill-over and quickly achieve human-to-human (HH) transmission capability. While global surveillance efforts collect wild specimens from diverse hosts and geo-locations annually, our ability to objectively, reliably and scalably risk-rank individual strains remains limited[5]. The Center for Disease Control's (CDC) current solution to this problem is the Influenza Risk Assessment Tool (IRAT)[6], which relies on time-consuming proteomics and transmission assays and potentially subjective evaluations by subject matter experts, taking weeks to months to compile for each strain of concern. With tens of thousands of strains being sequenced annually, this results in a scalability bottleneck. Our specific aims are:

Aim 1: Develop a platform powered by novel pattern discovery and recognition algorithms to automatically parse out emergent evolutionary constraints operating on Influenza A viruses in the wild, to provide a less-heuristic theory-backed scalable solution to emergence prediction. We plan to show that this capability enables preempting strains which are expected to be in future human circulation, and approximate IRAT scores of non-human strains without experimental assays or SME scoring, in second as opposed to weeks or months. Our approach automatically takes into account the time-sensitive variations in selection pressures as the background strain circulation changes over time, and will potentially be able to rank-order strains adapatively.

Aim 2: Validate our ability to predict future variations of viral proteins by showing that predicted variants of HA and NA fold correctly, and are functional, binding to the relevant human receptors in in-vivo laboratory experiments. Thus, bringing together rigorous data-driven modeling, and validation via tools from reverse genetics we plan to deliver an actionable and deployable platform that optimally exploits the current biosurvellance capacity.

The BioNORAD platform will enable proactive and actionable global surveillance for emerging pandemic threats from Influenza A. This importance of the ability to preempt pandemic risk to the national interest of the United States cannot be overstated, especially in the context of protecting DoD assest and personnel deployed in potentially high risk centers of emergence. Additionally, the BioNORAD will enable preemptive action including the inoculation of animal reservoirs before the first human infection, potentially eliminating the pandemic before it has a chance to trigger.

PIs for the proposed effort are Ishanu Chattopadhyay (lead, machine learning, UChicago) and Balaji Manicassamy (virology, Iowa State), and will be conducted over 2 years with a direct cost of 200K USD.

REFERENCES

[1] Dos Santos, G., Neumeier, E. & Bekkat-Berkani, R. Influenza: Can we cope better with the unpredictable? *Human vaccines & immunotherapeutics* **12**, 699–708 (2016).

[2] Huddleston, J. *et al.* Integrating genotypes and phenotypes improves long-term forecasts of seasonal influenza a/h3n2 evolution. *Elife* **9**, e60067 (2020).

[3] Reid, A. H. & Taubenberger, J. K. The origin of the 1918 pandemic influenza virus: a continuing enigma. *Journal of general virology* **84**, 2285–2292 (2003).

[4] Shao, W., Li, X., Goraya, M. U., Wang, S. & Chen, J.-L. Evolution of influenza a virus by mutation and re-assortment. *International journal of molecular sciences* **18**, 1650 (2017).

[5] Wille, M., Geoghegan, J. L. & Holmes, E. C. How accurately can we assess zoonotic risk? *PLoS biology* **19**, e3001135 (2021).

[6] CDC. Influenza risk assessment tool (irat) — pandemic influenza (flu) — cdc. https://www.cdc.gov/flu/pandemic-resources/national-strategy/risk-assessment.htm. (Accessed on 07/02/2021).