Fall Internship Report: A literature review on quantitative assays immunology

Jared Galloway

November 19, 2020

Abstract

Protection from rapidly evolving and spreading viruses is key to human health and survival. The molecular nature of infection and immune system defences provides us with a complex and noisy set of of problems to solve if we are to combat infectious disease and understand the nature of their evolution. Most notably, we need to understand the binding affinity of a virus to bind to a host cell via proteins on the expressed on the surface of the cell's outer membrane. These proteins allow the virus to enter a host cell, then proceed to hijack the cell's own machinery to replicate and propagate an infection. In the context of SARS-CoV-2, understanding the immune response with respect to those binding proteins is critical for prevention and prediction of disease severity. Additionally, understanding the evolution of these binding sites allows us to interpret the cross-species transmission and duration of immunity post-infection, or the potential mutational routes of binding escape. Neither of these are trivial problems to solve. Fortunately, recent advances in next generation sequencing (NGS), oligonucleotide synthesis (ONS), and PCR-induced ¡?¿ have driven the development of quantitative assays and given us the ability to explore and quantify fitness of particular sequences in the context of protein interaction. These methods have laid the foundation for exploration and development of advanced vaccines providing protection against deadly pathogens from causing serious illness and even death. In this literature review, we explore the advanced quantitative assays, Phage Immunoprecipitation Sequencing and Deep Mutational Scanning, which allow us to measure and explore these complex and noisy problems. Concretely, we will observe the results and methods from Shrock et al. [2020] and Starr et al. [2020], two papers that focus on the binding properties of the novel betacoronavirus, SARS-CoV-2.

Introduction

Modern mammalian immune systems are constituted by the aggregate of proteins and cells which defend against unwanted invaders (pathogens). These defences keep the pathogens from harming the delicate and complex biological systems which keep us alive and healthy. However, deadly pathogenic outbreaks which rapidly spread among humans and other species can often harm or kill a large percentage of populations [Wu et al., 2020]. In the case of viruses, replication as a function of fitness drives pathogens to evolve much in the same way we do – often meaning the most potent and infectious pathogens prevail as a product of their genome evolution [Twiddy et al., 2003, Felsenstein, 1981]. Fighting fire with fire, the adaptive immune system works through similar processes of mutation and selection, inside our own body, to evolve along-side these pathogens and confer specialized immunity - in many cases lasting throughout the lifetime of an individual. Incredibly, the combinatorial effects of specialized (VDJ) recombination results in enough diversity to select upon that evolution of specialized cells takes place in mere days (often a week or so) when encountering a new pathogen [Jung and Alt, 2004].

In contrast to all other forms of evolution (often on ecological timescales), The process of generating specific antibodies to ward off an infection is incredibly fast. Unfortunately, the symptoms of an infection during that time frame can still make an individual very ill, or even be fatal. The ability of viruses to hijack our cell's own machinery to replicate itself in order to propagate the infection make them effecient and deadly. Luckily, the process of producing antibodies need not occur everytime we ecounter the same virus. Rather, once an individual has encountered a pathogen and created the necessary cells needed to fend off the virus and infected cells, the defences that were used are stored in a sort of "immuno-memory" – using another type specialized cell. Upon contact with a pathogen the individual has encountered in the past, then, the immune system has the infrastructure in place to elicit a fast and effective response, Having this cellular machinery is what's known as immunity in an individual - and is key to survival in a world filled with microbial pathogens. One of the most impactful developments in human

health has been our ability to provoke immunity to common viruses without actually infecting us with a deadly disease causing pathogen. These biologically prepared agents are known as *vaccines*, and according to the center for disease control (CDC.gov) will have prevented over $21,000,000$ hospitalizations and roughly $750,000$ deaths among children born within the last 20 years – in the U.S alone. While this is an extreme success, the rate at which vaccines can be produced are a function of our ability to observe the physical properies of a virus. To date, the fastest a vaccine has been succesfully developed was during the mumps outbreak in 1969 and took 4 years from start to finish. Facing a more deadly pathogen, of which we are certain exists, this slow rate of development could pose an existential threat to the human race. Quantitative assays in Immunology, particularly in the last decade, have laid the foundation for measuring interaction between a virus and host cells at a rate and scale far greater than previously thought possible [Fowler and Fields, 2014, Thyagarajan and Bloom, 2014].

Commonly, a vaccine for some particular virus essentially models the virus - without any of the harmful properties. This can be thought of as giving your immune system a molecular picture of the virus so that it is prepared when the real thing is encountered. Anything that elicits an immune response is known as an *antigen*, and the antigen for a particular pathogen is known as the *epitope*. Knowing the epitope for any virus is key in modeling it for vaccines. Inferring a particular antigen is no trivial process, the number of possible peptides chains forming a protein which constitute the epitope for any particular virus are nearly infinitesimal. To date there is no direct way to isolate which proteins are expressed on a virus, and which constitute an antigen. To complicate further, little is known about how sequence variation affects protein function. As viruses evolve, we would like to know how possible mutations impact our immuno-defences; In the case of the novel coronavirus, SARS-CoV-2, high mutation rates have already been found in the region which binds to our cells. In order to predict or understand how long immunity will last in the face of evolution, we must explore all variants of the epitope and their respective binding affinity relative to the wild type sequence.

Fortunately, recent advances in next generation sequencing (NGS), oligonucleotide synthesis (ONS), and PCR-induced

Here, we dig into the benefits and limitations of two such quantitative assays, Phage Immunoprecipitation Sequencing (PhIP-Seq), and Deep Mutational Scanning (DMS) along with the analysis techniques we use to query the resulting data from these protocols.

Starr et al. [2020]

Quantitative assays

Phage-Immuno Precipitation Sequencing

Deep Mutational Scanning

Analysis and modeling

Future Perspectives and Conclusions

References

J Felsenstein. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.*, 17(6): 368–376, 1981. ISSN 0022-2844. URL `http://www.ncbi.nlm.nih.gov/pubmed/7288891`.

Douglas M. Fowler and Stanley Fields. Deep mutational scanning: a new style of protein science. *Nature Methods*, 11(8):801–807, Aug 2014. ISSN 1548-7105. doi: 10.1038/nmeth.3027. URL `https://doi.org/10.1038/nmeth.3027`.

David Jung and Frederick W. Alt. Unraveling v(d)j recombination: Insights into gene regulation. *Cell*, 116(2): 299–311, Jan 2004. ISSN 0092-8674. doi: 10.1016/S0092-8674(04)00039-X. URL `https://doi.org/10.1016/S0092-8674(04)00039-X`.

Ellen Shrock, Eric Fujimura, Tomasz Kula, Richard T. Timms, I-Hsiu Lee, Yumei Leng, Matthew L. Robinson, Brandon M. Sie, Mamie Z. Li, Yuezhou Chen, Jennifer Logue, Adam Zuiani, Denise McCulloch, Felipe J. N. Lelis, Stephanie Henson, Daniel R. Monaco, Meghan Travers, Shaghayegh Habibi, William A. Clarke, Patrizio

Caturegli, Oliver Laeyendecker, Alicja Piechocka-Trocha, Jon Li, Ashok Khatri, Helen Y. Chu, , Alexandra-Chloé Villani, Kyle Kays, Marcia B. Goldberg, Nir Hacohen, Michael R. Filbin, Xu G. Yu, Bruce D. Walker, Duane R. Wesemann, H. Benjamin Larman, James A. Lederer, and Stephen J. Elledge. Viral epitope profiling of covid-19 patients reveals cross-reactivity and correlates of severity. *Science*, 2020. ISSN 0036-8075. doi: 10.1126/science.abd4250. URL `https://science.sciencemag.org/content/early/2020/09/28/science.abd4250`.

Tyler N. Starr, Allison J. Greaney, Sarah K. Hilton, Daniel Ellis, Katharine H.D. Crawford, Adam S. Dingens, Mary Jane Navarro, John E. Bowen, M. Alejandra Tortorici, Alexandra C. Walls, Neil P. King, David Veesler, and Jesse D. Bloom. Deep mutational scanning of sars-cov-2 receptor binding domain reveals constraints on folding and ace2 binding. *Cell*, 182(5):1295 – 1310.e20, 2020. ISSN 0092-8674. doi: https://doi.org/10.1016/j.cell.2020.08.012. URL `http://www.sciencedirect.com/science/article/pii/S0092867420310035`.

Bargavi Thyagarajan and Jesse D Bloom. The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin. *eLife*, 3:e03300, jul 2014. ISSN 2050-084X. doi: 10.7554/eLife.03300. URL `https://doi.org/10.7554/eLife.03300`.

S. Susanna Twiddy, Edward C. Holmes, and Andrew Rambaut. Inferring the Rate and Time-Scale of Dengue Virus Evolution. *Molecular Biology and Evolution*, 20(1):122–129, 01 2003. ISSN 0737-4038. doi: 10.1093/molbev/msg010. URL `https://doi.org/10.1093/molbev/msg010`.

Fan Wu, Su Zhao, Bin Yu, Yan-Mei Chen, Wen Wang, Zhi-Gang Song, Yi Hu, Zhao-Wu Tao, Jun-Hua Tian, Yuan-Yuan Pei, Ming-Li Yuan, Yu-Ling Zhang, Fa-Hui Dai, Yi Liu, Qi-Min Wang, Jiao-Jiao Zheng, Lin Xu, Edward C. Holmes, and Yong-Zhen Zhang. A new coronavirus associated with human respiratory disease in china. *Nature*, 579(7798):265–269, Mar 2020. ISSN 1476-4687. doi: 10.1038/s41586-020-2008-3. URL `https://doi.org/10.1038/s41586-020-2008-3`.