EBioMedicine 66 (2021) 103333

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Commentary

'Geno-to-pheno' SARS-CoV-2 genome-COVID-19 association studies

Julian W Tang[a,b,*]

[a] Clinical Microbiology, University Hospitals of Leicester NHS Trust, Leicester Royal Infirmary, Level 5 Sandringham Building, Infirmary Square, Leicester LE1 5WW, UK
[b] Department of Respiratory Sciences, University of Leicester, LE1 7RH, UK

ARTICLE INFO

Article History:
Received 23 March 2021
Accepted 24 March 2021
Available online xxx

Disentangling viral genotypes and their correlation to potential clinical disease phenotypes is a long-standing problem - made worse by the different terminologies specific to individual viruses, often coined by specialists working on those individual viruses. For example, HIV 'genotyping' traditionally refers to HIV drug resistance testing, whereas hepatitis *B* and *C* 'genotyping' refer to identifying just the particular strain of these viruses present in a population, without reference to any drug resistance mutations.

To confuse things further, viral drug resistance testing can be based on two types of methodologies and databases - a 'genotypic' database where just the presence of well-characterized specific amino acid mutations can confer drug resistance, but also a 'phenotypic' database where *in vitro* viral culture methods are used to assess the growth capabilities of a specific patient's virus in the presence of individual antiviral drugs at different concentrations [1,2]. These can be combined into mega databases to aid drug resistance interpretation, such as the Stanford HIV database [https://hivdb.stanford.edu/].

These are examples of specific viral genotypic mutations that affect treatment outcomes. Let's refer to these as 'hard' genotypic correlates, i.e. if these viral mutations are present, that drug will not work against that virus, such as some of the mutations found in HIV [3].

Other types of viral genotypes can impact on vaccine efficacy, i.e. how well a vaccine-induced immune response will protect against infection or disease caused by a specific virus. A topical example of this is the emergence of various SARS-COV-2 variants that can reduce the efficacy (in clinical trials) and effectiveness (in real-life) of some of the COVID-19 vaccines currently being rolled out globally. *In vitro* (laboratory) studies have shown that the B.1.351/501Y.V2 South African variant of SARS-COV-2 reduces the ability of the Pfizer-BioNTech and Moderna vaccine-induced antibodies to protect against this variant, but not entirely. However, there is still some protection offered by these vaccines against severe disease and death [4,5].

Let's refer to these as 'medium' genotypic correlates, i.e. if these viral mutations are present it reduces the effect of the vaccine but does not negate its effect completely. Some drug resistance mutations are similar and will reduce viral replication rather than halt it completely, such as drug resistance mutations to HIV protease, CMV DNA polymerase and influenza neuraminidase inhibitors [1,2,6].

Given the above, what would we classify as 'soft' genotypic correlates? Here, I would define these as viral mutations that may have an impact on a particular clinical phenotype, like enhanced transmissibility or severity, but where these correlations are subject to multiple potential confounders and therefore may not be very robust.

One recent example of this might be the estimates of enhanced transmissibility and severity of some SARS-COV-2 variants, like the UK 'Kent' B.1.1.7/501Y.V1 [7,8]. These estimates may still be subject to some potential confounders including, but not limited to, incomplete sampling and sampling bias, seasonal factors such as the well-recognized winter exacerbations of comorbidities, indoor crowding and the enhanced social contact rates that come with festive celebrations, all of which can enhance viral transmissibility and disease severity without any intrinsic contribution from any new viral mutations. But whilst these confounders may potentially weaken the correlation, they may not necessarily completely negate it.

It is within this 'soft' genotypic correlate context that this multi-site Singaporean study falls. Young *et al.* compare the clinical severity and immune response to different SARS-COV-2 clades, in a patient dataset collated during the very early part of the COVID-19 pandemic (January to April 2020). The authors' main finding was that patients infected with the L/V clade of SARS-CoV-2 exhibited more severe illness with an enhanced cytokine response [9].

Their analysis is necessarily highly complex and statistical to take into account all the clinical and laboratory parameters to explore this correlation, without falling foul of potential confounding. Unfortunately, due to the time taken to write/submit/review/revise this article, their *specific findings* have become somewhat outdated, as the epidemiology of SARS-COV-2 has now moved away from the older 19A/B (L/O/V/S) clades and evolved more into 20A/B/C (G/GR/GH)-related lineages [10].

However, the *principle and approach* behind this type of analysis will always be relevant, especially with any new emerging pathogen. Once this type of 'geno-to-pheno' analysis is completed, it may yield

DOI of original article: http://dx.doi.org/10.1016/j.ebiom.2021.103319.
* Corresponding author at: Clinical Microbiology, University Hospitals of Leicester NHS Trust, Leicester Royal Infirmary, Level 5 Sandringham Building, Infirmary Square, Leicester LE1 5WW, UK.
E-mail address: julian.tang@uhl-tr.nhs.uk

https://doi.org/10.1016/j.ebiom.2021.103333

wider benefits in terms of how future COVID-19 patients are managed. For example, in the acute context of this Singaporean study, assuming that a rapid bedside test can be eventually developed to identify exactly with which viral genotype patients have been infected, this will allow patients to be stratified and managed accordingly, as potentially more or less severe cases of COVID-19, to improve overall patient outcomes.

Contributor

Julian W Tang was the sole author and contributor to this article.

Declaration of Competing Interest

None.

Acknowledgments

None.

References

[1] Pattery T, Verlinden Y, De Wolf H, et al. Development and performance of conventional HIV-1 phenotyping (Antivirogram®) and genotype-based calculated phenotyping assay (virco®TYPE HIV-1) on protease and reverse transcriptase genes to evaluate drug resistance. Intervirology 2012;55:138–46.
[2] Lurain NS, Chou S. Antiviral drug resistance of human cytomegalovirus. Clin Microbiol Rev 2010;23:689–712.
[3] Wensing AM, Calvez V, Ceccherini-Silberstein F, et al. 2019 update of the drug resistance mutations in HIV-1. Top Antivir Med 2019;27:111–21.
[4] Tada T, Dcosta BM, Samanovic-Golden M, et al. Neutralization of viruses with European, South African, and United States SARS-CoV-2 variant spike proteins by convalescent sera and BNT162b2 mRNA vaccine-elicited antibodies. bioRxiv 2021 [Preprint]. 2021 Feb 702.05.430003. doi: 10.1101/2021.02.05.430003.
[5] Edara VV, Norwood C, Floyd K, et al. Reduced binding and neutralization of infection and vaccine-induced antibodies to the B.1.351 (South African) SARS-CoV-2 variant. bioRxiv 2021 [Preprint]Feb 22:2021.02.20.432046. doi: 10.1101/2021.02.20.432046.
[6] World Health Organization. Summary of neuraminidase amino acid substitutions associated with reduced inhibition by neuraminidase inhibitors. Updated 26 April 2018. https://www.who.int/influenza/gisrs_laboratory/antiviral_susceptibility/NAI_Reduced_Susceptibility_Marker_Table_WHO.pdf?ua (accessed 20 March 2021)
[7] Davies NG, Abbott S, Barnard RC, et al. Estimated transmissibility and impact of SARSCoV- 2 lineage B.1.1.7 in England. Science 2021;3:eabg3055. doi: 10.1126/science.abg3055.
[8] Davies NG, Jarvis CI, et al. CMMID COVID-19 Working group. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. Nature 2021;15. doi: 10.1038/s41586-021-03426-1.
[9] Young BE, Wycliffe EW, Fong SW, et al. Association of SARS-CoV-2 clades with clinical, inflammatory and virologic outcomes: an observational study. EbioMedicine 2021. doi: 10.1016/j.ebiom.2021.103319.
[10] Alm E, Broberg EK, Connor T, et al. Geographical and temporal distribution of SARS-CoV- 2 clades in the WHO European region, January to June 2020. Eurosurveillance 2020;25:2001410. doi: 10.2807/1560-7917.ES.2020.25.32.2001410.