Analysis of Serologic Cross-Reactivity Between Common Human Coronaviruses and SARS-CoV-2 Using Coronavirus Antigen Microarray

Authors

Saahir Khan (1)*
Rie Nakajima (2)
Aarti Jain (2)
Rafael Ramiro de Assis (2)
Al Jasinskas (2)
Joshua M. Obiero (2)
Oluwasanmi Adenaiye (3)
Sheldon Tai (3)
Filbert Hong (3)
Donald K. Milton (3)
Huw Davies (2)
Philip L. Felgner (2)
Prometheus Study Group

*Corresponding Author: saahirk@uci.edu

Affiliations

1. Division of Infectious Diseases, Department of Medicine, University of California Irvine Health, Orange, CA
2. Department of Physiology and Biophysics, School of Medicine, University of California Irvine, Irvine, CA
3. Institute for Applied Environmental Health, School of Public Health, University of Maryland, College Park, MD
Abstract

The current practice for diagnosis of SARS-CoV-2 infection relies on PCR testing of nasopharyngeal or respiratory specimens in a symptomatic patient at high epidemiologic risk. This testing strategy likely underestimates the true prevalence of infection, creating the need for serologic methods to detect infections missed by the limited testing to date. Here, we describe the development of a coronavirus antigen microarray containing immunologically significant antigens from SARS-CoV-2, in addition to SARS-CoV, MERS-CoV, common human coronavirus strains, and other common respiratory viruses. A preliminary study of human sera collected prior to the SARS-CoV-2 pandemic demonstrates overall high IgG reactivity to common human coronaviruses and low IgG reactivity to epidemic coronaviruses including SARS-CoV-2, with some cross-reactivity of conserved antigenic domains including S2 domain of spike protein and nucleocapsid protein. This array can be used to answer outstanding questions regarding SARS-CoV-2 infection, including whether baseline serology for other coronaviruses impacts disease course, how the antibody response to infection develops over time, and what antigens would be optimal for vaccine development.
Background

The 2019 novel coronavirus strain (SARS-CoV-2) originating in Wuhan, China has become a worldwide pandemic with significant morbidity and mortality estimates up to 2% of confirmed cases. The current case definition for confirmed COVID-19 due to SARS-CoV-2 infection relies on PCR-positive nasopharyngeal or respiratory specimens, with testing largely determined by presence of fever or respiratory symptoms in an individual at high epidemiologic risk. However, this case definition likely underestimates the prevalence of SARS-CoV-2 infection, as individuals who develop subclinical infection that does not produce fever or respiratory symptoms are unlikely to be tested, and testing by PCR of nasopharyngeal or respiratory specimens is unlikely to be 100% sensitive in detecting subclinical infection. Widespread testing within the United States is also severely limited by the lack of available testing kits and testing capacity limitations of available public and private laboratories. Therefore, the true prevalence of SARS-CoV-2 infection is currently unknown, and the sensitivity of PCR to detect infection is also unknown.

Serology can play an important role in defining both the prevalence of and sensitivity of PCR for SARS-CoV-2 infection, particularly for subclinical infection. This point is demonstrated by analogy with influenza virus, for which a meta-analysis of available literature measured the fraction of asymptomatic infections detected by PCR as approximately 16%, while the fraction of asymptomatic infections detected by seroconversion was measured as approximately 75% \textsuperscript{1}. The seroprevalence of common human coronaviruses is known to increase throughout childhood to near 100% by adolescence \textsuperscript{2}. Thus, any serologic methodology to estimate prevalence of SARS-CoV-2 needs to identify and rule out cross-reactivity with these common human coronavirus strains.

One challenge in applying serology to SARS-CoV-2 is that the choice of antigen and choice of assay is less well defined for coronavirus than more well studied viruses such as influenza. However, prior approaches to serologic detection of infection with emerging coronaviruses including SARS and MERS have focused on the S1 domain of the spike (S) glycoprotein and the nucleocapsid (N) protein, which are considered the immunodominant antigens for these viruses \textsuperscript{3}. In particular, the S1 domain is strain-specific, while the N protein shows cross-reactivity across strains.

The assay methodologies used for serologic detection of coronavirus infection include binding and neutralization assays. These methodologies have been shown to be well correlated \textsuperscript{4}. However, neutralization assays require viral culture, which must be performed in high-level biosafety containment units for emerging coronaviruses with high epidemic potential such as SARS-CoV-2. Conversely, binding assays such as ELISA can be readily performed with widely available reagents and equipment so are field deployable and suitable for point of care testing.

The protein microarray methodology has been widely used to simultaneously perform binding assays against hundreds of antigens printed onto a nitrocellulose-coated slide for detection of multiple antibody isotypes \textsuperscript{5}. This methodology was recently demonstrated for simultaneous measurement of IgG and IgA antibodies against over 250 antigens from diverse strains and subtypes of influenza \textsuperscript{6}. This methodology has previously been applied to detect antibodies against the S1 domains of SARS and MERS coronaviruses \textsuperscript{7}. 

\textsuperscript{1} The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.
Methodology

Specimen Collection
The blood specimens used in this study were collected for a larger study where residents of a college resident community in an Eastern university were monitored prospectively to identify acute respiratory infection (ARI) cases using questionnaires and RT-qPCR, so as to characterize contagious phenotypes including social connections, built environment, and immunologic phenotypes. From among de-identified blood specimens for which future research use authorization was obtained, five specimens that showed high IgG reactivity against human coronaviruses in the larger study were chosen for validation of the coronavirus antigen microarray.

Antigen Microarray
The coronavirus antigen microarray used in this investigation includes 67 antigens across subtypes expressed in either baculovirus or HEK-293 cells (see Tables 1-3). These antigens were provided by Sino Biological Inc. (Wayne, PA) as either catalog products, or service products. The antigens were printed onto microarrays, probed with human sera, and analyzed as previously described (Figure 1). Briefly, lyophilized antigens were reconstituted to a concentration of 0.1 mg/mL in phosphate-buffered saline (PBS) with 0.001% Tween-20 (T-PBS) and then printed onto nitrocellulose-coated slides from Grace Bio Labs (GBL, Bend, OR) using an OmniGrid 100 microarray printer (GeneMachines). The microarray slides were probed with human sera diluted 1:100 in 1x GVS Fast Blocking Buffer (Fischer Scientific) overnight at 4°C, washed with T-TBS buffer (20 mM Tris-HCl, 150 mM NaCl, 0.05% Tween-20 in ddH2O adjusted to pH 7.5 and filtered) 3 times for 5 minutes each, labeled with secondary antibodies to human IgG conjugated to quantum dot fluorophore for 2 hours at room temperature, and then washed with T-TBS 3 times for 5 minutes each and dried. The slides were imaged using ArrayCam imager (Grace Bio Labs, Bend, OR) to measure background-subtracted median spot fluorescence. Non-specific binding of secondary antibodies was subtracted using saline control. Mean fluorescence of the 4 replicate spots for each antigen was used for analysis.

Statistical Analyses
Descriptive statistics were used to summarize the IgG reactivity as measured by mean fluorescence across antigen replicates. T-test or F-test were used to test for the mean differences in continuous variables across infection groups. All statistical analyses were conducted using R version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria).

Results
Overall, the 5 sera tested on the coronavirus antigen microarray all showed high IgG seroreactivity to antigens from common human coronaviruses and other respiratory viruses with known seasonal circulation versus low IgG seroreactivity to antigens from epidemic viruses that were not circulating at time of collection (Figure 2). Specifically, 4 of the 5 sera showed high IgG seroreactivity across the 4 common human coronaviruses,
while all of the sera showed low IgG seroreactivity to SARS-CoV-2, SARS-CoV, and MERS-CoV. All 5 sera showed high IgG seroreactivity to RSV and parainfluenza viruses, while 3 of the 5 sera showed high IgG seroreactivity to adenoviruses. For influenza, all 5 sera showed high IgG seroreactivity to H1N1 and H3N2 influenza A and influenza B strains but low IgG seroreactivity to H5N1 and H7N9 influenza A strains. With respect to specific antigens, the S1 domain of the spike protein including the receptor-binding domain (RBD) demonstrates very low cross-reactivity between epidemic coronaviruses and common human coronaviruses, whereas the S2 domain of the spike protein and the nucleocapsid protein (NP) show low-level cross-reactivity between these coronavirus subtypes. Similarly, the head domain of influenza hemagglutinin (HA1) is not cross-reactive between seasonal and avian influenza strains, whereas the stalk domain (HA2) is cross-reactive between influenza virus subgroups, as seen between H1N1 and H5N1 influenza viruses.

**Discussion**

This pilot study yields several insights into cross-reactivity of common human coronavirus antibodies for SARS-CoV-2 antigens. The antibodies to the S1 and RBD domains of spike protein are highly subtype-specific, consistent with the high variability in these sequences between different human coronaviruses. Conversely, the S2 domain of spike protein and NP protein are more cross-reactive, consistent with these sequences being highly conserved across coronaviruses. SARS-CoV-2 has caused a worldwide pandemic despite likely pre-existing cross-reactive antibodies to S2 domain and NP protein in most people, indicating that these antibodies are likely not protective, whereas antibodies to S1 and RBD domains are more likely to be protective. This observation favors a vaccination strategy based on S1 or RBD domains of spike protein over a vaccination strategy that also includes S2 domain or NP protein. In addition, S1 and RBD domains are more likely to generate subtype-specific serologic tests for population surveillance studies.

In addition, a key unexplained finding during the SARS-CoV-2 epidemic has been the low incidence of infection in children aged 15 and younger. This observation generates two related hypotheses: adults may have pre-existing antibodies against antigenically distinct coronaviruses that produce an ineffective humoral response to SARS-CoV-2 infection (antibody-dependent enhancement as demonstrated for dengue virus), or children younger than 15 may have initially encountered a coronavirus that is more closely related to SARS-CoV-2 so are more protected against this infection (immunologic imprinting or original antigenic sin as demonstrated for influenza virus). Both of these hypotheses would be informed by comparing the level of cross-reactive coronavirus antibodies in pediatric and adult cohorts and correlating these antibodies with incidence of severe disease.

**Conclusions**

A coronavirus antigen microarray has been constructed with antigens from epidemic coronaviruses including SARS-CoV-2 and common human coronaviruses, in addition to other common respiratory viruses. A pilot study of 5 naïve human sera shows high IgG
seroreactivity to common human coronaviruses but low IgG seroreactivity to SARS-CoV-2, with some cross-reactivity seen for S2 domain of spike protein and nucleocapsid protein. Further studies are needed including with SARS-CoV-2 convalescent sera to fully realize the potential of this novel methodology to characterize the seroprevalence of SARS-CoV-2 and the impact of pre-existing cross-reactive antibodies on the disease course.

Acknowledgements

Saahir Khan is supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant KL2 TR001416. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Prometheus-UMD was sponsored by the Defense Advanced Research Projects Agency (DARPA) BTO under the auspices of Col. Matthew Hepburn through agreements N66001-17-2-4023 and N66001-18-2-4015 (PI: Milton). This study was funded in part by the Defense Threat Reduction Agency via grants HDTRA1-18-1-0036 (PI: Davies) and HDTRA1-18-1-0035 (PI: Felgner). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position or policy of the funding agencies and no official endorsements should be inferred.
Figures

Figure 1. Schematic of antigen microarray printing, probing, imaging, and analysis. Reprinted with permission.

Figure 2. IgG seroreactivity as measured by mean fluorescence intensity of 5 serum specimens from naïve population on coronavirus antigen microarray.
**Tables**

### Table 1. Content of coronavirus antigen microarray.

| Virus            | Subtypes | Strain | Protein    | UniProt/GenBank | Expression | Synthesis | Construct | Catalogue No |
|------------------|----------|--------|------------|-----------------|------------|-----------|-----------|--------------|
| Coronavirus      | NL03     | NL03   | S1         | Q0ZME7          | HEK293     | N-(AA18-717)His-C | Sino Biological | 40060-V08B     |
| Coronavirus      | NL03     | NL03   | S1+S2      | A0A1L2Y186      | HEK293     | N-(AA18-132)His-C | Sino Biological | 40064-V08B     |
| Coronavirus      | 229E     | 229E   | S1         | A0A1L2Y186      | HEK293     | N-(AA18-536)His-C | Sino Biological | 40061-V08B     |
| Coronavirus      | 229E     | 229E   | S1+S2      | A0A1L2Y186      | Insect Cells | N-(AA18-1115)His-C | Sino Biological | 40065-V08B     |
| Coronavirus      | HKU1     | HKU1   | S1         | Q0ZME7          | HEK293     | N-(AA17-760)His-C | Sino Biological | 40021-V08B     |
| Coronavirus      | HKU1     | HKU1   | S1+S2      | Q0ZME7          | HEK293     | N-(AA13-756)His-C | Sino Biological | 40062-V08B     |
| Coronavirus      | HKU1     | HKU1   | S1         | Q0ZME7          | HEK293     | N-(AA13-1295)His-C | Sino Biological | 40066-V08B     |
| Coronavirus      | HKU1     | HKU1   | HE         | Q0ZME7          | HEK293     | N-(AA16-394)His-C | Sino Biological | Custom         |
| Coronavirus      | HKU23/356F | NP    | AVR40344.1 | HEK293         | N-(AA1-448)His-C | 40458-V08B |
| Coronavirus      | OC43     | OC43   | S1         | AVR40344.1      | HEK293     | N-(AA13-533)His-C | Sino Biological | Custom         |
| Coronavirus      | OC43     | OC43   | S1+S2      | AVR40344.1      | Insect Cells | N-(AA13-1304)His-C | Sino Biological | 40067-V08B     |
| Coronavirus      | OCM5     | OCM5   | HE         | ATN38879.2      | HEK293     | N-(AA16-394)His-C | Sino Biological | 40063-V08B     |
| Coronavirus      | MERS     | MERS   | S1-RBD     | AF888936.1      | Insect Cells | N-(AA338-505)F-C | Sino Biological | 40071-V08B     |
| Coronavirus      | MERS     | MERS   | S1-RBD     | AF888936.1      | Insect Cells | N-(AA338-502)F-C | Sino Biological | 40071-V31B     |
| Coronavirus      | MERS     | MERS   | S1-RBD     | AF888936.1      | Insect Cells | N-(AA367-606)F-C | Sino Biological | 40071-V31B     |
| Coronavirus      | MERS     | MERS   | S1         | AF888936.1      | HEK293     | N-(AA1-725)His-C | Sino Biological | 40069-V08B     |
| Coronavirus      | MERS     | MERS   | S1+S2      | AF888936.1      | Insect Cells | N-(AA1-725)His-C | Sino Biological | 40069-V08B     |
| Coronavirus      | MERS     | MERS   | S2         | AF888936.1      | HEK293     | N-(AA726-1296)His-C | Sino Biological | 40070-V08B     |
| Coronavirus      | MERS     | MERS   | NP         | AF888943.1      | Insect Cells | N-(AA1-413)His-C | Sino Biological | 40068-V08B     |
| Coronavirus      | SARS     | SARS   | S1-RBD     | AAX16192.1      | Insect Cells | N-(AA306-527)F-C | Sino Biological | 40150-V31B     |
| Coronavirus      | SARS     | SARS   | S1-RBD     | AAX16192.1      | Insect Cells | N-(AA306-527)F-C | Sino Biological | 40150-V08B     |
| Coronavirus      | SARS     | SARS   | S1         | AAX16192.1      | Insect Cells | N-(AA1-667)His-C | Sino Biological | 40150-V08B     |
| Coronavirus      | SARS     | SARS   | NP         | 828858.1       | Insect Cells | N-(AA1-422)His-C | Sino Biological | 40143-V08B     |
| Coronavirus      | SARS     | SARS   | Plpro      | AAX16193.1      | E.coli     | N-(AA1541-1859)His-C | Sino Biological | 40524-V08E     |
| Coronavirus      | 2019-nCoV | 2019-nCoV | S1-RBD   | HEK293          | Sino Biological | N-(AA1-709)His-C | 40059-V08B     |
| Coronavirus      | 2019-nCoV | 2019-nCoV | S1         | HEK293          | Sino Biological | N-(AA1-709)His-C | 40059-V08B     |
| Coronavirus      | 2019-nCoV | 2019-nCoV | S1         | HEK293          | Sino Biological | N-(AA1-709)His-C | 40059-V08B     |
| Coronavirus      | 2019-nCoV | 2019-nCoV | S2         | HEK293          | Sino Biological | N-(AA1-709)His-C | 40059-V08B     |
| Coronavirus      | 2019-nCoV | 2019-nCoV | S1+S2      | HEK293          | Sino Biological | N-(AA1-709)His-C | 40059-V08B     |
| Coronavirus      | 2019-nCoV | 2019-nCoV | NP         | HEK293          | Sino Biological | N-(AA1-709)His-C | 40059-V08B     |

### Table 2. Coronavirus antigens on microarray.
| Virus | Subtype | Strain | Protein | LigPro|GenBank | Expression | Synthesis | Genome | Catalogue No |
|-------|---------|--------|---------|-------|---------|-----------|-----------|--------|--------------|
| RSV   | A       | LA2/84/2013 | F       | AA076254X | Insect Cells | Sino Biological | N(AA1-529)-His-C | Custom |
| RSV   | A       | LA2/84/2013 | G       | AA076254X | HEK293 | Sino Biological | N(AA4-321)-His-C | Custom |
| RSV   | A       | A2       | F       | BAF133 | Insect Cells | Sino Biological | N(AA1-529)-His-C | 11049-V08B |
| RSV   | A       | nPbT134  | G       | BAF133 | HEK293 | Sino Biological | N(AA66-297)-His-C | 11070-V08H |
| RSV   | A       | RS-2     | F       | BAF133 | Insect Cells | Sino Biological | N(AA1-529)-His-C | 40037-V08B |
| RSV   | B       | TH-10526/2014 | F       | K7WU9 | Insect Cells | Sino Biological | N(AA1-525)-His-C | Custom |
| RSV   | B       | TH-10526/2014 | G       | AAIA242MLK4 | HEK293 | Sino Biological | N(AA66-310)-His-C | Custom |
| RSV   | B       | B1       | G       | HEK293 | HEK293 | Custom | N(AA67-299)-His-C | 13026-V08H |
| hMPV  | A       | PER/CIFO0320/2010/A | G       | HEK293 | Sino Biological | 52N-228N-His | Custom |
| hMPV  | B       | PER/CIFO0320/2010/B | G       | HEK293 | Sino Biological | 52D-238S-His | Custom |
| hMPV  | B       | PER/CIFO0320/2010/A | F       | HEK293 | Sino Biological | 2800-490G-His | Custom |
| Parainfluenza 1 | 12O3 | F       | AAIA1061X6 | Insect Cells | Sino Biological | N(AA2-497)-His-C | Custom |
| Parainfluenza 1 | 12O3 | H       | AAIA1061X6 | Insect Cells | Sino Biological | N(AA9-575)-C | Custom |
| Parainfluenza 1 | 12O3 | H       | AAIA1061X6 | Insect Cells | Sino Biological | N(AA9-575)-C | Custom |
| Parainfluenza 4 | hPIV-4b10-H2/2016 | F       | AAIA1061X6 | Insect Cells | Sino Biological | N(AA2-496)-His-C | Custom |
| Parainfluenza 4 | hPIV-4b10-H2/2016 | H       | AAIA1061X6 | Insect Cells | Sino Biological | N(AA4-575)-C | Custom |
| Adenovirus 3 | hAdV-3/45659 | Fiber | P04801 | E. coli | Sino Biological | N(His-Pro1)-C | Custom |
| Adenovirus 4 | hAdV-4/28280 | Fiber | P04801 | E. coli | Sino Biological | N(His-Pro1)-C | Custom |
| Adenovirus 7 | Adeno7 10519 | Fiber | P04801 | E. coli | Sino Biological | N(His-Pro1)-C | Custom |
| Influenza H1N1 | A/Beijing/22808/2009 | HA1 | A0042031 | HEK293 | Sino Biological | N(AA1-344)-His-C | 40035-V08H |
| Influenza H1N1 | A/Beijing/22808/2009 | HA1 | A0042031 | HEK293 | Sino Biological | N(AA1-344)-His-C | 40035-V08H |
| Influenza H3N2 | A/Texas/200/2012 | HA1 | A0042031 | HEK293 | Sino Biological | N(AA1-344)-His-C | 40035-V08H |
| Influenza H3N2 | A/Texas/200/2012 | HA1 | A0042031 | HEK293 | Sino Biological | N(AA1-344)-His-C | 40035-V08H |
| Influenza A | B/Malaysia/2506/2004 | HA1 | A0042031 | HEK293 | Sino Biological | N(AA1-344)-His-C | 11716-V08H |
| Influenza B | B/Malaysia/2506/2004 | HA1 | A0042031 | HEK293 | Sino Biological | N(AA1-344)-His-C | 11716-V08H |
| Influenza B | B/Phuket/2073/2013 | HA1 | B202452304 | HEK293 | Sino Biological | N(AA1-344)-His-C | 40498-V08H |
| Influenza H5N1 | A/Vietnam/1203/2004 | HA1 | B202452304 | HEK293 | Sino Biological | N(AA1-344)-His-C | 40498-V08H |
| Influenza H5N1 | A/Vietnam/1203/2004 | HA1 | B202452304 | HEK293 | Sino Biological | N(AA1-344)-His-C | 40498-V08H |
| Influenza H7N9 | A/Anhui/1/2013 | HA1 | AGJ519531 | HEK293 | Sino Biological | N(AA1-344)-His-C | 40103-V08H |
| Influenza H7N9 | A/Anhui/1/2013 | HA1 | AGJ519531 | HEK293 | Sino Biological | N(AA1-344)-His-C | 40103-V08H |

Table 3. Non-coronavirus respiratory virus antigens on microarray.
References

1. Leung, N. H., Xu, C., Ip, D. K. & Cowling, B. J. Review Article: The Fraction of Influenza Virus Infections That Are Asymptomatic: A Systematic Review and Meta-analysis. *Epidemiology* **26**, 862-872, doi:10.1097/EDE.0000000000000340 (2015).

2. Zhou, W., Wang, W., Wang, H., Lu, R. & Tan, W. First infection by all four non-severe acute respiratory syndrome human coronaviruses takes place during childhood. *BMC Infect Dis* **13**, 433, doi:10.1186/1471-2334-13-433 (2013).

3. Agnihothram, S. *et al.* Evaluation of serologic and antigenic relationships between middle eastern respiratory syndrome coronavirus and other coronaviruses to develop vaccine platforms for the rapid response to emerging coronaviruses. *J Infect Dis* **209**, 995-1006, doi:10.1093/infdis/jit609 (2014).

4. Chan, C. M. *et al.* Examination of seroprevalence of coronavirus HKU1 infection with S protein-based ELISA and neutralization assay against viral spike pseudotyped virus. *J Clin Virol* **45**, 54-60, doi:10.1016/j.jcv.2009.02.011 (2009).

5. Davies, D. H. *et al.* Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery. *Proc Natl Acad Sci U S A* **102**, 547-552, doi:10.1073/pnas.0408782102 (2005).

6. Khan, S. *et al.* Use of an Influenza Antigen Microarray to Measure the Breadth of Serum Antibodies Across Virus Subtypes. *J Vis Exp*, doi:10.3791/59973 (2019).

7. Reusken, C. *et al.* Specific serology for emerging human coronaviruses by protein microarray. *Euro Surveill* **18**, 20441, doi:10.2807/1560-7917.es2013.18.14.20441 (2013).

8. Zhu, S. *et al.* Ventilation and laboratory confirmed acute respiratory infection (ARI) rates in college residence halls in College Park, Maryland. *Environ Int* **137**, 105537, doi:10.1016/j.envint.2020.105537 (2020).

9. Jain, A. *et al.* Evaluation of quantum dot immunofluorescence and a digital CMOS imaging system as an alternative to conventional organic fluorescence dyes and laser scanning for quantifying protein microarrays. *Proteomics* **16**, 1271-1279, doi:10.1002/pmic.201500375 (2016).

10. Nakajima, R. *et al.* Protein Microarray Analysis of the Specificity and Cross-Reactivity of Influenza Virus Hemagglutinin-Specific Antibodies. *mSphere* **3**, doi:10.1128/mSphere.00592-18 (2018).