Substantial Differences in SARS-CoV-2 Antibody Responses Elicited by Natural Infection and mRNA Vaccination

Rafael Assis, Aarti Jain, Rie Nakajima, Al Jasinskas, Saahir Kahn, Anton Palma, Daniel M. Parker, Anthony Chau, Amanda Leung, Christina Grabar, Fjolla Muqolli, Ghali Khalil, Jessica Colin Escobar, Jenny Ventura, D. Huw Davies, Bruce Albala, Bernadette Boden-Albala, Sebastian Schubl, Philip L. Felgner

University of California Irvine, School of Medicine and the Vaccine R&D Center
University of California Irvine, School of Medicine, Department of Neurology (BA, BBA) and Center for Clinical Research (BA)
University of California Irvine, Program in Public Health, Department of Health Society and Behavior(BBA), Department of Population Health Disease Prevention and Department of Epidemiology and Biostatistics (DP)

Abstract:
We analyzed data from two ongoing serologic surveys, a longitudinal cohort of health care workers (HCW) from the University of California Irvine Medical Center (Orange County, CA, USA), collected from May and December 2020 through March 2021, and a cross sectional county-wide study in July 2020 (actOC; Orange County, CA) and a more focused community study in the city of Santa Ana (Santa Ana Cares; Orange County, CA, USA), collected in December 2020 - in order to compare the antibody responses to SARS-CoV-2 natural infection and vaccination. In addition, we serially tested 9 volunteers at multiple time points to analyze the time course of vaccine-induced antibody response in more detail. In May 2020, 1060 HCW were enrolled and had finger stick samples collected. Finger stick samples were again collected in December 2020, before vaccination, as well as January, February and March 2021 during vaccination campaign. A total of 8,729 finger stick blood specimens were probed and analyzed for IgG and IgM antibodies using a coronavirus antigen microarray (COVAM).The microarray contained 10 SARS-CoV-2 antigens including nucleocapid protein (NP) and several varying fragments of the spike protein, as well as 4 SARS, 3 MERS, 12 Common CoV, and 8 Influenza antigens.Based on a random forest based prediction algorithm, between May and December, prior to vaccine rollout, we observed that seropositivity in the HCW cohort increased from 4.5% to 13%. An intensive vaccination campaign with mRNA vaccines was initiated on December 16, 2020 and 6,724 healthcare workers were vaccinated within 3 weeks. The observed seropositivity of the HCW specimens taken in the last week of January 2021 jumped to 78%, and by the last week in February it reached 93%, and peaked at 98% seropositive in March. The antibody profile induced by natural exposure differed from the profile induced after mRNA vaccination. Messenger RNA vaccines induced elevated antibody (Ab) reactivity levels against the Receptor Binding Domain (RBD) domain of SARS-CoV-2 spike, and cross-reactive responses against SARS and MERS RBD domains. Nucleocapsid protein (NP), which is an immunodominant antigen induced after natural exposure, is

not present in the vaccine and can be used as a biomarker of past exposure. The results show that naturally-exposed individuals mount a stronger anti-spike response upon vaccination than individuals that were not previously exposed. Longitudinal specimens taken at approximately weekly intervals from 9 individuals show variation in the response to the mRNA vaccine, with some showing a vigorous response to the first dose (prime) and others requiring a subsequent dose (boost) to reach high anti-SARS-CoV-2 levels. Antibody titers determined by serial dilution of the specimens were used to accurately compare antibody levels in these samples. mRNA vaccinees after the boost have higher Ab titers (up to 10 times higher) than convalescent plasmas from donors who recovered from natural infection. The results of this study exemplify the time course and outcomes expected from similar mRNA mass vaccination campaigns conducted in other institutions.

Introduction

Protective efficacy of SARS-CoV-2 spike mRNA vaccines reported by the developers, Pfizer and Moderna, has been spectacular, showing convincing evidence of protection within only 14 days of the first immunization. [1, 2] To further understand the mRNA vaccine induced immune response we were interested in comparing the antibody response induced by the vaccine against responses induced by natural exposure to SARS-CoV-2. Here we show results using a multiplex solid phase immunofluorescent assay for quantification of human antibodies against 37 antigens from SARS-CoV-2, other novel and common coronaviruses, and influenza viruses that are causes of respiratory infections. [3-7] This assay uses a small volume of blood derived from a finger stick, does not require the handling of infectious virus, quantifies the level of different antibody types in serum and plasma and is amenable to scaling-up. Finger stick blood collection enables large scale epidemiological studies to define the risk of exposure to SARS-CoV-2 in different settings. [8] Since the assay requires 1 microliter of blood it is also practical for monitoring immunogenicity in neonates, children and small animal models.

Our results show that the nucleic acid vaccines in use in this setting are remarkably effective at elevating Ab levels against SARS-CoV-2 antigens, rapidly converting seronegative individuals into seropositives. The observed seroconversion level and breadth induced by the mRNA vaccines is much greater than that induced by natural infection. After probing more than 8,729 pre- and post-vaccination specimens our results confirm that the mRNA vaccines can be used in an aggressive and targeted vaccination campaign to immunize large groups within a matter of weeks.

Methods

COVID seroprevalence surveys in Orange County, California

Here we analyzed data from ongoing serologic surveys of healthcare workers (HCW) from the University of California Irvine Medical Center (UCIMC, Orange County, CA, USA) and from residents of the Orange County community. The first community survey (actOC) conducted in July of 2020, was county-wide, and recruitment was done via a proprietary phone list. This survey of 2,979 individuals was meant to be representative of the age, ethnicity, and socio-economic makeup of the county (detailed in [8]). The results of this county-wide survey indicated that the city of Santa Ana was a COVID-19 hotspot, especially on the Hispanic population. Surveillance of reported cases and test positivity corroborated this finding. A second, seroprevalance survey was then conducted in Santa Ana as the Santa Ana Cares study in December of 2020. Recruitment of 3347 individuals for this second survey was done using randomized house sampling within cenus tracts coupled with a community engaged campaign with support from Latino Health Access (a community based health organization that has been based in Santa Ana for over 2 decades, https://www.latinohealthaccess.org/). Analysis of the second seroprevalence survey is ongoing. While the first survey was county-wide, the serological test positivities reported in this analysis come from zip codes in Santa Ana alone.

Samples were also collected from the UCIMC longitudinal HCW study in May and December 2020. An aggressive and comprehensive mRNA vaccination campaign started at UCIMC on December 16 2020 and 6,724 HCW were vaccinated in 3 weeks. Three additional cross sectional samples were taken at end of January, February, and March 2021.

A Coronavirus Antigen Microarray (COVAM) was used to measure antibody levels against 37 antigens from coronaviruses and influenza. COVAM measurements taken at a single dilution of plasma can be used as a parameter to compare relative Ab titers between individual specimens against each of the individual 37 antigens. The COVAM contained 10 SARS-CoV-2, 4 SARS, 3 MERS, 12 Common CoV and 8 influenza antigens.(Figure 2) Samples were probed and analyzed on the COVAM and each individual was provided with the results of their test (Supplementary Section) according to the IRB protocol. (de Assis RR *et al.,* 2021).(Figure 1)

Results

mRNA vaccination achieves 99% seropositivity within 3 months after initiating an aggressive and inclusive vaccination campaign

This study was designed to track the seroprevalence at UCIMC since May 2020 and in the Orange County community that is served by the hospital system starting in July (Figure 2). In July the crude calculations of seroprevalence in Santa Ana zip codes was 18%, and in December 26%.(Figure 3A). Prior to the vaccination campaign, which

launched in late December, 2020 the 13% seroprevalence measured at UCIMC was lower than the 26% Santa Ana cross sectional results suggesting that strict transmission control measures enforced at the hospital played a role in keeping COVID-19 exposure levels low. On December 16, 2020 the vaccination campaign started at the hospital and seroprevalence jumped from 13% (early December) to 78% in January, 93% in February, and 98.7% in the last week of March 2021. (Figure 3B) This observation strongly corroborates the high efficacy of the nucleic acid vaccine in stimulating an antibody response and also highlights the success of the vaccination campaign that immunized 6724 HCW from 12/16/2020-1/05/2021, and 10,000 more since then.

Differences were noted in the Ab responses induced by the vaccine compared to natural exposure. (Figure 3) The nucleocapsid protein in particular is an immunodominant antigen for which the antibody response increases in concordance with natural exposure. (Figure 3A) However nucleocapsid is not a component of the mRNA vaccines and consequently there is no vaccine-induced increase in Ab against this antigen. Accordingly, anti-spike antibody levels increased in vaccinees while the nucleocapsid protein Ab level remained constant between Jan and March 2021. (Figure 3B) This suggested to us that anti-nucleocapsid antibodies can be used as a biomarker of prior natural exposure within a population of seropositive vaccinees.

Natural exposure and mRNA induced antibody profiles; anti-nucleocapsid Ab biomarker of natural exposure

Data from 3,347 specimens collected from Santa Ana residents in December 2020 are shown in the heatmap Figure 3A. The level of antibody measured in each specimen against each antigen is recorded as Mean Fluorescence Intensity (MFI) according to the graduated scale from 0 to 60,000. In order assess the seroreactivity, we utilized a Random Forest based prediction algorithm that used data from a well characterized training set (well characterized pre-CoV seronegatives collected in 2019 and PCR-confirmed positive cases) to classify the samples as seroreactive or not seroreactive (de Assis, R.R. 2021[a, b]). In summary, these alorithms were constructed to classify SARS-CoV-2 serostatus using reactivity of multiple antigens to maximize sensitivity and specificity. With these machine learning algorithms, the samples were classified as either SARS-CoV-2 seropositive, clustered to the left, or seronegative and clustered to the right, with a high accuracy. Seropositive specimens recognize nucleoprotein and full-length spike. RBD segments are recognized less well.

The heatmap in Figure 4B shows reactivity of specimens from 750 UCIMC HCW, 93% were seropositive, of whom most were vaccinated. The anti-SARS-CoV-2 Ab reactivity induced by vaccination (panel B) differs from the Ab profile induced by natural exposure (panel A). The vaccine induces higher Ab levels against the RBD containing segments compared to the level induced by natural exposure in the Santa Ana cohort.

Since all adults in these cohorts are exposed to seasonal colds and influenza virus infections and seasonal vaccinations, all the individuals have baseline Ab levels against common-cold CoV and influenza. Thus, background Ab levels against all Common CoV and influenza antigens are equivalent in both the Santa Ana and HCW groups irrespective of whether they are COVID seropositive or not.

The heatmaps show that seropositive vaccinees in the HCW cohort can be classified into two groups, according to whether they are seropositive for SARS-CoV-2 NP or not (Figure 4B), whereas most of the naturally exposed population (Figure 4A) shows high reactivity to both SARS-CoV-2 NP and full-length spike (S1+S2) .

mRNA vaccines induce higher Ab levels and greater Ab breadth than natural exposure to infection

Mean MFI signals for each of the novel coronavirus antigens in the natural exposure (actOC and Santa Ana Cares) and vaccination (HCW) groups are plotted in Figure 5A. Natural exposure in seropositive people induces Abs against NP and all of the spike fragments (RBD, S1, S2 and S1+S2) with highest levels against NP, full-length spike (S1+S2) and the S2 domain. The S1 and RBD domains` antibody reactivity are lower in naturally exposed individuals.

Vaccinated individuals have high Ab levels against full-length spike and the S2 domain of SARS-CoV-2 Spike, and significantly higher levels of Ab against S1 and the RBD domains. In natural exposure there was no significant cross-reactivity against SARS S1 or RBD domains. However the vaccine induced significant cross-reactive Abs against the SARS spike. Cross-reactivity against SARS NP and full-length MERS S protein is evident in both the natural exposure and vaccinated groups. These results show that the Ab responses against Spike RBD variants are significantly elevated in vaccinated individuals compared to those naturally exposed. Vaccination induces more robust Ab response than natural exposure alone, suggesting that those who have recovered from COVID may well benefit from the vaccination.

Nucleocapsid protein is a biomarker associated with natural exposure

Unlike the natural exposure group that reacts uniformly to both nucleoprotein and full-length spike, vaccinees can be separated into two distinct groups of those who react to NP and those who do not. Natural exposure induces a dominant Ab response against the nucleocapsid protein (NP), but since NP is not in the vaccine, there is no vaccine induced response against it. In this way vaccinated people who had a prior natural exposure can be classified because they have Abs to NP. Vaccinated people who were never previously exposed lack Abs against NP. In this way seropositive HCW vaccinees can be separated into NP negative and NP positive groups.

The results in Figure 5B compare the Ab responses against the novel coronavirus antigens between the NP positive and NP negative vaccinees. The mean MFI for NP from SARS –CoV-2, SARS and MERS are low because the groups were selected on that basis. All of the other spike variants are elevated in the NP positive group compared to the NP negative group. This data further supports the advice that people who were previously exposed will still benefit from getting vaccinated.

Progression of the prime and boost responses differ between individuals

Figure 6 shows results of longitudinal specimens taken at regular intervals from 9 individuals pre- and post-mRNA vaccination. Each person differs in time course, and the response to the Prime and the Boost. Subject #1 had a weak response to the prime and a stronger response to the boost. #2 responded with a strong reactivity to both the prime and the boost with a clear increase in antibody levels for the spike variants. #3 is a recovered confirmed COVID-19 case. As expected, this individual showed an elevated baseline Ab reactivity against NP and all of the SARS-CoV-2 variants. After the fist dose,the individual showed an increase in antibody reactivity, however, no further increase was observed after the boost dose. #4 responded slowly to the prime. Subjects #7, #8 and #9 had elevated NP at baseline and responded rapidly to the prime without significant further increase after the boost.

Five individuals had low baseline NP reactivity that did not change post-vaccination. Four individuals had elevated NP reactivity at baseline which did not change significantly post-vaccination, and one of these individuals was a confirmed recovered COVID case. These results suggest that the peak of antibody production occurred after the boost dose (at least after day 35 post primer / day 5 post boost) and, although a elevated antibody response was observed after the prime, for may individuals, a more robust response was only observed after the boost.

Anti-spike Ab titers induced by the mRNA vaccine are higher than those induced by natural exposure

COVAM measurements taken at a single dilution of plasma can be used as a parameter to compare relative Ab titers between individual specimens. Individual samples can also be titered by serial dilution to obtain a more quantitative measurement. In Figure 7A, 2 convalescent plasmas from recovered COVID cases, and pre- and post-boost vaccination plasmas from Subject #5 were titered. The curves are generated by making 8 half-log serial dilutions of the plasmas before probing the COVAM arrays. These curves highlight the observation that high titers against NP are present in convalescent plasma that are lacking in the vaccinees.

Figure 7B plots the midpoint titers of 10 SARS-CoV-2 antigens in 4 convalescent plasmas and pre- and post-boost plasmas from 2 vaccinees. Convalescent plasma vary

in their titers against NP and full-length spike. The vaccinees lack Ab against NP and have significantly higher titers after the boost against all of the spike antigens compared to convalescent plasma. A summary of the midpoint titers is available in supplementary Table 1.

Discussion

In this study, we compared antibody responses induced by SARS-CoV-2 natural exposure with the responses induced by the mRNA vaccines. Pre-vaccine natural exposure data was obtained from specimens from large serial cross-sections from residents of Orange County and the city of Santa Ana, CA, [8] and from mRNA vaccinated healthcare workers at the UCI Medical Center participating in an aggressive vaccination campaign. Within weeks of administration, the mRNA vaccines induced higher Ab levels against spike proteins than observed after natural exposure. These results coincide with equally remarkable clinical trial data showing rapid induction of mRNA protective efficacy on a similar timescale. [1, 2]

The UCI Medical Center achieved a very rapid introduction of the vaccine beginning on December 16, 2021. Within 5 weeks 78% of the individuals tested were seropositive for spike and 3 months later 99% of a March 2021 cross sectional sample was positive. These results illustrate the high vaccine uptake and the extent of antibody response to the vaccine in this population.

mRNA vaccines induce higher Ab levels and greater Ab breadth than natural exposure to infection and differences were particularly notable against the RBD domain. Out of a collection of 3,473 specimens collected from the Santa Ana Cares study in December 2020 we classified 920 as seropositive due to natural exposure before the vaccine was introduced. In February we had a similar number of vaccine induced seropositive healthcare workers. The virus uses the spike RBD domain that binds to the ACE2 receptor on respiratory cells to enter and infect them. Vaccinated individuals had significantly elevated Ab levels against RBD domain segments, supporting the protective immunity induced by this vaccine as previously published. [1, 2] To account for this difference between natural exposure and the vaccine, the virus may have evolved to conceal the RBD epitope to evade immune recognition. The mRNA vaccine produces a protein conformation that better exposes the RBD epitope to the immune system.

In addition to inducing increased Ab levels again SARS-CoV-2 RBD, the mRNA vaccine induced cross-reactive responses against SARS spike and SARS RBD. Conversely, natural exposure did not induce a cross-reactive response against the SARS spike and SARS RBD. The weak anti-RBD response induced by natural exposure may provide a mechanism for new variants to enter the population. This result can be interpreted based on immune selection pressure. Importantly, the mRNA vaccine induces a marked cross-reactive response against SARS spike, indicating that the mRNA vaccine adopts

a conformation that presents cross-reactive epitopes to the immune system. This effect of the mRNA vaccine to induce cross-reactivity against diverse CoV strains is encouraging, providing further evidence that it may be effective against emerging virus variants.

Antibody recognition induced by natural exposure against the NP from SARS-CoV-2 and SARS is concordant with an R^2 value of 0.85. This may indicate a relative lack of selective pressure on this antigen during evolution of these two CoV species. Conversely, the anti-spike response induced by natural exposure does not cross-react against SARS spike or SARS RBD domain indicating immune selection pressure across these strains because of the importance of this epitope in the infection process.

Anti-nucleocapsid Ab is a biomarker of natural exposure to SARS-CoV-2 and can be used to distinguish individuals in a vaccinated population who have been previously exposed to the virus. The nucleoprotein is not present in currently used vaccines. Our data also suggests that people who have had a prior exposure to the virus mount a stronger immune response to the vaccine than those whose immune response has not yet been primed by a previous exposure or vaccination.

These results may also have relevance for both the dose response hypothesis and with regard to herd immunity. Several authors have suggested that disease outcomes may be related to the dose inoculum, with individuals being exposed to inocula with higher virus loads potentially having more severe disease outcomes. [9] While the currently used vaccines in this setting do not rely on viral materials, they do offer a glimpse into controlled high level exposure to proteins that are specific to SARS-CoV-2. Our results show that individuals who have been vaccinated mount higher across-the-board antibody responses than those who have been exposed to variable viral inocula (i.e. through natural exposure). Second, the variable antibody responses among the pre-vaccine population may also indicate that immune responses to natural infections are not as strong as those among individuals who have been vaccinated. This could also indicate that immunity from naturally acquired infections is not as strong as that acquired from vaccination, with potential relevance for reaching and maintaining herd immunity. We should not assume that previously infected individuals are immune or that they cannot transmit the virus.

The original influenza nucleic acid vaccination report used the nucleoprotein antigen because it was conserved across influenza subtypes and it would therefore be a more universal vaccine [10]. The experiment was successful, it was universally effective across diverse strains, and it implicated a cell mediated component, killing of infected cells, in the observed efficacy. As reported for influenza, a more universal SARS CoV vaccine may include the nucleocapsid protein antigen.

Individuals differ in the progression of response to the mRNA prime and boost. Some have a weak response to the prime and experience a substantial effect of the boost. To account for these differences, the group of vaccinees that are NP positive also have

significantly higher vaccine induced responses than the NP negative individuals. This effect is also evident from the small sample of longitudinal specimens we collected from lab members, those with elevated baseline NP reacted more rapidly against the antigens. In the small sample of logitudinal specimens, anti-spike Ab titers induced by the mRNA vaccine are higher than those induced by natural exposure

Serological assays for SARS-CoV-2 are of critical importance to identify highly reactive human donors for convalescent plasma therapy, to investigate correlates of protection, and to measure vaccine efficacy and durability. Here we describe results using a multiplex solid phase immunofluorescent assay for quantification of human antibodies against 37 antigens from SARS-CoV-2, other novel and common coronaviruses, and influenza viruses that are causes of respiratory infections. This assay uses a small volume of blood derived from a finger stick, does not require the handling of infectious virus, quantifies the level of different antibody types in serum and plasma and is amenable to scaling. Finger stick blood collection enables large scale epidemiology studies to define the risk of exposure to SARS-CoV-2 in different settings. Since the assay requires 1 microliter of blood it is also practical for monitoring immunogenicity in small animal models. After probing more than 8,000 pre- and post-vaccination specimens our results confirm that the mRNA vaccine can be used in an aggressive and targeted vaccination campaign to immunize large groups within a matter of weeks.

There are stark differences between actionable interpretation of molecular PCR results and the serological results like those reported here. PCR tests answer the question whether a person has virus in their respiratory secretions as a confirmatory test accounting for the cause of COVID symptoms. It is a useful test in settings where there is high incidence of active infection, patients experiencing symptoms, household contacts, and for contact tracing. Serological tests address different questions of whether the individual has an immune response to the virus, could I have immunity to the COVID 19 virus, how long does it last, do I need the vaccine if I had COVID, can I go to work yet, which vaccine is better, and when do I need another shot.

The concept of nucleic acid vaccines appeared 30 years ago after it was shown that plasmid DNA and RNA could be injected into mouse skeletal muscle tissue in vivo and the encoded transgenes were expressed at the injection site. [11, 12] After intramuscular (IM) injection of a plasmid encoding HIV gp120, induction of anti-gp120 Abs was reported[13]. That was followed by a 1993 report showing efficacy of an influenza nucleic acid vaccine in a rodent model[10]. This was a nucleocapsid based nucleic acid vaccine that induced cross-subtype protection against both group 1 and group 2 viruses (A/PR/8/34 (H1N1) and A/HK/68 (H3N2)). The utility of cationic lipids for gene delivery was discovered and reported in 1987 [14] and synthetic self-assembling lipoplexes for gene delivery described[15-17]. These results spawned a branch of gene therapy science, and an NIH study section, Genes and Drug Delivery (GDD) was established in 2002 that continues to support this research emphasis. Since then

synthetic gene delivery system research and nucleic acid vaccine science has flourished.

DNA vaccines were the first nucleic acid vaccines to be manufactured and tested on a pharmaceutical scale [18, 19]. The mRNA vaccines that are being distributed so widely today may seem to have suddently emerged, but there has been 30 years of scientific discovery, discourse and development, work from hundreds of scientists, numerous biotechnology companies and billions of public and private dollars invested enabling this effective response with a vaccine at this moment.

Supplementary Methods

Coronavirus Antigen Microarray (CoVAM) Report

This document describes the pipeline used to analyze the COVAM array and generate the individual reports.

Step 1: Data pre-processing

The first step of the analysis is importing all data into the R environment. The sample set containing the known negative and known positive controls, here named “Control Set”, is loaded separately from the sample set being analyses.

Following this step, to prevent errors when addressing specific columns, or samples, all spaces are removed both from the column names from all data sets imported, as well as from the Unique sample IDs reference from the meta data files.

On the data processing steps, the following are performed:

From the raw data, the signal to noise ration (SNR) is calculated. The SNR is calculated as the median signal intensity of a given spot divided by the background signal of the vicinity surrounding area. For the quality check purposes, the mean SNR is Calculated only for spots with MFI over 20,000. Samples with a mean SNR below 2 are flagged for further visual inspection or for reprobing.

After calculating the men SNR, the control spots are then assessed. First, for each sample, and each antigen (printed in triplicates), the first and third quartile as well as interquartile range (IQR) are calculated for the control spots. The a upper MFI limit of 1.5 times the IQR over the third quartile and a lower limit of 1.5 times the IQR bellow the first quartile are defined. Spots outside this range are removed and replaced with the mean MFI of the remaining replicates of the spot.

Next, a similar approach is applied to flag samples for which the overall control spots distribution is out of range (2*IQR + third Quartile for the upper limit and first quartile – 2*IQR for the lower limit). For this, all controls spots of a given sample are used. Out of range samples are flagged for further visual inspection or reprobing.

Finally, the printing buffer background reactivity is subtracted from each spot and the samples are normalized.

Step 2: Normalization

Data normalization is performed in two steps. First The control spots are normalized against the training set using the Quantile Normalization method. This allows to calculate a normalization factor that will be used to rescale the data to match the training set and preserving the individual reactivity diversity. After normalizing the control spots, their sum is calculated. A rescaling factor is calculated by dividing the sum of the normalized control spots of the training set by the sum of the normalized control spots of each sample. The resulting factor is then multiplied by the reactivity of each spot resulting in a rescaled data frame. The mean reactivity of the normalized data is then calculated.

Step 3 a: Prediction models

Previous to the sample analysis, the prediction models were constructed using a sample set composed by samples with known diagnosis for COVID-19. These samples are both Negative controls (samples collected before the pandemic) and Positive controls (Samples from individuals diagnosed for COVID-19 by PCR). This control set is heer referred to as Training Set.

The Construction of the prediction models was performed as following.

1. Data is pre-processed and normalized as described above.
2. The reference data set was decomposed into a vector using the function 'unmatrix' from the package gData (version 2.18.0).
3. A mixture model is calculated for the vector using the function 'normalmixEM' from the package 'mixtools' (version 1.2.0).
4. A cutoff is then calculated as 3 standard deviations over the mean of the negative signal curve.
5. Wilcox test for each antigen was performed comparing the positive controls and negatives control, considering significant, antigens with $p < 0.05$.

following the selection of seropositive antigens, an optimal predictive combination of these antigens was selected. (that left us with 7 antigens as seropositive for IgG, and 8 for IgM).

The selection was performed as follows:

1. For every possible combination of the seropositive SARS-CoV-2 antigens from 1 all (7 for IgG and 8 for IgM), the reference set was randomly divided into a training and a testing sets at a 70%/30% ratio.
2. A logistic regression was generated using the reference set. The regression was generated using the function 'glm' of the 'stats' package (version 4.0.0).and a ROC curve was calculated (package pROC version 1.16.2).
3. The optimal coordinates of the ROC curve were obtained based on the 'youden index', by prioritizing the specificity.
4. The coordinates were obtainded using the function 'coords' from the pROC library. The coordinates are obtainded in a table format with each

row containing a regression threshold and its related specificity and sensitivity.

5. The coordinates were then subset to represent specificities of 0.95 or higher. A threshold was then defined as the threshold on the coordinate with the highest specificity on the subset.
6. A logistic regression was then calculated using the testing set and each sample classified as negative or positive by comparison with the threshold.
7. A confusion matrix was calculated by comparing the predicted outcomes and the known classifications (“known negative” or “Known positive”) and the prediction specificity and sensitivity stored into a vector.
8. This analysis was repeated 1000 times and the sensitivity and sensitivity calculated as the mean predicted performance of all repetitions.

The performance outcome for each antigen combination was analyzed and a selection of the best performing combinations was made based on the specificity and sensitivity. The selected candidates were then tested using the full reference sample set. The test was performed as follows:

1. A logistic regression for each antigen combination candidate using the full reference set. Then a ROC curve was calculated and the coordinate table with all curve points was obtained.
2. The coordinates of each candidate were compared in order to select the candidate with the highest sensitivity, given a fixed specificity of 1 (100%).

In addition to the logistic regression model, a Random Forest model was constructed using all reactive antigens.

Step 3 b: Reports.

After Data Normalization, the predictions models, constructed as described above, are loaded and reactivity predictions are performed using Random Forest and Logistic Regression for the multi antigen combinations. In addition to the multi antigen predictions, a prediction for each single SARS-CoV-2 antigen was performed for every sample, for both IgG, and IgM. These predictions were performed using the threshold calculated using the optimal ‘youden’ index. Every sample can be classified as reactive or not reactive for each single SARS-CoV-2 antigen.

The report phase consists on the output of single pdf files with the individual subject predictions and interpretation. The file consists on a brief explanation of the array on the first page, as well as some information on the performance of the array with the current settings. In addition, on the first page there is a short disclaimer of the scope and limitations of the assay.

The second page consists of a table for all the SARS-CoV-2 antigens with their ROC predictions. These predictions are for a qualitative understanding of one’s reactivity and may not directly correlate with the multi antigen prediction.

The Multi antigen prediction, or the sample classification into the three reactive groups, is presented also on a short table displaying the prediction of IgG and IgM separately.

The overall sero-reactivity of the sample to all antigens is depicted on two graphs on the second page. One showing the reactivity for IgG and one for IgM.
On each graph, the individual`s reactivity is represented as dots with its standard errors. For reference, a red line representing the positive control mean reactivity with it`s confidence interval, as well as a blue line representing the negative controls mean reactivity with it`s confidence interval are also plotted.

References

1. Polack, F.P., et al., *Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine.* N Engl J Med, 2020. **383**(27): p. 2603-2615.
2. Baden, L.R., et al., *Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine.* N Engl J Med, 2021. **384**(5): p. 403-416.
3. Khan, S., et al., *Use of an Influenza Antigen Microarray to Measure the Breadth of Serum Antibodies Across Virus Subtypes.* J Vis Exp, 2019(149).
4. de Assis, R.R., et al., *Analysis of SARS-CoV-2 antibodies in COVID-19 convalescent blood using a coronavirus antigen microarray.* Nat Commun, 2021. **12**(1): p. 6.
5. Assis, R., et al., *Distinct SARS-CoV-2 antibody reactivity patterns in coronavirus convalescent plasma revealed by a coronavirus antigen microarray.* Sci Rep, 2021. **11**(1): p. 7554.
6. Hedde, P.N., et al., *A modular microarray imaging system for highly specific COVID-19 antibody testing.* Lab Chip, 2020. **20**(18): p. 3302-3309.
7. Nakajima, R., et al., *Protein Microarray Analysis of the Specificity and Cross-Reactivity of Influenza Virus Hemagglutinin-Specific Antibodies.* mSphere, 2018. **3**(6).
8. Bruckner, T.A., et al., *Estimated seroprevalence of SARS-CoV-2 antibodies among adults in Orange County, California.* Sci Rep, 2021. **11**(1): p. 3081.
9. Van Damme, W., et al., *COVID-19: Does the infectious inoculum dose-response relationship contribute to understanding heterogeneity in disease severity and transmission dynamics?* Med Hypotheses, 2021. **146**: p. 110431.
10. Ulmer, J.B., et al., *Heterologous protection against influenza by injection of DNA encoding a viral protein.* Science, 1993. **259**(5102): p. 1745-9.
11. Wolff, J.A., et al., *Direct gene transfer into mouse muscle in vivo.* Science, 1990. **247**(4949 Pt 1): p. 1465-1468.
12. Malone, R.W., P.L. Felgner, and I.M. Verma, *Cationic liposome-mediated RNA transfection.* Proc Natl Acad Sci U S A, 1989. **86**(16): p. 6077-81.
13. Felgner, P.L. and G. Rhodes, *Gene therapeutics.* Nature, 1991. **349**(6307): p. 351-2.
14. Felgner, P.L., et al., *Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure.* Proc Natl Acad Sci U S A, 1987. **84**(21): p. 7413-7.
15. Felgner, P.L., et al., *Nomenclature for synthetic gene delivery systems.* Hum Gene Ther, 1997. **8**(5): p. 511-2.

16. Felgner, J.H., et al., *Enhanced gene delivery and mechanism studies with a novel series of cationic lipid formulations.* J. Biol. Chem., 1994. **269**(4): p. 2550-2561.
17. Felgner, P.L., *Nonviral strategies for gene therapy.* Sci Am, 1997. **276**(6): p. 102-6.
18. San, H., et al., *Safety and short-term toxicity of a novel cationic lipid formulation for human gene therapy.* Hum Gene Ther, 1993. **4**(6): p. 781-8.
19. Felgner, P.L., *DNA vaccines.* Curr. Biol., 1998. **8**(16): p. R551-553.

Acknowledgements

This work was supported by the HDTRA1-16-C-0009, HDTRA1-18-1-0035, HDTRA1-18-1-0036 and a University of California, Irvine CRAFT-COVID grants. 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.

Author contributions

The coronavirus antigen microarray (COVAM) was designed by S. Khan and P. Felgner and was constructed by R. Nakajima, A. Jasinskas and R. Assis at UCI. The specimens probing on the COVAM was performed by A. Jain at UCI. Data analysis was performed by R. de Assis at UCI. The manuscript and figures were prepared by Philip Felgner and R. de Assis with input and approval from all other authors.

Competing Interests

The coronavirus antigen microarray is intellectual property of the Regents of the University of California that is licensed for commercialization to Nanommune Inc. (Irvine, CA), a private company for which Philip L. Felgner is the largest shareholder and several co-authors (de Assis, Jain, Nakajima, Jasinskas, Davies, and Khan) also own shares. Nanommune Inc. has a business partnership with Sino Biological Inc. (Beijing, China) which expressed and purified the antigens used in this study. The other authors have no competing interests.

Figure Legends

Figure 1. Study Design. Finger stick blood specimens were collected at weekly intervals from drive-through locations around Orange County and from healthcare workers at the University of California Medical Center. Individual samples were probed on the COVAM, quantified and analyzed. Personalized serology reports were generated and linked to individual QR codes for everyone to access their own report.

Figure 2. The content of the Coronavirus Antigen Microarray is shown. There are 10 SARS-CoV-2 antigens, 3 SARS, 3 MERS, 12 Common COV, and 8 influenza antigens. Each antigen is printed in triplicate and organized as shown on the images with Orange boxes around the SARS-CoV-2 antigens, Blue SARS, Green MERS, Yellow Common CoV, and Purple for Influenza. Three different samples are shown, a Negative Pre-CoV, Natural Infection (actOC), and a sample from an mRNA vaccinee (HCW). The Pre-CoV sample has negligible reactivities to SARS-CoV-2, SARS and MERS, whereas Natural Infection and the vaccinees have significant Abs against the novel CoV. The red-white arrows point to the nucleocapsid protein which detects antibodies in naturally exposed people but not in the vaccinees.

Figure 3. A. Finger stick blood specimens were collected from Orange County, Santa Ana in July (2,979 specimens) and Santa Ana in December (3,347 specimens), and seroprevalence measured on the COVAM array. B. Seroprevalence in cross-sections from the UCI Medical Center was measured by COVAM analysis at 2 time points before the start of the mRNA vaccination campaign on December 16, 2020 and monthly intervals in 2021. The gray bar is the COVAM seroprevalence prediction and the blue bar is the nucleocapsid protein seropositivity.

Figure 4. The heat maps show all of the IgG reactivity data from 3,347 pre-vaccination specimens collected from Santa Ana in December 2020 (A), and 750 post-vaccination specimens collected from the UCIMC in February (B). The 37 antigens are in rows and the specimens are in 3,347 columns for panel A and 750 columns for panel B. The level of antibody measured in each specimen against each antigen is recorded as Mean Fluorescence Intensity (MFI) according to the graduated scale from 0 to 60,000. Red is a high level, white a low level and black is in between. Samples are classified as either SARS-CoV-2 seropositive clustered to the left (orange bar) or seronegative and clustered to the right (blue bar). Panel A. Seropositive specimens recognize nucleoprotein and full-length spike. RBD segments are recognized less well. Panel B. Reactivity of specimens from 750 UCIMC HCW, 93% were vaccinated and seropositive. Panel B. Heatmap shows that seropositive vaccinees in the HCW cohort can be classified into two group whether they are seropositive for nucleoprotein or not, whereas the naturally exposed population (panel A) is uniformly seropositive for both nucleoprotein and full-length spike.

Figure 5A. Mean MFI signals for each of the novel coronavirus antigens in the natural exposure cohort from Santa Ana in December 2020 (actOC) and the February 2021 vaccination group (HCW) are plotted. The figure shows that Ab responses against Spike RBD variants are significantly elevated in mRNA vaccinated people compared to naturally exposed individuals. Vaccination induces a broader and higher titer Ab response than natural exposure alone, so those who have recovered from COVID can be expected to benefit from the vaccination.

Figure 5B. Unlike the natural exposure group that reacts uniformly to both nucleoprotein and full-length spike, vaccinees can be separated into two distinct groups

of those who react to NP and those who do not. Natural exposure induces a dominant Ab response against the nucleocapsid protein (NP), but since NP is not in the vaccine, there is no vaccine induced response against it. In this way vaccinated people who had a prior natural exposure can be classified because they have Abs to NP. Vaccinated people who were never previously exposed lack Abs against NP. The results in Figure 4B compare the Ab responses against the novel coronavirus antigens between the NP positive and NP negative vaccinees. This data further supports the directive that people who are previously exposed will benefit by getting a boost against RBD.

Figure 6. Longitudinal specimens taken at weekly intervals from 9 individuals pre- and post-mRNA vaccination. Individual differ substantially in their response to the prime. Five individuals had low baseline NP reactivity that did not change post-vaccination. Four individuals had elevated NP reactivity at baseline which also did not change significantly post-vaccination; subject #3 was a recovered confirmed COVID case. In this small group, higher baseline NP predicts a higher response after the prime. These results support a directive to get the boost in order to achieve more uniform protection within a population of individuals.

Figure 7A. Convalescent plasmas from 2 recovered COVID cases, and pre- and post-boost specimens from Subject #5 were titered and the titration curves are shown. The curves are generated by making 8 half log serial dilutions of the plasmas before probing 8 separate COVAM arrays. These curves highlight the observation that high titers against NP are present in convalescent plasma that are lacking in the vaccinees. (Red Arrow)

Figure 7B. The midpoint titers of 10 SARS-CoV-2 antigens from 4 convalescent plasmas and plasmas from 2 vaccinees after the prime and after the boost are plotted Convalescent plasma vary in their titers against NP and full-length spike. The vaccinees lack Ab against NP and have significantly higher titers after the boost against all of the spike antigens compared to convalescent plasma.

Supplementary Figure Legends

Supplementary Figure 1.The general analysis pipeline consists of three main steps: the preprocessing, the normalization and then the statistical prediction analysis. The preprocessing include steps like calculation the Signal to Noise Ratio (SNR) and determine if a sample needs to be further checked or re-assayed (due to the background reactivity levels). If successful, samples are successful analyzed for their SNR, the controls spots are checked to remove outlier spots that could skew normalization. Then, the distribution of the control spots is analyzed and low-quality samples (for which the control spots deviate from the expected) are flagged to be re-assayed. Then the samples are normalized, and the mean fluorescence intensity calculated from the average of the 3 replicates in the array. After normalization, a machine learning based algorithm is used to classify each sample as reactive or not reactive to SARS-CoV-2 (using multiple antigens) as well as to individual antigens. Then, individual reports are generated for each sample (this can be in the form of individual pdf files that may be delivered to the subject).

Supplementary Figure 2. After the machine learning classification of each sample individual pdf files containing the results can be generated. The panels in the figure are representative of a typical negative (or non-reactive) result (left panel) and of a typical positive (Reactive) sample (on the right). The data printed on the reports are basic reactivity classification for the SARS-CoV-2 antigens (Only reactive and Non-reactive denominations are given). As well as the machine learning classification (multi antigen classification) denominations. For the multi antigen classification, the results from the logistic regression as well as the results from random forest, as well as the random forest probabilities are given. The multi antigen classification is the main result and is the one used to classify an individual as exposed, or reactive to SARS-CoV-2 as individual antigens alone have a much lower performance in the classification.

Finally, since the COVAM is composed of multiple viruses, the reactivity to the entire array is given to both IgG and IgM. This reactivity is given as the normalized mean florescence intensities and as a reference, the confidence intervals of a known control set of samples (known positives red line and red bands and known negatives blue line and blue bands) are given. Although these reports give a much more comprehensive view of an individual`s reactivity status to SARS-CoV-2, they are intended mainly as a guidance as the COVAM array is not approved by the FDA as a diagnostic test.

Figure 1

Longitudinal Study Design, Sample Collection, and Assay Parameters

Samples tested

Collection	Number	Date
Orange County	2,979	July ’20
Santa Ana	3,347	Dec ‘20
UCI Healthcare Workers	1,060	May '20
UCI Healthcare Workers	313	Dec '20
Vaccination Start Date	***December 16, 2020***	
UCI Healthcare Workers	140	Jan '21
UCI Healthcare Workers	750	Feb '21
UCI Healthcare Workers	157	Mar ‘21
Total	**8,746**	

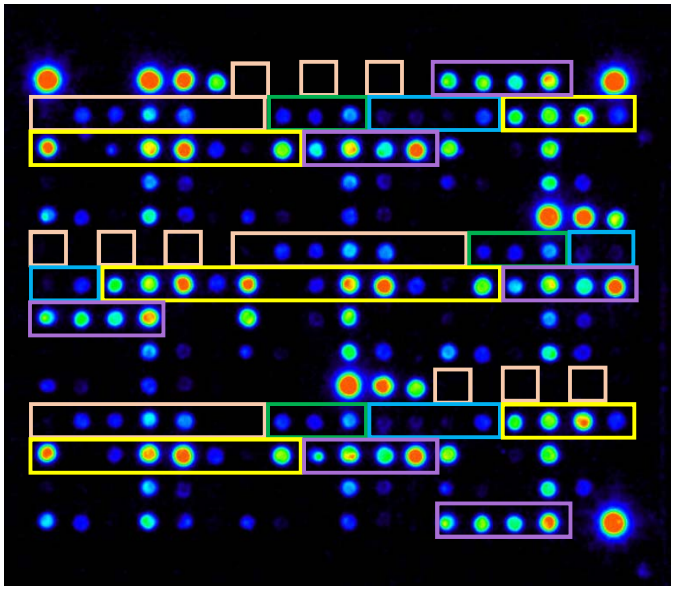
Measurements

Virus	Antigen #
SARS-CoV-2	10
SARS	4
MERS	3
Common CoV	12
Influenza A/B	8
Total	37
Triplicate	111
IgG&IgM	222
Specimens	**8,746**
Measurement#	**1,941,612**

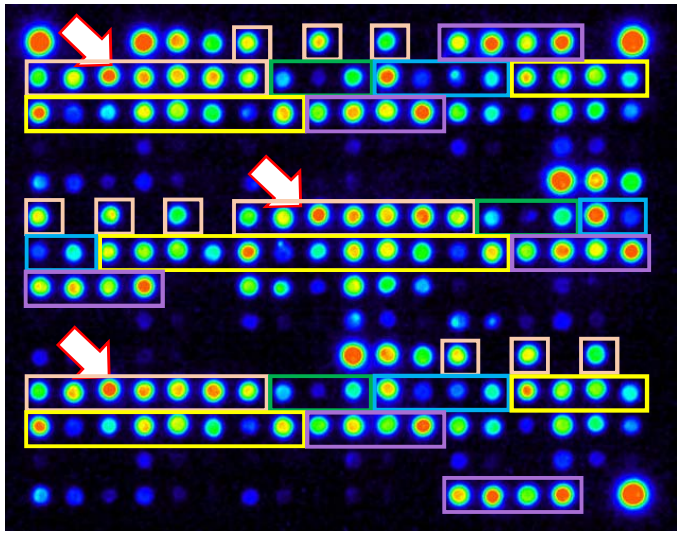
Figure 2 **COVAM** *Coronavirus Antigen Microarray*

		COVAM 4 Antigen Content
SARS-CoV-2	1	Nucleocapsid-His
	2	Spike S1 + S2-His
	3	Spike S2 ECD-His
	4	Spike S1-mFc
	5	Spike S1-His (Bac)
	6	Spike S1-His (HEK)
	7	Spike RBD-mFc
	8	Spike RBD-rFc
	9	Spike RBD-His (Bac)
	10	Spike RBD-His (HEK)
SARS	11	Nucleocapsid-His
	12	Spike S1-His
	13	Spike RBD-His
	14	Spike RBD-rFc
MERS	15	Nucleocapsid-His
	16	Spike S1+S2 ECD-His
	17	Spike RBD-rFc
Common CoV	18	HKU23-Nucleocapsid-His
	19	229E- Spike S1-His
	20	229E- Spike S1+S2 ECD-His
	21	HKU1-Nucleocapsid-His
	22	HKU1-Spike S1-His
	23	HKU1-Spike S1+S2 ECD-His
	24	NL63-Nucleocapsid-His
	25	NL63-Spike S1-His
	26	NL63-Spike S1+S2 ECD-His
	27	OC43-HA esterase-His
	28	OC43-Nucleocapsid-His
	29	OC43-Spike S1+S2 ECD-His
Influenza	30	B/Malaysia-HA1-His
	31	B/Malaysia-HA0-His
	32	B/PHUKET-HA1-His
	33	B/PHUKET-HA0- His
	34	A/Beijing-H1N1-HA1-His
	35	A/Beijing-H1N1-HA0- His
	36	A/Texas-H3N2-HA1-His
	37	A/Texas-H3N2-HA0-His

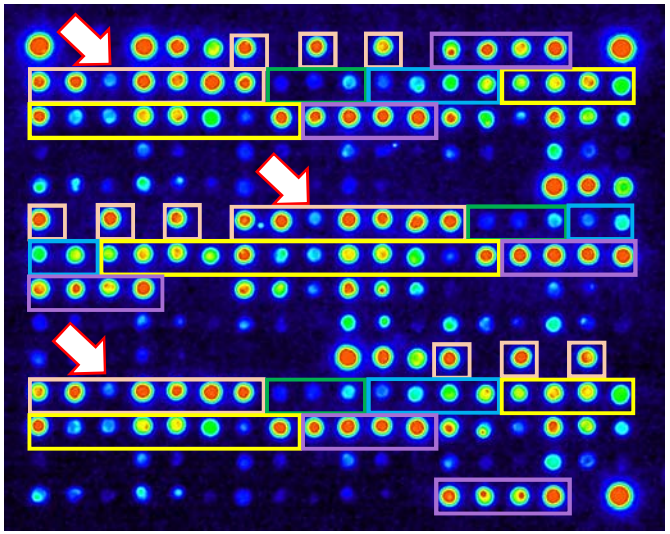
Negative Pre-CoV



actOC Natural Infection



HCW Vaccinated



Antigens

Virus	Antigen #
SARS-CoV-2	10
SARS	4
MERS	3
Common CoV	12
Influenza A/B	8

⇨ Red-white arrows = SARS-CoV-2-NP

Figure 3

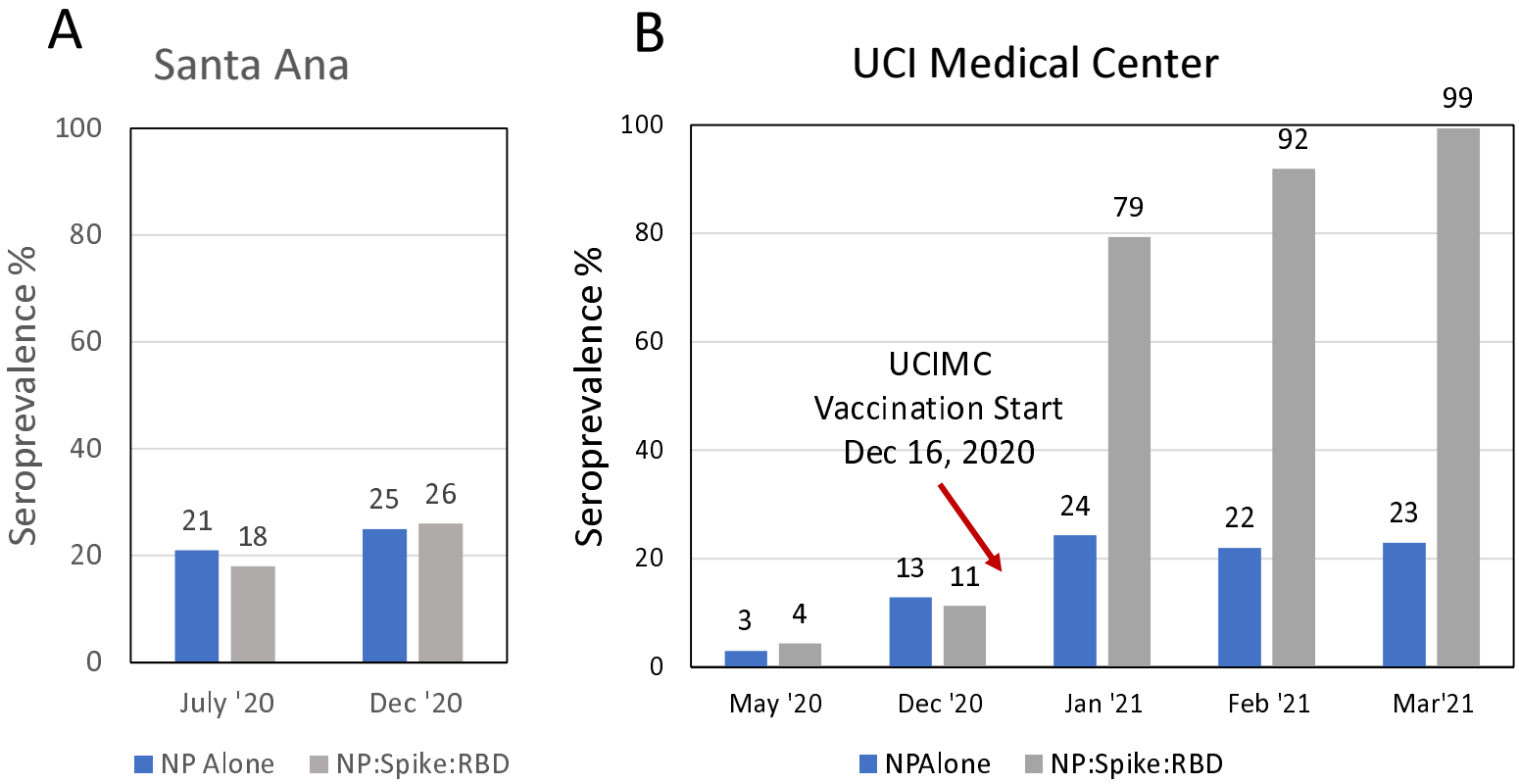

A
Santa Ana
Seroprevalence %
100
80
60
40
20
0
21
18
25
26
July '20
Dec '20
NP Alone
NP:Spike:RBD
B
UCI Medical Center
Seroprevalence %
100
80
60
40
20
0
3
4
13
11
24
79
22
92
23
99
UCIMC
Vaccination Start
Dec 16, 2020
May '20
Dec '20
Jan '21
Feb '21
Mar'21
NPAlone
NP:Spike:RBD

Figure 4

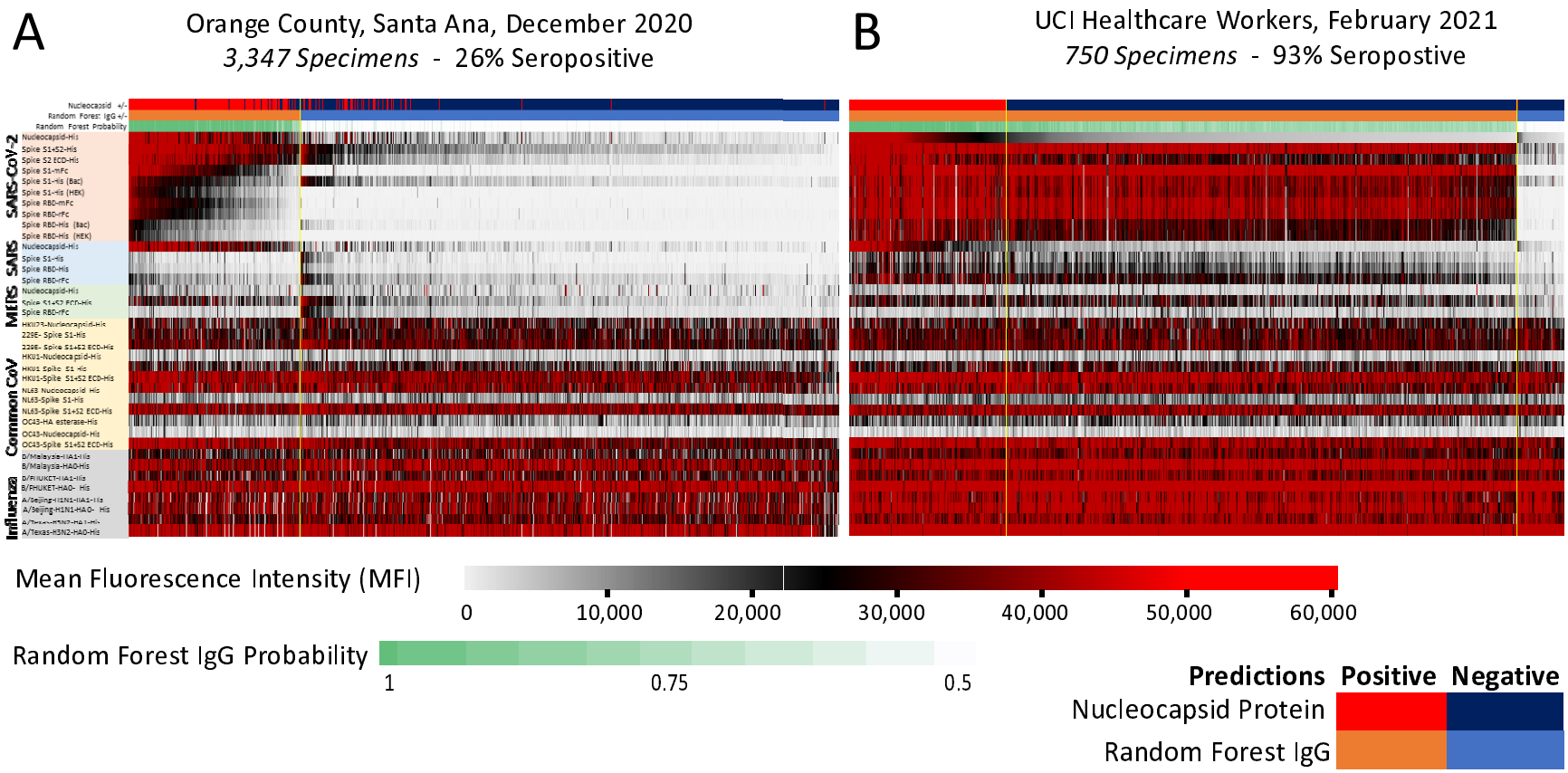

A
Orange County, Santa Ana, December 2020
3,347 Specimens - 26% Seropositive
B
UCI Healthcare Workers, February 2021
750 Specimens - 93% Seropostive
Nucleocapsid +/-
Random Forest: IgG +/-
Random Forest Probability
SARS-CoV-2
Nucleocapsid-His
Spike S1+S2-His
Spike S2 ECD-His
Spike S1-mFc
Spike S1-His (Bac)
Spike S1-His (HEK)
Spike RBD-mFc
Spike RBD-rFc
Spike RBD-His (Bac)
Spike RBD-His (HEK)
SARS
Nucleocapsid-His
Spike S1-His
Spike RBD-His
Spike RBD-rFc
MERS
Nucleocapsid-His
Spike S1+S2 ECD-His
Spike RBD-rFc
Common CoV
HKU73-Nucleocapsid-His
229E- Spike S1-His
229E- Spike S1+S2 ECD-His
HKU1-Nucleocapsid-His
HKU1-Spike S1-His
HKU1-Spike S1+S2 ECD-His
NL63-Nucleocapsid-His
NL63-Spike S1-His
NL63-Spike S1+S2 ECD-His
OC43-HA esterase-His
OC43-Nucleocapsid-His
OC43-Spike S1+S2 ECD-His
Influenza
B/Malaysia-HA1-His
B/Malaysia-HA0-His
B/PHUKET-HA1-His
B/PHUKET-HA0- His
A/Beijing-H1N1-HA1-His
A/Beijing-H1N1-HA0- His
A/Texas-H3N2-HA0-His
Mean Fluorescence Intensity (MFI)
0
10,000
20,000
30,000
40,000
50,000
60,000
Random Forest IgG Probability
1
0.75
0.5
Predictions
Positive
Negative
Nucleocapsid Protein
Random Forest IgG

Figure 5A

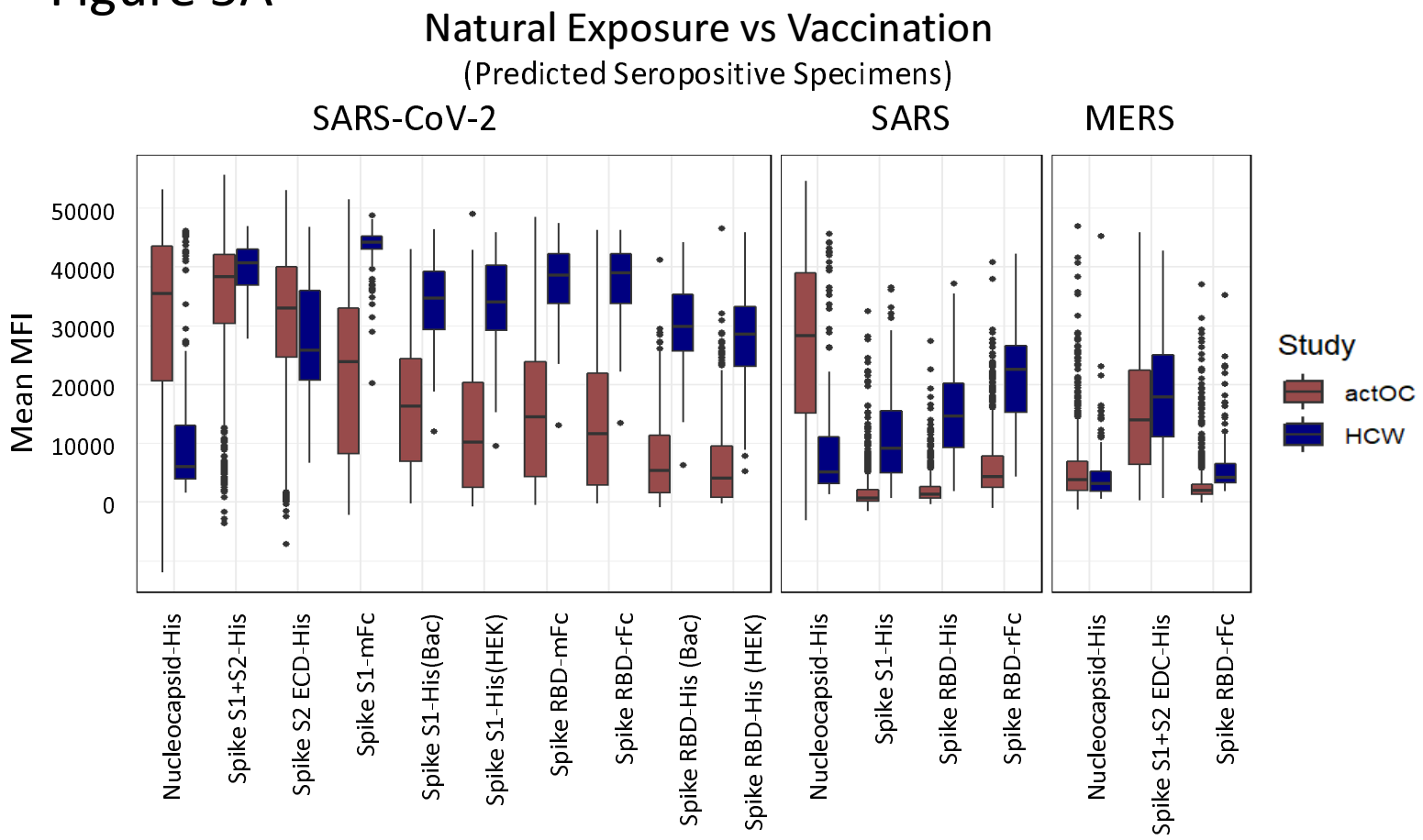

Natural Exposure vs Vaccination
(Predicted Seropositive Specimens)
SARS-CoV-2
SARS
MERS
Mean MFI
50000
40000
30000
20000
10000
0
Nucleocapsid-His
Spike S1+S2-His
Spike S2 ECD-His
Spike S1-mFc
Spike S1-His(Bac)
Spike S1-His(HEK)
Spike RBD-mFc
Spike RBD-rFc
Spike RBD-His (Bac)
Spike RBD-His (HEK)
Nucleocapsid-His
Spike S1-His
Spike RBD-His
Spike RBD-rFc
Nucleocapsid-His
Spike S1+S2 EDC-His
Spike RBD-rFc
Study
actOC
HCW

Figure 5B

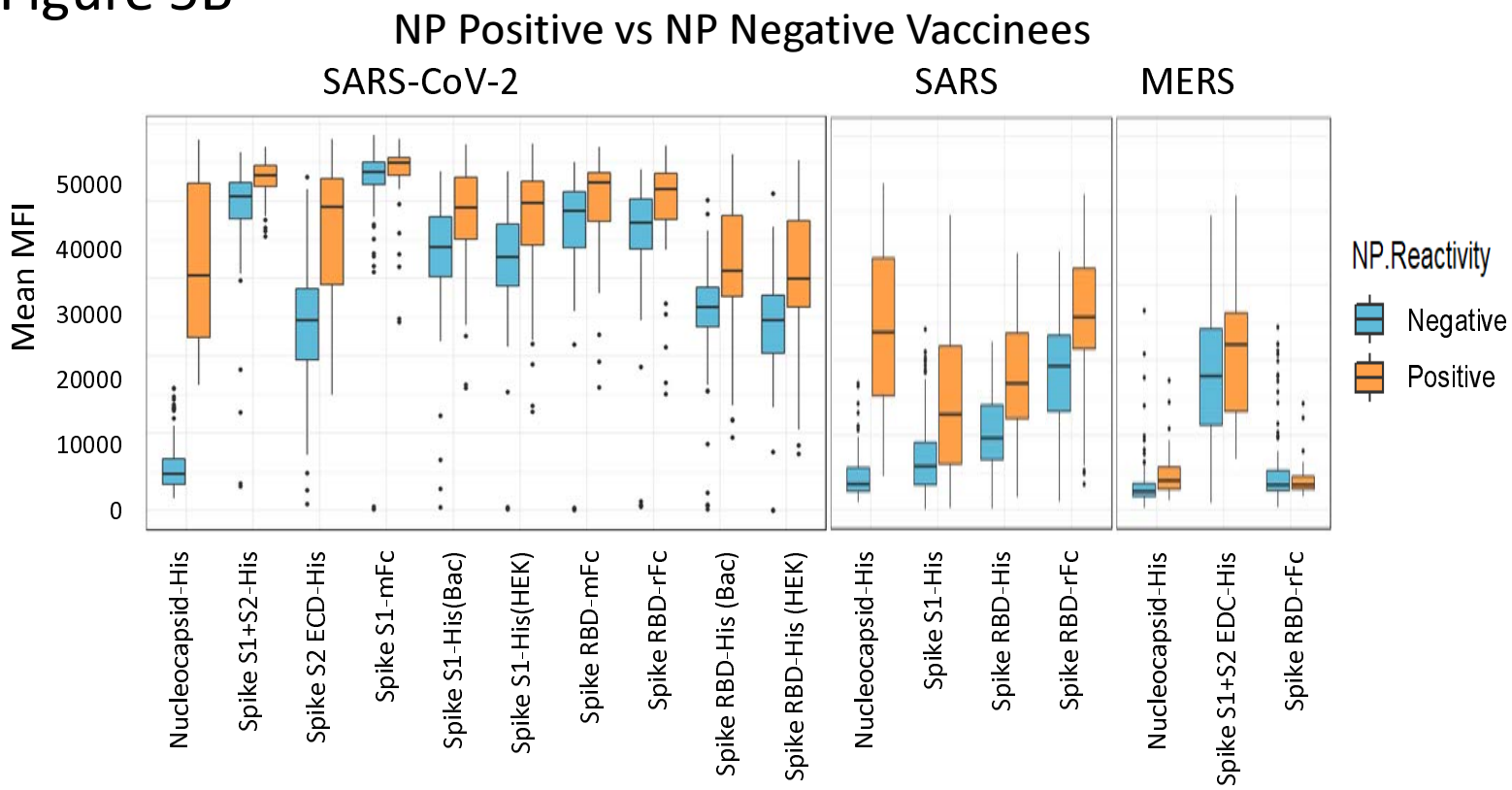

NP Positive vs NP Negative Vaccinees
SARS-CoV-2
SARS
MERS
Mean MFI
50000
40000
30000
20000
10000
0
Nucleocapsid-His
Spike S1+S2-His
Spike S2 ECD-His
Spike S1-mFc
Spike S1-His(Bac)
Spike S1-His(HEK)
Spike RBD-mFc
Spike RBD-rFc
Spike RBD-His (Bac)
Spike RBD-His (HEK)
Nucleocapsid-His
Spike S1-His
Spike RBD-His
Spike RBD-rFc
Nucleocapsid-His
Spike S1+S2 EDC-His
Spike RBD-rFc
NP.Reactivity
Negative
Positive

Figure 6

mRNA Vaccine Induced Ab Response

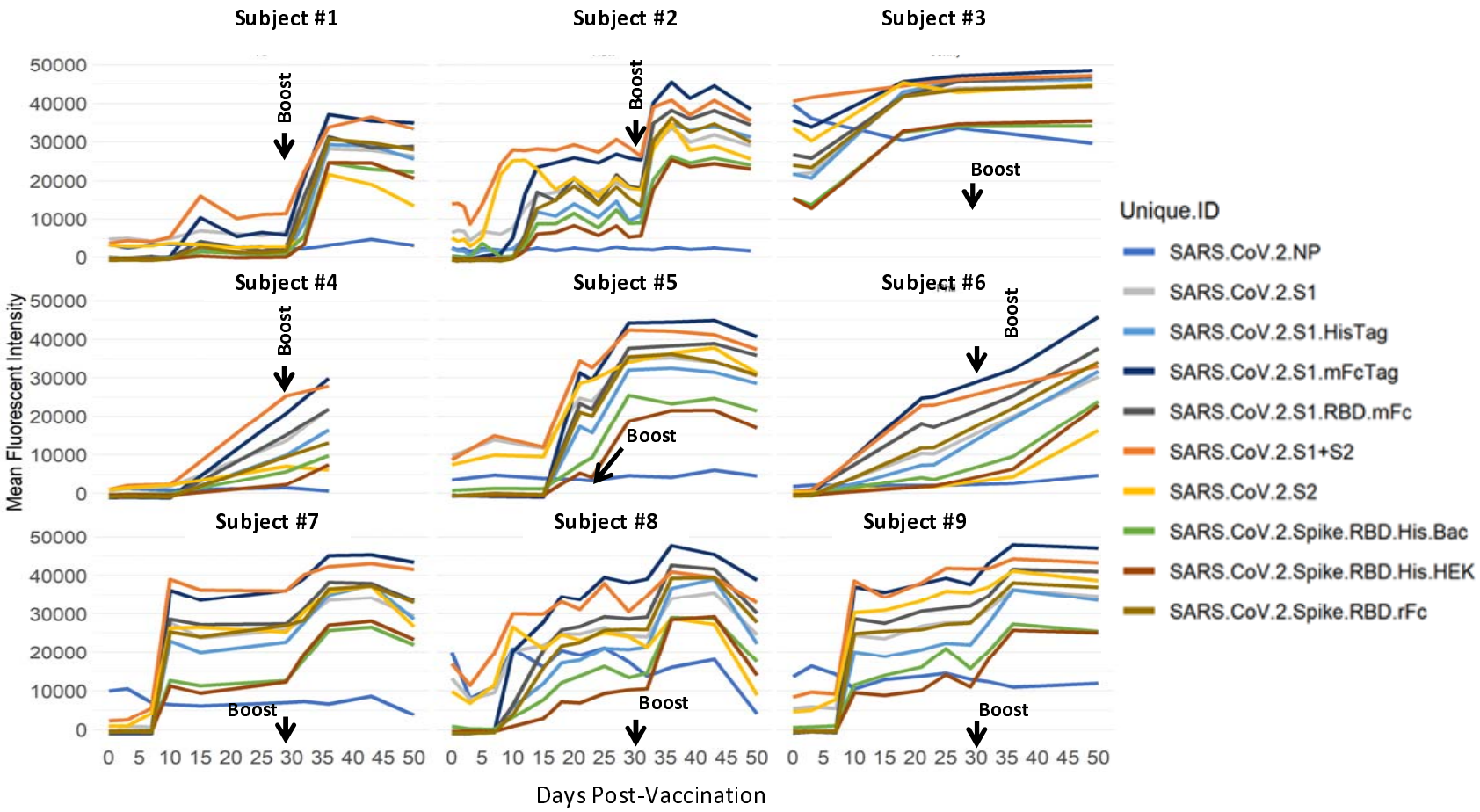
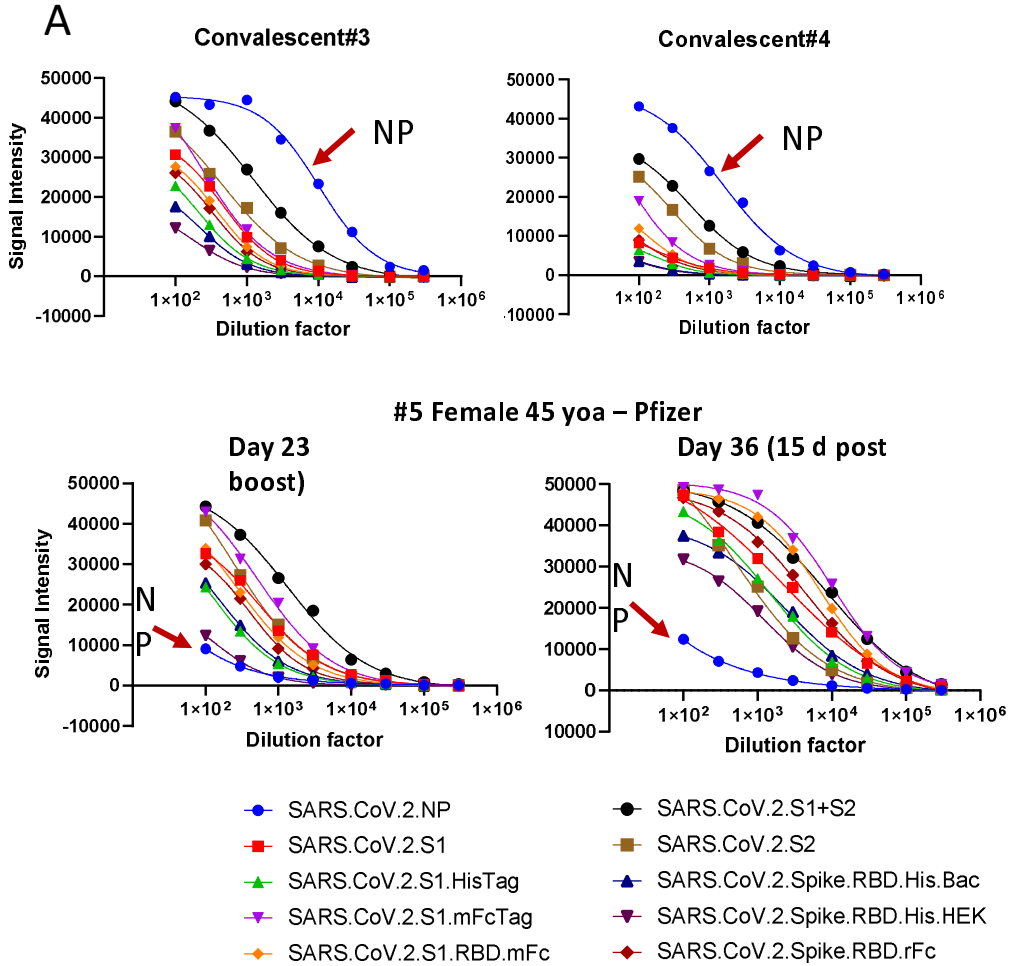
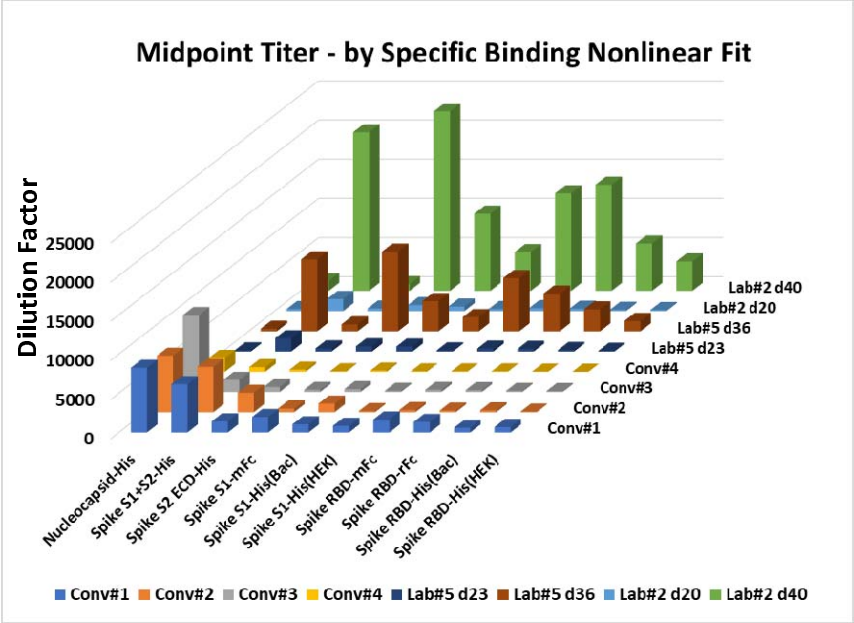


Figure 7

A

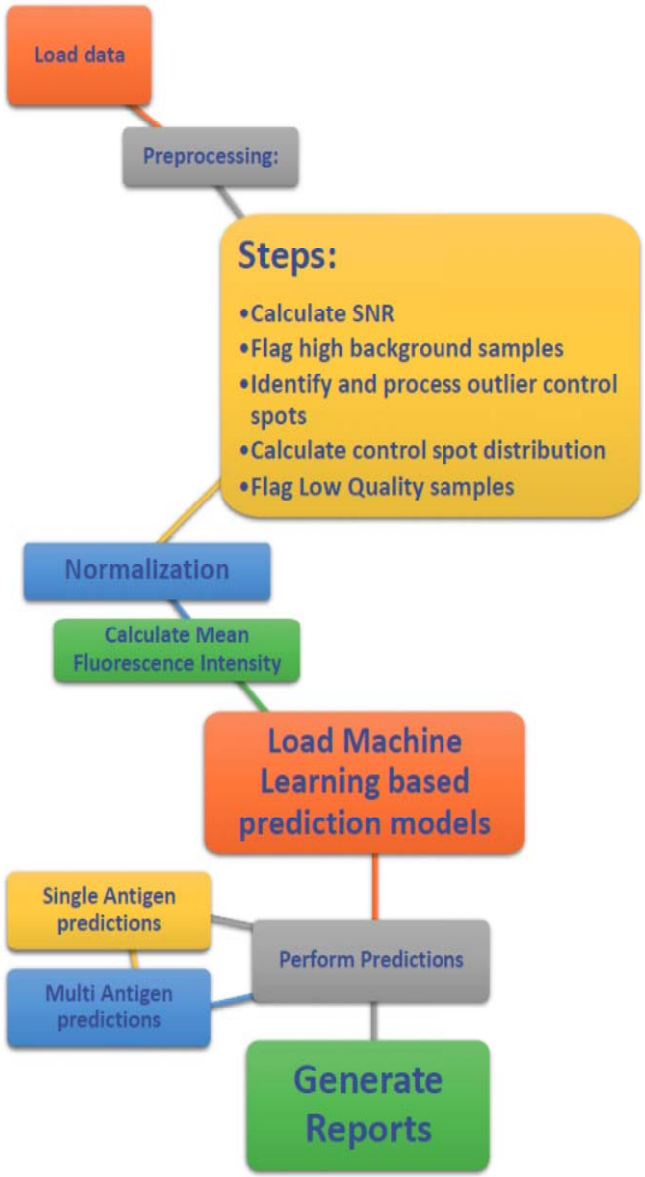

Convalescent#3
Convalescent#4
Signal Intensity
Dilution factor
NP
#5 Female 45 yoa – Pfizer
Day 23 boost)
Day 36 (15 d post
N
P
SARS.CoV.2.NP
SARS.CoV.2.S1
SARS.CoV.2.S1.HisTag
SARS.CoV.2.S1.mFcTag
SARS.CoV.2.S1.RBD.mFc
SARS.CoV.2.S1+S2
SARS.CoV.2.S2
SARS.CoV.2.Spike.RBD.His.Bac
SARS.CoV.2.Spike.RBD.His.HEK
SARS.CoV.2.Spike.RBD.rFc

B


Midpoint Titer - by Specific Binding Nonlinear Fit
Dilution Factor
Nucleocapsid-His
Spike S1+S2-His
Spike S2 ECD-His
Spike S1-mFc
Spike S1-His(Bac)
Spike S1-His(HEK)
Spike RBD-mFc
Spike RBD-rFc
Spike RBD-His(Bac)
Spike RBD-His(HEK)
Conv#1
Conv#2
Conv#3
Conv#4
Lab#5 d23
Lab#5 d36
Lab#2 d20
Lab#2 d40

Supplementary Figure 1

Analysis method and report generation



Supplementary Figure 2

10,000 Individual Patient Specific Reports

Subject: 754551

Single Antigen Logistic Regression

Antigen	IgG	IgM
SARS.CoV.2.NP	Not Reactive	Not Reactive
SARS.CoV.2.S1	Not Reactive	Not Reactive
SARS.CoV.2.S1.HisTag	Not Reactive	Not Reactive
SARS.CoV.2.S1.mFcTag	Not Reactive	Not Reactive
SARS.CoV.2.S1.RBD.mFc	Not Reactive	Not Reactive
SARS.CoV.2.S1+S2	Not Reactive	Not Reactive
SARS.CoV.2.S2	Not Reactive	Not Reactive
SARS.CoV.2.Spike.RBD.His.Bac	Not Reactive	Not Reactive
SARS.CoV.2.Spike.RBD.His.HEK	Not Reactive	Not Reactive
SARS.CoV.2.Spike.RBD.rFc	Not Reactive	Not Reactive

SARS−CoV−2 reactivity profile Classification:
(see detailed explanation on previous page)

Multi Antigen prediction

	IgG	IgM
Logistic Reg.	Non−Reactive	Non−Reactive
RandomForest (RF)	Non−Reactive	Non−Reactive
*RF Probability **	26%	15.2%

*Random Forest algorthm probability of classification as 'Reactive' the probability cutoff is currently set at 70%(IgG)/85%(IgM)

Subject: 755334

Single Antigen Logistic Regression

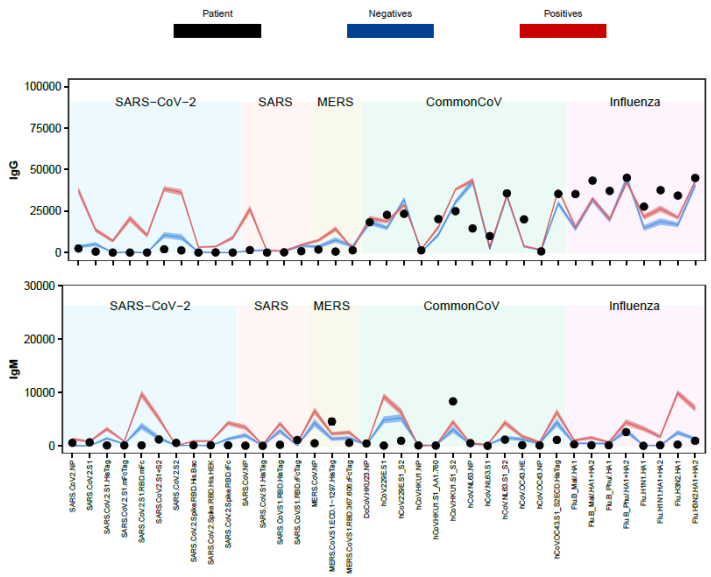
Antigen	IgG	IgM
SARS.CoV.2.NP	Not Reactive	**Reactive**
SARS.CoV.2.S1	**Reactive**	**Reactive**
SARS.CoV.2.S1.HisTag	**Reactive**	Not Reactive
SARS.CoV.2.S1.mFcTag	**Reactive**	**Reactive**
SARS.CoV.2.S1.RBD.mFc	**Reactive**	Not Reactive
SARS.CoV.2.S1+S2	**Reactive**	**Reactive**
SARS.CoV.2.S2	**Reactive**	**Reactive**
SARS.CoV.2.Spike.RBD.His.Bac	**Reactive**	**Reactive**
SARS.CoV.2.Spike.RBD.His.HEK	**Reactive**	Not Reactive
SARS.CoV.2.Spike.RBD.rFc	**Reactive**	**Reactive**

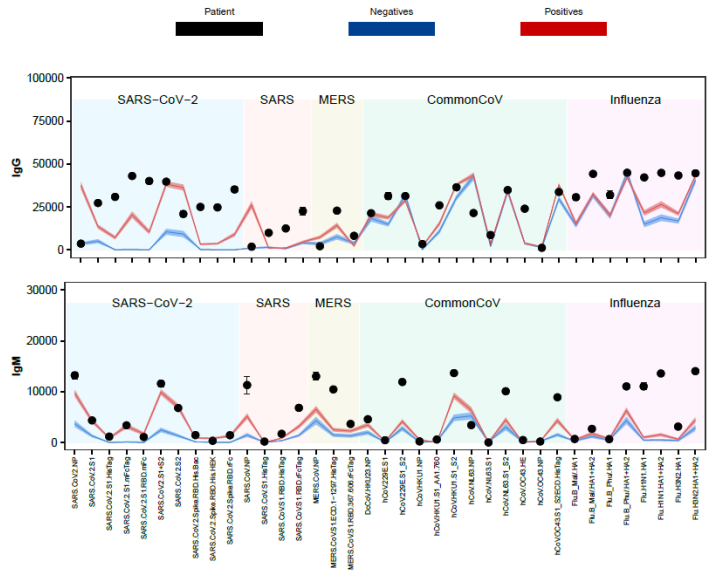
SARS−CoV−2 reactivity profile Classification:
(see detailed explanation on previous page)

Multi Antigen prediction

	IgG	IgM
Logistic Reg.	**Reactive**	**Reactive**
RandomForest (RF)	**Reactive**	**Reactive**
*RF Probability **	77.8%	96.2%

*Random Forest algorthm probability of classification as 'Reactive' the probability cutoff is currently set at 70%(IgG)/85%(IgM)





Supplementary Table 1

Midpoint titer – Specific Binding Saturation Nonlinear Fit

(Dilution Factor)

	Conv#1	Conv#2	Conv#3	Conv#4	Lab#5 d23	Lab#5 d36	Lab#2 d20	Lab#2 d40
Nucleocapsid-His	8251	7210	9814	1763	158	328	266	1285
Spike S1+S2-His	6161	5790	1549	588	1695	9183	1542	20239
Spike S2 ECD-His	1411	2480	631	267	399	884	333	871
Spike S1-mFc	1971	468	276	59	668	10166	760	22910
Spike S1-His(Bac)	1036	1097	366	151	635	3855	576	9843
Spike S1-His(HEK)	835	150	146	53	153	1807	270	4950
Spike RBD-mFc	1645	278	272	52	376	6789	384	12442
Spike RBD-rFc	1341	213	232	54	346	4815	302	13543
Spike RBD-His(Bac)	600	234	139	19	191	2801	171	6039
Spike RBD-His(HEK)	692	71	116	3	85	1323	136	3757