Oral SARS-CoV-2 host responses predict the early COVID-19 disease course

Running title: Implications of oral SARS-CoV-2 detection and antibody responses

Authors: William T Seaman, PhD[1,9,10], Olive Keener[2], Wenwen Mei[8,10], Katie R Mollan, PhD[8,10], Corbin D Jones, PhD[3,10], Audrey Pettifor, PhD[4,10], Natalie M Bowman, MD, MPH[5,10], UNC OBSc Working Group[5,10], Frank Wang, PhD[6], and Jennifer Webster-Cyriaque, DDS, PhD[1,7,9,10]

1 Division of Oral and Craniofacial Health Sciences, University of North Carolina, Chapel Hill, NC
2 North Carolina School of Math and Science, Durham, NC,
3 Department of Biology, University of North Carolina, Chapel Hill, NC
4 Department of Epidemiology, University of North Carolina, Chapel Hill, NC,
5 Division of Infectious Diseases, University of North Carolina, Chapel Hill, NC
6 Biomedomics Inc, Morrisville, NC
7 Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, NC
8 Department of Biostatistics, University of North Carolina, Chapel Hill, NC
9 National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD
10 School of Medicine, University of North Carolina, Chapel Hill, NC

Address for Correspondence:
Dr. Jennifer Webster-Cyriaque DDS PhD
NIAID/NIH
Jennifer.Webster-cyriaque@nih.gov

Abstract (250)

Objectives: Oral fluids provide ready detection of Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and host responses. This study sought to determine relationships between oral virus, oral anti-SARS-CoV-2-specific antibodies, and symptoms.

Methods: Saliva/throat wash (saliva/TW) were collected from asymptomatic and symptomatic, nasopharyngeal (NP) SARS-CoV-2 RT-qPCR+, subjects (n=47). SARS-CoV-2 RT-qPCR, N-antigen detection by immunoblot and lateral flow assay (LFA) were performed. RT-qPCR targeting viral subgenomic RNA (sgRNA) was sequence confirmed. SARS-CoV-2-anti-S protein RBD LFA assessed IgM and IgG responses. Structural analysis identified host salivary molecules analogous to SARS-CoV-2-N-antigen. Statistical analyses were performed.

Results: At baseline, LFA-detected N-antigen was immunoblot-confirmed in 82% of TW. However, only 3/17 were saliva/TW qPCR+. Sixty percent of saliva and 83% of TW demonstrated persistent N-antigen at 4 weeks. N-antigen LFA signal in three negative subjects suggested potential cross-detection of 4 structurally analogous salivary RNA binding proteins (alignment 19-29aa, RMSD 1-1.5 Angstroms). At entry, symptomatic subjects demonstrated replication-associated sgRNA junctions, were IgG+ (94%/100% in saliva/TW), and IgM+ (75%/63%). At 4 weeks, SARS-CoV-2 IgG (100%/83%) and IgM (80%/67%) persisted. Oral IgG correlated 100% with NP+PCR status. Cough and fatigue severity (p=0.0008 and 0.016), and presence of nausea, weakness, and composite upper respiratory symptoms (p=0.005, 0.037 and 0.017) were negatively associated with oral IgM. Female oral IgM levels were higher than male (p=0.056).

Conclusion:
Important to transmission and disease course, oral viral replication and persistence showed clear relationships with select symptoms, early Ig responses, and gender during early infection. N-antigen cross-reactivity may reflect mimicry of structurally analogous host proteins.

Keywords: SARS-CoV-2, COVID-19, saliva, throat wash, antibody response, LFA, sgRNA

Word count - 2992

Introduction

SARS-CoV-2 emergence became the source of the largest coronavirus-driven pandemic. This positive-strand RNA virus, and COVID-19 etiologic agent, infects and replicates in the upper respiratory tract, respiratory mucosa, oral mucosa and salivary glands[1]. The presence of Ace2 receptor+ oral cells, SARS-CoV-2 RNA detection, and detection of virus capable of inducing cytoplasmic effect suggests that oral viral replication and production occur[1]. SARS-CoV-2 was consistently detected in unstimulated whole mouth fluid (WMF), oropharyngeal WMF, and gingival crevicular fluid [2-6]. The CDC and the European Center for Disease Prevention and Control recommend nasal/oral swabs or saliva testing[7]. Nasopharyngeal swab reverse transcriptase-polymerase chain reaction (NP-RT-qPCR) is the gold standard for SARS-CoV-2 detection, however, adequacy of harvested material and collection time relative to disease onset can result in low methodologic sensitivity[8, 9]. Meta-analysis demonstrated high concordance, 92.5% (95%CI: 89.5-94.7), between saliva and nasopharyngeal/oropharyngeal swabs (NPS/OPS) with lower sensitivities in saliva than nasopharyngeal/oropharyngeal swabs, 86.5% (95%CI: 83.4-89.1) and 92.0% (95%CI: 89.1-94.2), respectively[7].

Cochrane assessment of SARS-CoV2 antigen tests (n=48), including lateral flow assays (LFA), demonstrated varied sensitivity between symptomatic and asymptomatic participants with highest sensitivity closest to symptom onset[10]. While LFA-assessed oral SARS-CoV-2-targeted immune responses can reflect systemic responses[11], oral biomarkers as prognostic COVID-19 indicators have not been significantly explored. Here, RT-qPCR, LFA, and immunoblot were used for oral SARS-CoV-2 detection. Longitudinal assessment of symptomatic participants suggested oral viral persistence. Disease severity and symptoms were associated with oral SARS-CoV-2 host responses and viral presence. Collectively, these findings provide novel insights to oral markers of prognosis, persistence, and transmission.

Materials and Methods

Enrollment 47 ambulatory participants provided IRB-approved written informed consent. Participants were part of the 1) longitudinal symptomatic group (n=17), at entry SARS-CoV-2 NP-RT-qPCR+ with stratified symptoms (mild, moderate, severe) based on NIH COVID-19 treatment guidelines, 2) blinded asymptomatic seropositive/seronegative cross-sectional group (n=30) [12], 3) Archived, preCOVID-19 era saliva and throat wash (TW)[13].

Biospecimens NPS in transport medium, saliva, and TW were collected at the same visit. Unstimulated WMF (saliva) and TW gargle with 10 ml of saline were collected. Blood was collected by venipuncture for serology. Samples were transported, stored at 4ºC, and processed within 24 hours.

Nucleic acid extraction and RT-qPCR. Virus was heat/chemically inactivated. RNA was extracted using Trizol (Life Technologies) and reverse transcribed using Superscript III (Life Technologies), according to manufacturer's instructions. First strand synthesis was primed with random hexamers (Life Technologies) and reactions used in SYBRGreen-based qPCR. The SARS-CoV-2 genome was targeted using 3 regional primer sets described in Figure 3A[14, 15]. A SARS-CoV-2 reference RNA-based (ATCC VR-3276SD) standard curve allowed absolute copy number quantitation. Subgenomic RNA (sgRNA) detection across viral junctions was performed using the forward primer, CoV25UTR with gene-specific reverse primers (**Figure 3A**). Percent sgRNA was calculated using the equation: $2^{-(sgRNA\ Ct - NRNA\ Ct)}/100$. Positive reactions were treated with ExoSapIt according to manufacturer's instructions (USB, Cleveland, Ohio) and Sanger sequencing performed using SARS-CoV-2-specific primers (Eton Bioscience, Research Triangle Park, NC).

Recombinant His-tagged SARS-CoV-2 Nucleocapsid (N-antigen). E. coli BL21 (DE3) bacteria were plasmid transformed with SARS-CoV-2 N coding region ligated in frame to the pET30 6X histidine-tagged coding region. After IPTG (1 mM) induction, bacterial pellets were lysed and

protein purified with NTA-Ni agarose beads (Life Technologies). Eluates were subjected to SDS-PAGE followed by Coomassie staining for retrieval and protein purity.

Ectopic expression of SARS-CoV-2 proteins in oral keratinocytes. Full-length subgenomic PCR products encoding SARS-CoV-2 S, E, M or N were generated from SARS-CoV-2+ salivary total RNA. cDNAs were cloned into pCMV-myc and transfected into immortalized human oral keratinocytes (NOK) using Fugene-6 (Promega). Forty-eight hours post transfection, media was subjected to SDS-PAGE/ immunoblot to detect myc-tagged proteins.

Immunoblot. Total protein was isolated using Trizol according to the manufacturer's instructions. Protein pellets were resuspended in sample buffer and proteins electroblotted to PVDF membrane. Blots were blocked followed by primary antibody incubation, incubation with secondary HRP-conjugated antibody, then protein detection by ECL Prime (GE) and imaging using GE ImageQuant LAS4000.

Lateral Flow Assay. SARS-CoV-2 (N-antigen, anti-SARS-CoV-2 spike protein IgG/IgM) were detected in TW/saliva using LFA cartridges according to manufacturer's instructions (BioMedomics, Research Triangle Park, NC)[16, 17]. Bands were visualized and quantitated using ImageJ. Control bands for each detection strip were used to normalize antigen-specific bands and IgG/IgM intensity. Quantitation was determined relative to preCOVID-19, archived saliva/TW signal.

NCBI Structural analysis. The SARS-CoV-2 N-antigen N-terminal RNA binding domain (RBD) crystal structure (PDB ID: 6M3M) was compared to other publicly available crystal structures in the Molecular Modeling Database (MMDB) using Vast+ (https://www.ncbi.nlm.nih.gov). 3D structures of superimposed biological assemblies were visualized using the web-based 3D viewer, iCn3D version 4.3.1[18].

Statistical methods. Given small sample sizes, non-parametric tests were used to test correlations and compare data. The Kendall rank correlation test assessed the existence of monotonic relationships of ordinal or continuous variables using the ranks of the data. The Mann–Whitney U test was used where the independent variables were binary and determined whether the two groups were from the same population when the groups were independent. Mann–Whitney U and Wilcoxon signed rank tests were used to compare response variables between two groups. When groups were dependent, Wilcoxon signed rank tests assessed differences in the medians of matched groups.

Results

Assessment of the SARS-CoV-2 N-Antigen Lateral Flow Assay. N-antigen LFA was validated using lysates from immortalized normal oral keratinocytes (NOK) transfected with a myc-tagged SARS-CoV-2 N-antigen expression vector. N-antigen was detected by anti-myc antibody immunoblot (**Figure 1A**) or LFA (**Figure 1B**). Bacteria-produced, his-tagged N-antigen was purified (**Figure 1C**), used for N-LFA quantification, and migrated at 55 KD, matching the predicted molecular weight of SARS-CoV-2 N-antigen, while pET30 vector negative control ran at ~9 KD (**Figure 1C**). LFA limit of detection was determined using two-fold serial dilutions of preCOVID-19 saliva spiked with recombinant N-antigen and detected his-tagged N-antigen at 43-680 picograms (pg). A faint band at 85 pg and robust detection >170 pg demonstrates LFA's semi-quantitative nature (**Figure 1D**). No band was detected in saliva spiked with pET30 recombinant protein, indicating signal specificity. LFA allowed longitudinal detection of N-antigen in saliva from a representative NP-RT-qPCR+ participant at symptom onset 14, and 28 days, suggesting active persistent infection, while signal was undetected in two PreCOVID-19 saliva samples (**Figure 1E).**

Assessment of the SARS-CoV-2 Anti-Spike RBD IgM/IgG Lateral Flow Assay). Recombinant anti-Spike RBD IgG/IgM in LFA assays detected anti-SARS-CoV-2 Spike-specific IgM and IgG oral immune responses. Two-fold dilution allowed immunoglobulin LFA validation. Mixtures of Spike RBD-specific recombinant human IgM/IgG in preCOVID-19 saliva detected analytical sensitivity down to 0.63 microliters for IgM, however, IgG was undetected after a single 2-fold dilution (**Figure 1F**). Archived saliva, spiked with anti-SARS-CoV-2 antibody demonstrated IgM signal but not IgG, perhaps reflecting IgM detection of a pre-2019 human coronavirus. Unspiked demonstrated no signal (**Figure 1G**). A symptomatic participant demonstrated IgG at baseline and at 14 days but not IgM (**Figure 1G**).

Symptomatic and Asymptomatic Assessment

Ambulatory groups who participated in survey and biospecimen collection were: 1) deidentified-symptomatic, NP-RT-qPCR+, 2) blinded-asymptomatic seropositive or 3) blinded-asymptomatic seronegative (**Figure 2A**). Symptomatic participants (n=17) were assessed at baseline, 14 and 28 days. The mean age was 40.2 years with equivalent gender distribution. The cohort was 64.7% Caucasian, 20% Hispanic/Latino, and 6.4% each of Black and Asian participants (**Figure 2B**). RT-qPCR targeted three SARS-CoV-2 genomic regions (orf1, E and N) in saliva/TW. Known amounts of N viral RNA were detected by standard curve, allowing absolute copy number quantification, that was more sensitive than relative quantitation of E/orf1(**Figure 3F**). All full-length and subgenomic (sgRNA) viral RNAs are amplified by 2019-nCoV_N1-F/2019-nCoV_N1-R allowing relative quantitation (**Figure 3A**). Reactions containing >10 copies were consistently reproducible (**Figure 3B, F**). PreCOVID-19 saliva/TW RT-qPCR, did not produce signal >10 copies/reaction. Oral samples containing >10 copies/reaction were considered SARS-CoV-2 RNA positive. Using these criteria, 59% of saliva samples (10/17) and 64% (9/14) of TW samples from the symptomatic group were SARS-CoV-2 RNA+ (**Figure 3F**).

Oral N-antigen was assessed by immunoblot. Media from NOKs, transfected with myc-tagged SARS-CoV-2 structural protein expression vectors, was subject to myc-specific immunoblot, demonstrating readily detectible N-antigen. Detection of myc-tagged spike (S) and matrix (M) proteins in the medium depended on N expression suggesting virus-like particle (VLP) production (**Figure 3C**). Immunoblot analysis, using a commercially available SARS-CoV-2 N-antigen antibody, detected salivary N-antigen at the expected migration of 55 KD in 10/10 subjects assayed, similar to myc-tagged recombinant N-antigen. pCMV-myc transfected cell medium and preCOVID-19 saliva samples, A and B, were negative controls (**Figure 3D**). TW immunoblot detected N-antigen in 7/10 subjects with 2/7 displaying faint bands (**Figure 3E**). The salivary protein, alpha-amylase served as a loading control. There was striking concordance between NP positivity (reference), salivary N-antigen immunoblot detection, and IgG detection - salivary (93%)/TW (100%) (**Figure3F**). IgM concordance with NP-RT-qPCR+ status was 75% in saliva and 50% in TW.

The asymptomatic group (n=30) was composed of seropositive (n=15) and seronegative/uninfected (n=15) participants. The mean age was 44, with 14 males and 16 females and was 72.4% Caucasian, 4% Asian, 7% Hispanic and 3.4% Black. Of 15 seropositive participants, only one, subject 8A, was NP-RT-qPCR+, IgG positive, and IgM positive in the saliva (**Figure 3G**). Two others (1A and 4A) within the seropositive-asymptomatic group were LFA positive (**Figure 3G**). Among seronegative subjects, salivary IgG and IgM were not detected however, 3/7 subjects demonstrated N-antigen by LFA (**Figure 3H**). To determine potential targets capable of cross-reactive binding, host salivary proteins structurally analogous to N were assessed (**Figure 4**). The vector alignment search tool, VAST+, localized macromolecular structures similar in 3D shape to SARS-CoV-2 N-antigen[19]. VAST+ detected N-antigen complete or partial matches, that included SARS-CoV-1 and human RNA binding proteins Line1Orf1p, hnRNP H, RBM7, and 17S U2 snRNP. Salivary proteome members

RBM7 and 17S U2 snRNP, demonstrated 29aa structural alignment at 1.47 angstroms and 27aa structural alignment at 1.52 angstroms, respectively. Line1ORF1p (19aa at 1.00 angstroms) and hnRNP H (28aa at 1.50 angstroms) have salivary proteome isoforms[20]. SARS-CoV-1 served as a reference control demonstrating structural alignment of 95aa at 0.58 angstroms (**Figure 4**).

Detection of sgRNAs, arguably only produced during active viral replication, confirmed TW LFA positivity (**Figure 5,3F**). Detection of overlapping junction sequences between the end of the 5' UTR and the beginning of the S, orf3a, or N coding regions in oral fluids (**schematic, Figure 5A**) was confirmed by direct sequencing of qPCR products shown as qPCR+/number assayed - 20% (2 /10), 82% (9/11), and 27% (3/11) subjects, respectively (**Figures 5B-D**). No RT-qPCR signal for sgRNAs were detected in preCOVID-19 RNA samples/water negative control. Overall, 90-100% of those who were RNA+ in oral fluids (sal/tw) demonstrate orf3a sgRNA, were LFA N-antigen positive, immunoblot confirmed, and IgG positive (**Figure 3F**).

Viral metrics in the oral fluids over time and between genders

LFA test and control band intensity was measured by ImageJ to provide relative measures of N-antigen and immunoglobulin presence in oral fluids. An insignificant downward trend in salivary N-antigen detection from baseline to 28 days, suggested persistence in oral fluids (Wilcoxon Signed Rank Test) (**Figure 6A**). While quantitatively less N-antigen was detected in symptomatic TW at 28 days compared to baseline (relative band intensity, median=2070 vs 2310), this decrease was not significant (Wilcoxon Signed Rank Test, p=0.9). In the symptomatic group, significant differences were detected between preCOVID-19 archived saliva, baseline, and 28 days (Wilcoxon Rank Sum Test, p=0.003) (**Figure 6B**). At baseline, 56% of the symptomatic were N-antigen positive in saliva and 91% were N-antigen positive in TW by LFA. Using NP-RT-qPCR as the gold standard, LFA assay sensitivity was 0.563 (CI,

0.299, 0.802) in saliva and 0.91 (CI, 0.587,0.998) in TW (**Figure 6C**). Salivary IgG levels were significantly different between baseline and preCOVID-19 archived samples (Wilcoxon Rank Sum Test, p=0.005). IgG trended upward over time and IgM levels, approximately a log lower than the IgG levels, remained the same over 28 days (p=0.625) (**Figure 6A**). While gender differences were not detected in oral N-antigen or TW IgG, females demonstrated higher levels of IgM than males (Mann-Whitney U test, p=0.056) (**Figure 6D**).

SARS-CoV-2 oral outcomes, COVID-19 symptoms and symptom severity, and oral persistence. Self-reported COVID-19 symptom severity at baseline (absent, mild, moderate, or severe) and presence of SARS-CoV-2 oral N-antigen or antibody in saliva/TW were assessed. COVID-19 symptom (weakness, muscle ache, nausea, loss of taste/smell and upper respiratory tract symptoms which encompassed cough, shortness of breath, sore throat, nasal obstruction, nasal discharge) presence or absence at baseline was reported by participants as yes/no. Kendall Rank Correlation tests determined associations between severity of cough and fatigue with salivary IgM (p=0.008 and 0.016 respectively) (**Figure 7A and B**). At baseline, salivary IgM was consistently associated with weakness (p=0.037), muscle ache (p=0.019), nausea (p=0.005) and upper respiratory symptoms (p=0.017), but not loss of taste/smell (p=0.458) (Mann-Whitney U test). Oral N-antigen and IgG detection were not associated with COVID-19 symptoms in oral fluids (Mann-Whitney U test) (**Figure 8A)**. While longitudinal assessment of oral antigen and immunoglobin responses detected no clear directional trends, there was consistent detection of IgM, IgG, and oral N-antigen at baseline and over time (n=11) (**Figure 8B**).

Discussion

In this study, we consistently detect SARS-CoV-2 antigen and antibody in oral fluids during symptomatic COVID-19. N-antigen detection was immunoblot- and sgRNA-confirmed in

NP-RT-qPCR+ subjects providing significant implications for oral transmission. SARS-CoV-2-targeted oral IgG responses were highly correlated with nasopharynx positivity and oral N-antigen detection. Oral IgM levels indicated both symptom presence and severity.

This study, and others, detected salivary SARS-CoV-2 RNA with distinct viral shedding dynamics compared to NP[1, 21]. Prior assessments detected a 3-fold lower positive detection rate in saliva than NPS[21], leading some to question the role of oral virus. Here, cumulative data suggest oropharyngeal viral replication. SARS-CoV-2 was detected in oral fluids from NP-RT-qPCR+ participants using several RT-qPCR-based viral detection methods, 1) three distinct primer pairs targeting 3 regions of the viral genome, 2) absolute RNA copy number determination using a standard curve, and 3) sgRNA, shown to be a marker of active replication during early symptomatic infection[22, 23]. Orf3A sgRNA was detected in 82% of those tested by sgRTqPCR, (fast, sensitive, economically feasible, and reliable) and is a viable marker of viral replication based on its contribution to viral titer and disease in hACE+ mice[24]. SARS-CoV-2 protein was consistently detected by two distinct methods, LFA and immunoblot. We are unaware of other studies demonstrating immunoblot confirmed viral antigen detection in oral fluids. Importantly, persistent N-antigen detection provides significant implications for continued potential oral transmission (**Figure 3F,8B**).

One critique of rapid LFA was false positives. Here, we detect positive LFA results in two seropositive (1A,4A), and three uninfected subjects (16A,17A,18A). During asymptomatic infection it is impossible chronical where subjects are in their infection cycle. Seropositive/NP-subjects could demonstrate oral infection (1A,4A) or oral infection may subside in seropositive/NP+ (8A). Alternatively, cross-reactive host proteins might be detected. Analysis determined potential cross-reactivity with structurally analogous host RNA binding proteins (**Figure 5**). 17S U2 snRNP and RBM7, are salivary proteome members and two others, LINE1 Orf1p and hnRNP H, have salivary proteome isoforms (LINE1 Orfp1 and hnRNP A2/B1, hnRNP

M, hnRNPK)[25]. This raises the possibility that N-antigen-specific LFA can detect conformationally-similar proteins, perhaps reflecting the potential for N-antigen mimicry through RNA binding domains. An example implication is SARS-CoV-2 host genome integration. The LINE1 ORF1p analogue mediates retrotranposable element genome integration[26]. RNA virus sequences have been detected across vertebrate genomes, with several integration signals consistent with LINE retrotransposon germline integration of viral cDNA copies[27]. Recently, subgenomic sequences, derived from SARS-CoV-2's 3' end, were shown to be integrated into host cell DNA[28]. Hence, host interactions with N-antigen or similar actions by N-antigen might facilitate SARS-CoV-2 genome integration.

Anti-spike RBD-specific antibodies during natural infection and vaccination were detected in saliva[29, 30] with temporal kinetics that reflect blood. While anti-spike-RBD IgG levels were negatively correlated with salivary viral load in the Silva study[31], here in 16/16 symptomatic and 1/1 asymptomatic participants, anti-spike-RBD IgG was 100% positively correlated with SARS-CoV-2 NP-RT-qPCR+ status, and symptomatic individuals consistently demonstrated salivary N-antigen by immunoblot. The NP+ status and IgG relationship was also shown by Pisanic et,al. with IgG positive responses in 24/24 RT-qPCR-confirmed COVID-19 cases[11]. Together, this suggests that salivary IgG detection indicates concurrent nasopharynx positivity. Detection of simultaneous oral virus and oral host immune responses at baseline suggests a highly infectious pre-symptom state followed by concurrent symptoms, persistent oral infection, and virus-targeted host responses.

Oral fluids hold promise as SARS-CoV-2 prognostic indicators. Huang et al. suggested a correlation between salivary viral RNA burden and COVID-19 symptoms, including taste loss[1]. Others showed positive associations between salivary viral load (VL) and COVID-19-related pro-inflammatory markers; IL-6, IL-18, IL-10, and CXCL10[31]. Here, oral viral RNA detection did not correlate with COVID-19 symptoms. Relationships between IgG, VL, and disease course

were not detected[21, 32]. However, we detected statistically significant relationships between salivary IgM and multiple symptoms including weakness, muscle ache, nausea, and upper respiratory symptoms but not taste/smell. Relative IgM levels were significantly associated with degree of fatigue and cough severity suggesting that oral IgM signifies an active early immune response associated with milder disease course.

Gender differences are detected in COVID-19-related morbidity and mortality. Here, gender differences were detected, with women having consistently higher oral IgM levels than men. Systematic review of gender differences determined that women were less likely to present with severe disease and be admitted to the intensive care unit (ICU) than men (OR 0.75 [0.60–0.93] $p<0.001$ and OR 0.45 [0.40–0.52] $p<0.001$, respectively[33]. While higher smoking rates and reluctance to seek health care may contribute to male disease, the well-described evolution of greater humoral immunity in females could contribute to these differences[34, 35]. Here, women had markedly higher IgM levels, fewer symptoms, and milder symptom severity. Differences in IgM production may provide biologic underpinning, contributing to the COVID-19-related morbidity and mortality gender gap.

These findings provide novel insights to oral SARS-CoV-2 infection. Oral IgM detection predicted milder disease and female predilection. The potential for cross-detection of N-antigen and structurally similar host RNA binding factors, suggests viral mimicry. Oral persistence has significant implications for viral transmission in the absence of mask use and vaccination.

References

1. Huang N, Perez P, Kato T, et al. SARS-CoV-2 infection of the oral cavity and saliva. Nat Med 2021;27(5):892-903.
2. Kannian P, Lavanya C, Ravichandran K, et al. SARS-CoV2 antigen in whole mouth fluid may be a reliable rapid detection tool. Oral Dis; 2021.
3. Chen JH, Yip CC, Poon RW, et al. Evaluating the use of posterior oropharyngeal saliva in a point-of-care assay for the detection of SARS-CoV-2. Emerg Microbes Infect 2020;9(1):1356-59.
4. Azzi L, Carcano G, Gianfagna F, et al. Saliva is a reliable tool to detect SARS-CoV-2. J Infect 2020.
5. To KK, Tsang OT, Leung WS, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis 2020.
6. Gupta S, Mohindra R, Chauhan PK, et al. SARS-CoV-2 Detection in Gingival Crevicular Fluid. J Dent Res 2021;100(2):187-93.
7. Ibrahimi N, Delaunay-Moisan A, Hill C, et al. Screening for SARS-CoV-2 by RT-PCR: Saliva or nasopharyngeal swab? Rapid review and meta-analysis. PLoS One 2021;16(6):e0253007.
8. Higgins RL, Rawlings SA, Case J, et al. Longitudinal SARS-CoV-2 antibody study using the Easy Check COVID-19 IgM/IgG lateral flow assay. PLoS One 2021;16(3):e0247797.
9. Williams E, Isles N, Chong B, et al. Detection of SARS-CoV-2 in saliva: implications for specimen transport and storage. J Med Microbiol 2021;70(2).
10. Dinnes J, Deeks JJ, Berhane S, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. Cochrane Database Syst Rev 2021;3:CD013705.
11. Pisanic N, Randad PR, Kruczynski K, et al. COVID-19 Serology at Population Scale: SARS-CoV-2-Specific Antibody Responses in Saliva. J Clin Microbiol 2020;59(1).
12. Pettifor A, DiPrete BL, Shook-Sa BE, et al. A prospective study of asymptomatic SARS-CoV-2 infection among individuals involved in academic research under limited operations during the COVID-19 pandemic. PLoS One 2022;17(4):e0267353.
13. Valentine J, Saladyanant T, Ramsey K, et al. Impact of periodontal intervention on local inflammation, periodontitis, and HIV outcomes. Oral Dis 2016;22 Suppl 1:87-97.
14. Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill 2020;25(3).
15. CDC: Centers for Disease Control and Prevention. "https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html". Accessed September 17, 2020.
16. Liu L, Meyers K, Purpura LJ, et al. Development and performance of a point-of-care rapid antigen test for detection of SARS-COV-2 variants. J Clin Virol Plus 2022;2(3):100080.
17. Naranbhai V, Chang CC, Beltran WFG, et al. High Seroprevalence of Anti-SARS-CoV-2 Antibodies in Chelsea, Massachusetts. J Infect Dis 2020;222(12):1955-59.
18. Wang Y, Geer LY, Chappey C, Kans JA, Bryant SH. Cn3D: sequence and structure views for Entrez. Trends Biochem Sci 2000;25(6):300-2.

19. Madej T, Marchler-Bauer A, Lanczycki C, Zhang D, Bryant SH. Biological Assembly Comparison with VAST. Methods Mol Biol 2020;2112:175-86.
20. Lau WW, Hardt M, Zhang YH, Freire M, Ruhl S. The Human Salivary Proteome Wiki: A Community-Driven Research Platform. J Dent Res 2021:220345211014432.
21. Li L, Tan C, Zeng J, et al. Analysis of viral load in different specimen types and serum antibody levels of COVID-19 patients. J Transl Med 2021;19(1):30.
22. Chen Z, Ng RWY, Lui G, et al. Profiling of SARS-CoV-2 Subgenomic RNAs in Clinical Specimens. Microbiol Spectr 2022;10(2):e0018222.
23. Long S. Correction: Long, S. SARS-CoV-2 Subgenomic RNAs: Characterization, Utility, and Perspectives. Viruses 2020, 13, 1923. Viruses 2022;14(7).
24. Silvas JA, Vasquez DM, Park JG, et al. Contribution of SARS-CoV-2 Accessory Proteins to Viral Pathogenicity in K18 Human ACE2 Transgenic Mice. J Virol 2021;95(17):e0040221.
25. Lau WW, Hardt M, Zhang YH, Freire M, Ruhl S. The Human Salivary Proteome Wiki: A Community-Driven Research Platform. J Dent Res 2021;100(13):1510-19.
26. Huang CR, Burns KH, Boeke JD. Active transposition in genomes. Annu Rev Genet 2012;46:651-75.
27. Horie M, Honda T, Suzuki Y, et al. Endogenous non-retroviral RNA virus elements in mammalian genomes. Nature 2010;463(7277):84-7.
28. Zhang L, Richards A, Barrasa MI, et al. Reverse-transcribed SARS-CoV-2 RNA can integrate into the genome of cultured human cells and can be expressed in patient-derived tissues. Proc Natl Acad Sci U S A 2021;118(21).
29. Isho B, Abe KT, Zuo M, et al. Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. Sci Immunol 2020;5(52).
30. Ketas TJ, Chaturbhuj D, Portillo VMC, et al. Antibody Responses to SARS-CoV-2 mRNA Vaccines Are Detectable in Saliva. Pathog Immun 2021;6(1):116-34.
31. Silva J, Lucas C, Sundaram M, et al. Saliva viral load is a dynamic unifying correlate of COVID-19 severity and mortality. medRxiv 2021.
32. Meinberger D, Koch M, Roth A, et al. Analysis of IgM, IgA, and IgG isotype antibodies Directed against SARS-CoV-2 spike glycoprotein and ORF8 in the course of COVID-19. Sci Rep 2021;11(1):8920.
33. Lakbar I, Luque-Paz D, Mege JL, Einav S, Leone M. COVID-19 gender susceptibility and outcomes: A systematic review. PLoS One 2020;15(11):e0241827.
34. Fink AL, Klein SL. The evolution of greater humoral immunity in females than males: implications for vaccine efficacy. Curr Opin Physiol 2018;6:16-20.
35. Markle JG, Fish EN. SeXX matters in immunity. Trends Immunol 2014;35(3):97-104.

Funding sources

The work here was supported by NCI U54CA260543-01

This work was supported by the North Carolina Policy Collaboratory through appropriation from the North Carolina General Assembly (NCGA) in support of research on treatment, community testing, and prevention of COVID-19 (as mandated by the NCGA in subdivision (23) of Section 3.3 of Session Law 2020-4).

Potential conflicts of interest: FW is employed by BioMedomics, Inc, the manufacturer of the N-antigen and anti-SARS-CoV-2 Spike RBD protein IgG/IgM LFA cartridges used in this study. No other authors declare any conflicts of interest.

Acknowledgements

The UNC Delta Core Facility performed sample processing

Biomedomics provide the N-antigen and anti-S-CoV-2 Spike RBD protein IgG/IgM LFA cartridges used in this study.

Institutional Review Board numbers associated with samples collected and used in this manuscript

UNC IRB study #20-0792 – Sample collection from Symptomatic participants
UNC IRB study #20–1771 – Sample collection from Asymptomatic particpants
UNC IRB study # 07-1431 – Sample collection from Archived, PreCovid prticipants

Figure Legends

Figure 1. Assessment of Lateral flow assays for detection of SARS-CoV-2 N-antigen and anti-SARS-CoV-2 spike RBD domain targeted antibodies A. Expression of myc-tagged N-antigen in NOK cells by immunoblot. Vector lane indictates cells transfected with the empty plasmid, pCMV-myc. N-antigen detected with anti-myc antibody is indicated by the arrow. **B.** LFA detected N-antigen in the lysates of NOKs transfected with myc-tagged N expression plasmid. Both, control and N-antigen bands, are indicated. Only the control band is detected in lysates from cells transfected with empty vector. **C.** Expression and purification of His-tagged SARS-CoV-2 N-antigen in E. coli. A band consistent with the size of N-antigen (55 kD) can be seen in crude lysates from cells transformed with His-tagged, N-antigen expression plasmid. This band is eluted from Ni-beads during purification. His-tagged protein expressed in bacteria transformed with empty vector (pET30) served as a negative control. **D.** Dilution series of His-tagged N-antigen in preCOVID-19 saliva and detection using Biomedomics N-antigen LFA. **E.** Detection of N-antigen in saliva of preCOVID-19 and symptomatic SARS-CoV-2 NP+ subjects at entry and post entry 14D and 28D by LFA. The detection of N-antigen band is indicated. **F.** Two-fold dilution series of anti-SARS-CoV-2 Spike RBD (S-RBD) domain IgG/IgM antibodies in pre COVID-19 saliva and detection using Biomedomics IgG/IgM lateral flow assay cartridges. Control, anti-S-RBD IgM and S-RBD IgG are indicated. The 0 lane indicates unspiked, preCOVID-19 saliva only.

Figure 2. Study Schema and Asymptomatic/symptomatic Participant Demographics. A. Enrollment inclusion criteria for subjects who were COVID-19 negative (n=15), had asymptomatic COVID-19 (n=15), or had symptomatic COVID-19 infection (n=17) are listed within the figure. Subsequent to enrollment, participants answered surveys and provided biospecimens (saliva, TW, blood) that were assessed for SARS-CoV-2 detection and antibody

responses. B. Symptomatic SARS-CoV-2 NP+ and asymptomatic participant demographics are indicated.

Figure 3. Oral virologic characterization of study participants at entry. All subjects were assayed by qPCR for detection of viral RNA in the nasopharynx. Saliva/TW was assayed by qPCR for viral RNA and for LFA-based detection of N-antigen, anti–spike IgG and anti-spike IgM. **A.** Oligonucleotide sequences for primers used for RT-qPCR and production of cDNA for construction of expression plasmids. E_Sarbeco_F and E_Sarbeco_R have been described. The sequences of 2019-nCoV_N1-F and 2019-nCoV_N1-R, were obtained from the CDC. The SARS-CoV-2 reference sequence GenBank:MN908947 was used to determine oligonucleotide sequences and map coordinates. SfiI (ggccatggaggcc) and SalI (gtcgac) restriction enzyme sites used for cloning into the expression vector, pCMV-myc, are underlined in expression primers. The start codon and stop codons in expression primers are in bold. **B.** Standard curves were generated by RT-qPCR using know copy numbers of SARS-CoV-2 RNA and primers targeting the nucleocapsid (N) coding region to quantitate viral RNA isolated from patient saliva and TW. **C.** Immunoblot was used to detect SARS-CoV-2 structural proteins in cell culture media obtained from NOKs transfected with individual expression vectors encoding myc-tagged, SARS-CoV-2 structural proteins (S, E, M and N) or cotranfected with all myc-tagged expression vectors. Medium from cells transfected with empty vector (pCMV-myc) was used as a negative control. Proteins were detected using myc-specific antibody (mouse anti-myc, sc-40, Santa Cruz Biotechnology). **D and E.** Protein from subset of saliva and TW samples (respectively) was used for the detection of viral proteins. Immunoblot detection of SARS-CoV-2 N-antigen was assessed in a subset of symptomatic individuals (n=10) at baseline (top panel) using N-specific antibody (PA5-114448, LifeTechnologies, Inc). Detection of salivary alpha-amylase (anti-amylase, A8273, Sigma-Aldrich) served as a loading control (bottom panel). Saliva or TW from archived preCOVID-19 patients A and B were included as negative controls. Cell culture

medium from NOK cells transfected with myc-tagged SARS-CoV-2 N-antigen expression construct was used as a positive control. Cell culture medium from NOKs transfected with empty vector was included as a negative control. N-antigen was detected at 55 KD and alpha amylase, the loading control, was detected at 58.4 KD. **F.** Table of symptomatic subjects (n=17) who were all nasopharynx RT-qPCR positive. RT-qPCR and LFA was also performed on saliva and TW of these subjects. Viral RNA copy number (molecules of virus RNA) in 25 microliters of saliva or TW was determined using N-specific primers. Green highlight indicates a positive result. Red highlight indicates a negative result. ND indicates not determined. **G.** Asymptomatic participants (n=15) were all anti spike IgG positive in their blood. **H.** COVID-19 negative participants (n=15) were anti-spike negative and nasopharynx negative.

Figure 4. The SARS-CoV-2 nucleocapsid mimics host RNA binding proteins that are expressed within the salivary proteome and may be responsible for cross reactivity in LFA assays. VAST+ was used to generate RNA binding protein structures demonstrating 3D similarity to the SARS-CoV-2 N-antigen. Simultaneous alignment generated molecular protein pairs between N-antigen (first column) and host RNA binding proteins (second column). Non-aligned N-antigen amino acids are rendered in green, conserved aligned amino acids are rendered in red and aligned, non-conserved amino acids are rendered in blue. Non-aligned host protein amino acids proteins are rendered in olive green, conserved aligned amino acids are rendered in pink and aligned, non-conserved amino acids are rendered in light blue. The spatial arrangement of overlapping amino acids can be seen in the merged sequence alignment (third column). The number of structurally aligned amino acids (conserved and non-conserved) is shown. The root mean square deviation (RMSD) is used as a measurement between atoms in the backbone of the molecular structures and is listed in angstroms above each merged image and detection of the host protein within the salivary proteome is displayed as yes in the fourth column while detection of a related isoform within the salivary proteome is displayed as ISO.

Figure 5. SARS-CoV-2 subgenomic RNAs (sgRNA) are detected the saliva of symptomatic COCID patients. A. Schematic diagram showing the RNA genome of SARS-CoV-2 (~30 kb) and the subgenomic RNAs encoding major viral structural proteins (S, E, M and N) as well as orf3a. Detection of SARS-CoV-2 RNA by RT-qPCR was performed using primer pairs corresponding to orf1 (blue arrows), E (peach arrows) and N (red arrows) coding regions. SgRNA was detected using a forward primer corresponding to the 5'UTR (black arrow) and either S-specific (yellow), orf3a-specific (green) or N-specific (red arrow) reverse primers. The RT-qPCR signal generated with N1 primers (red arrows) were used to calibrate levels of sgRNA. **B, C and D.** The percent levels of S, orf3a or N sgRNA determined using the 5'UTR primer and the specific sgRNA reverse primer are shown. Percent sgRNA levels relative to total viral RNA levels are determined as follows: 2^-(sgRNA Ct/N Ct) X 100. Direct sequencing of RT-qPCR amplified products was performed to determine the junction sequence of the 5'UTR and the coding region of each sgRNA. Sequencing information obtained from samples is indicated with an asterisk. The overlap sequence between the 5'UTR and each specific sgRNA is indicated.

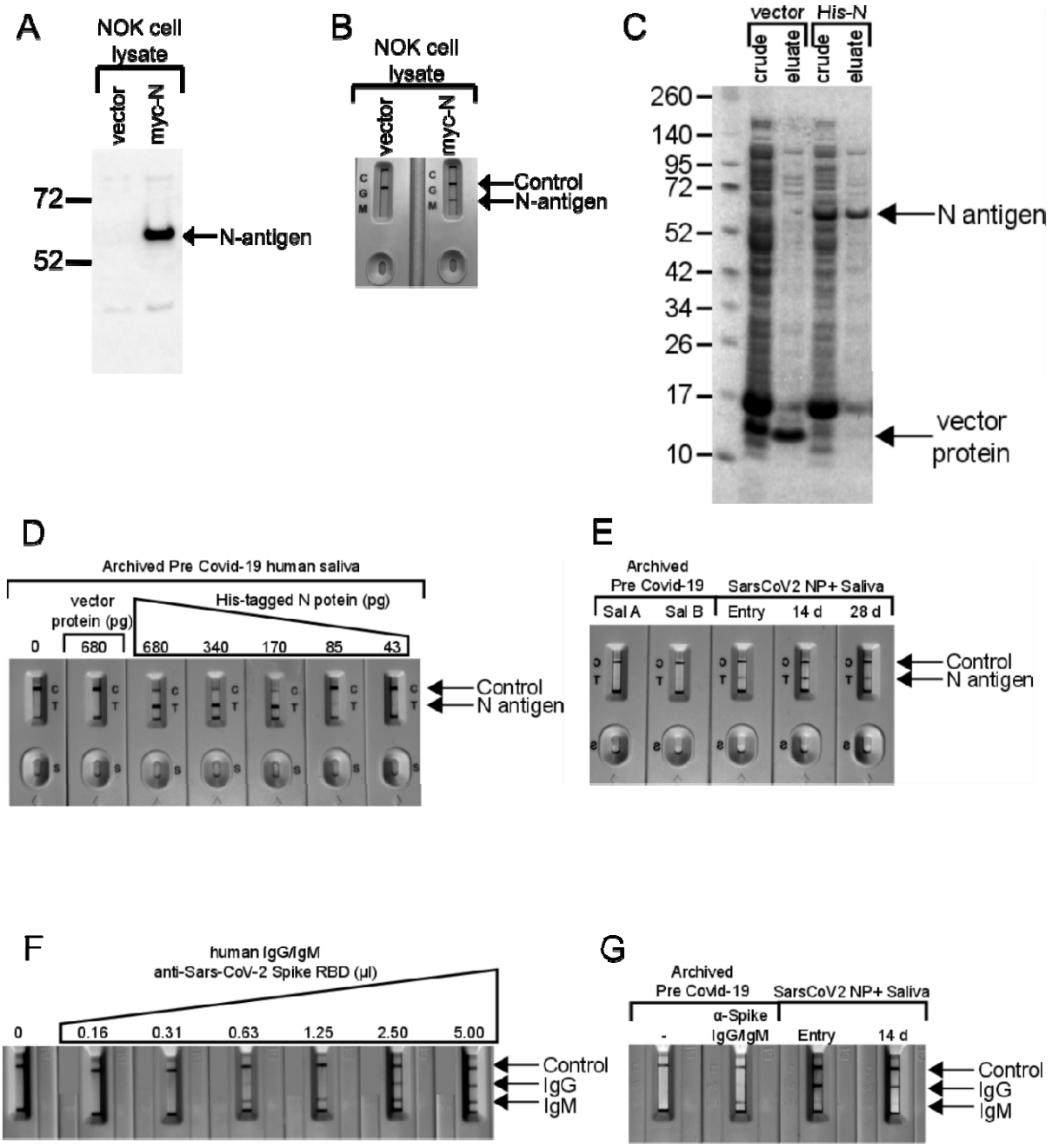
Figure 6. Longitudinal semi-quantitative analysis of SARS-CoV-2 N-antigen and anti-SARS-CoV-2 specific IgG/IgM antibodies in oral fluids of infected individuals demonstrates persistence and gender differences A. Detection of N-antigen (left panel), anti-Spike IgG (middle panel), and anti-Spike IgM (right panel) in TW of symptomatic subjects at entry (BL), 14 days (14D), and 28 days (28D) post entry compared to preCOVID-19, archived (Pre) TW. **B.** LFA detection of N-antigen in TW of symptomatic subjects at entry (BL) and 28 days (28D) post entry compared to preCOVID-19, archived TW. The Wilcoxon Rank Sum test assessed difference between preCOVID-19 and entry and Wilcoxon Signed Rant Test assessed significant differences between matched entry and 28 day samples. **C.** Saliva and TW N-antigen, anti-Spike IgG and anti-Spike IgM LFA assay sensitivity was determined relative to NP-

RT-qPCR+ results in symptomatic participants. **D.** Detection of N-antigen (left panel), anti-Spike IgG (middle panel), and anti-Spike IgM (right panel) in TW of female vs male symptomatic subjects at entry demonstrates differences in relative IgM levels.

Figure 7. Salivary IgM predicted a more favorable disease course. The relationship between symptom severity and semi-quantitative SARS-CoV-2 protein and anti-spike antibody levels in saliva and TW of symptomatic infected individuals with cough and fatigue was assessed at baseline using LFA. Kendall Rank Correlation test detected a decreasing relationship between salivary IgM detection and the severity of fatigue and cough.

Figure 8. Salivary IgM was associated with the absence of most symptoms. A. The relationship between symptom presence (weakness, muscle ache, nausea, loss of taste and smell and upper respiratory symptoms [including cough, shortness of breath, sore throat, nasal obstruction (stuffy nose), nasal discharge [runny nose]) and semi-quantitative SARS-CoV-2 N-antigen and anti-spike antibody levels in saliva and TW of symptomatic infected individuals was assessed at baseline as determined by LFA. The Mann-Whitney U test detected a relationship between the absence of symptoms and higher oral IgM levels in saliva. **B**. Longitudinal oral virologic characterization of symptomatic participants detects viral persistence with concurrent antibody responses suggesting presentation post-acute infection. Saliva (triangle) and TW (circle) were assayed by LFA for detection of N-antigen (red), anti –spike IgG (green) and anti-spike IgM (blue) at entry, 14d and 28d (n=12).

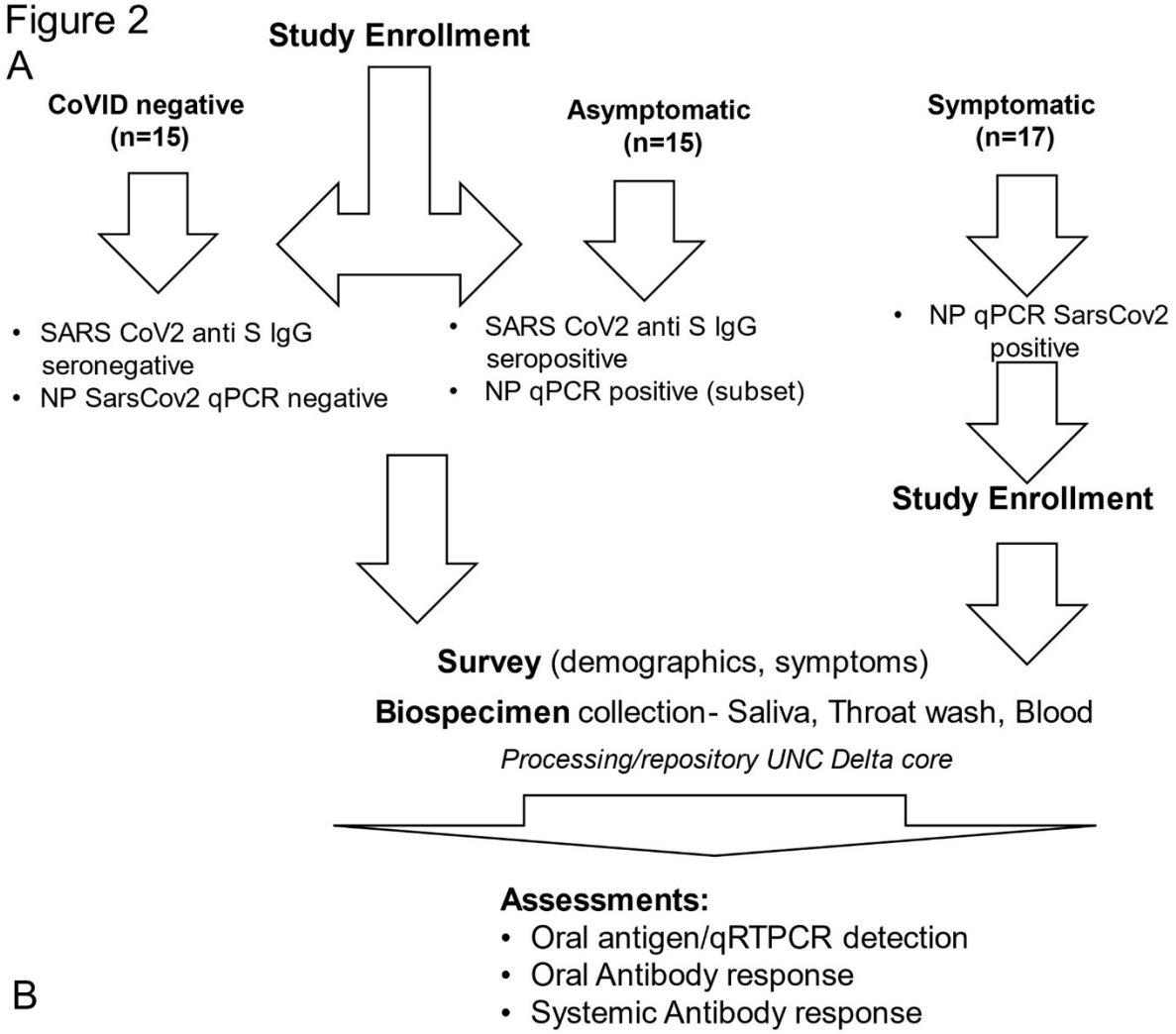
Figure 1


A
NOK cell lysate
vector
myc-N
72
52
N-antigen
B
NOK cell lysate
vector
myc-N
Control
N-antigen
C
vector
His-N
crude
eluate
crude
eluate
260
140
95
72
52
42
34
26
17
10
N antigen
vector protein
D
Archived Pre Covid-19 human saliva
vector protein (pg)
His-tagged N potein (pg)
0
680
680
340
170
85
43
Control
N antigen
E
Archived Pre Covid-19
SarsCoV2 NP+ Saliva
Sal A
Sal B
Entry
14 d
28 d
Control
N antigen
F
human IgG/IgM anti-Sars-CoV-2 Spike RBD (µl)
0
0.16
0.31
0.63
1.25
2.50
5.00
Control
IgG
IgM
G
Archived Pre Covid-19
SarsCoV2 NP+ Saliva
-
α-Spike IgG/IgM
Entry
14 d
Control
IgG
IgM

575

Figure 2

A



B

Observational		N=17	Asymptomatic/ Covid negative		N=30
Age	mean (min,max)	40.29 (38,62)	**Age**	mean (min,max)	44 (18,65)
Sex	M, F, not provided	(8,8,1)	**Sex**	M, F	14,16
Race			**Race**		
	White	12 (70.6%)		White	21 (72.4%)
	Black	1 (5.9%)		Black	1 (3.45%)
	Asian	1 (5.9%)		Asian	4 (13.79%)
	Missing	1 (5.9%)		Hispanic	2 (6.9%)
	Unknown	2 (11.8%)		Other/multi races	1 (3.45%)
Ethnicity			**Ethnicity**		
	Hispanic/ Latino	3 (20%)		Hispanic/ Latino	2 (6.9%)
	Non Hispanic /Latino	11 (73.3%)		Non Hispanic/Latino	27 (93.1%)

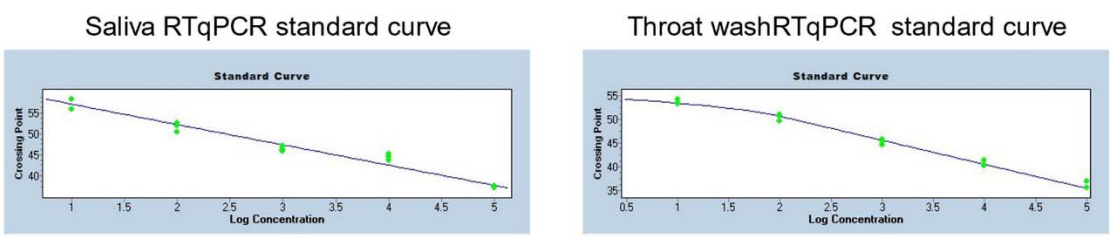
576

Figure 3

A

Experiment	Forward primer	sequence	Reverse primer	sequence
RT-qPCR	MN908947orf1F 16364-16386	TGTCTGTTAATCCGTATGTT TGC	MN908947orf1R comp. 16411-16437	ACCTCCTAAGTAAAGTTG AGTCACATC
RT-qPCR	E_Sarbeco_F 26269-26294	ACAGGTACGTTAATAGTTAA TAGCGT	E_Sarbeco_R compl. 26360-26381	ATATTGCAGCAGTACGCA CACA
RT-qPCR	2019-nCoV_N1-F 28283-28302	GACCCCAAAATCAGCGAA AT	2019-nCoV_N1-R comp. 28331-28354	TCTGGTTACTGCCAGTTG AATCTG
SsgRNA RT-qPCR	CoV25UTF 1-25	ATTAAAGGTTTATACCTTCC CAGGT	SRTqPCRR comp. 21622-21643	TGCAGGGGGTAATTGAGT TCTG
orf3asgRNA RT-qPCR	CoV25UTF 1-25	ATTAAAGGTTTATACCTTCC CAGGT	orf3aRTqPCRR Comp. 25474-25493	GTTGCAGTAGCGCGAACA AA
NsgRNA RT-qPCR	CoV25UTF 1-25	ATTAAAGGTTTATACCTTCC CAGGT	2019-nCoV_N1-R comp. 28331-28354	TGCAGGGGGTAATTGAGT TCTG
S gene expression	SexpF 21563-21587	ggggccatggaggcc**ATG**TTTGTT TTTCTTGTTTTATTGCC	SexpR comp. 25363-25384	cccgtcgac**TTA**TGTGTAATG TAATTTGACT
E gene expression	EexpF 26245-26265	ggggccatggaggcc**ATG**TACTC ATTCGTTTCGGAA	EexpR comp. 26449-26472	cccgtcgac**TTA**GACCAGAA GATCAGGAACTCT
M gene expression	MexpF 26523-26542	ggggccatggaggcc**ATG**GCAGA TTCCAACGGTAC	MexpR comp. 27167-27191	cccgtcgac**TTA**CTGTACAAG CAAAGCAATATTG
N gene expression	NexpF 28274-28297	ggggccatggaggcc**ATG**TCTGA TAATGGACCCCAAAAT	NexpR comp. 29512-29533	cccgtcgac**TTA**GGCCTGAGT TGAGTCAGCA

B



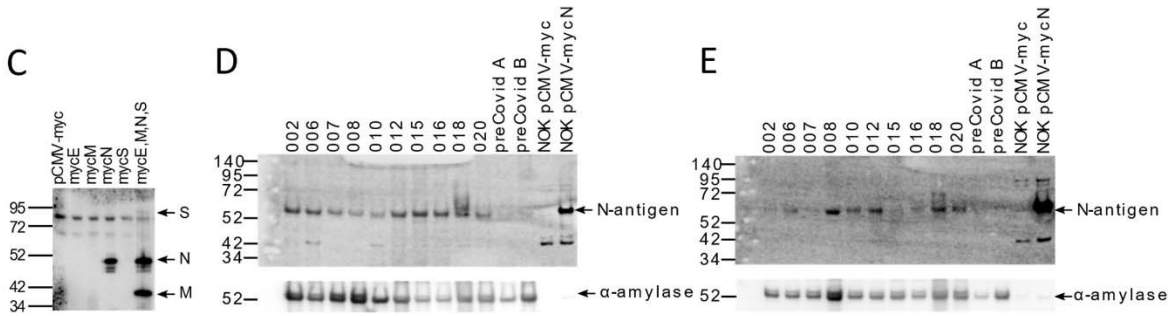


Figure 3

F

Symptomatic																
	RT-qPCR	Copy # N RNA	Copy # N RNA	ViralRNA saliva		sgRNA			N antigen LFA		N Antigen immunoblot		Anti-S IgG LFA		Anti-S IgM LFA	
Participant	NP	S	TW	E	ORF1	S	3a	N	S	TW	S	TW	S	TW	S	TW
002	+	32	34	-	-	-	-	-	-	+	+	-	+	+	-	-
006	+	216	104	-	-	-	+	-	-	+	+	+	+	+	+	-
007	+	-	46	-	-	-	+	-	++	++	+	-/+	++	+	-	+
008	+	352	122	-	-	-	+	-	+	++	+	+	+	++	-	+
010	+	-	78	+	-	-	+	-	-	+	+	+	+	+	+	-
012	+	420	ND	ND	ND	ND	ND	ND	++	ND	+	+	++	ND	+	ND
015	+	25	183	-	-	-	+	-	+	++	+	-	+	++	-	+
016	+	23440	116	+	+	+	+	+	+	+	+	-/+	+	+	+	-
017	+	48	253	+	-	-	+	-	+	+	ND	ND	+	+	+	+
018	+	15	6547	+	+	+	+	+	-	+	+	+	+	+	-	+
019	+	-	ND	ND	ND	ND	ND	ND	-	ND	ND	ND	-	ND	-	ND
020	+	28	-	-	-	-	-	-	+	+	+	-	++	ND	+	ND
022	+	-	-	ND	ND	ND	ND	ND	++	-	ND	ND	+	-	+	-
024	+	-	-	ND	ND	ND	ND	ND	-	++	ND	ND	+	+	+	+
025	+	-	-	ND	ND	ND	ND	ND	-	++	ND	ND	+	+	+	-
027	+	-	-	ND	ND	ND	ND	ND	++	-	ND	ND	++	+	+	+
0517	+	21233	ND	ND	ND	ND	+	+	ND	ND	ND	ND	ND	ND	ND	ND

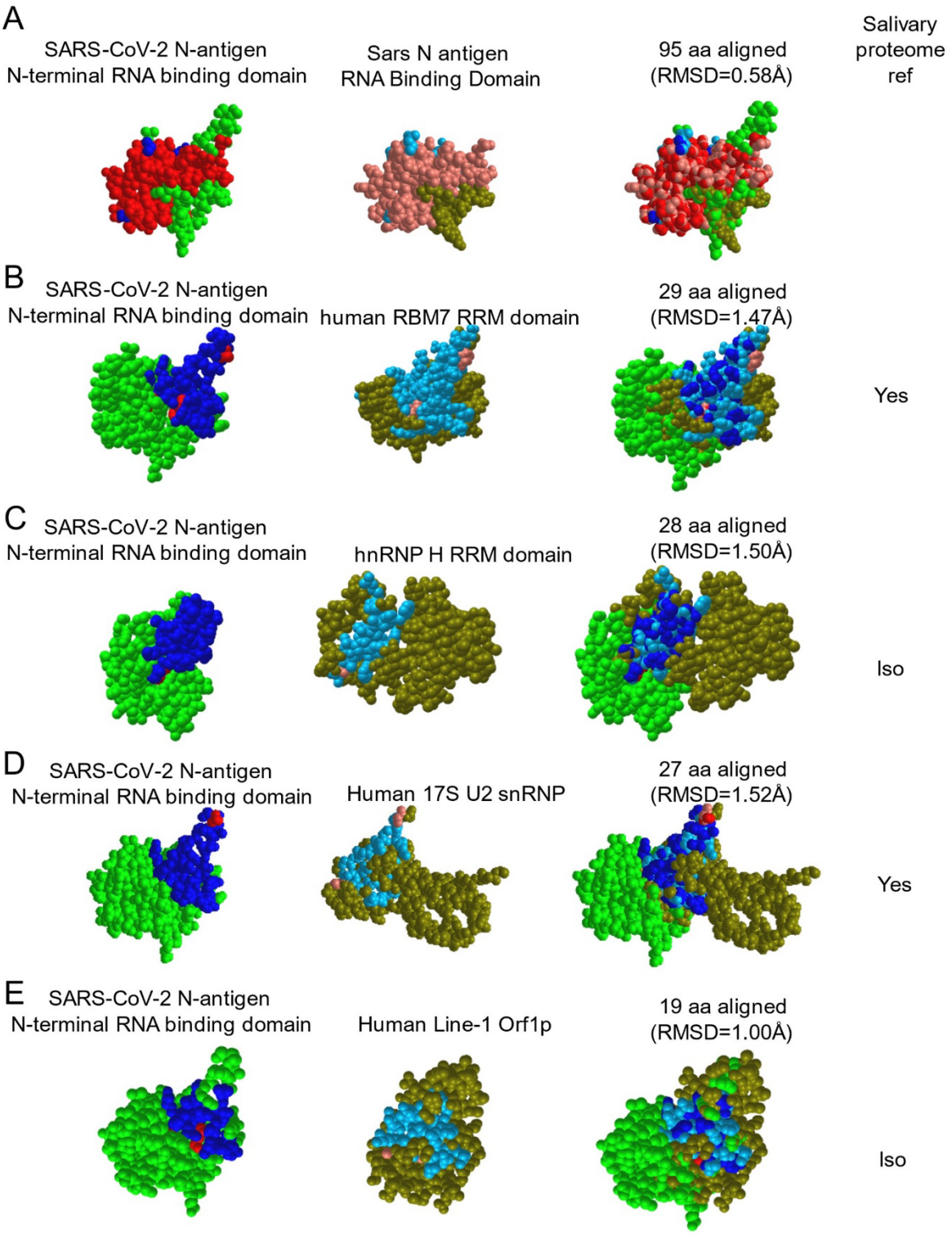
G

Asymptomatic	Seropositive				
	NP	Blood	Saliva		
Participant	RT-PCR	Anti S serostatus	N antigen	Anti S IgG	Anti S IgM
1A	-	+	+	-	-
2A	-	+	-	-	-
3A	-	+	-	-	-
4A	-	+	+	-	-
5A	-	+	-	-	-
6A	-	+	-	-	-
7A	-	+	-	-	-
8A	+	+	-	+	+
9A	-	+	ND	-	-
10A	-	+	ND	-	-
11A	-	+	ND	-	-
12A	-	+	ND	-	-
13A	-	+	ND	-	-
14A	-	+	ND	-	-
15A	-	+	ND	-	-

H

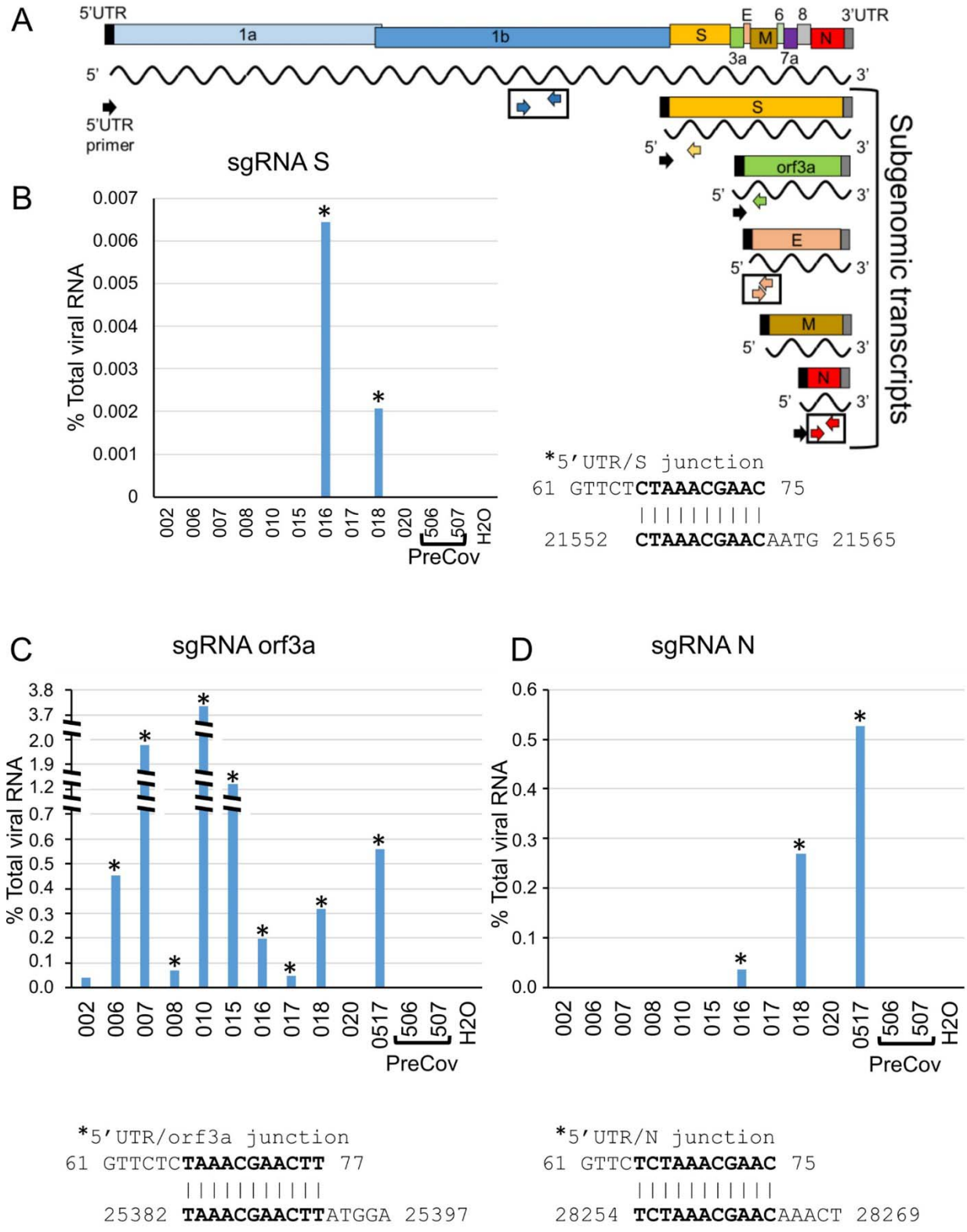
CoVID (-)	Seronegative				
	NP	Blood	Saliva		
Participant	RT PCR	Anti S serostatus	N antigen	Anti S IgG	Anti S IgM
16A	-	-	+	-	-
17A	-	-	+	-	-
18A	-	-	+	-	-
19A	-	-	-	-	-
20A	-	-	-	-	-
21A	-	-	-	-	-
22A	-	-	-	-	-
23A	-	-	ND	-	-
24A	-	-	ND	-	-
25A	-	-	ND	-	-
26A	-	-	ND	-	-
27A	-	-	ND	-	-
28A	-	-	ND	-	-
29A	-	-	ND	-	-
30A	-	-	ND	-	-

Figure 4


A
SARS-CoV-2 N-antigen
N-terminal RNA binding domain
Sars N antigen
RNA Binding Domain
95 aa aligned
(RMSD=0.58Å)
Salivary
proteome
ref
B
SARS-CoV-2 N-antigen
N-terminal RNA binding domain
human RBM7 RRM domain
29 aa aligned
(RMSD=1.47Å)
Yes
C
SARS-CoV-2 N-antigen
N-terminal RNA binding domain
hnRNP H RRM domain
28 aa aligned
(RMSD=1.50Å)
Iso
D
SARS-CoV-2 N-antigen
N-terminal RNA binding domain
Human 17S U2 snRNP
27 aa aligned
(RMSD=1.52Å)
Yes
E
SARS-CoV-2 N-antigen
N-terminal RNA binding domain
Human Line-1 Orf1p
19 aa aligned
(RMSD=1.00Å)
Iso

579

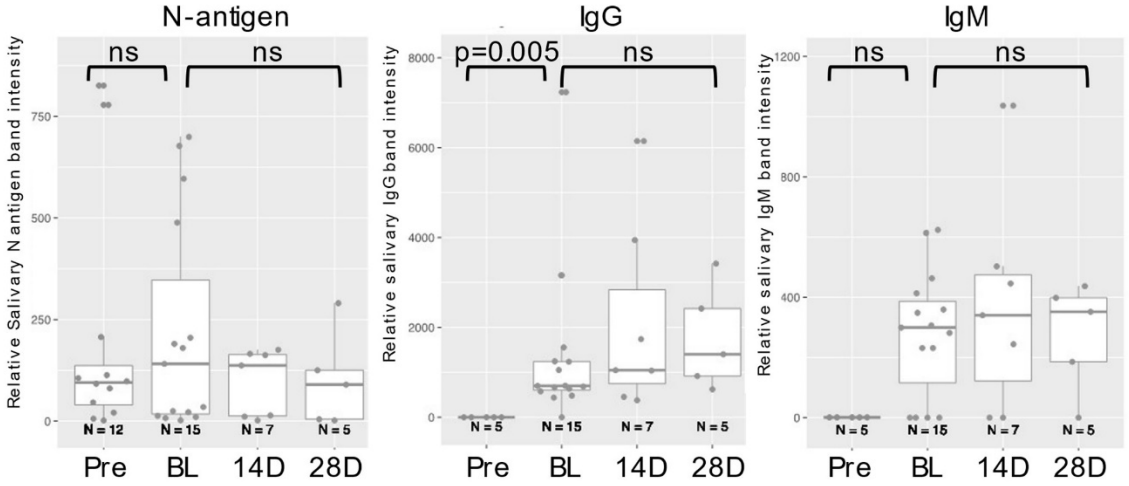
Figure 5


A
5'UTR
1a
1b
S
E
6
8
3'UTR
M
N
3a
7a
5'
3'
5'UTR primer
S
orf3a
E
M
N
Subgenomic transcripts
B
sgRNA S
% Total viral RNA
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
002
006
007
008
010
015
016
017
018
020
506
507
H2O
PreCov
*5'UTR/S junction
61 GTTCTCTAAACGAAC 75
21552 CTAAACGAACAATG 21565
C
sgRNA orf3a
% Total viral RNA
3.8
3.7
2.0
1.9
1.2
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
002
006
007
008
010
015
016
017
018
020
0517
506
507
H2O
PreCov
*5'UTR/orf3a junction
61 GTTCTCTAAACGAACTT 77
25382 TAAACGAACTTATGGA 25397
D
sgRNA N
% Total viral RNA
0.6
0.5
0.4
0.3
0.2
0.1
0.0
002
006
007
008
010
015
016
017
018
020
0517
506
507
H2O
PreCov
*5'UTR/N junction
61 GTTCTCTAAACGAAC 75
28254 TCTAAACGAACAAACT 28269

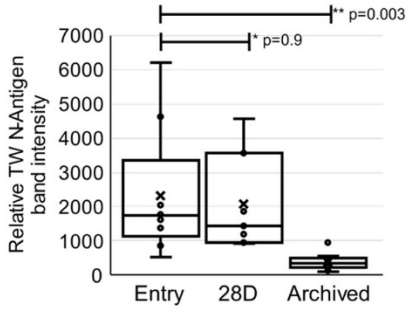
580

Figure 6

A



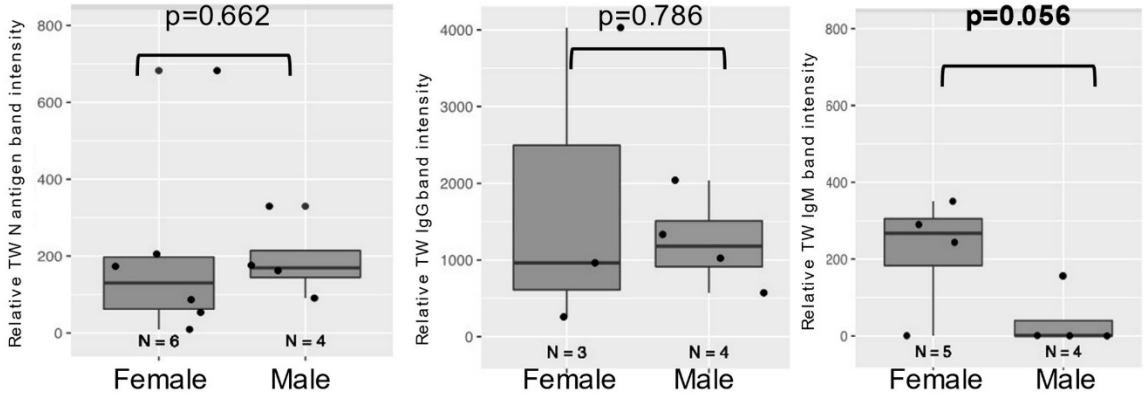
B



C

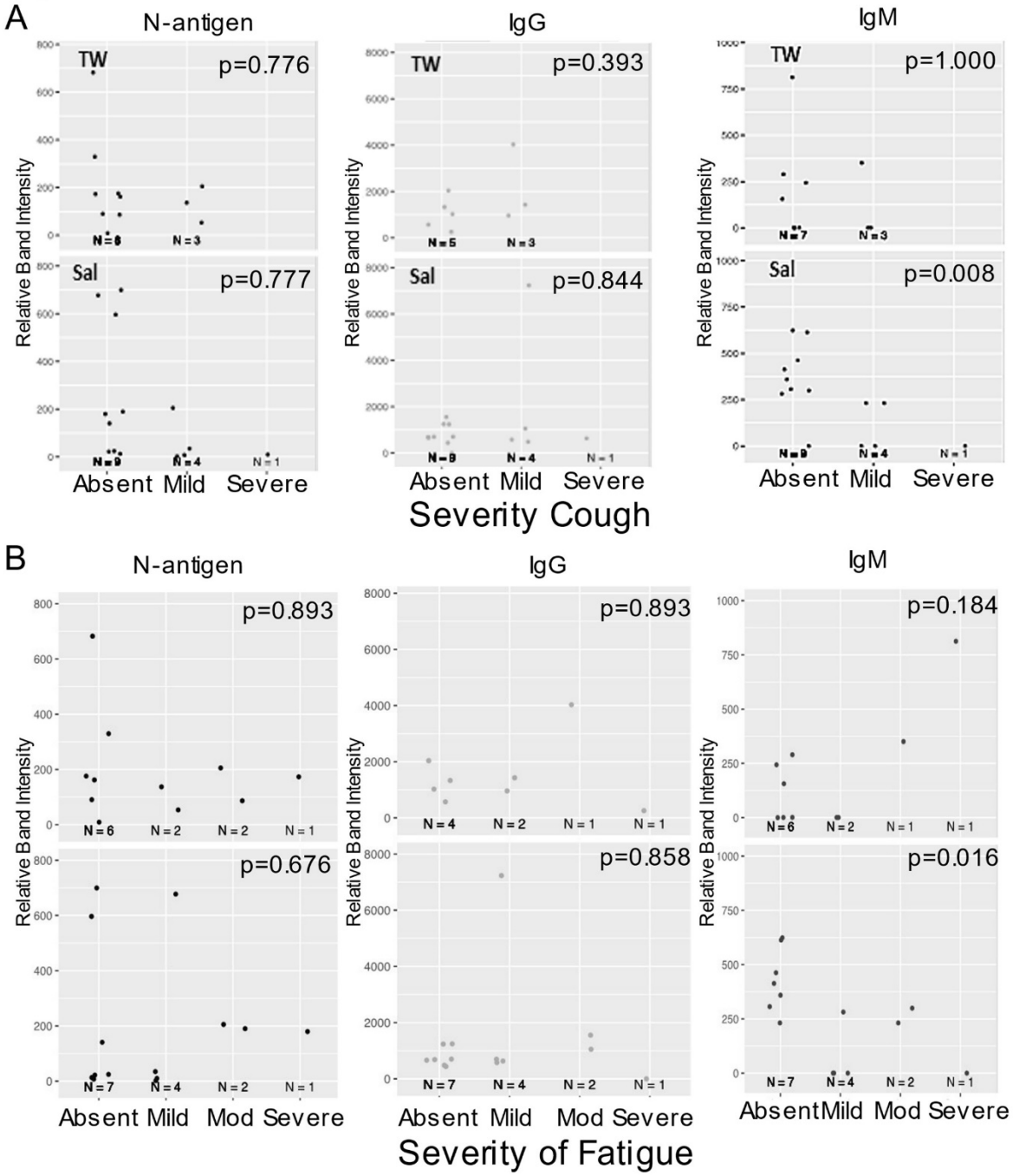
Type of Test	Detection	# Tested positive	Measure	Confidence Interval
Saliva	N Antigen (LFA)	9 / 16	0.563	(0.299, 0.802)
	N Antigen (IB)	10/10	1.000	
	IgG Antibody	15 / 16	0.938	(0.698, 0.998)
	IgM Antibody	12 / 16	0.750	(0.476, 0.927)
Throat-wash	N Antigen	10 / 11	0.910	(0.587, 0.998)
	N Antigen (IB)	8/10	0.800	
	IgG Antibody	8 / 8	1.000	(0.631, 1.000)
	IgM Antibody	5 / 10	0.500	(0.187, 0.813)

D



581

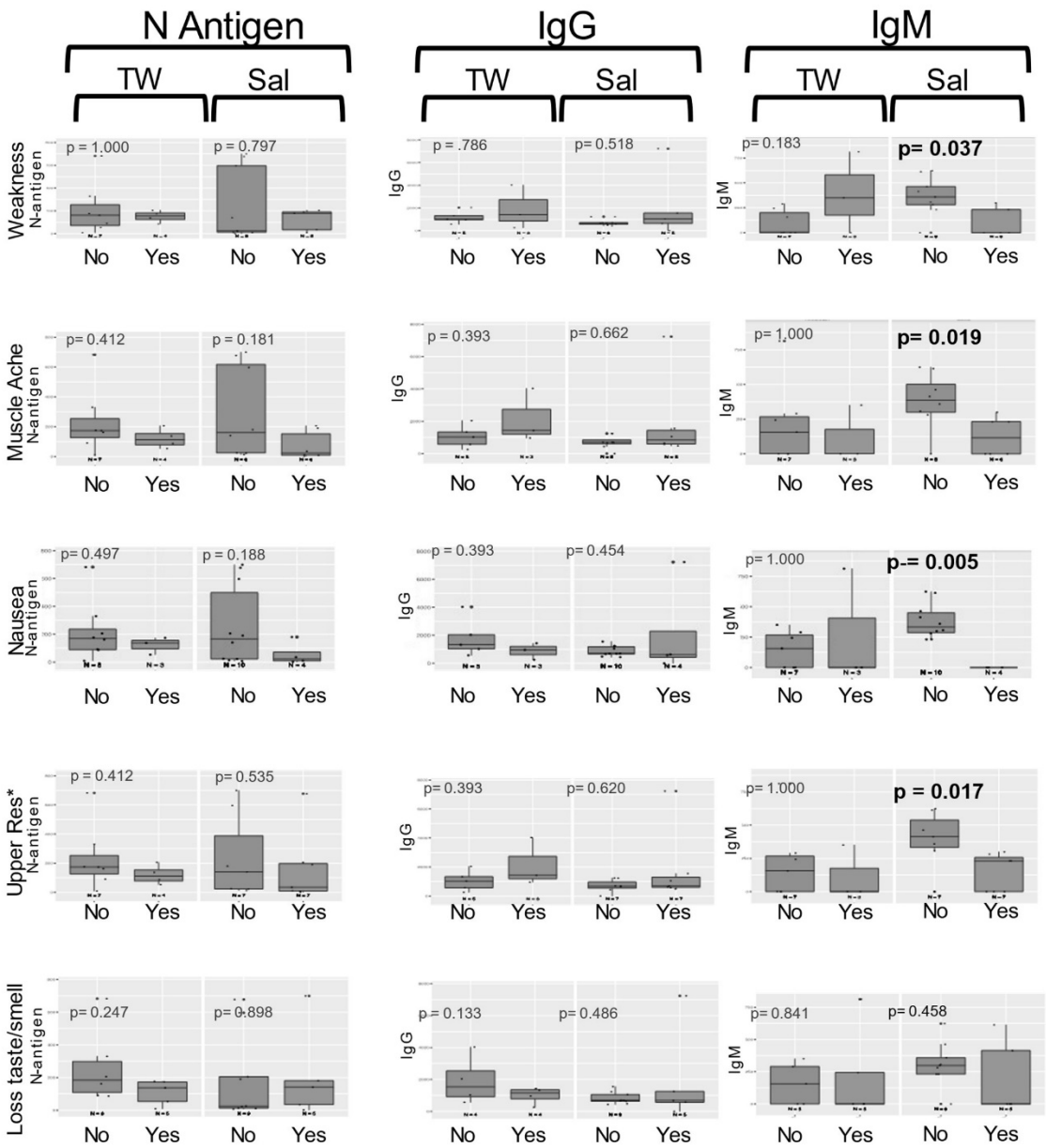
Figure 7


A
N-antigen
IgG
IgM
TW
p=0.776
p=0.393
p=1.000
Sal
p=0.777
p=0.844
p=0.008
Relative Band Intensity
Absent Mild Severe
Severity Cough
B
N-antigen
IgG
IgM
p=0.893
p=0.893
p=0.184
p=0.676
p=0.858
p=0.016
Absent Mild Mod Severe
Severity of Fatigue

582

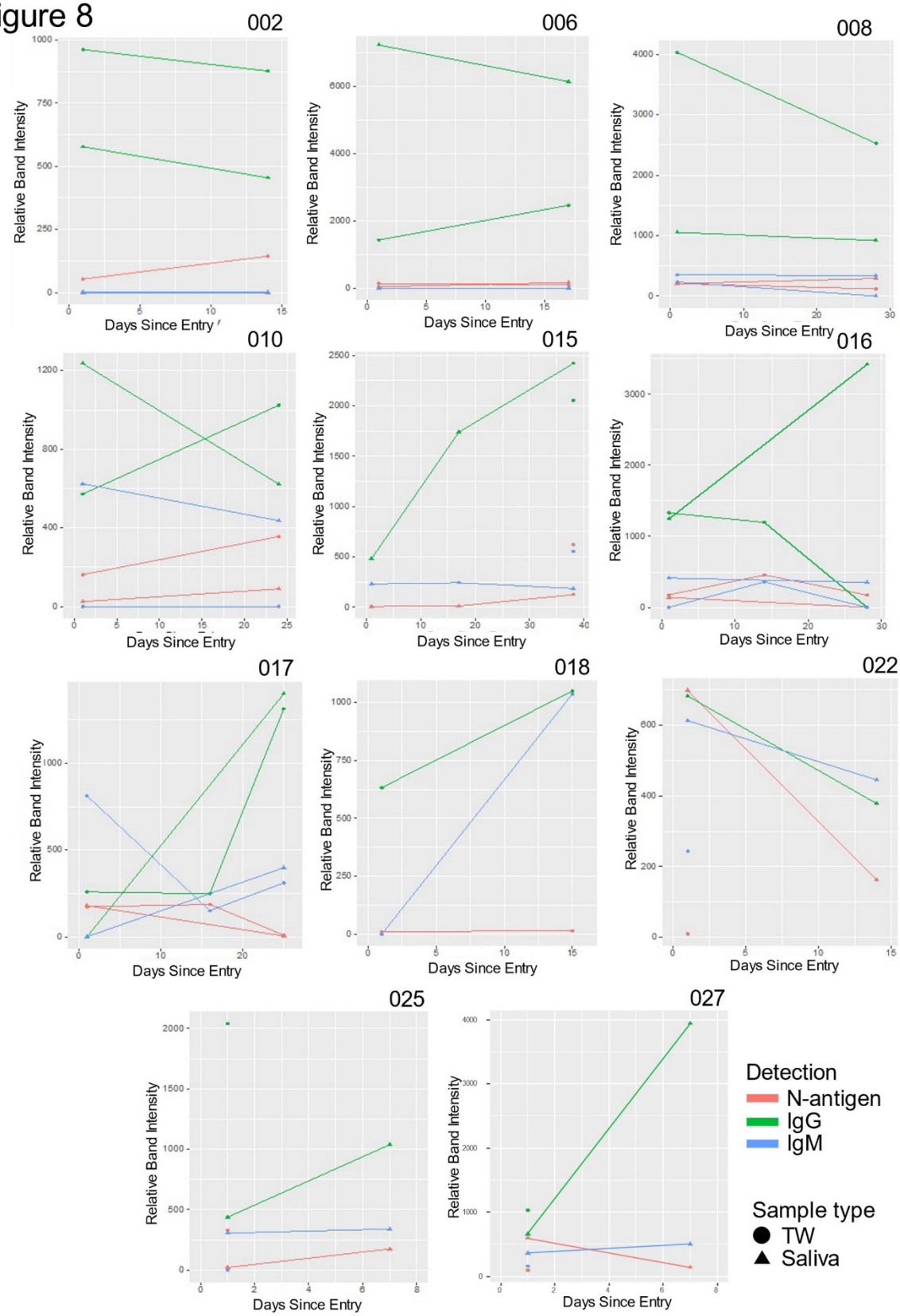
Figure 8

A



583

Figure 8
B



584

Figure 1

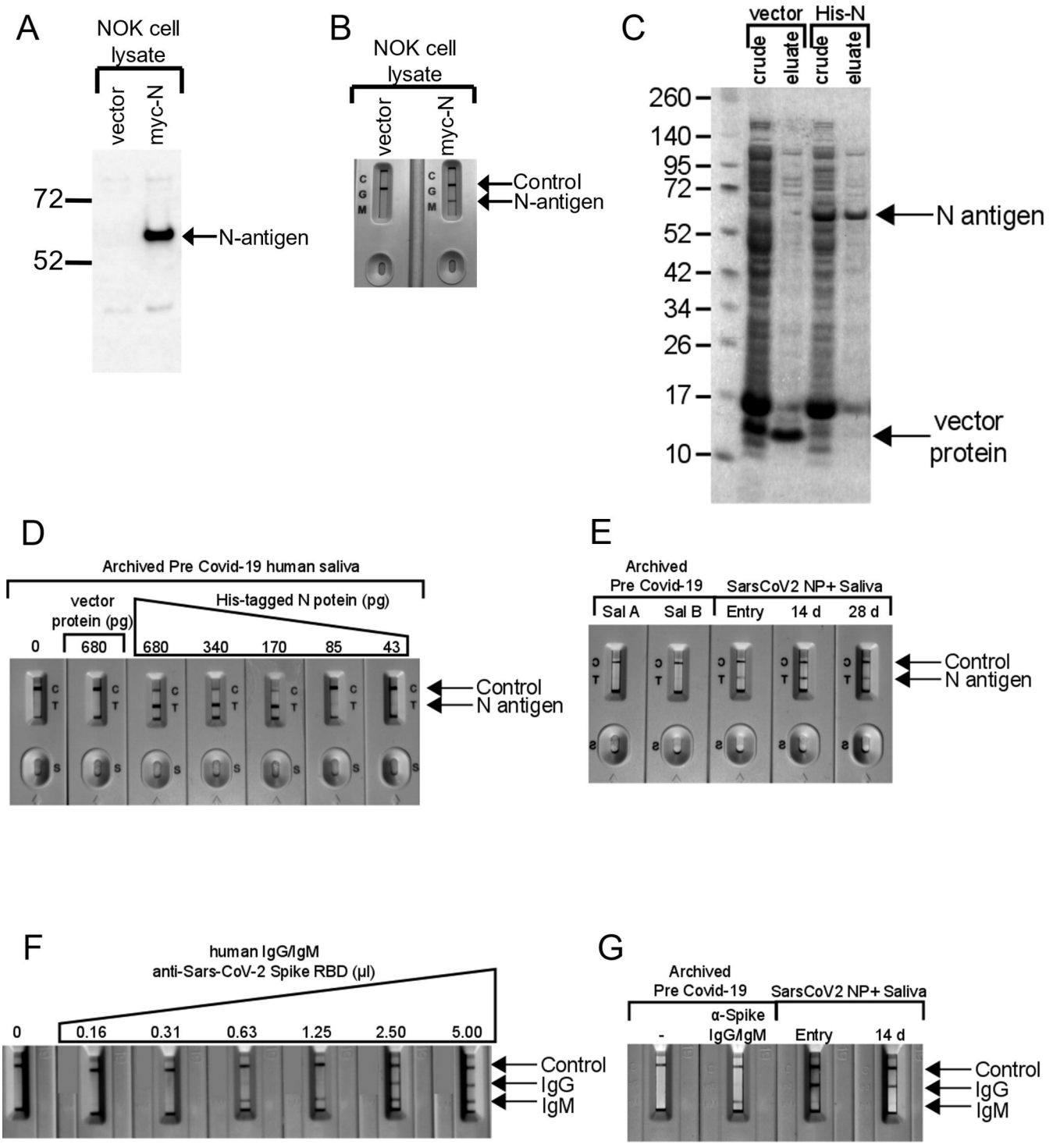
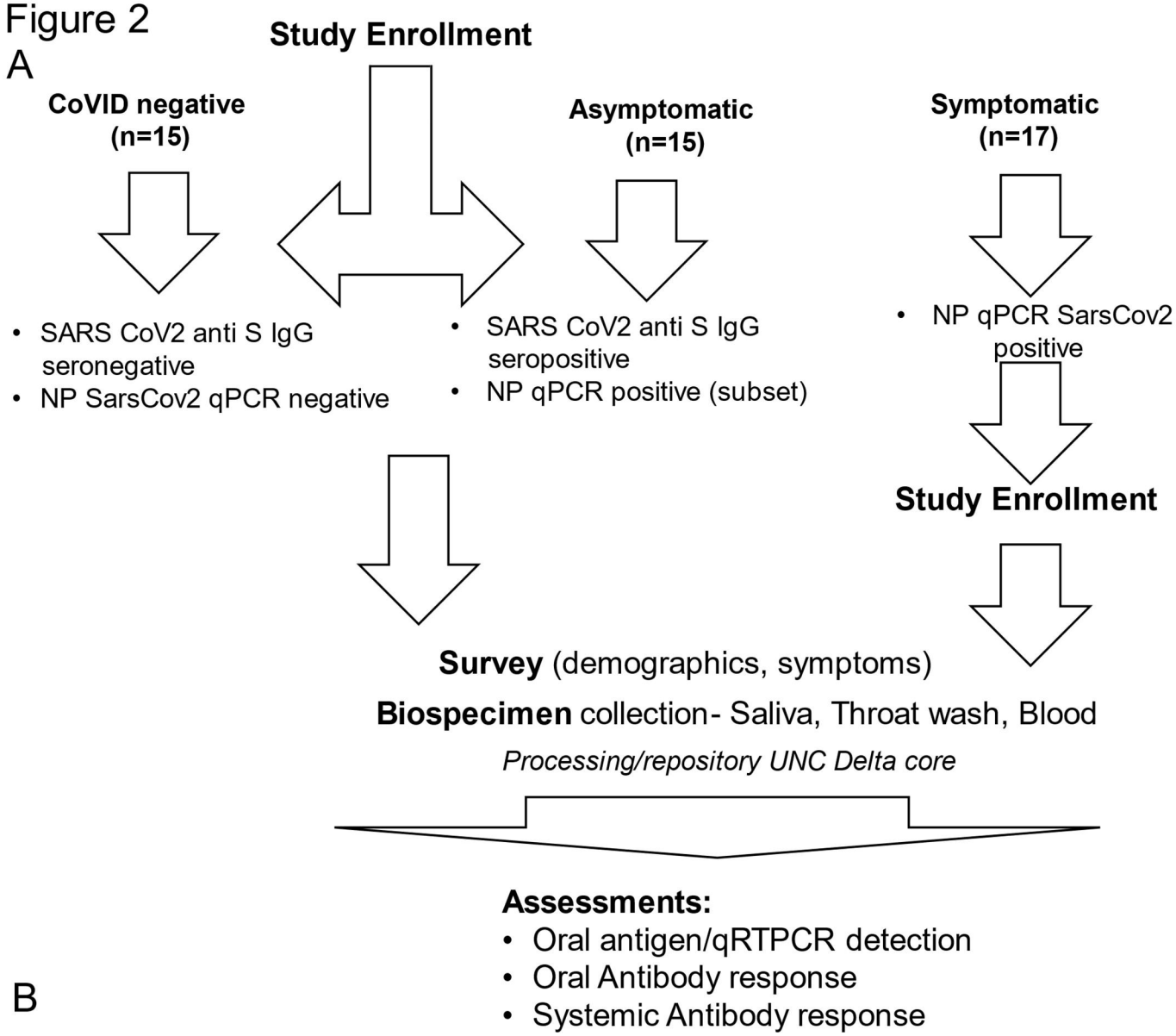

A
NOK cell lysate
vector
myc-N
72
52
N-antigen
B
NOK cell lysate
vector
myc-N
C G M
Control
N-antigen
C
vector
His-N
crude
eluate
crude
eluate
260
140
95
72
52
42
34
26
17
10
N antigen
vector protein
D
Archived Pre Covid-19 human saliva
vector protein (pg)
His-tagged N potein (pg)
0
680
680
340
170
85
43
C T
S
Control
N antigen
E
Archived Pre Covid-19
SarsCoV2 NP+ Saliva
Sal A
Sal B
Entry
14 d
28 d
Control
N antigen
F
human IgG/IgM anti-Sars-CoV-2 Spike RBD (µl)
0
0.16
0.31
0.63
1.25
2.50
5.00
Control
IgG
IgM
G
Archived Pre Covid-19
SarsCoV2 NP+ Saliva
-
α-Spike IgG/IgM
Entry
14 d
Control
IgG
IgM

Figure 2

A



B

Observational		N=17
Age	mean (min,max)	40.29 (38,62)
Sex	M, F, not provided	(8,8,1)
Race		
	White	12 (70.6%)
	Black	1 (5.9%)
	Asian	1 (5.9%)
	Missing	1 (5.9%)
	Unknown	2 (11.8%)
Ethnicity		
	Hispanic/ Latino	3 (20%)
	Non Hispanic /Latino	11 (73.3%)

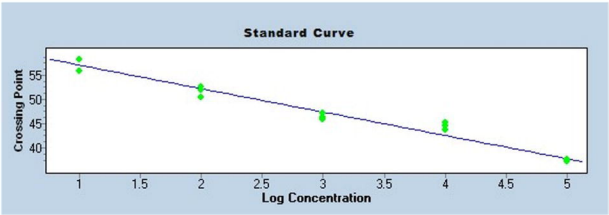
Asymptomatic/ Covid negative		N=30
Age	mean (min,max)	44 (18,65)
Sex	M, F	14,16
Race		
	White	21 (72.4%)
	Black	1 (3.45%)
	Asian	4 (13.79%)
	Hispanic	2 (6.9%)
	Other/multi races	1 (3.45%)
Ethnicity		
	Hispanic/ Latino	2 (6.9%)
	Non Hispanic/Latino	27 (93.1%)

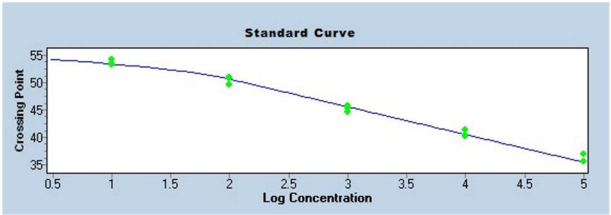
Figure 3

A

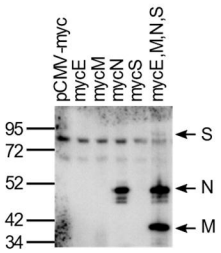
Experiment	Forward primer	sequence	Reverse primer	sequence
RT-qPCR	MN908947orf1F 16364-16386	TGTCTGTTAATCCGTATGTT TGC	MN908947orf1R comp. 16411-16437	ACCTCCTAAGTAAAGTTG AGTCACATC
RT-qPCR	E_Sarbeco_F 26269-26294	ACAGGTACGTTAATAGTTAA TAGCGT	E_Sarbeco_R compl. 26360-26381	ATATTGCAGCAGTACGCA CACA
RT-qPCR	2019-nCoV_N1-F 28283-28302	GACCCCAAAATCAGCGAA AT	2019-nCoV_N1-R comp. 28331-28354	TCTGGTTACTGCCAGTTG AATCTG
SsgRNA RT-qPCR	CoV25UTF 1-25	ATTAAAGGTTTATACCTTCC CAGGT	SRTqPCRR comp. 21622-21643	TGCAGGGGGTAATTGAGT TCTG
orf3asgRNA RT-qPCR	CoV25UTF 1-25	ATTAAAGGTTTATACCTTCC CAGGT	orf3aRTqPCRR Comp. 25474-25493	GTTGCAGTAGCGCGAACA AA
NsgRNA RT-qPCR	CoV25UTF 1-25	ATTAAAGGTTTATACCTTCC CAGGT	2019-nCoV_N1-R comp. 28331-28354	TGCAGGGGGTAATTGAGT TCTG
S gene expression	SexpF 21563-21587	ggggccatggaggcc**ATG**TTTGTT TTTCTTGTTTTATTGCC	SexpR comp. 25363-25384	cccgtcgac**TTA**TGTGTAATG TAATTTGACT
E gene expression	EexpF 26245-26265	ggggccatggaggcc**ATG**TACTC ATTCGTTTCGGAA	EexpR comp. 26449-26472	cccgtcgac**TTA**GACCAGAA GATCAGGAACTCT
M gene expression	MexpF 26523-26542	ggggccatggaggcc**ATG**GCAGA TTCCAACGGTAC	MexpR comp. 27167-27191	cccgtcgac**TTA**CTGTACAAG CAAAGCAATATTG
N gene expression	NexpF 28274-28297	ggggccatggaggcc**ATG**TCTGA TAATGGACCCCAAAAT	NexpR comp. 29512-29533	cccgtcgac**TTA**GGCCTGAGT TGAGTCAGCA

B

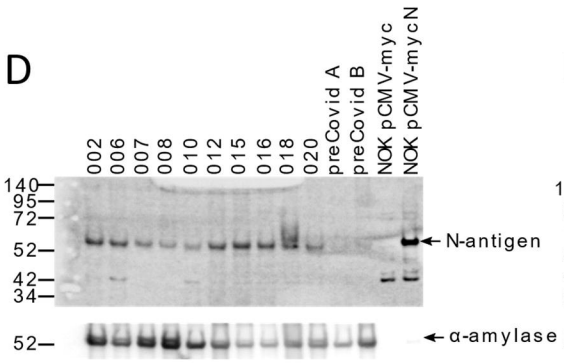




C



D



E

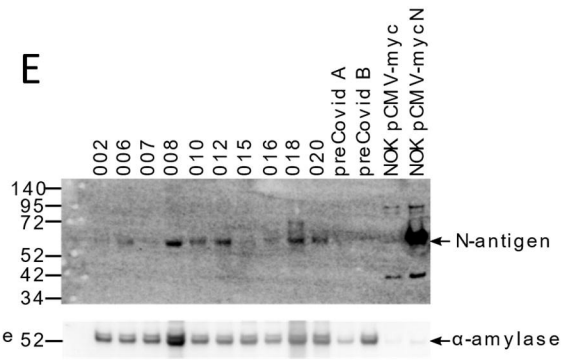


Figure 3

F

Symptomatic																
	RT-qPCR	Copy# N RNA	Copy# N RNA	ViralRNA saliva		sgRNA			N antigen LFA		N Antigen immunoblot		Anti-S IgG LFA		Anti-S IgM LFA	
Participant	NP	S	TW	E	ORF1	S	3a	N	S	TW	S	TW	S	TW	S	TW
002	+	32	34	-	-	-	-	-	-	+	+	-	+	+	-	-
006	+	216	104	-	-	-	+	-	-	+	+	+	+	+	+	-
007	+	-	46	-	-	-	+	-	++	++	+	-/+	++	+	-	+
008	+	352	122	-	-	-	+	-	+	++	+	+	+	++	-	+
010	+	-	78	+	-	-	+	-	-	+	+	+	+	+	+	-
012	+	420	ND	ND	ND	ND	ND	ND	++	ND	+	+	++	ND	+	ND
015	+	25	183	-	-	-	+	-	+	++	+	-	+	++	-	+
016	+	23440	116	+	+	+	+	+	+	+	+	-/+	+	+	+	-
017	+	48	253	+	-	-	+	-	+	+	ND	ND	+	+	+	+
018	+	15	6547	+	+	+	+	+	-	+	+	+	+	+	-	+
019	+	-	ND	ND	ND	ND	ND	ND	-	ND	ND	ND	-	ND	-	ND
020	+	28	-	-	-	-	-	-	+	+	+	-	++	ND	+	ND
022	+	-	-	ND	ND	ND	ND	ND	++	-	ND	ND	+	-	+	-
024	+	-	-	ND	ND	ND	ND	ND	-	++	ND	ND	+	+	+	+
025	+	-	-	ND	ND	ND	ND	ND	-	++	ND	ND	+	+	+	-
027	+	-	-	ND	ND	ND	ND	ND	++	-	ND	ND	++	+	+	+
0517	+	21233	ND	ND	ND	ND	+	+	ND	ND	ND	ND	ND	ND	ND	ND

G

Asymptomatic	Seropositive				
	NP	Blood	Saliva		
Participant	RT-PCR	Anti S serostatus	N antigen	Anti S IgG	Anti S IgM
1A	-	+	+	-	-
2A	-	+	-	-	-
3A	-	+	-	-	-
4A	-	+	+	-	-
5A	-	+	-	-	-
6A	-	+	-	-	-
7A	-	+	-	-	-
8A	+	+	-	+	+
9A	-	+	ND	-	-
10A	-	+	ND	-	-
11A	-	+	ND	-	-
12A	-	+	ND	-	-
13A	-	+	ND	-	-
14A	-	+	ND	-	-
15A	-	+	ND	-	-

H

CoVID (-)	Seronegative				
	NP	Blood	Saliva		
Participant	RT PCR	Anti S serostatus	N antigen	Anti S IgG	Anti S IgM
16A	-	-	+	-	-
17A	-	-	+	-	-
18A	-	-	+	-	-
19A	-	-	-	-	-
20A	-	-	-	-	-
21A	-	-	-	-	-
22A	-	-	-	-	-
23A	-	-	ND	-	-
24A	-	-	ND	-	-
25A	-	-	ND	-	-
26A	-	-	ND	-	-
27A	-	-	ND	-	-
28A	-	-	ND	-	-
29A	-	-	ND	-	-
30A	-	-	ND	-	-

Figure 4

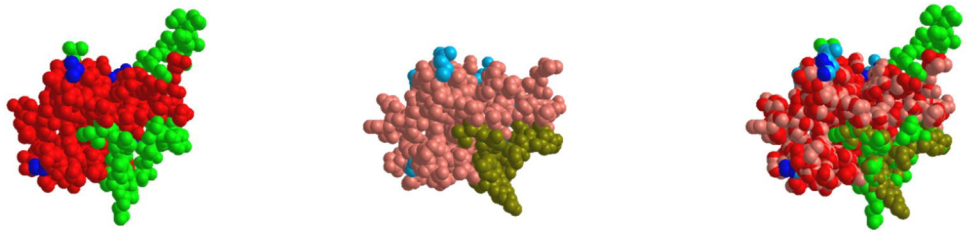
A

SARS-CoV-2 N-antigen
N-terminal RNA binding domain

Sars N antigen
RNA Binding Domain

95 aa aligned
(RMSD=0.58Å)

Salivary proteome ref

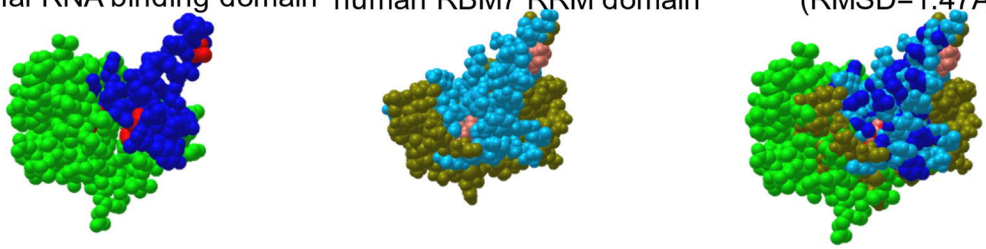


B

SARS-CoV-2 N-antigen
N-terminal RNA binding domain

human RBM7 RRM domain

29 aa aligned
(RMSD=1.47Å)



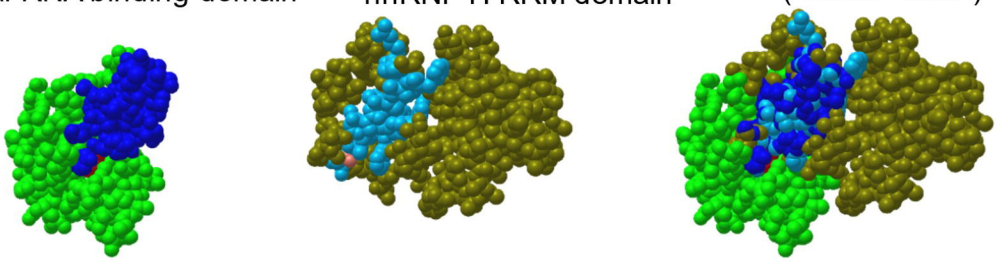
Yes

C

SARS-CoV-2 N-antigen
N-terminal RNA binding domain

hnRNP H RRM domain

28 aa aligned
(RMSD=1.50Å)



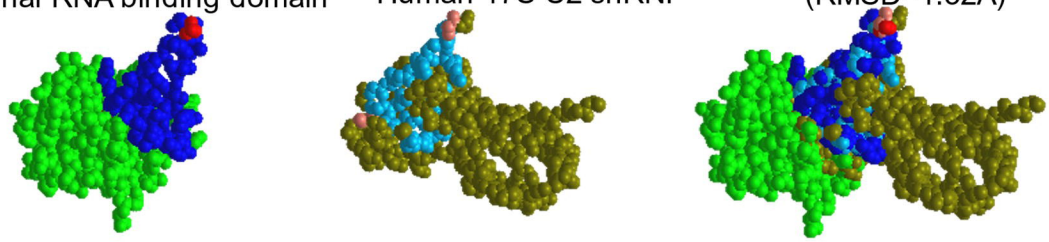
Iso

D

SARS-CoV-2 N-antigen
N-terminal RNA binding domain

Human 17S U2 snRNP

27 aa aligned
(RMSD=1.52Å)



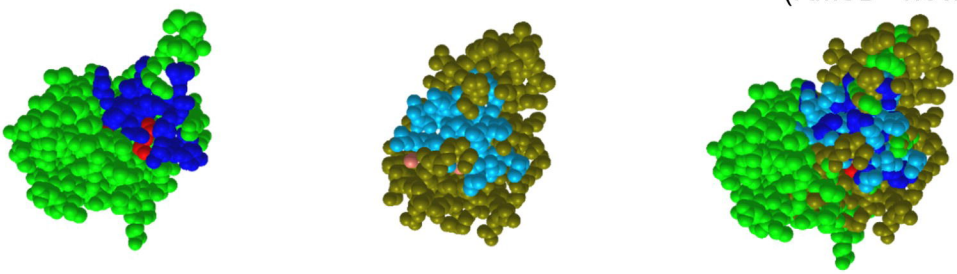
Yes

E

SARS-CoV-2 N-antigen
N-terminal RNA binding domain

Human Line-1 Orf1p

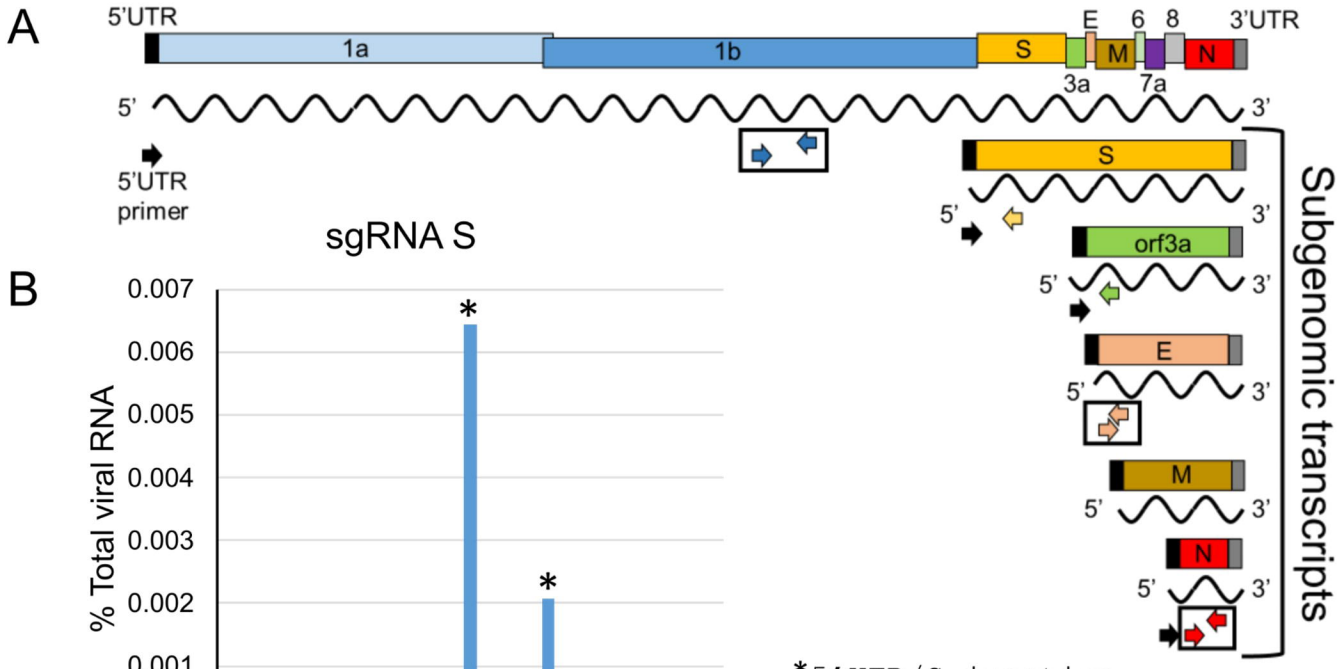
19 aa aligned
(RMSD=1.00Å)



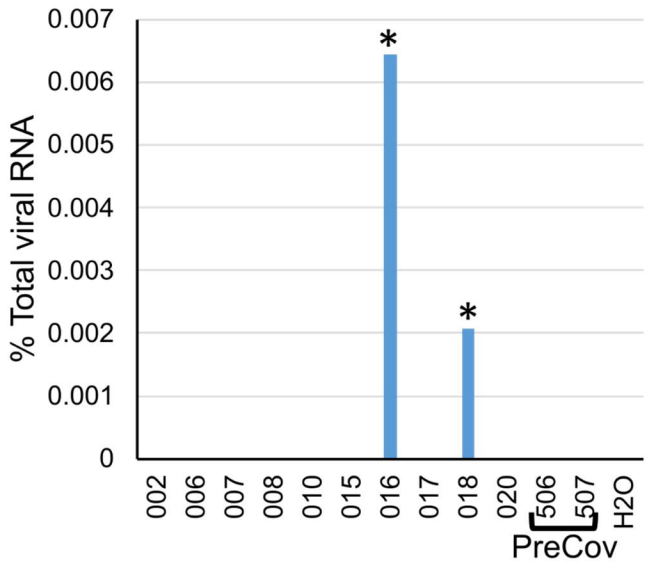
Iso

Figure 5

A

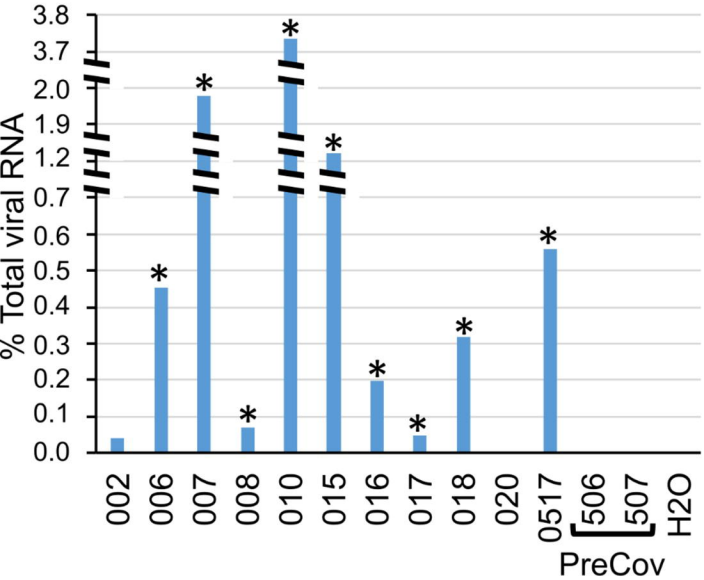


B

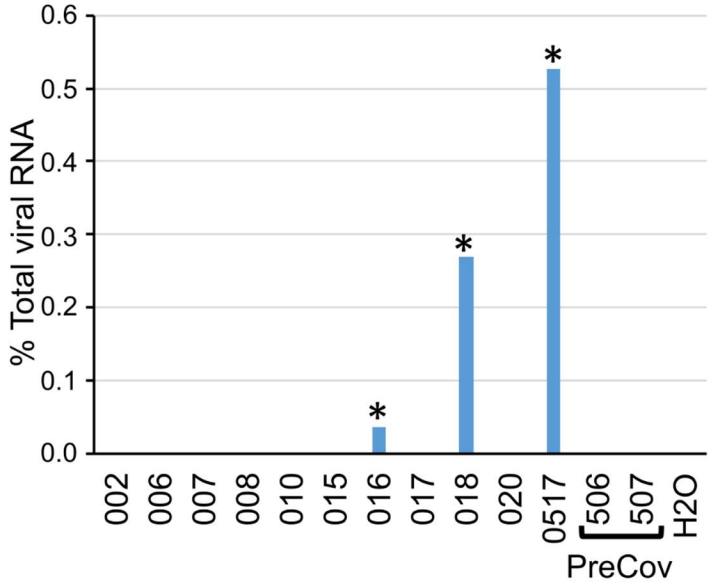


*5'UTR/S junction
61 GTTCT**CTAAACGAAC** 75
||||||||||
21552 **CTAAACGAAC**AATG 21565

C



D

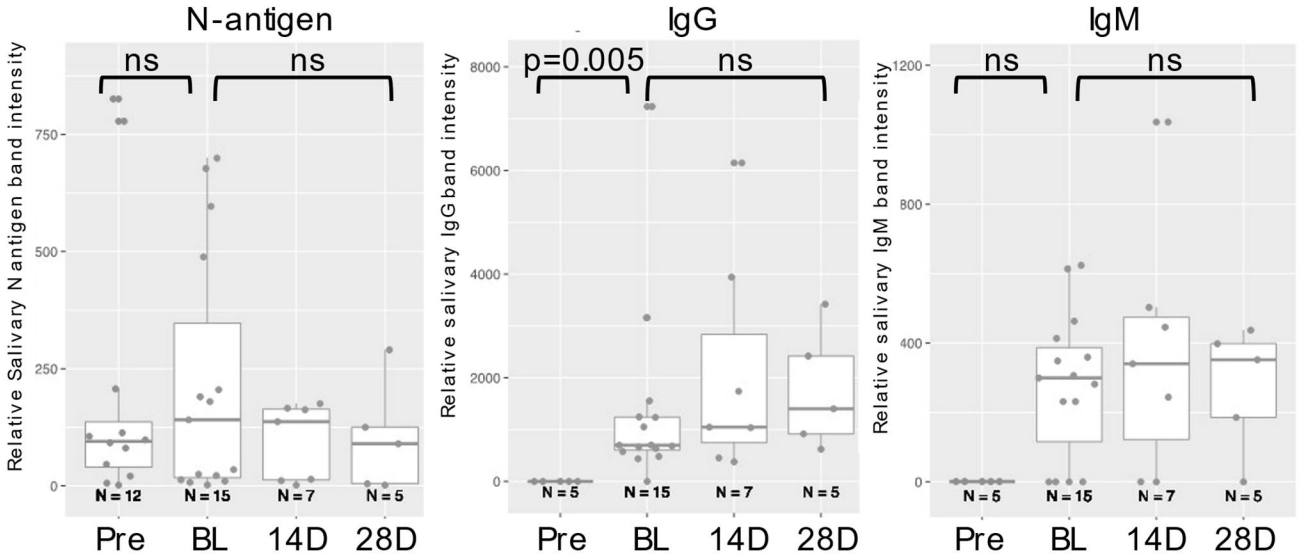


*5'UTR/orf3a junction
61 GTTCTC**TAAACGAACTT** 77
||||||||||||
25382 **TAAACGAACTT**ATGGA 25397

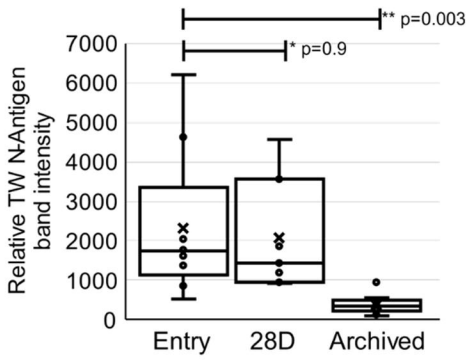
*5'UTR/N junction
61 GTTC**TCTAAACGAAC** 75
||||||||||||
28254 **TCTAAACGAAC**AAACT 28269

Figure 6

A



B



C

Type of Test	Detection	# Tested positive	Measure	Confidence Interval
Saliva	N Antigen (LFA)	9 / 16	0.563	(0.299, 0.802)
	N Antigen (IB)	10/10	1.000	
	IgG Antibody	15 / 16	0.938	(0.698, 0.998)
	IgM Antibody	12 / 16	0.750	(0.476, 0.927)
Throat-wash	N Antigen	10 / 11	0.910	(0.587, 0.998)
	N Antigen (IB)	8/10	0.800	
	IgG Antibody	8 / 8	1.000	(0.631, 1.000)
	IgM Antibody	5 / 10	0.500	(0.187, 0.813)

D

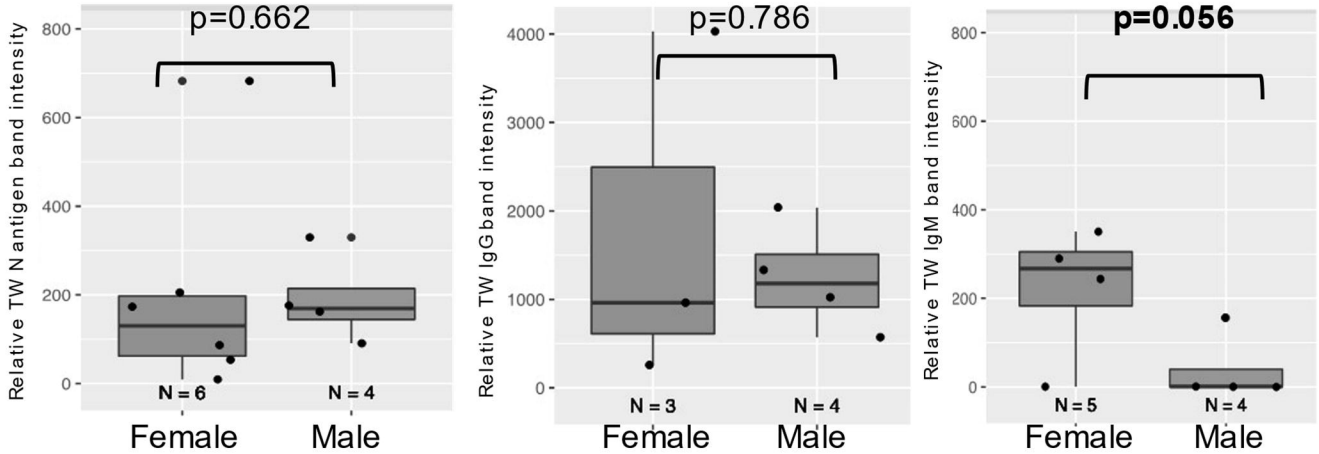


Figure 7

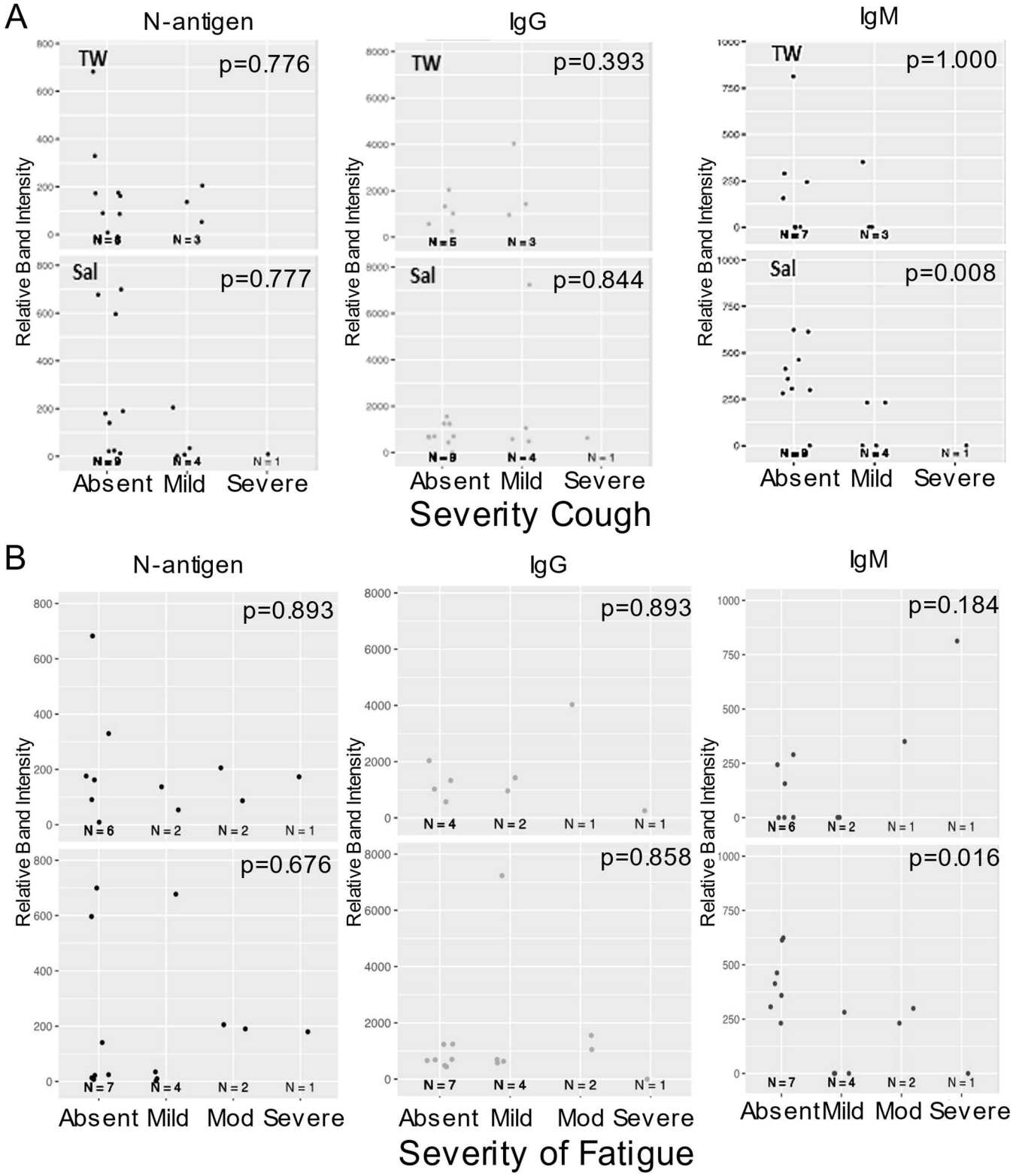

A
N-antigen
IgG
IgM
TW
Sal
p=0.776
p=0.777
p=0.393
p=0.844
p=1.000
p=0.008
Relative Band Intensity
Absent Mild Severe
Severity Cough
B
p=0.893
p=0.676
p=0.893
p=0.858
p=0.184
p=0.016
Absent Mild Mod Severe
Severity of Fatigue

Figure 8

A

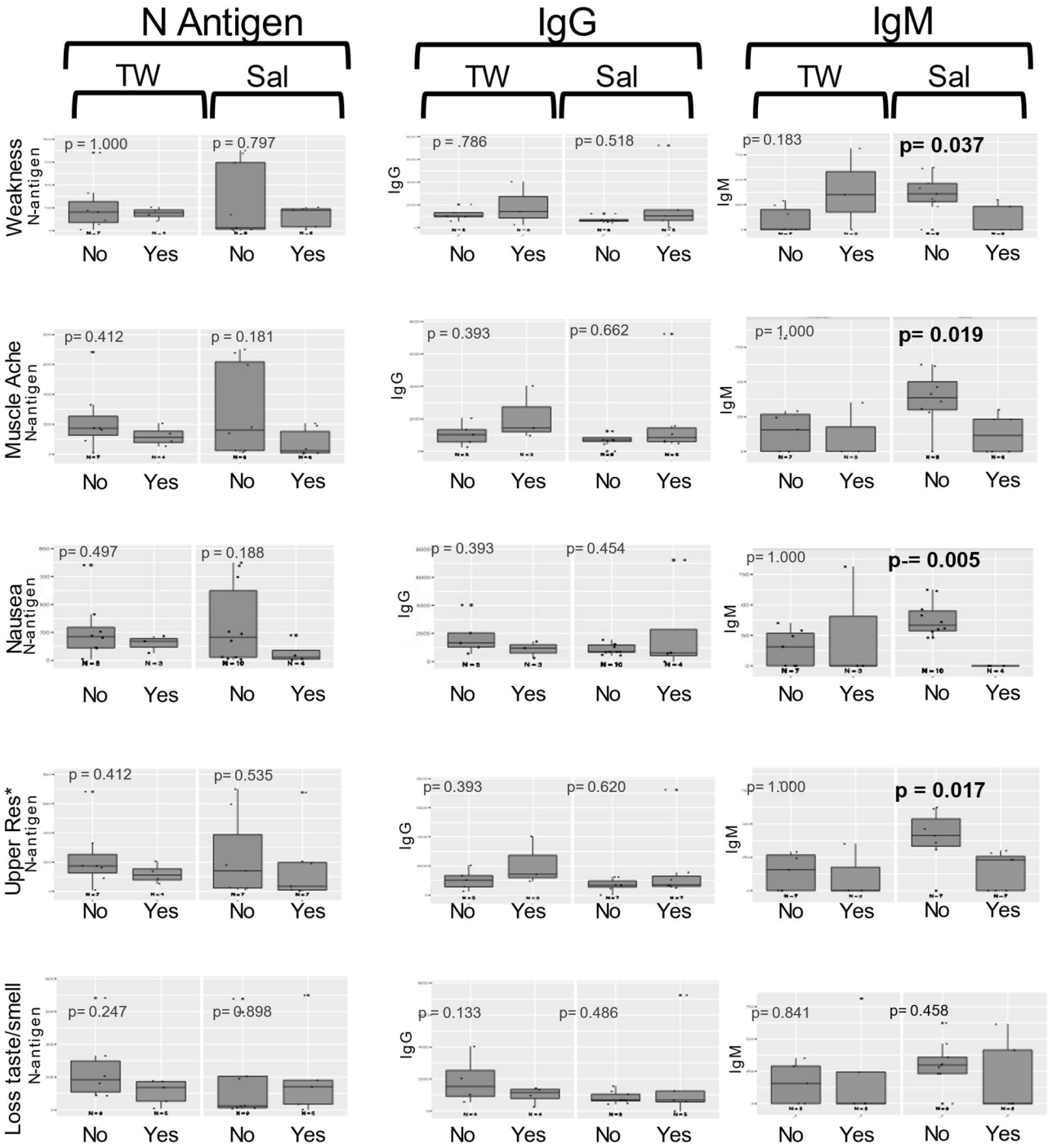
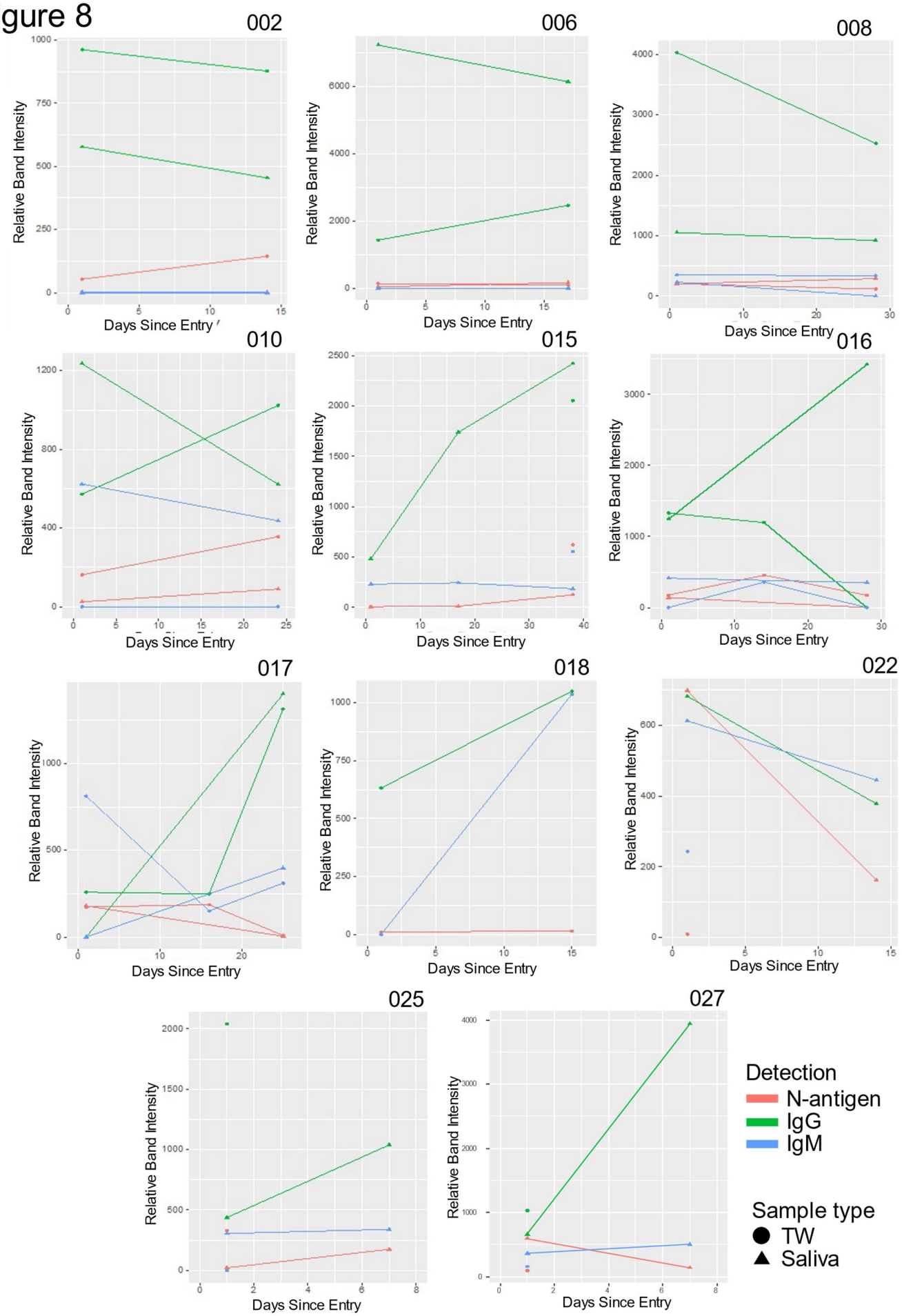

N Antigen
IgG
IgM
TW
Sal
Weakness
N-antigen
p = 1.000
p = 0.797
p = .786
p= 0.518
p= 0.183
p= 0.037
Muscle Ache
N-antigen
p= 0.412
p = 0.181
p = 0.393
p= 0.662
p= 1.000
p= 0.019
Nausea
N-antigen
p= 0.497
p = 0.188
p = 0.393
p= 0.454
p= 1.000
p-= 0.005
Upper Res*
N-antigen
p = 0.412
p= 0.535
p= 0.393
p= 0.620
p= 1.000
p = 0.017
Loss taste/smell
N-antigen
p= 0.247
p= 0.898
p = 0.133
p= 0.486
p= 0.841
p= 0.458
IgG
IgM
No
Yes

Figure 8
B


002
006
008
010
015
016
017
018
022
025
027
Relative Band Intensity
Days Since Entry
Detection
N-antigen
IgG
IgM
Sample type
TW
Saliva