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

William T Seaman

National Institute of Dental and Craniofacial Research, National Institutes of Health

Olive Keener

University of North Carolina

Wenwen Mei

University of North Carolina

Katie R Mollan

University of North Carolina

Corbin D Jones

University of North Carolina

Audrey Pettifor

University of North Carolina

Natalie M Bowman

University of North Carolina

Frank Wang

Biomedomics Inc

Jennifer Webster-Cyriaque (✉ Jennifer.Webster-cyriaque@nih.gov)

National Institute of Dental and Craniofacial Research, National Institutes of Health

Article

Keywords:

Abstract

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.

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. A total of 47 Covid-19 era participants were assessed. Inclusion criteria for entry into the symptomatic longitudinal observational cohort (n = 17) required subjects to be NP SARS-CoV-2 RT-qPCR positive. COVID-19 + patients were recruited after written informed consent and were stratified with mild, moderate or severe symptoms at the time of presentation based on the NIH criteria (NIH COVID-19 treatment guidelines). For the asymptomatic cohort (n = 30), participants were either asymptomatic, anti-Sars-CoV-2 spike seronegative (n = 15) or asymptomatic, Sars-CoV-2 spike seropositive (n = 15) institution campus dwellers including students, staff and faculty who provided written informed consent. Archived, pre-COVID pandemic saliva and throat wash samples were used as SARS-CoV-2 negative controls. All samples were obtained and stored under IRB approval (Symptomatic participants, UNC IRB study #20−0792; Asymptomatic participants, UNC IRB study #20−1771; Archived, PreCovid participants, UNC IRB study #07-1431). All experimental protocols used to assess specimens were performed in accordance with approval by the University of North Carolina IRB (Symptomatic participants, UNC IRB study #20−0792; Asymptomatic participants, UNC IRB study #20−1771; Archived, PreCovid participants, UNC IRB study #07-1431). All methods used to assess specimens were performed in accordance with the relevant guidelines put forth by the University of North Carolina.

Biospecimens For RT-qPCR, NPS, saliva and throat wash samples were collected concurrently. The NPS samples were collected in 3ml of viral transport medium (VTM). Unstimulated whole mouth fluid (WMF) samples (saliva) and throat wash gargle with 10 ml of normal saline (throat wash) were each collected in sterile 50mL wide-mouthed screw-capped containers. Samples were immediately transported for processing where the NPS, saliva, throat wash, and blood samples were stored at 4°C and processed within 24 hours. After processing 1 ml aliquots of samples were stored at -80°C. Similarly, 1 ml aliquots of archived, preCovid-19 saliva and throat wash samples have been stored at -80°C for > 5 years.

Nucleic acid extraction and RT-qPCR. Saliva and throat wash from COVID-19 patients were used as a source for the detection of SARS-CoV-2 RNA. Trizol (Life Technologies, Carlsbad, CA) was used to inactivate virus and RNA was extracted according to the manufacturer's instructions. Briefly, 750

microliters of Trizol were added to 250 microliters of oral fluid (saliva or throat wash). Following chloroform addition and phase separation, the aqueous phase was collected. Glycogen (2 micrograms) was added to the aqueous phase and RNA was precipitated with isopropanol. RNA pellets were dried and resuspended in water.

RNA was reverse transcribed using Superscript III (Life Technologies) according to the manufacturer's instructions. First strand synthesis was primed with random hexamers (Life Technologies). First strand synthesis reactions were used in SybrGreen-based qPCR. Three regions of the SARS-CoV-2 genome were targeted using virus-specific primers (Fig. 3A): WTSorf1F/WTSorf1R, E_Sarbeco_F/E_Sarbeco_R[14] and 2019-nCoV_N1-F/2019-nCoV_N1-R[15]. A standard curve was generated using SARS-CoV-2 reference RNA (ATCC VR-3276SD) for absolute copy number quantitation.

Subgenomic RNA (sgRNA) detection was performed using the forward primer, CoV25UTR (nucleotides 1–25; GenBank:MN908947) with the reverse primers, SRTqPCRR, orf3aRTqPCRR or 2019-nCoV_N1-R (Fig. 3A). Percent sgRNA was calculated using the equation: $2^{-(sgRNA\ Ct\ -\ N\ Ct)}$ X 100 where sgRNA Ct values were signal obtained using the 5'UTR/specific sgRNA primer pair and the N Ct value was signal obtained using the 2019-nCoV_N1-F/2019-nCoV_N1-R primer pair. Following qPCR, reactions from positive wells were removed, treated with ExoSapIt according to the manufacturer's instructions (USB, Cleveland, Ohio) and Sanger sequencing was performed (Eton Bioscience, Research Triangle Park, NC) using gene-specific primers: SRTqPCRR for SsgRNA, orf3aRTqPCRR for orf3asgRNA or 2019-nCoV_N1-R for NsgRNA.

Ectopic expression of SARS-CoV-2 proteins in transfected oral keratinocytes. Total RNA was isolated from the saliva of SARS-CoV-2 -infected individuals. RT-qPCR was used to amplify cDNA encoding SARS-CoV-2 Spike (S), Envelope (E) or N-antigen. The generated cDNAs were digest with SfiI and SalI and cloned into SfiI/SalI sites of pCMV-myc eukaryotic expression plasmid. Expression in eukaryotic cells would result in the production of viral proteins with an amino-terminal c-myc tag. Immortalized human oral keratinocytes (NOK) were grown in keratinocyte serum-free media (Life Technologies, Carlsbad, CA). Cells were transfected with expression vectors using Fugene-6 (Promega, Madison, WI). Forty-eight hours post transfection, media was removed and subjected to SDS-PAGE and immunoblot analysis to detect myc-tagged proteins.

To obtain lysates from whole saliva, TW or transfected NOK, protein was isolated from the organic phase obtained using Trizol (Life Technologies, Carlsbad, CA) according to the manufacturer's instructions. Protein was precipitated after the addition of isopropanol and centrifugation at 12,000 X G. Pellets were washed twice with 95% ethanol/0.3 M guanidine-HCl followed by a final wash with 100% ethanol. Each wash was performed for 20 minutes. Final protein pellets were dried and resuspended in 1X sample buffer (250 mM Tris, 8.0, 2% SDS, 20% glycerol, 50 mM DTT, 0.1% bromophenol blue). Pellets in buffer were incubated at 65° C until dissolved.

Immunoblot analysis of NOK cell lysates and media. Twenty-five micrograms of whole lysate or 10 microliters of transfected cell culture media were loaded onto precast NuPage 4–12% Bis-Tris gels (Life Technologies, Carlsbad, CA). Proteins were electroblotted to PVDF membrane using western blot transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) for 2 hours at 200 mamps, constant current. Following transfer, blots were blocked by incubation with 5% nonfat dry milk in 1X PBS containing 0.1% Tween-20 (1X PBS-T) at room temperature. Following blocking, blots were incubated with primary antibody overnight at 4° C. Blots were washed twice with 1X PBS-T followed by incubation with HRP-conjugated secondary antibody (Promega) diluted 1:10,000 in PBS-T, 5% milk for 1 hour at room temperature. Blots were washed twice with PBS-T and protein bands were detected by ECL after incubation with ECL Prime (GE Healthcare, Chicago, Il) according to the manufacturer's instructions. Blots were imaged using a GE ImageQuant LAS4000 system. The following primary antibodies were used to detect proteins on western blots: rabbit anti SARS-CoV-2 nucleocapsid (PA5-114448, Life Technologies), mouse anti-c-myc (sc-40, Santa Cruz Biotechnology, Dallas, TX), anti-amylase (A8273, Sigma-Aldrich, St. Louis, MO).

Overexpression of His-tagged SARS-CoV-2 Nucleocapsid in E. coli. The NcoI/SalI fragent from pCMV-mycN containing the entire coding region of N was removed and inserted into the NcoI/SalI sites of pET30a. The resulting plasmid, containing the SARS-CoV-2 nucleocapsid (N) coding region in frame with the 6X histidine tagged coding region of pET30, was used to transform E. coli BL21 (DE3) bacteria. Bacteria with pET30a empty vector were used as a source of negative control his-tagged protein. Bacteria containing pET30a will produce an ~ 9 kD his-tagged protein after induction with IPTG. Single colonies were grown in LB medium containing kanamycin (50 micrograms/ml) to an OD A600 of 1.0. To induce gene expression, IPTG (1 mM final concentration) was add and cultures were incubated at 30° C for 2 hours. Following centrifugation of cultures, bacterial pellets were resuspended in 1 ml NPI-10 (300 mM NaCl, 50 mM sodium phosphate, pH8.0, 10 mM imidazole, 1 mg/ml lysozyme). Lysates were incubated with NTA-Ni agarose beads (Life Technologies) to purify his-tagged proteins. Beads were washed with NPI-20 (300 mM NaCl, 50 mM sodium phosphate, pH8.0, 20 mM imidazole). Protein was eluted by resuspending beads in 250 microliters NPI-500 (300 mM NaCl, 50 mM sodium phosphate, pH8.0, 500 mM imidazole). Eluates were subjected to SDS-PAGE followed by Coomassie staining to assess retrieval and protein purity. Protein concentration was determined by Bradford assay.

Lateral Flow Assay. Detection of SARS-CoV-2 N-antigen or anti-SARS-CoV-2 spike protein IgG and IgM was accomplished using lateral flow chambers (BioMedomics, Research Triangle Park, NC). For N-antigen 50 microliters of sample (saliva or throat wash) was diluted with 50 microliter of lysis buffer (supplied by the manufacturer). Samples in lysis buffer were incubated at room temperature. Following incubation, 80 microliters of sample was applied to the lateral flow chamber. Bands were visualized and quantitated using ImageJ software. The control band for each detection cartridge was used to normalize antigen-specific band. For detection of anti-spike IgG/IgM, 50 microliters of sample (saliva or throat wash) was added directly to the lateral flow chamber. Two drops of COVID-19 IgG/IgM rapid test buffer (supplied by manufacturer) were added and bands were allowed to develop. Control bands were used

normalize IgG and IgM intensity. Quantitation was determined relative to signal produced by pre-COVID-19 archived saliva and throat wash samples.

NCBI Structural analysis

The crystal structure of the N-terminal RNA binding domain of SARS-CoV-2 N antigen (PDB ID: 6M3M) was compared to other publicly available crystal structures in the Molecular Modeling Database (MMDB) using Vector Alignment Search Tool (Vast+) within the Domains & Structures module on the National Center for Biotechnology Information website (https://www.ncbi.nlm.nih.gov). Vast + makes geometric structural comparisons between macromolecules in the absence of sequence similarity. The 3D structures of superimposed biological assemblies were visualized using the web-based 3D viewer iCn3D version 4.3.1 to support the visualization style[11].

Statistical methods

Given the small sample sizes, non-parametric tests were used to test correlations and compare data. Kendall rank correlation and Mann – Whitney U tests were used to test the correlation between the independent and response variables. 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 was used to determine 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 the groups were dependent, Wilcoxon signed rank tests were used to assess, 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 (Fig. 1A) or LFA (Fig. 1B). Bacteria-produced, his-tagged N-antigen was purified (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 1G). A symptomatic participant demonstrated IgG at baseline and at 14 days but not IgM (Fig. 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 (Fig. 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 (Fig. 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(Fig. 3F). All full-length and subgenomic (sgRNA) viral RNAs are amplified by 2019-nCoV_N1-F/2019-nCoV_N1-R allowing relative quantitation (Fig. 3A). Reactions containing > 10 copies were consistently reproducible (Fig. 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+ (Fig. 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 (Fig. 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 (Fig. 3D). TW immunoblot detected N-antigen in 7/10 subjects with 2/7 displaying faint bands (Fig. 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%) (Fig. 3F). 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 (Fig. 3G). Two others (1A and 4A) within the seropositive-asymptomatic group were LFA positive (Fig. 3G). Among seronegative subjects, salivary IgG and IgM were not detected however, 3/7 subjects demonstrated N-antigen by LFA (Fig. 3H). To determine

potential targets capable of cross-reactive binding, host salivary proteins structurally analogous to N were assessed (Fig. 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 (Fig. 4).

Detection of sgRNAs, arguably only produced during active viral replication, confirmed TW LFA positivity (Fig. 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, Fig. 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 (Figs. 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 (Fig. 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) (Fig. 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) (Fig. 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 (Fig. 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) (Fig. 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) (Fig. 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) (Fig. 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) (Fig. 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) (Fig. 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 (Fig. 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 (Fig. 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.

Declarations

Data availability

The data supporting the results in this manuscript are available in the manuscript and in the supplementary data files.

References

1. Huang, N. *et al.* SARS-CoV-2 infection of the oral cavity and saliva. *Nat Med* **27**, 892-903, doi:10.1038/s41591-021-01296-8 (2021).
2. Kannian, P. *et al.* in *Oral Dis* (2021).
3. Chen, J. H. *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* **9**, 1356-1359, doi:10.1080/22221751.2020.1775133 (2020).

4. Azzi, L. *et al.* Saliva is a reliable tool to detect SARS-CoV-2. *J Infect*, doi:10.1016/j.jinf.2020.04.005 (2020).

5. To, K. K. *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*, doi:10.1016/s1473-3099(20)30196-1 (2020).

6. Gupta, S. *et al.* SARS-CoV-2 Detection in Gingival Crevicular Fluid. *J Dent Res* **100**, 187-193, doi:10.1177/0022034520970536 (2021).

7. Ibrahimi, N. *et al.* Screening for SARS-CoV-2 by RT-PCR: Saliva or nasopharyngeal swab? Rapid review and meta-analysis. *PLoS One* **16**, e0253007, doi:10.1371/journal.pone.0253007 (2021).

8. Higgins, R. L. *et al.* Longitudinal SARS-CoV-2 antibody study using the Easy Check COVID-19 IgM/IgG lateral flow assay. *PLoS One* **16**, e0247797, doi:10.1371/journal.pone.0247797 (2021).

9. Williams, E. *et al.* Detection of SARS-CoV-2 in saliva: implications for specimen transport and storage. *J Med Microbiol* **70**, doi:10.1099/jmm.0.001285 (2021).

10. Dinnes, J. *et al.* Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev* **3**, CD013705, doi:10.1002/14651858.CD013705.pub2 (2021).

11. Pisanic, N. *et al.* COVID-19 Serology at Population Scale: SARS-CoV-2-Specific Antibody Responses in Saliva. *J Clin Microbiol* **59**, doi:10.1128/JCM.02204-20 (2020).

12. Pettifor, A. *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* **17**, e0267353, doi:10.1371/journal.pone.0267353 (2022).

13. Valentine, J. *et al.* Impact of periodontal intervention on local inflammation, periodontitis, and HIV outcomes. *Oral Dis* **22 Suppl 1**, 87-97, doi:10.1111/odi.12419 (2016).

14. Corman, V. M. *et al.* Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* **25**, doi:10.2807/1560-7917.ES.2020.25.3.2000045 (2020).

15. CDC. <https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html> (

16. Liu, L. *et al.* Development and performance of a point-of-care rapid antigen test for detection of SARS-COV-2 variants. *J Clin Virol Plus* **2**, 100080, doi:10.1016/j.jcvp.2022.100080 (2022).

17. Naranbhai, V. *et al.* High Seroprevalence of Anti-SARS-CoV-2 Antibodies in Chelsea, Massachusetts. *J Infect Dis* **222**, 1955-1959, doi:10.1093/infdis/jiaa579 (2020).

18. Wang, Y., Geer, L. Y., Chappey, C., Kans, J. A. & Bryant, S. H. Cn3D: sequence and structure views for Entrez. *Trends Biochem Sci* **25**, 300-302, doi:10.1016/s0968-0004(00)01561-9 (2000).

19. Madej, T., Marchler-Bauer, A., Lanczycki, C., Zhang, D. & Bryant, S. H. Biological Assembly Comparison with VAST. *Methods Mol Biol* **2112**, 175-186, doi:10.1007/978-1-0716-0270-6_13 (2020).

20. Lau, W. W., Hardt, M., Zhang, Y. H., Freire, M. & Ruhl, S. The Human Salivary Proteome Wiki: A Community-Driven Research Platform. *J Dent Res*, 220345211014432, doi:10.1177/00220345211014432 (2021).

21. Li, L. *et al.* Analysis of viral load in different specimen types and serum antibody levels of COVID-19 patients. *J Transl Med* **19**, 30, doi:10.1186/s12967-020-02693-2 (2021).

22. Chen, Z. *et al.* Profiling of SARS-CoV-2 Subgenomic RNAs in Clinical Specimens. *Microbiol Spectr* **10**, e0018222, doi:10.1128/spectrum.00182-22 (2022).

23. Long, S. Correction: Long, S. SARS-CoV-2 Subgenomic RNAs: Characterization, Utility, and Perspectives. Viruses 2020, 13, 1923. *Viruses* **14**, doi:10.3390/v14071406 (2022).

24. Silvas, J. A. *et al.* Contribution of SARS-CoV-2 Accessory Proteins to Viral Pathogenicity in K18 Human ACE2 Transgenic Mice. *J Virol* **95**, e0040221, doi:10.1128/JVI.00402-21 (2021).

25. Lau, W. W., Hardt, M., Zhang, Y. H., Freire, M. & Ruhl, S. The Human Salivary Proteome Wiki: A Community-Driven Research Platform. *J Dent Res* **100**, 1510-1519, doi:10.1177/00220345211014432 (2021).

26. Huang, C. R., Burns, K. H. & Boeke, J. D. Active transposition in genomes. *Annu Rev Genet* **46**, 651-675, doi:10.1146/annurev-genet-110711-155616 (2012).

27. Horie, M. *et al.* Endogenous non-retroviral RNA virus elements in mammalian genomes. *Nature* **463**, 84-87, doi:10.1038/nature08695 (2010).

28. Zhang, L. *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* **118**, doi:10.1073/pnas.2105968118 (2021).

29. Isho, B. *et al.* Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. *Sci Immunol* **5**, doi:10.1126/sciimmunol.abe5511 (2020).

30. Ketas, T. J. *et al.* Antibody Responses to SARS-CoV-2 mRNA Vaccines Are Detectable in Saliva. *Pathog Immun* **6**, 116-134, doi:10.20411/pai.v6i1.441 (2021).

31. Silva, J. *et al.* Saliva viral load is a dynamic unifying correlate of COVID-19 severity and mortality. *medRxiv*, doi:10.1101/2021.01.04.21249236 (2021).

32. Meinberger, D. *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* **11**, 8920, doi:10.1038/s41598-021-88356-8 (2021).

33. Lakbar, I., Luque-Paz, D., Mege, J. L., Einav, S. & Leone, M. COVID-19 gender susceptibility and outcomes: A systematic review. *PLoS One* **15**, e0241827, doi:10.1371/journal.pone.0241827 (2020).

34. Fink, A. L. & Klein, S. L. The evolution of greater humoral immunity in females than males: implications for vaccine efficacy. *Curr Opin Physiol* **6**, 16-20, doi:10.1016/j.cophys.2018.03.010 (2018).

35. Markle, J. G. & Fish, E. N. SeXX matters in immunity. *Trends Immunol* **35**, 97-104, doi:10.1016/j.it.2013.10.006 (2014).

Figures

Figure 1

A NOK cell lysate

B NOK cell lysate

C

D Archived Pre Covid-19 human saliva

E

F human IgG/IgM anti-Sars-CoV-2 Spike RBD (µl)

G
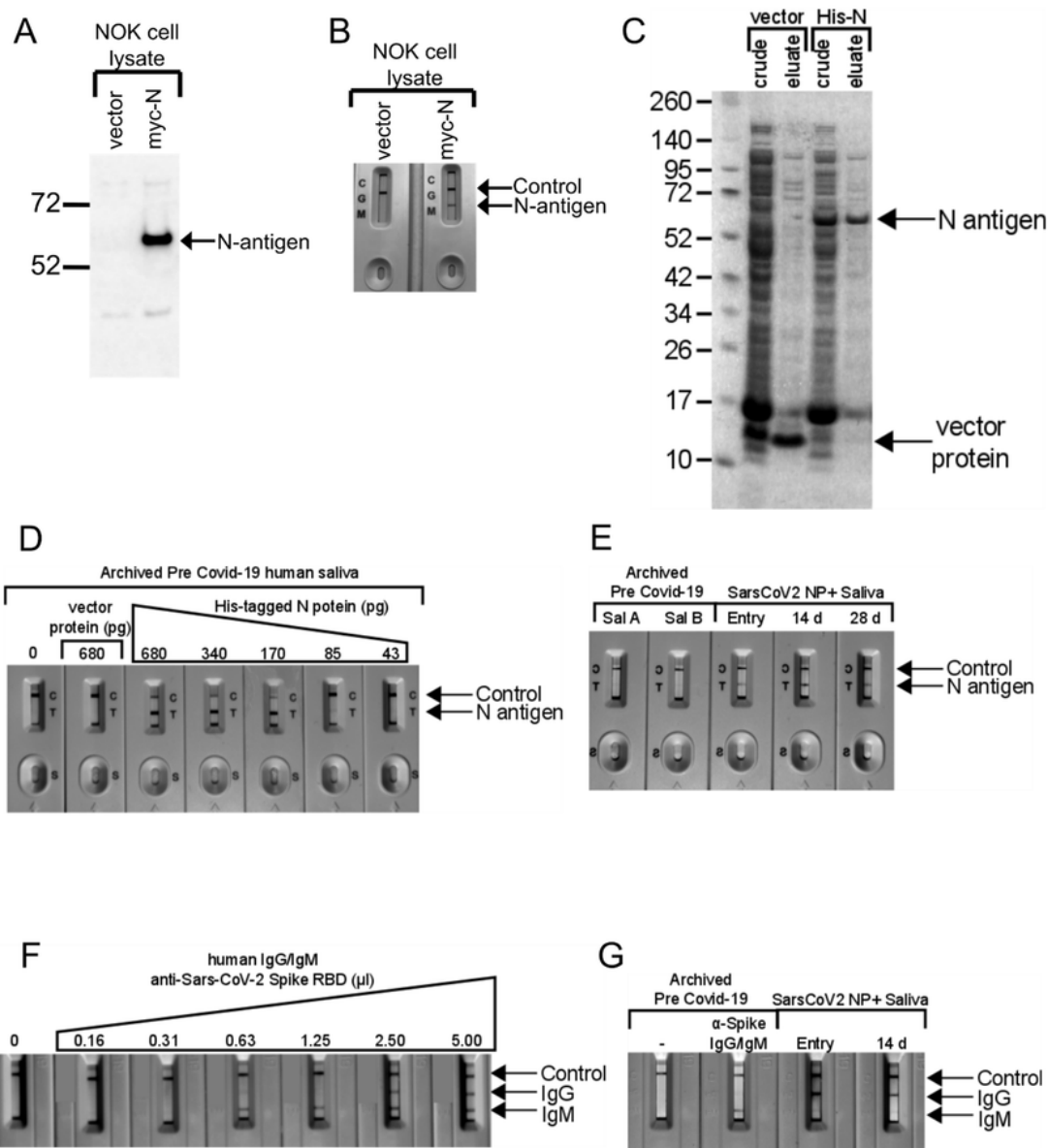
Figure 1

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

A

Study Enrollment

CoVID negative (n=15)
- SARS CoV2 anti S IgG seronegative
- NP SarsCov2 qPCR negative

Asymptomatic (n=15)
- SARS CoV2 anti S IgG seropositive
- NP qPCR positive (subset)

Symptomatic (n=17)
- NP qPCR SarsCov2 positive

Study Enrollment

Survey (demographics, symptoms)
Biospecimen collection- Saliva, Throat wash, Blood
Processing/repository UNC Delta core

Assessments:
- Oral antigen/qRTPCR detection
- Oral Antibody response
- Systemic Antibody response

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%)

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

A

Experiment	Forward primer	sequence	Reverse primer	sequence
RT-qPCR	MN908947orf1F 16364-16386	TGTCTGTTAATCCGTATGTTTGC	MN908947orf1R comp. 16411-16437	ACCTCCTAAGTAAAGTTGAGTCACATC
RT-qPCR	E_Sarbeco_F 26269-26294	ACAGGTACGTTAATAGTTAATAGCGT	E_Sarbeco_R compl. 26360-26381	ATATTGCAGCAGTACGCACACA
RT-qPCR	2019-nCoV_N1-F 28283-28302	GACCCCAAAATCAGCGAAAT	2019-nCoV_N1-R 28331-28354	TCTGGTTACTGCCAGTTGAATCTG
SsgRNA RT-qPCR	CoV25UTF 1-25	ATTAAAGGTTTATACCTTCCCAGGT	SRTqPCRR comp. 21622-21643	TGCAGGGGGTAATTGAGTTCTG
orf3asgRNA RT-qPCR	CoV25UTF 1-25	ATTAAAGGTTTATACCTTCCCAGGT	orf3aRTqPCRR Comp. 25474-25493	GTTGCAGTAGCGCGAACAAA
NsgRNA RT-qPCR	CoV25UTF 1-25	ATTAAAGGTTTATACCTTCCCAGGT	2019-nCoV_N1-R comp. 28331-28354	TGCAGGGGGTAATTGAGTTCTG
S gene expression	SexpF 21563-21587	ggggccatggaggccATGTTTGTTTTTCTTGTTTTATTGCC	SexpR comp. 25363-25384	ccggtcgacTTATGTGTAATGTAATTTGACT
E gene expression	EexpF 26245-26265	ggggccatggaggccATGTACTCATTCGTTTCGGAA	EexpR comp. 26449-26472	ccggtcgacTTAGACCAGAAGATCAGGAACTCT
M gene expression	MexpF 26523-26542	ggggccatggaggccATGGCAGATTCCAACGGTAC	MexpR comp. 27167-27191	ccggtcgacTTACTGTACAAGCAAAGCAATATTG
N gene expression	NexpF 28274-28297	ggggccatggaggccATGTCTGATAATGGACCCCAAAAT	NexpR comp. 29512-29533	ccggtcgacTTAGGCCTGAGTTGAGTCAGCA

B

Saliva RTqPCR standard curve — Throat wash RTqPCR standard curve

C — D — E
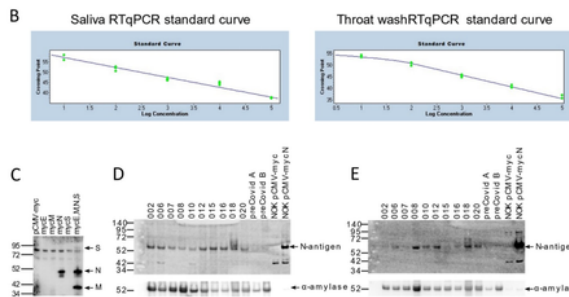
(Immunoblots)

Figure 3

F — Symptomatic

Participant	RT-qPCR NP	Copy# N RNA S	Copy# N RNA TW	ViralRNA saliva E	ViralRNA saliva ORF1	sgRNA S	sgRNA 3a	sgRNA N	N antigen LFA S	N antigen LFA TW	N Antigen immunoblot S	N Antigen immunoblot TW	Anti-S IgG LFA S	Anti-S IgG LFA TW	Anti-S IgM LFA S	Anti-S IgM LFA 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	-	-	-	-	-	-	+	+	+	-	++	-	+	-
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

Participant	NP RT-PCR	Blood Anti S serostatus	Saliva N antigen	Saliva Anti S IgG	Saliva 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

Participant	NP RT PCR	Blood Anti S serostatus	Saliva N antigen	Saliva Anti S IgG	Saliva 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 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

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)
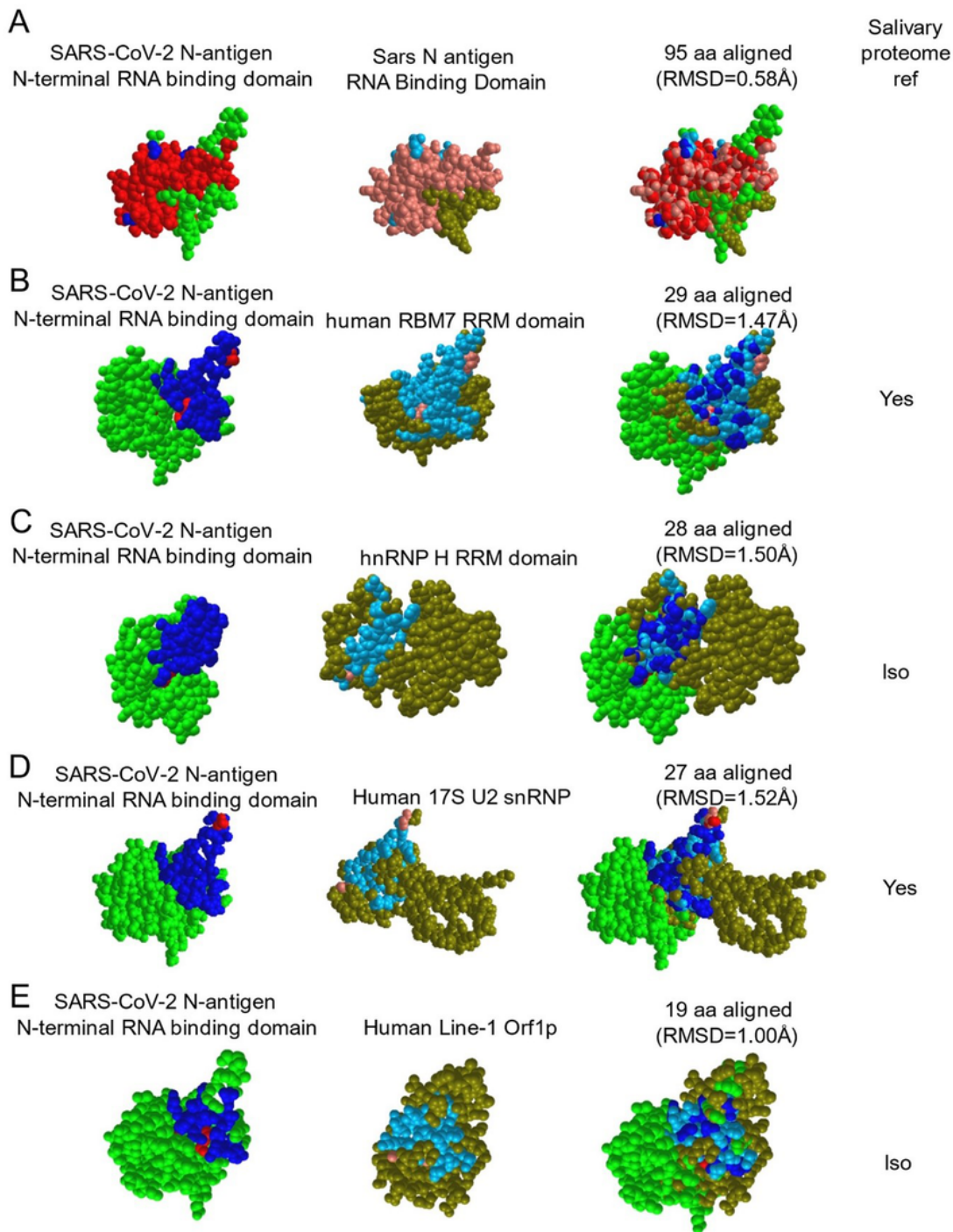
Iso

Figure 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.
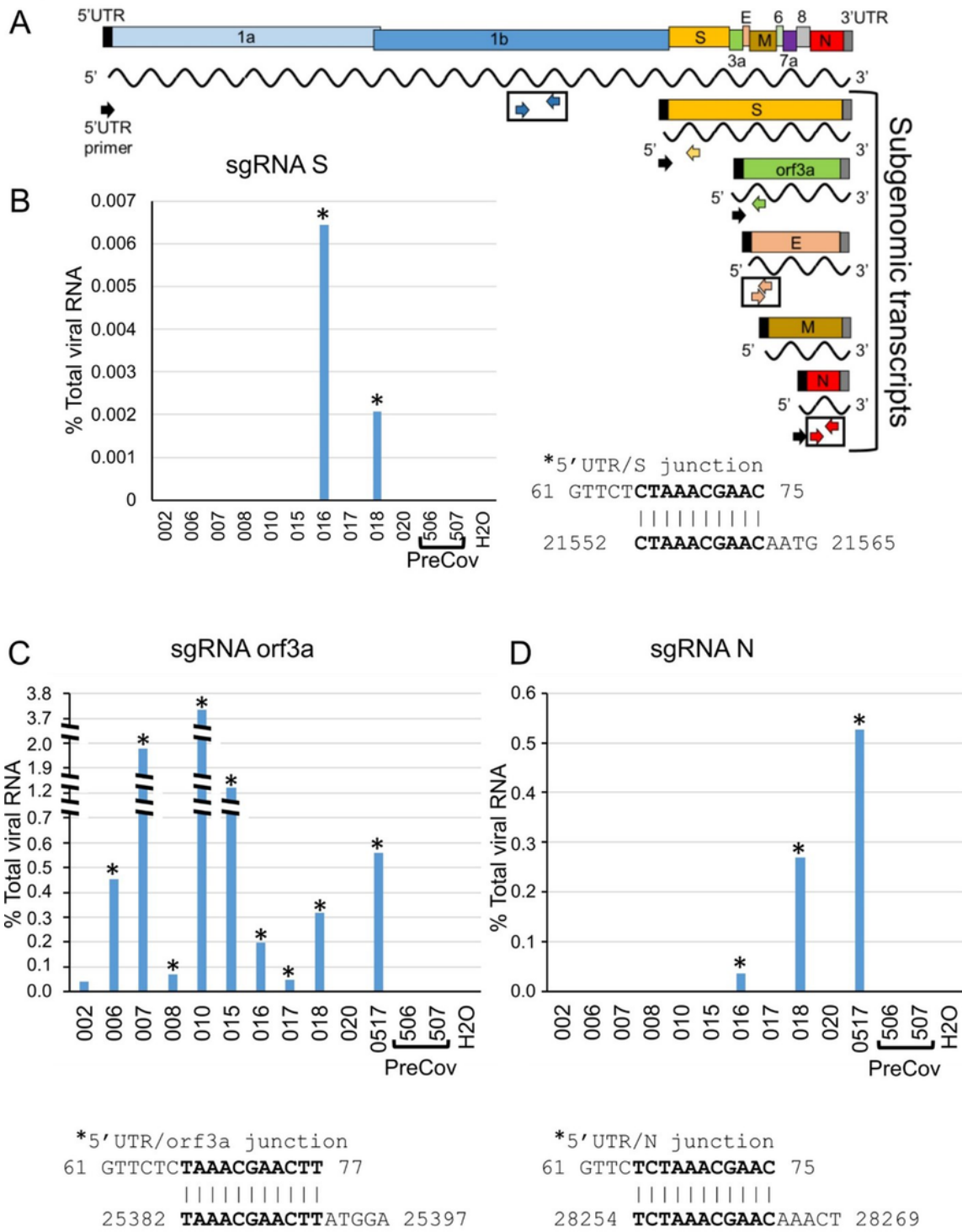
Figure 5

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: $2\text{^-(sgRNA Ct/N Ct)} \times 100$. 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

A

N-antigen | IgG | IgM

(boxplots: Pre, BL, 14D, 28D)

N-antigen: ns, ns; N=12, N=15, N=7, N=5
IgG: p=0.005, ns; N=5, N=15, N=7, N=5
IgM: ns, ns; N=5, N=15, N=7, N=5

B

Relative TW N-Antigen band intensity — Entry, 28D, Archived
** p=0.003
* p=0.9

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

Relative TW N antigen band intensity: p=0.662; Female N=6, Male N=4
Relative TW IgG band intensity: p=0.786; Female N=3, Male N=4
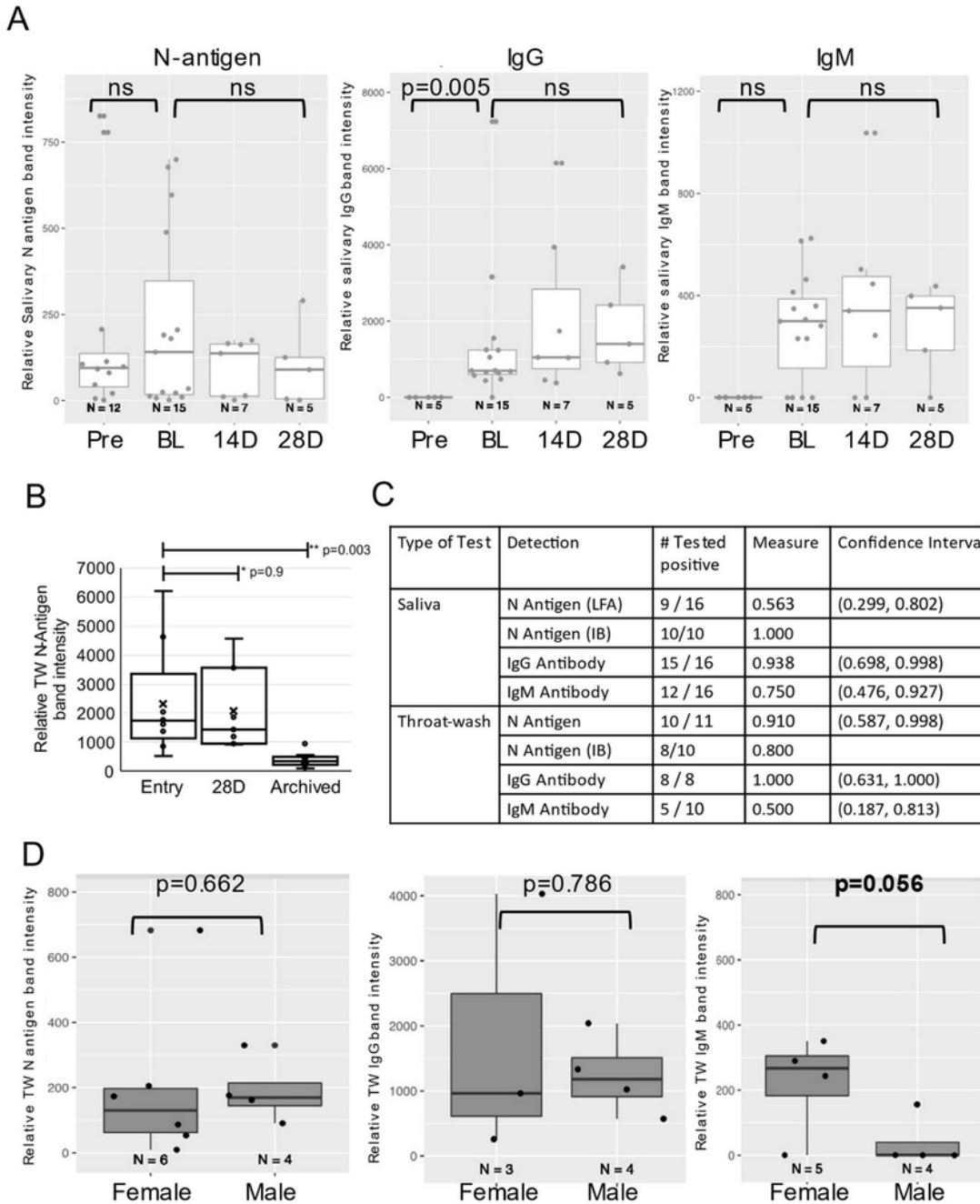Relative TW IgM band intensity: **p=0.056**; Female N=5, Male N=4

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.
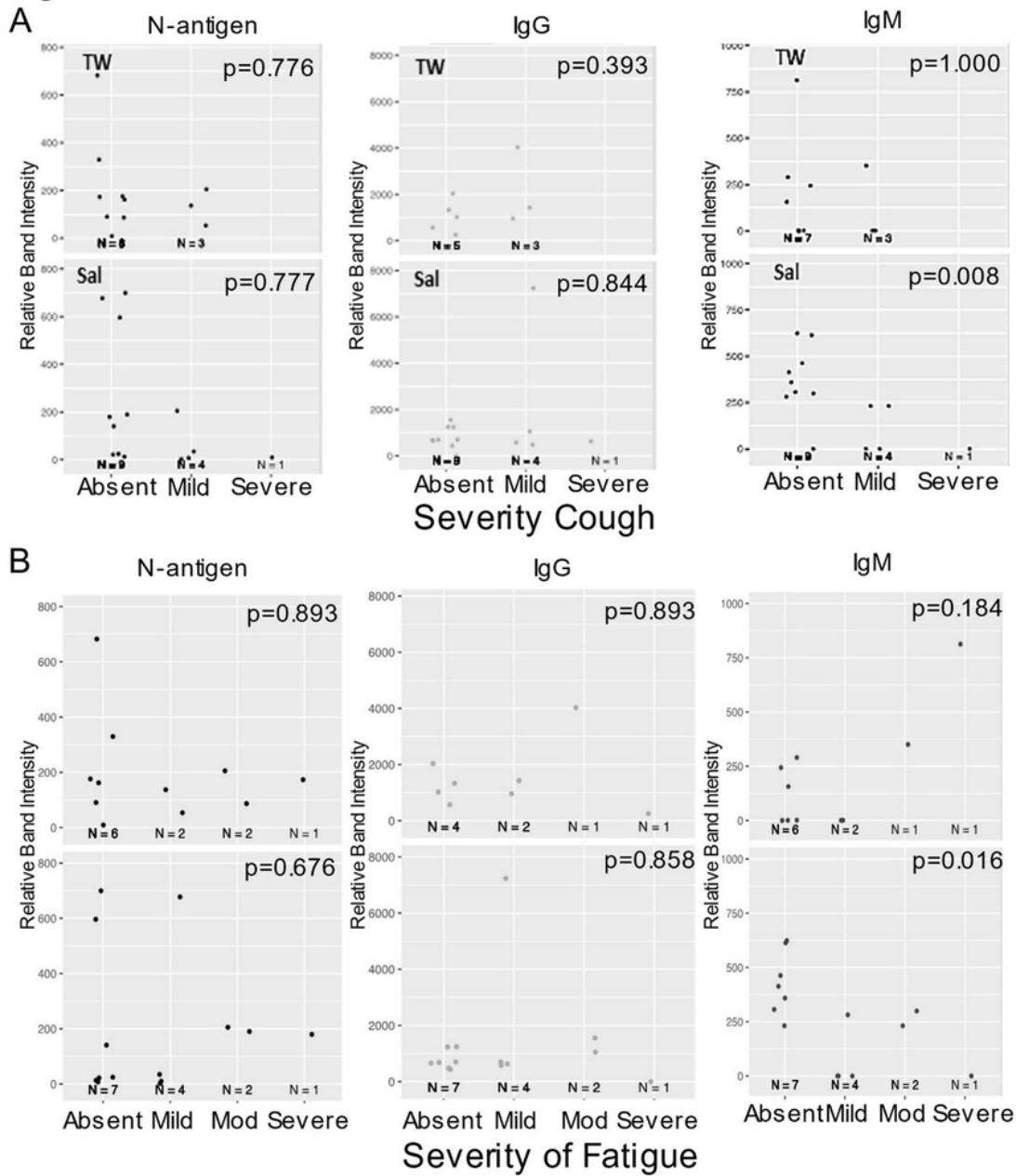

Figure 7

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.
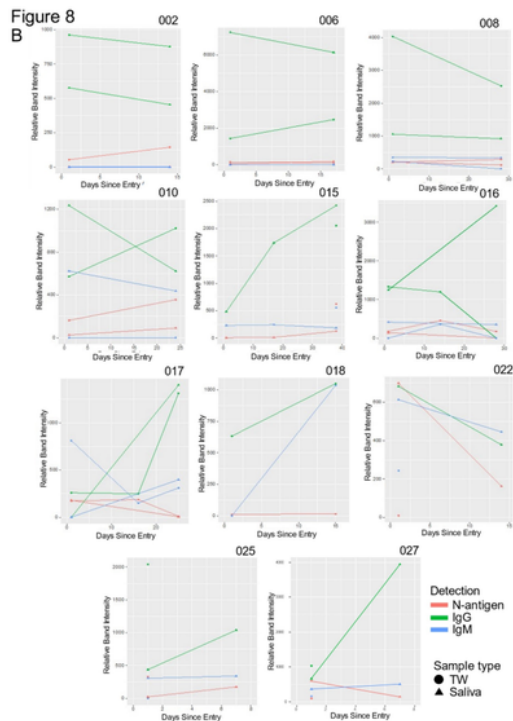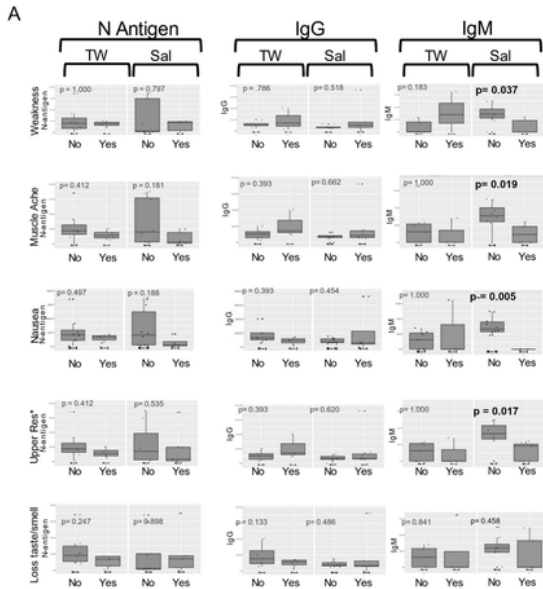

Figure 8

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).