

## Supporting Information for Publication

### Specific and Rapid SARS-CoV-2 Identification Based on LC-MS/MS analysis

Ofir Schuster<sup>1\*#</sup>, Anat Zvi<sup>2</sup>, Osnat Rosen<sup>3</sup>, Hagit Achdout<sup>1</sup>, Amir Ben-Shmuel<sup>1</sup>, Ohad Shifman<sup>2</sup>, Shmuel Yitzhaki<sup>4</sup>, Orly Laskar<sup>1</sup> and Liron Feldberg<sup>5\*#</sup>

<sup>1</sup> Department of Infectious Diseases, Israel Institute for Biological Research, Ness Ziona, Israel

<sup>2</sup> Department of Biochemistry and Molecular Genetics, Israel Institute for Biological Research, Ness Ziona, Israel

<sup>3</sup> Department of Biotechnology, Israel Institute for Biological Research, Ness Ziona, Israel

<sup>4</sup> Israel Institute for Biological Research, Ness Ziona, Israel

<sup>5</sup> Department of Analytical Chemistry Israel Institute for Biological Research, Ness Ziona, Israel

\* Corresponding author

# Equal contribution

#### **Figure S1: S and N proteins sequences and their predicted tryptic peptides**

##### **S protein**

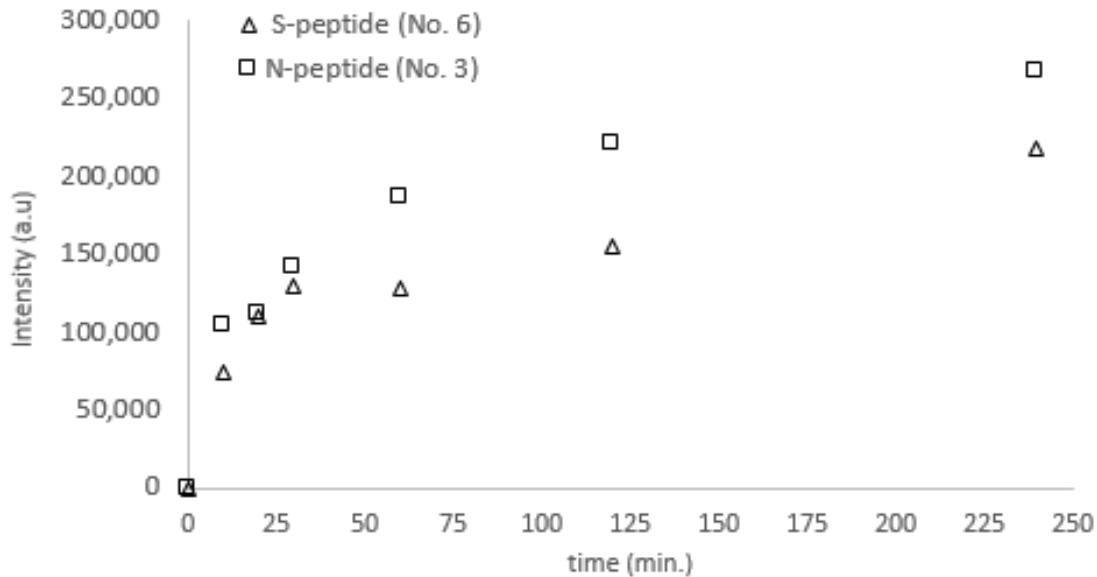
MFVFLVLLPLVSSQCVNLTR**TQLPPAYTNSFTR****RGVYYPDK**VFRSSVLHSTQDLFLPFFSN  
VTWFHAIHVSNGTKRFDNPVLPFNDGVYFASTEKSNIIR**GWIFGTTLDSK**TQSLLI VNN  
ATNVVIKVCFEQFCNDPFLGVYYHKNNK**SWMESEFR**VYSSANNCTFEYVSQPFLMDLEGKQ  
GNFKNLREFVFK**NIDGYFK**IYSK**HTPINLVR**DLPQGFSALEPLVDLPIGINITR**FQTLAL**  
**HR**SYLTPGDSSSGWTAGAAAYVGYLQPRFTLLKYNENGTITDAVDCALDPLSETKCTLK**S**  
**FTVEK****GIYQTSNFR****VQPTESIVR**FPNITNLCPFGEFNATR**FASVYAWNR**KRISNCVADYSV  
LYNSASFSTFK**CYGVSPK**LNDLCFTNVYADSFVIRGDEV**QIAPGQTGK****IADYNYK**LPDD  
FTGCVIAWNSNNLDSK**VGGNYNYLYR**LFRK**SNLKPFR**DI STEIYQAGSTPCNGVEGFNCF  
PLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKK**STNLVK**NKCVNFNFNGLTGTGV  
LTESNKK**FLPFQQFGR****DIADTTDAVR**DPQTL EILDITPCSFGGVS VITPGTNTSNQVAVLY  
QDVNCTEVPVAIHADQLTPTWR**VYSTGSNVFQTR**AGCLIGAEHVNNSYECDIPIGAGICAS  
YQTQTN SPRRARSVASQSIIAYTMSLGAENSVAYSNSIAIPTNFTISVTTEILPVSMTKT  
SVDCTMYICGDSTEC SNLLLQYGSFCTQLNR**ALTGIAVEQDK****NTQEVFAQVK**QIYKTPPIK  
DFGGFNFSQILPDP SKPSKR**SFIEDLLFNK****VTLADAGFIK****QYGDCLGDIAAR****DLICAQK**FN  
GLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY  
ENQK**LIANQFNSAIGKI****QDSLSTASALGK**LQDVVNQNAQALNTLVKQLSSNFGAISSVLN  
DILSRDK**VEAEVQIDR**LITGR**LQSLQTYVTQQLIR**AAEIR**ASANLAATK****MSECVLGQSK**R  
**VDFCGK**GYHLMSFPQSAPHGVVFLHVTVYVPAQEK**NFTTAPAICHGK**AHFPRREGV FVSNGT  
HWFVTQRNFYEPQIITDNTFVSGNCDVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS  
PDVDLGDISGINASVVNIQKEIDR**LNEVAK****NLNESLIDLQELGK**YEQYIKWPWYIWLGFIA  
GLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPV LKGVKLHYT

## N protein

**MSDNGPQNQR**NAPRITFGGSPDSTGSNQNNGERSGARSKQRRPQGLPNNTASWFTALTQHGK  
EDLKFFPRGQGVPIINTNSSPDDQIGYRRATRIRIRGGDGKMKDLSPRWYFYLLGTGPEAGLP  
YGANKDGI I WVATEGALNTPK **DHIGTR**NPANNAIIVLQLPQGTTLPK **GFYAEGSRGGSQAS**  
**SR**SSSRSRNSSR **NSTPGSSR****GTSPAR**MAGNGGDAALALLLLDR **LNQLESK**MSGK **GQQQQGQ**  
**TVTK****KSAAEASK**KPRQKRTATK **AYNVTQAFGR**RGPEQTQGNFGDQELIR **QGTDYK**HWPQIA  
QFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDKDPNFK **DQVILLNK****HIDAYKTFPPT**  
**EPK**KDKKKK **ADETQALPQR**QKKQQVTLLPAADLDDFSK **QLQSMSSADSTQA**

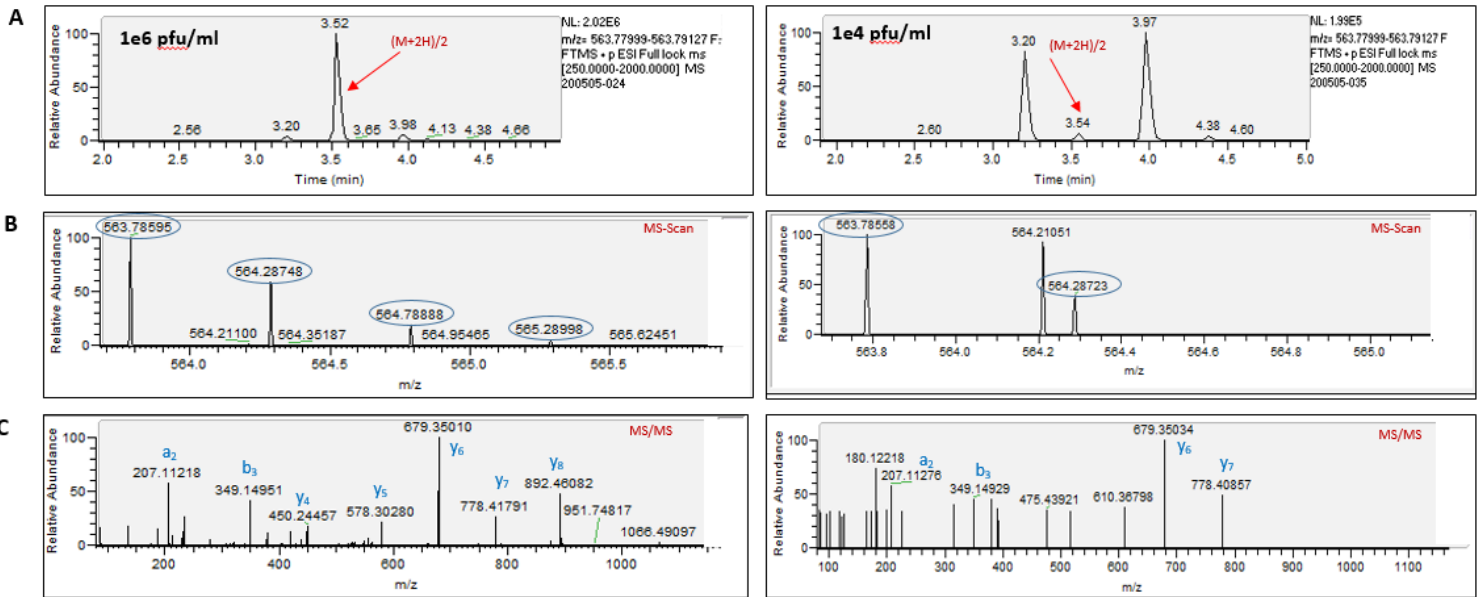
**Figure S1:** The predicted peptides following tryptic digest presented on top of the parental spike protein (A) and nucleocapsid protein (B). All peptide sequences are shown in bold and are underlined; highlighted in yellow are peptides specific to SARS-CoV-2 or SARS. Cleavage sites are highlighted in blue. Proteins sequence origins: GenPept QHD43416 and GenBank accession YP\_009724397.2, for S and N proteins, respectively (see materials and methods, section 3.2).

## Figure S2: SARS-CoV-2 kinetic profile



**Figure S2:** Kinetic profiles for tryptic digest of SARS-CoV-2 of a representative marker for S protein and N protein (peptides number 6 and 3, respectively). In each profile, each point represents the intensity at a specific time point.

**Figure S3:** LC-MS/MS chromatogram and spectrum of a representative marker



**Figure S3:** Identification of SARS-COV-2 representative marker, derived from N protein (AYNVTQAFGR). For SARS-CoV-2 identification will be required a specific marker with proper chromatographic retention time, accurate masses (< 2ppm) and at list two MS/MS fragments. (A) Extracted ion chromatogram (EIC) of m/z 563.7856 from full scan LC-MS run of SARS-CoV-2 after precipitation (1e6 PFU/ml-left and 1e4 PFU/ml-right). (B) Mass spectrum of the expected chromatographic peak (mass accuracy< 2ppm. 1e6 PFU/ml-left and 1e4 PFU/ml-right). (C) MS/MS spectrum of m/z 563.7856 (7 identified fragments in 1e6 PFU/ml- left, 4 identified fragments in 1e4 PFU/ml- right).

**Table S1:** LC-MS analysis of N protein markers

Peptide No.	Tryptic peptide	Molecular ion (Multiple charge)	Retention time (min)	Purified N protein			Cell-cultured SARS-CoV-2			SARS-CoV-2 precipitation		
				Signal Intensity	Precision (%)	Linearity	Signal Intensity	Precision (%)	Linearity (PFU/ml)	Signal Intensity	Precision (%)	Linearity (PFU/ml)
1	GQQQQGQIVTK	601.8098 (M+2H)/2	2.3	4.8e6	11	10ppb-10ppm	3.1e4	5	10 <sup>5</sup> -10 <sup>6</sup>	4.0e5	8	10 <sup>4</sup> -10 <sup>6</sup>
2	ADETQALPQR	564.78583 (M+2H)/2	2.8	4.7e6	13	10ppb-10ppm	4.0e4	2	10 <sup>5</sup> -10 <sup>6</sup>	3.8e5	11	10 <sup>4</sup> -10 <sup>6</sup>
3	AYNVTQAFGR	563.78563 (M+2H)/2	3.5	2e7	28	10ppb-10ppm	2.1e5	9	10 <sup>4</sup> -10 <sup>6</sup>	1.9e6	12	10 <sup>3</sup> -10 <sup>6</sup>
4	DQVILLNK	471.7845 (M+2H)/2	3.5	8e6	30	10ppb-10ppm	9.7e4	20	10 <sup>4</sup> -10 <sup>6</sup>	6.7e5	8	10 <sup>4</sup> -10 <sup>6</sup>
5	TFPPTPEPK	458.7423 (M+2H)/2	3.3	1.5e7	22	10ppb-10ppm	7.8e4	3	10 <sup>5</sup> -10 <sup>6</sup>	7.8e5	13	10 <sup>4</sup> -10 <sup>6</sup>
6	GFYAEGSR	443.7063 (M+2H)/2	2.9	1.4e7	22	10ppb-10ppm	1.4e5	8	10 <sup>4</sup> -10 <sup>6</sup>	6.8e5	30	10 <sup>4</sup> -10 <sup>6</sup>
7	NSTPGSSR	403.1935 (M+2H)/2	1.9	7e6	15	10ppb-10ppm	2.2e4	3	10 <sup>5</sup> -10 <sup>6</sup>	1.9e5	4	10 <sup>4</sup> -10 <sup>6</sup>
8	GGSQASSR	375.1804 (M+2H)/2	1.2	2.6e5	4	—	2.5e3	30	—	5.8e4	14	10 <sup>5</sup> -10 <sup>6</sup>

**Table S1:** LC-MS analysis of tryptic peptides derived from the Nucleocapsid (N) protein. The peptide analytical performance was evaluated according to ESI response, precision (triplicate) and linearity observed from tryptic digestion of pure proteins and cell cultured SARS-CoV-2 with and without prior precipitation.

**Table S2:** LC-MS analysis of S protein markers

Peptide No.	Tryptic peptide	Molecular ion (Multiple charge)	Retention time (min)	Purified S protein			Cell-cultured SARS-CoV-2			SARS-CoV-2 precipitation		
				Signal Intensity	Precision (%)	Linearity	Signal Intensity	Precision (%)	Linearity (PFU/ml)	Signal Intensity	Precision (%)	Linearity (PFU/ml)
1	LQSLQTYVTQQLIR	564.32107 (M+3H)/3	4.3	1.2e6	12	10ppb-10ppm	9.8e4	3	10 <sup>3</sup> -10 <sup>6</sup>	2.1e5	3	10 <sup>4</sup> -10 <sup>6</sup>
2	TQLPPAYTNSFTR	748.3806 (M+2H)/2	3.6	3.2e5	67	1ppm-10ppm	3.8e4	14	10 <sup>3</sup> -10 <sup>6</sup>	1.2e4	6	—
3	IQDLSSTASALGK	689.3646 (M+2H)/2	3.4	2e6	18	10ppb-10ppm	4.9e4	11	10 <sup>4</sup> -10 <sup>6</sup>	1.1e5	7	10 <sup>4</sup> -10 <sup>6</sup>
4	VYSTGSNVFQTR	679.8386 (M+2H)/2	3.3	8e5	5	100ppb-10ppm	4.2e4	19	10 <sup>3</sup> -10 <sup>6</sup>	1.9e5	11	10 <sup>3</sup> -10 <sup>6</sup>
5	LIANQFNSAIGK	638.3564 (M+2H)/2	3.6	2e6	60	100ppb-10ppm	8e4	5	10 <sup>3</sup> -10 <sup>6</sup>	1.7e5	17	10 <sup>3</sup> -10 <sup>6</sup>
6	SFIEDLLFNK	613.32680 (M+2H)/2	4.7	2.5e6	15	10ppb-10ppm	2.8e5	5	10 <sup>4</sup> -10 <sup>6</sup>	4.5e5	30	10 <sup>4</sup> -10 <sup>6</sup>
7	GWIFGTTLDSK	612.8166 (M+2H)/2	4.3	4.3e5	17	100ppb-10ppm	1.2e5	14	10 <sup>4</sup> -10 <sup>6</sup>	2e5	21	10 <sup>4</sup> -10 <sup>6</sup>
8	VGGNYNYLYR	609.7987 (M+2H)/2	3.5	6.6e5	11	100ppb-10ppm	4.9e4	6	10 <sup>3</sup> -10 <sup>6</sup>	1.6e5	23	10 <sup>3</sup> -10 <sup>6</sup>
9	NTQEVFAQVK	582.3064 (M+2H)/2	3.3	1.9e5	34	100ppb-10ppm	6.4e3	30	—	5.2e4	3	10 <sup>3</sup> -10 <sup>6</sup>
10	ALTGIAVEQDK	572.8140 (M+2H)/2	3.3	3.0e5	53	100ppb-10ppm	5.0e3	46	—	5.8e4	2	10 <sup>3</sup> -10 <sup>6</sup>
11	FLPFQQFGR	570.3035 (M+2H)/2	4.2	9.9e5	28	100ppb-10ppm	1.7e5	20	10 <sup>4</sup> -10 <sup>6</sup>	6.3e5	30	10 <sup>4</sup> -10 <sup>6</sup>
12	FASVYAWNR	557.2774 (M+2H)/2	3.7	8e5	8	100ppb-10ppm	4.1e4	30	10 <sup>3</sup> -10 <sup>6</sup>	1.1e5	17	10 <sup>4</sup> -10 <sup>6</sup>
13	FQTLALHR	366.8854 (M+3H)/3	3.6	1.7e6	7	100ppb-10ppm	1.6e5	24	10 <sup>4</sup> -10 <sup>6</sup>	3.9e5	30	10 <sup>4</sup> -10 <sup>6</sup>
14	GIYQTSNFR	543.2724 (M+2H)/2	3.2	3.6e5	30	100ppb-10ppm	3.7e4	29	10 <sup>3</sup> -10 <sup>6</sup>	1.0e5	23	10 <sup>3</sup> -10 <sup>6</sup>
15	MSECVLGQSK	541.2544 (M+2H)/2	3.1	1.3e5	68	—	—	—	—	—	—	—
16	DIADTTDAVR	538.7646 (M+2H)/2	3	2.9e4	77	1ppm-10ppm	7.9e3	32	10 <sup>3</sup> -10 <sup>6</sup>	3.5e4	15	10 <sup>3</sup> -10 <sup>6</sup>

Peptide No.	Tryptic peptide	Molecular ion (Multiple charge)	Retention time (min)	Purified S protein			Cell-cultured SARS-CoV-2			SARS-CoV-2 precipitation		
				Signal Intensity	Precision (%)	Linearity	Signal Intensity	Precision (%)	Linearity (PFU/ml)	Signal Intensity	Precision (%)	Linearity (PFU/ml)
17	SWMESEFR	536.2318 (M+2H)/2	3.7	3.6e5	14	100ppb-10ppm	4.6e4	1	10 <sup>5</sup> -10 <sup>6</sup>	8.9e4	15	10 <sup>5</sup> -10 <sup>6</sup>
18	VTLADAGFIK	517.7977 (M+2H)/2	3.9	2.2e6	18	10ppb-10ppm	2.2e5	13	10 <sup>4</sup> -10 <sup>6</sup>	1.4e5	29	10 <sup>4</sup> -10 <sup>6</sup>
19	VQPTESIVR	514.7904 (M+2H)/2	3.1	9.7e5	50	100ppb-10ppm	9.6e4	1	10 <sup>4</sup> -10 <sup>6</sup>	2e5	7	10 <sup>4</sup> -10 <sup>6</sup>
20	SNLKPFER	330.8504 (M+3H)/3	2.8	4.7e5	13	100ppb-10ppm	5.3e4	8	10 <sup>5</sup> -10 <sup>6</sup>	6.5e4	13	10 <sup>5</sup> -10 <sup>6</sup>
21	HTPINLVR	475.2825 (M+2H)/2	3	1.5e6	10	100ppb-10ppm	2.2e5	13	10 <sup>4</sup> -10 <sup>6</sup>	3.5e5	16	10 <sup>4</sup> -10 <sup>6</sup>
22	QIAPGQTGK	450.2509 (M+2H)/2	2.5	8.1e5	28	100ppb-10ppm	5.1e4	18	10 <sup>5</sup> -10 <sup>6</sup>	2e5	7	10 <sup>4</sup> -10 <sup>6</sup>
23	IADYNYK	443.71889 (M+2H)/2	2.9	3.6e5	16	100ppb-10ppm	4.8e4	1	10 <sup>5</sup> -10 <sup>6</sup>	8.3e4	80	10 <sup>5</sup> -10 <sup>6</sup>
24	ASANLAATK	423.7376 (M+2H)/2	2.5	2.9e6	17	10ppb-10ppm	5e4	6	10 <sup>5</sup> -10 <sup>6</sup>	1.8e5	25	10 <sup>4</sup> -10 <sup>6</sup>
25	DLICAQK	395.7100 (M+2H)/2	3.1	6.6e4	64	1ppm-10ppm	—			—		

**Table S2:** LC-MS analysis of tryptic peptides derived from the Spike (S) protein. The peptide analytical performance was evaluated according to ESI response, precision (triplicate) and linearity observed from tryptic digestion of pure proteins and cell cultured SARS-CoV-2 with and without prior precipitation.

**Table S3:** Universality of S3 and S4 markers

<b>peptide S3</b>		<b>peptide S4</b>	
<b>peptide sequence</b>	<b>#occurrences (in 56,940)</b>	<b>peptide sequence</b>	<b>#occurrences (in 56,940)</b>
FKTLLALHR	11	HTPIBLVR	1
FQILLALHR	8	HTPINL	1
FQPLLALHR	1	HTPINLAR	9
FQTLFALHR	9	HTPINLLR	27
FQTLA	2	HTPINLV	50
FQTLLAFHR	1	<b>HTPINLVR</b>	<b>55519</b>
FQTLLALH	9	HTPINLXR	3
FQTLLALHK	3	HTPISLVR	1
<b>FQTLLALHR</b>	<b>55519</b>	HTPITLVR	1
FQTLLALRR	1	HTPIXLVR	2
FQTLLALXR	7	HTPIYLVR	2
FQTLLALYR	24	HTPJNLVR	1
FQTLLAXHR	4	HTPXNLVR	2
FQTLLAXXR	1	HTXINLVR	4
FQTLLPFHR	2	PINLVR	14
FQTLLSLHR	9	QTPINLVR	2
FQTLLVLHR	5	TPINL	1
FQTLLXLHR	8	TPINLVR	14
FQTLLXXHR	4	TPXNLVR	1
FQTLPALHR	5	<i>ND</i>	1285*
FQTLXALHR	1		
FQTLXXLHR	10		
FQTLXXXHR	1		
FQTXLALHR	3		
FQTXALHR	2		
FQTXXLHR	2		
FQXLLALH	1		
FQXLLALHR	5		
FRTLLALHR	3		
FXTLLALHR	2		
LLALHR	3		
QTLLALH	1		
QTLLALHR	7		
TLLALHR	11		
YQTQTNSHR	14		
<i>ND</i>	1241*		

**Table S3:** Universality of markers S3 and S4. Variant sequences of peptides S3 and S4, and the extent of their occurrences in the 56,940 genomes retrieved from GISAID, present a high level of sequence identity (>97.5%). \* Truncated sequences in which the peptide region is missing

**Table S4:** Markers retention time stability

Matrix	Concentration	AYNVTQAFGR	GFYAEGSR	SFIEDLLFNK	FLPFQQFGR	FQTLALHR	HTPINLVR
purified proteins	10ppm	3.55	2.92	4.68	4.16	3.59	3.01
	1 ppm	3.52	2.88	4.69	4.18	3.61	3.03
	0.1 ppm	3.53	2.89	4.70	4.17	3.61	3.03
SARS-CoV-2 spiked into buffer	1e6 PFU/ml	3.52	2.87	4.70	4.16	3.58	3.00
	1e5 PFU/ml	3.53	2.92	4.71	4.18	3.62	3.03
	1e4 PFU/ml	3.53	2.88	4.71	4.20	3.60	3.04
SARS-CoV-2 spiked into different NP samples	1e5 PFU/ml	3.56	2.86	4.75	4.19	3.59	3.04
	1e5 PFU/ml	3.49	2.85	4.68	4.19	3.58	3.01
	1e6 PFU/ml	3.48	2.84	4.67	4.15	3.57	3.00
stdv		0.03	0.03	0.02	0.02	0.02	0.02
mean		3.52	2.88	4.70	4.18	3.59	3.02
precision		0.7	1.0	0.5	0.4	0.5	0.5

**Table S4:** Retention time stability of markers in different matrix at different days and concentrations