Detection of SARS-CoV-2 Proteins in Wastewater Samples by Mass Spectrometry

Linda R. Lara-Jacobo, Golam Islam, Jean-Paul Desaulniers, Andrea E. Kirkwood, and Denina B. D. Simmons*

ABSTRACT: The recent COVID-19 pandemic overwhelmed the health system worldwide, and there was a need to track outbreaks and try to use this information as an early warning system. Wastewater-based epidemiology (WBE) enabled detection of the SARS-CoV-2 virus in wastewater treatment plant influents. Until now, the most used technique for this detection has been the quantitative polymerase chain reaction (qPCR)-based quantification of SARS-CoV-2 RNA. This study proposes a mass spectrometry (MS)-based method that detected specific SARS-CoV-2 proteins in wastewater, 5 and 6 days ahead of the case data for two municipalities. We identified unique peptides of eight proteins related to the SARS-CoV-2 virus and COVID-19 infection. We detected the nonstructural protein (NSP) pp1ab (transcribed after host cell infection) most frequently in all of the samples. As a result, we suspect that in the active cases of COVID-19, the pp1ab protein is present in high abundance in the urine and feces and that this protein could be used as an alternative biomarker. These data were collected before mass vaccination occurred in the population.

KEYWORDS: SARS-CoV-2, COVID-19, liquid chromatography–mass spectrometry, wastewater, proteomics

INTRODUCTION

On March 12, 2020, the World Health Organization (WHO) declared COVID-19 as a global pandemic. COVID-19 is produced by a novel β-coronavirus known as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Several studies have detected SARS-CoV-2 RNA in urine, feces, and other bioluids from both symptomatic and asymptomatic people with COVID-19, suggesting that SARS-CoV-2 RNA could be detected in human wastewater. Thus, wastewater-based epidemiology (WBE) is now used as an approach to monitor COVID-19 prevalence in many different places around the world. Polymerase chain reaction-based quantification (PCR-based quantification) is the most common SARS-CoV-2 RNA detection method utilized in WBE, but there are other methods for viral biomolecule detection that could work as well as including liquid chromatography–tandem mass spectrometry (LC–MS/MS) and immunochemical methods. The aim of this study was to evaluate the presence of SARS-CoV-2 proteins in untreated wastewater (WW) influents collected from two wastewater treatment plants (WWTPs), located in Durham Region, Ontario, Canada, using an LC–MS/MS-based proteomics approach.

SARS-CoV-2 contains a single RNA genomic strand (+) that encodes at least 29 proteins, including four structural proteins (present as part of the viral particle): spike glycoprotein (S), membrane (M), the envelope (E), and nucleocapsid (N). The remaining proteins are known as nonstructural proteins (NSP), which are produced by the host cell after viral infection. The S protein, with a size of 180–200 kDa, is a transmembrane domain protein anchored in the viral membrane, where its main function is to allow the virus to fuse with the host cell. To avoid detection by the host immune system, these spike proteins are coated with polysaccharides (hence the term glycoprotein) that allow the virus to camouflage. The M protein is the most abundant structural protein of SARS-CoV-2, and while it is thought to also be a glycoprotein, its function is not fully understood. There are three additional transmembrane proteins that are similarly not as well studied as the S protein and are named after their open reading frame location: ORF3a, ORF7a, and ORF8. The ORF3a protein regulates the subcellular environment in the host and plays an important role in the defense against infection by inducing apoptosis. Both ORF7a and ORF8 are accessory proteins whose functions are not...
determined.\textsuperscript{18,19} Finally, the E protein forms a protein–lipid ion-transporting pore and is involved in virus morphogenesis, assembly, and the induction of apoptosis.\textsuperscript{20,23} Replicase polyprotein 1ab (ppb1ab) and replicase polyprotein 1a (ppb1a) are nonstructural proteins (NSPs) involved in the transcription and replication of the virus and are produced after host cell infection by the virus. These proteins inhibit host translation and maintain optimal cellular conditions for the replication of SARS-CoV-2.\textsuperscript{2,24}

In wastewater-based studies of SARS-CoV-2, the main detection technique has been PCR-based quantification, which has been used to measure SARS-CoV-2 RNA concentrations in untreated wastewaters in Australia,\textsuperscript{11} Canada,\textsuperscript{6,7} Germany,\textsuperscript{9} Netherlands,\textsuperscript{11} Spain,\textsuperscript{10} and USA.\textsuperscript{1} In these studies, scientists detected the concentration of the nucleocapsid (N) gene (number of copies/L) (which encodes the nucleocapsid protein(s)).\textsuperscript{5,7,10,21} At the present time, there are no published reports of liquid chromatography–tandem mass spectrometry (LC–MS/MS)-based SARS-CoV-2 protein detection in wastewater, but there are a couple of published studies that use LC–MS/MS to measure SARS-CoV-2 proteins in nasopharyngeal and oropharyngeal clinical samples.\textsuperscript{24,25}

Measuring viral proteins using LC–MS/MS offers great advantages, including the considerably shorter run time,\textsuperscript{25} potentially improved sensitivity,\textsuperscript{26} the ability to multiplex and perform simultaneous nontargeted measurements, and a significantly reduced cost per sample compared to PCR-based quantification.\textsuperscript{24,26} In the present study, our goal was to determine if we could detect the presence of SARS-CoV-2 proteins in untreated wastewater using a bottom-up proteomics approach with LC–MS/MS and then to evaluate the potential use of this methodology as a routine method for the detection of SARS-CoV-2 in wastewater.

\section*{Materials and Methods}

\textbf{Protein Precipitation and Digestion.} Twenty-four-hour composite influent samples from each of the two wastewater treatment plants (WWTP) from Durham Region (Ontario, Canada) were obtained over the course of 15 weeks (from October 22, 2020, to April 4, 2021) for a total of 115 samples across all sampling times and locations. Wastewater from two municipalities flows into these two WWTPs, although the boundaries of the wastewater catchments do not perfectly correspond to a single municipality. The WWTP and municipalities were related as follows: WWTP1—Whitby (population = 135,556) and WWTP2—Oshawa (population = 170,071). In the present study, all of the samples were collected as “fresh” (unpasteurized) 24 h composite samples and were stored at 4°C for less than 48 h until processing. We added 50 mL of acetone at 4°C (CAS number 67-64-1, Fisher Scientific) to 50 mL of the wastewater and stored overnight at −20°C to precipitate proteins from the liquid phase. The samples were then centrifuged for 15 min at 3405×g. After centrifugation, we carefully removed and discarded the supernatant, and the pellet was transferred to a low-retention 1.5 mL microcentrifuge tube and then left open in the fume hood for 60 min to allow any remaining acetone to evaporate. We then resuspended the pellet in 50 μL of 2 mM TCEP (tris(2-carboxyethyl)phosphine hydrochloride) solution (CAS number 51805-45-9, Sigma-Aldrich) in 100 mM ammonium bicarbonate buffer (CAS number 15715-58-9 1.0 M solution, Sigma-Aldrich), gently vortexed, and incubated the sample at room temperature for 45 min. We then added iodoacetamide (CAS number 144-48-9, Sigma-Aldrich) solution to a final concentration of 4 mM and incubated the samples in the dark at room temperature for 45 min. Finally, we performed formic acid chemical digestion with 10% formic acid (CAS number 64-18-6, Sigma-Aldrich) and incubated at 115°C for 30 min. This method was previously described by Simmons et al.\textsuperscript{26}

\textbf{LC–MS/MS Peptide Analysis.} Peptides were identified with an Agilent 1260 Infinity Binary Pump HPLC coupled to an Agilent 6545 Quadrupole Time-Of-Flight mass spectrometer. Five microliters of each sample were injected in duplicate onto the separation column (Agilent ZORBAX 300SB-C18, Micro Bore Rapid Resolution 1.0 × 50 mm² 3.5 μm) and eluted with a 50 min gradient (Solvent A = 0.1% formic acid in 95:5 water/acetonitrile; Solvent B = 0.1% formic acid in 95:5 acetonitrile/water; 0–2 min 2% Solvent B; 2–22 min 40% Solvent B; 22–27 min 60% Solvent B; 27–32 min 85% Solvent B; 32–50 min 2% Solvent B.). The column was kept at 4°C, with a flow rate set at 0.100 μL/min, and the autosampler compartment was kept at 6°C. Auto MS/MS (data-dependent) analysis was performed on the first set of samples, using a linear collision energy ramp and top 20 most abundant precursors with active exclusion, and after peptides were identified by database search, the m/z and retention time of each was added to an inclusion list for targeted MS/MS during our subsequent data-dependent acquisition runs. Specific instrumental settings and acquisition method details are provided in the Supporting Information. We also included an analytical blank injection and a bovine serum albumin tryptic digest standard to monitor baseline carry-over and instrumental performance. Blank spectral files were searched using the same search strategy as the samples. The instrument was calibrated before each run using an eight-point calibration standard with residuals less than 0.1 ppb for each m/z measured along the curve. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD026412.

\textbf{Database Search and Analysis of Identified Proteins.} For the viral protein identification, peptide spectral data was extracted, filtered, and searched against UniProtKB/Swiss-Prot SARS-CoV-2 database (downloaded September 2020). The database contained 14 SARS-CoV-2 proteins. The search was set in Spectrum Mill MS Proteomics Workbench (Agilent G2721AA/G2733AA) using an automated workflow to search for peptide sequence matches (scores defined in more detail in the Supporting Information). Briefly, protein hits were filtered using peptide database match score >6, and %SPI (scored peak intensity) >70% (Table S1). We then manually filtered all peptides to include only peptides with a false discovery rate (FDR) <1%, which was determined from a decoy reverse database search (Table S2).

\textbf{SARS-CoV-2 Detection by RT-qPCR.} Total RNA was extracted from 40 mL of raw influent samples using a commercially available QIAGEN Powermicrobiome kit (QIAGEN, Canada). One-step RT-qPCR was performed on the RNA samples using the Reliance One-Step Multiplex RT-PCR Supermix 4X (Bio-Rad Laboratories Inc, Canada). The 20 μL reaction volume was composed of 5 μL of RNA template input of each sample, 500 nM Forward (GAC CCC AAA ATC AGC GAA AT) and Reverse (TCT GGT TAC TGC CAG TTG AAT CTA CTG) primers (CDC N1 assay),\textsuperscript{27} 125 nM of custom TaqMan probe (FAM-ACC CCG CAT TAC GTT TGG TGC ACC-MGBNFQ) (Thermo Fisher Scientific, Environ. Sci. Technol. 2022, 56, 5062–5070).
MA) targeting the N gene, 8 μL of PCR grade H2O, and 5 μL of Reliance supermix. The RT-qPCR reactions were all run in triplicates and according to MIQE guidelines using a CFX Connect qPCR thermocycler (Bio-Rad, Hercules, CA), RT was performed at 50 °C, 10 min, followed by polymerase activation at 95 °C for 10 min, and 45 cycles of denaturation, annealing/extension at 95 °C/10 s, then 60 °C/30 s, respectively. Serial dilutions of the Twist Control 2 viral RNA standard (Twist Biosciences, South San Francisco, CA) were run on every 96-well PCR plate to produce standard curves used to quantify the copies of SARS-CoV-2 genes. The RT-qPCR analysis was validated using no-template controls (NTCs) containing water as template, no reverse transcriptase controls (NRTs), and the VetMax Xeno kit (Thermo Fisher Scientific, MA) as internal inhibition control to monitor the presence of PCR inhibition. No inhibition was observed for all samples that were utilized for the RT-qPCR analysis. All samples analyzed were quantified according to the MIQE recommendations using the standard curve method with primer efficiencies for each target ranging from 95 to 110%, the $R^2 \geq 0.98$, and slope $\sim$3.3 to 3.4. The dynamic range of our linear standard curve was between $1 \times 10^5$ and $1 \times 10^1$ copies/μL. The limit of detection for the SARS-CoV-2 N gene with 95% coefficient of variation was five copies/well. Any crossing threshold values above 40 cycles were identified as negative reactions, assuming no amplification/detection occurred.

**Normalization.** The normalization of both data (protein pp1ab and RT-qPCR) was determined by the daily flow and the population served by WWTP (see the equations below).

\[
N_i = \frac{V_{S_i} F}{P}
\]

With $N_i$ normalization for protein pp1ab (mean total intensity counts (TIC) of protein per inhabitant per day), $V_{S_i}$ viral signal (number of copies/m³), $F$ flow (m³/day), and $P$ population served (inhabitant).

**Results and Discussion.** In total, we identified 245 unique peptides (962 total peptides) from seven SARS-CoV-2 proteins. Figure 1 displays the mean abundance of all peptides for each protein in influent samples over the course of the sampling period as a heatmap for each WWTP.

\[
N_i = \frac{V_{S_i} F}{P}
\]
Specifically, we detected (1) structural proteins: M (6 peptides), S (15 peptides), and N (20 peptides); (2) transmembrane proteins: ORF3a (8 peptides), ORF7a (16 peptides), and ORF8 (10 peptides); and (3) the postinfection NSP pp1ab (170 peptides) (see Table S1 for complete protein identification details). The sequence coverage for each protein was as follows: 13.33% (M), 38.02% (S), 13.33% (N), 30.26% (ORF3a), 40.50% (ORF7a), 57.02% (ORF8), and the nonstructural protein 45.67% (pp1ab) (Figure 2). It was immediately striking that we detected the peptides from pp1ab the most often (in 97.82% of our positive samples) compared to the other 6 SARS-CoV-2 proteins. Thus, we determined that the pp1ab protein was the most frequently present in the Durham Region WWs.

Considering that the structural proteins will degrade as the SARS-CoV-2 virus travels through the human gut and the wastewater system, we think that it makes sense that we did not detect structural proteins with as much frequency or intensity as the pp1ab protein. In contrast, after LC−MS/MS analyses of nasal swabs and gargle samples collected from

---

Figure 2. Sequence coverage of SARS-CoV-2 structural and nonstructural proteins found in wastewater influents. The colors in each protein represent the peptide sequence that was detected.
positive COVID-19 patients, the most abundant peptides were from structurally abundant proteins (N most often, then the E, M, and S proteins). Those structural proteins are all part of the viral envelope, which makes them the easiest to detect in a freshly collected clinical sample, but they are also the proteins that are most exposed to environmental factors and will likely degrade first in the human gut and wastewaters known to contain many chemicals with the potential to degrade SARS-CoV-2.32

What is interesting to consider is the fate of the pp1ab protein. As we currently understand SARS-CoV-2 pathology, the pp1ab protein would not be found inside the viral envelope because it is made by human cellular machinery after SARS-CoV-2 infection, and so we wonder how it is possible that this protein is the most abundant in our wastewater samples? First, we considered the possibility that we were obtaining false positives; however, we filtered the peptides match with a local FDR <0.1%, and they only matched to SARS-CoV-2 viruses after a BLAST search against the entire Uniprot database (1000 Hits, default settings). Additionally, 170 peptides we detected and matched to pp1ab covered >45% of the full protein sequence in our combined dataset from 23 weeks of wastewater samples across two different wastewater catchments. Therefore, we are confident that we are not observing a false-positive result.

We posit that there are two main factors that could affect the persistence of the pp1ab protein in the environment: (1) the abundance of the protein in the feces and urine in SARS-CoV-2-infected people within the community and (2) the size and structure of the pp1ab protein. After extensive literature search, we could not find any reports about the abundance of the pp1ab protein in urine or feces, with the majority of papers that include pp1ab in the text describing only its general function in SARS-CoV-2 pathogenesis.33−35 Thus, it is difficult to discuss if the pp1ab protein is shed in high levels through urine and feces because we did not measure this directly. However, based on our results, we think that it is highly probable that this protein is highly abundant in human excrement. Pp1ab has been found in nasopharyngeal and oropharyngeal samples from a large clinical sample cohort using LC−MS/MS,25 and as more evidence emerges, we may learn more about this protein in other body fluids.

Additionally, the pp1ab protein is very large in comparison with the structural proteins of the virus (7096 amino acid residues in length and 794 058 Da in mass),36 which could be a second factor contributing to the abundance of pp1ab in wastewater. In infected cells, pp1ab is cleaved into multiple...
functional proteins by proteolytic autoprocessing. If the protein is released from cells before complete autoprocessing, its sheer size could make the pp1ab protein more resistant to complete degradation while traveling through the human gut and municipal wastewater system. We also observed much of the pp1ab sequence coverage in the nsp 2 and nsp 3 (3C-like proteinase) regions, which may indicate that these regions (or proteins) are protected within the tertiary structure of the protein, and thus more resistant to degradation than other regions in the pp1ab polyprotein sequence.

We employed a cross-correlation analysis comparing the mean total intensity counts (the average of the intensities of all mass spectral peaks for peptides belonging to the same protein) of pp1ab in the wastewater from each treatment plant to the number of new reported cases of COVID-19 by onset date ( Figures 3A,B and 4A,B). The daily flow normalized pp1ab signal coverage in the nsp 2 and nsp 3 (3C-like proteinase) regions, which may indicate that these regions (or proteins) are protected within the tertiary structure of the protein, and thus more resistant to degradation than other regions in the pp1ab polyprotein sequence.

We employed a cross-correlation analysis comparing the mean total intensity counts (the average of the intensities of all mass spectral peaks for peptides belonging to the same protein) of pp1ab in the wastewater from each treatment plant to the number of new reported cases of COVID-19 by onset date ( Figures 3A,B and 4A,B). The daily flow normalized pp1ab signal coverage in the nsp 2 and nsp 3 (3C-like proteinase) regions, which may indicate that these regions (or proteins) are protected within the tertiary structure of the protein, and thus more resistant to degradation than other regions in the pp1ab polyprotein sequence.

We employed a cross-correlation analysis comparing the mean total intensity counts (the average of the intensities of all mass spectral peaks for peptides belonging to the same protein) of pp1ab in the wastewater from each treatment plant to the number of new reported cases of COVID-19 by onset date ( Figures 3A,B and 4A,B). The daily flow normalized pp1ab signal coverage in the nsp 2 and nsp 3 (3C-like proteinase) regions, which may indicate that these regions (or proteins) are protected within the tertiary structure of the protein, and thus more resistant to degradation than other regions in the pp1ab polyprotein sequence.
events, population usage, degradation). Of interest was the fact that the protein signals were observed 5 and 6 days ahead (with a positive lag value), while the RNA signals measured by RT-qPCR were 4 and 10 days behind the cases by onset date (with a negative lag value). This could mean that when human cells are infected, pp1ab protein production increases immediately, which is then followed by an increase in shedding of viral RNA, observed a few days later—which makes sense as the pp1ab proteins are the means by which the SARS-CoV-2 virus replicates, as part of the viral life cycle.\(^1\)

To our knowledge, we are the first to use LC–MS/MS to measure SARS-CoV-2 proteins in wastewater. Neault et al.\(^3\) detected SARS-CoV-2 structural proteins in wastewater with immunoblotting and then measured their abundance with an immune-linked PCR method. They found that proteins were detected more frequently than RNA and concluded that proteins were present in higher abundance. As well, they found that protein and RNA were visually correlated, but no statistical correlation analyses were performed. We attempted to correlate our RNA and protein data, but there was no significant correlation to be found, even with cross correlation (data not shown). In general, weak correlations are likely the results of the complexity and dynamicity of the wastewater matrix—with so many different variables that could influence viral detection including climate and weather, population variation, and wastewater infrastructure differences. We think that a protein-based LC–MS/MS approach offers high utility as a routine WBE method because it can function as an early warning signal for SARS-CoV-2, the consumable supplies cost \(\sim\) less per sample compared to RT-qPCR (based upon our own expenditures), and because the method is efficient as it can detect more biomarkers at once (potentially hundreds). For example, LC–MS/MS can be utilized to measure many other human-protein biomarkers of SARS-CoV-2 infection and susceptibility,\(^40,41\) as well as other viruses, pathogens, and disease biomarkers simultaneously in the same sample using our shotgun approach. One limitation is the high level of variation in wastewater samples and the possible presence of many unknown latent variables within the highly complex matrix that is wastewater. This makes absolute quantification a challenge, and thus correlations with our semiquantitative approach can be weak and difficult to interpret. However, the objective of this study was to determine if we could use this approach to observe larger temporal trends that reflect population-level SARS-CoV-2 infections, and we feel that this method successfully accomplished our purpose.

Here, we have presented an MS-based method for WW samples that specifically detect SARS-CoV-2 proteins, 6 and 5 days ahead of the case data, with between 0 and 27 new cases per day in these two municipalities while monitoring the period. We were able to identify unique peptides of at least eight proteins related to the SARS-CoV-2 virus and COVID-19 infection. We observed the consistent presence of the NSP protein pp1ab in all influent samples, which is transcribed after host cell infection. Given the results from the present study, we suspect that the pp1ab protein is present in high abundance in the urine and (more likely) feces\(^2\) of active COVID-19 cases, and that this protein might make an excellent alternative biomarker for monitoring populations and to detect infection outbreaks. Our ultimate goal is to use this methodology for sensitive, accurate, and routine detection of SARS-CoV-2 population-level infection in WW samples and to expand into other types of WBE using LC–MS/MS, including other human pathogens and biomarkers of disease for future public health studies.

### ASSOCIATED CONTENT

#### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.1c04705.

Table S1. Proteins identified by nontargeted data-dependent acquisition and database search using Spectrum Mill (list of proteins identified by nontargeted method). Table S2. Proteins identified by nontargeted data-dependent acquisition and database search using Spectrum Mill with Global FDR <1% (list of proteins identified by targeted method and FDR <1%). Table S3. Inclusion list for targeted MS/MS data acquisition (PDF)

### AUTHOR INFORMATION

#### Corresponding Author
Denina B. D. Simmons — Faculty of Science, The University of Ontario Institute of Technology, Oshawa, Ontario, Canada L1G 0C5; Email: Denina.Simmons@ontariotechu.ca

#### Authors
Linda R. Lara-Jacobo — Faculty of Science, The University of Ontario Institute of Technology, Oshawa, Ontario, Canada L1G 0C5; orcid.org/0000-0001-9412-0916

Golam Islam — Faculty of Science, The University of Ontario Institute of Technology, Oshawa, Ontario, Canada L1G 0C5

Jean-Paul Desaulniers — Faculty of Science, The University of Ontario Institute of Technology, Oshawa, Ontario, Canada L1G 0C5; orcid.org/0000-0002-9596-4552

Andrea E. Kirkwood — Faculty of Science, The University of Ontario Institute of Technology, Oshawa, Ontario, Canada L1G 0C5

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.1c04705

#### Funding

MITACS and the Ontario Clean Water Agency.

#### Notes

The authors declare no competing financial interest.

#### ACKNOWLEDGMENTS

The authors would like to thank The Region of Durham Public Health and Works Departments and Ontario Tech University for their assistance and in-kind contributions to this project.

#### REFERENCES

1. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K. S. M.; Lau, E. H. Y.; Wong, J. Y.; Xiang, X.; Xiang, N.; Wu, Y.; Li, C.; Chen, Q.; Li, D.; Liu, T.; Zhao, J.; Liu, M.; Tu, W.; Chen, C.; Jin, L.; Yang, R.; Wang, Q.; Zhou, S.; Wang, R.; Liu, H.; Luo, Y.; Liu, Y.; Shao, G.; Li, H.; Tuo, Z.; Yang, Y.; Deng, Z.; Liu, B.; Ma, Z.; Zhang, Y.; Shi, G.; Lam, T. T. Y.; Wu, J. T.; Gao, G. F.; Cowling, B. J.; Yang, B.; Leung, G. M.; Feng, Z. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N. Engl. J. Med. 2020, 382, 1199–1207.

2. Wang, W.; Xu, Y.; Gao, R.; Lu, R.; Han, K.; Wu, G.; Tan, W. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. JAMA 2020, 2, 1843–1844.
SARS-CoV-2 RNA from Wastewater. Environ. Sci. Technol. 2021, 55, 4880–4888.
(33) Yan, S.; Wu, G. Spatial and temporal roles of SARS-CoV P(PLpro). A snapshot. FASEB J. 2021, 35, e21197.
(34) Naqvi, A. A. T.; Fatima, K.; Mohammad, T.; Fatima, U.; Singh, I. K.; Singh, A.; Atif, S. M.; Hariprasad, G.; Hasan, G. M.; Hassan, M. I. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. Biochim. Biophys. Acta 2020, 1866, No. 165878.
(35) V’Kovski, P.; Kratzel, A.; Steiner, S.; Stalder, H.; Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. Nat. Rev. Microbiol. 2021, 19, 155–170.
(36) P0DTD1 (RIAB_SARS2). UniProtKB. https://www.uniprot.org/uniprot/P0DTD1 (accessed September 19, 2019)
(37) Rusinol, M.; Zammit, I.; Itarte, M.; Fores, E.; Martinez-Puchol, S.; Girones, R.; Borrego, C.; Corominas, L.; Bofill-Mas, S. Monitoring waves of the COVID-19 pandemic: Inferences from WWTPs of different sizes. Sci. Total Environ. 2021, 787, No. 147463.
(38) Chen, Y.; Liu, Q.; Guo, D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. J. Med. Virol. 2020, 92, 418–423.
(39) Neault, N.; Baig, A. T.; Graber, T. E.; D’Aoust, P. M.; Mercier, E.; Alexandrov, I.; Crosby, D.; Baird, S.; Mayne, J.; Pounds, T.; MacKenzie, M.; Figeys, D.; MacKenzie, A.; Delatolla, R. SARS-CoV-2 Protein in Wastewater Mirrors COVID-19 Prevalence, medRxiv 2020.09.01.20185280.2020. DOI: 10.1101/2020.09.01.20185280.
(40) Li, Y.; Wang, Y.; Liu, H.; Sun, W.; Ding, B.; Zhao, Y.; Chen, P.; Zhu, L.; Li, Z.; Li, N.; Chang, L.; Wang, H.; Bai, C.; Xu, P. Urine Proteome of COVID-19 Patients. Urine 2020, 2, 1–8.
(41) Wendt, R.; Kalbitz, S.; Lubbert, C.; Kellner, N.; Macholz, M.; Schroth, S.; Ermisch, J.; Latosisnka, A.; Arnold, B.; Mischak, H.; Beige, J.; Metzger, J. Urinary Proteomics Associates with COVID-19 Severity: Pilot Proof-of-Principle Data and Design of a Multicentric Diagnostic Study. Proteomics 2020, 20, No. e2000202.
(42) Chen, Y.; Chen, L.; Deng, Q.; Zhang, G.; Wu, K.; Ni, L.; Yang, Y.; Liu, B.; Wang, W.; Wei, C.; Yang, J.; Ye, G.; Cheng, Z. The presence of SARS-CoV-2 RNA in the feces of COVID-19 patients. J. Med. Virol. 2020, 92, 833–840.