

Article

Investigating the Use of SARS-CoV-2 (COVID-19) Odor Expression as a Non-Invasive Diagnostic Tool—Pilot Study

Janet Crespo-Cajigas ^{1,†}, Vidia A. Gokool ^{1,†}, Andrea Ramírez Torres ¹, Liam Forsythe ², Benjamin S. Abella ², Howard K. Holness ¹, Alan T. Charlie Johnson ³, Richard Postrel ⁴ and Kenneth G. Furton ^{1,*}

¹ Global Forensic and Justice Center, Department of Chemistry and Biochemistry, Florida International University, Miami, FL 33199, USA

² Department of Emergency Medicine and Penn Acute Research Collaboration, University of Pennsylvania, Philadelphia, PA 19104, USA

³ Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

⁴ VOC Health, Inc., Miami Beach, FL 33140, USA

* Correspondence: furtonk@fiu.edu

† These authors contributed equally to this work.

Abstract: Since the beginning of the COVID-19 pandemic, there has been enormous interest in the development of measures that would allow for the swift detection of the disease. The rapid screening and preliminary diagnosis of SARS-CoV-2 infection allow for the instant identification of possibly infected individuals and the subsequent mitigation of the disease spread. Herein, the detection of SARS-CoV-2-infected individuals was explored using noninvasive sampling and low-preparatory-work analytical instrumentation. Hand odor samples were obtained from SARS-CoV-2-positive and -negative individuals. The volatile organic compounds (VOCs) were extracted from the collected hand odor samples using solid phase microextraction (SPME) and analyzed using gas chromatography coupled with mass spectrometry (GC-MS). Sparse partial least squares discriminant analysis (sPLS-DA) was used to develop predictive models using the suspected variant sample subsets. The developed sPLS-DA models performed moderately (75.8% (± 0.4) accuracy, 81.8% sensitivity, 69.7% specificity) at distinguishing between SARS-CoV-2-positive and negative -individuals based on the VOC signatures alone. Potential markers for distinguishing between infection statuses were preliminarily acquired using this multivariate data analysis. This work highlights the potential of using odor signatures as a diagnostic tool and sets the groundwork for the optimization of other rapid screening sensors such as e-noses or detection canines.

Keywords: SARS-CoV-2; COVID-19; odor signature; HS-SPME-GC-MS; machine learning; sPLS-DA modeling; non-invasive diagnostic tool



Citation: Crespo-Cajigas, J.; Gokool, V.A.; Ramírez Torres, A.; Forsythe, L.; Abella, B.S.; Holness, H.K.; Johnson, A.T.C.; Postrel, R.; Furton, K.G. Investigating the Use of SARS-CoV-2 (COVID-19) Odor Expression as a Non-Invasive Diagnostic Tool—Pilot Study. *Diagnostics* **2023**, *13*, 707. <https://doi.org/10.3390/diagnostics13040707>

Academic Editor: Anna Baraniak

Received: 9 December 2022

Revised: 16 January 2023

Accepted: 18 January 2023

Published: 13 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in the deaths of over 6 million people on a global scale since the start of the COVID-19 pandemic in 2019 [1]. Due to the widespread expansion of this disease, along with its high capacity for infection even among asymptomatic individuals, there have been multiple studies on the development of technology and tools that would aid in its rapid screening and diagnosis [2–5]. The most common tools currently used for confirmatory diagnosis are the reverse transcriptase–quantitative polymerase chain reaction (RT-qPCR) test, which has been known to have ~90% sensitivity, on average [6–8], and the antigen test, which varies widely in terms of sensitivity and can range from 35 to 72% depending on the brand [6,7,9]. Despite the high sensitivity RT-qPCR is known for, there are several disadvantaging factors that can lead to false negatives such as a degraded sample or inadequate sample collection, a limit of detection higher than the available viral content, viral mutations, and a poor performance in sample testing [10]. False negative tests can result in further spreading of

the disease by undiagnosed patients. Additionally, PCR-based tests seem to have a long turnaround time, with individuals receiving a notification of their test results 24–48 h after specimen collection, further enhancing viral circulation. On the other hand, antigen tests are more efficient for patient screening, as they produce swift responses, are less expensive, and allow for self-testing [7,9]. Unfortunately, the sensitivity of the assay is sacrificed in exchange for a rapid response time. The sensitivity is even further reduced in asymptomatic populations [9]. Moreover, both RT-qPCR and antigen diagnostic tests can require a marginally invasive and uncomfortable sample collection from the patient [11], which could discourage individuals from seeking diagnostic services. There remains a pressing need for effective screening and diagnostic testing for public health surveillance, as SARS-CoV-2 has been proven to cause severe health outcomes.

Some other methods of detection that have been preliminarily conceptualized include—but are not limited to—the use of analytical instrumentation [4,12], e-noses [13], and canines [4,12,14]. Each of the aforementioned research studies have implemented the analysis of volatile organic compounds (VOCs) that are characteristic of SARS-CoV-2-infected individuals. VOCs are emitted by humans and other living organisms through various bodily secretions such as breath, saliva, urine, and sweat [15,16]. These VOCs can change over time and are often influenced by alterations in metabolic conditions, such as those induced by diseases [17]. Multiple reviews have been published over the past decade on the scent of diseases and how odor expression can potentially be used to diagnose various ailments such as gastro-intestinal diseases, infected wounds, and cancer [17–19]. Recently, the diagnosis of SARS-CoV-2 infection through the analysis of exhaled breath VOCs has become more thoroughly explored as an alternative screening method [3,20–22]. In 2020, Ruszkiewicz et al. performed a preliminary study wherein they discovered that they were able to distinguish between individuals who were SARS-CoV-2-infected and those who were presenting similar symptoms using gas chromatography (GC) and ion mobility spectroscopy with approximately 80.7% accuracy [22]. They hypothesized that a series of compounds including methanol and various ketones could provide a foundation for the development of a breath test for SARS-CoV-2 infection. Another study in 2021 commented on the reduced concentrations of acetone in the breath samples of SARS-CoV-2-infected individuals [20]. Additionally, Abumeeiz et al. (2021) assessed the opportunities and challenges of using a breathalyzer for the diagnosis of SARS-CoV-2 infections. The authors remarked that while high upfront costs and standardization might become an issue, the implementation of this technology could provide a worthwhile, noninvasive, and rapid test that would enable better disease management and responses to community outbreaks [3]. The researchers also took on a different approach to collect VOC odor from patients in a less invasive fashion, that is, from the hands versus the breath. Numerous studies have been published on the availability of VOCs from the hands of individuals and the ability of these VOCs to differentiate between subjects [23–25].

Hand odor collection would allow for a more rapid screening of individuals for disease states versus breath analysis methodologies, which require large volumes of exhaled breath to be collected for disease state determination. The analysis of hand odor also proves less invasive than the nasopharyngeal swabbing incurred for various RT-qPCR and antigen diagnostic tests. Additionally, hand odor swabs can be treated as a “clean sample” requiring no sample cleanup or liquid extraction prior to analysis by analytical instrumentation.

For these reasons, this research seeks to leverage the VOC detection of human hand odor using headspace solid phase microextraction (HS-SPME) sampling and gas chromatography-mass spectrometry (GC-MS) analysis to identify chemical biomarkers that may distinguish between SARS-CoV-2-positive and -negative individuals. The use of chemometric tools allows for efficient data manipulation and machine learning to develop sPLS-DA predictive models. This is the first in a set of papers that lays the groundwork for a non-invasive diagnostic approach for SARS-CoV-2. It is the belief of the researchers that the further expansion and successful implementation of this approach will also result in a framework for the development of rapid diagnostic tools for other diseases.

2. Materials and Methods

2.1. Human Hand Odor Collection

2.1.1. Sample Collection Details

Human hand odor samples were collected at three different time periods corresponding to the height of the Delta variant and the two Omicron SARS-CoV-2 subvariants (Table 1). The samples collected during late 2021 (Delta variant's dominance) were repurposed from a separate task. As such, the sample sets collected during 2021 and 2022 were collected using different collection procedures.

Table 1. SARS-CoV-2-Positive and -Negative Samples.

SARS-CoV-2 Infection Status	Timeframe	Collection Timespan	Dominant SARS-CoV-2 Variant at Time of Collection	Number of Samples
Positive	Late 2021	June 2021–October 2021	Delta	20
	Early 2022	February 2022–May 2022	Omicron BA.2	13
	Mid 2022	July 2022–September 2022	Omicron BA.5	23
Negative	2021–2022	September 2021–October 2022	Delta/Omicron BA.2/ Omicron BA.5	46

All samples were collected by researchers at the Penn Acute Research Collaboration-Penn Presbyterian Medical Center (PARC-PPMC) in Philadelphia, PA. The samples were collected from a patient population present in the Emergency Department at the time of sampling. All samples were collected under the approval of the University of Pennsylvania's Institutional Review Board (IRB# 848819). The VOC analysis of the collected samples was conducted by researchers at Florida International University in Miami, Florida.

2.1.2. Late 2021 Hand Odor Samples

Prior to the sample collection, the patients were asked to “cleanse” their hands by wiping them with a WaterWipes® wipe. They were then instructed to dry their hands with a paper towel or air-dry them. Samples were collected using sterile medical gauze. The gauze packaging was opened by the researcher without touching the gauze. The patient was instructed to (1) grab the exposed gauze and ball it up in their hands, (2) hold it in between their closed palms for approximately 10 s, and then (3) wipe both hands with the gauze before (4) placing it into a plastic specimen bag.

The specimen bag was labeled with a patient identifier. The hand odor sample was collected in conjunction with an underarm odor sample and a method control (blank) sample. All samples were packaged into separate specimen bags. The hand odor and control samples were used in this study; the underarm odor samples were not. The samples were stored below 3 °C until they were transferred to Florida International University for SPME odor extraction and subsequent GC-MS analysis.

2.1.3. Early-2022 and Mid-2022 Hand Odor Samples

Samples were collected using pre-treated sterile gauze [26]. The researcher prepared for the collection procedure by opening a vial containing the gauze and removing it with a pair of clean tweezers. The gauze was placed into the patient's hand, and the patient was instructed to (1) wipe both palms fully, (2) hold the gauze between their closed palms for 30 s, and then (3) ball-up the gauze and push it through the opening of the vial (held by the researcher). The researcher then resealed the vial and placed it into an aluminum barrier bag. This bag and the vial were labeled with patient information and then stored at −20 °C until they were ready for analysis. The samples were transferred into FIU's custody for chemical analysis and were delivered along with three method blank samples per shipment. The samples were generally received monthly during active sample collection periods.

2.2. Confirmation of SARS-CoV-2 Infection

In this study, 95 individuals were tested and diagnosed for SARS-CoV-2 infection using the Roche cobas[®] SARS-CoV-2 Duo Test for use on the cobas[®] 6800/8800 systems (Roche Diagnostics, Basel, Switzerland). The healthcare provider collected nasopharyngeal swab samples from all 95 patients for this PCR diagnostic test, which is the standard SARS-CoV-2 diagnostic method used by PPMC. Positive results were reported, as per the FDA EUA instructions for use, and signify the presence of SARS-CoV-2 RNA [27]. Three (3) individuals were tested using the Abbott BinaxNOW antigen test (Abbott Laboratories, Chicago, IL, USA), which consisted of a self-collected anterior nasal swab specimen. The results of this test indicated whether the nucleocapsid protein antigen of SARS-CoV-2 was identified [28]. Finally, the remaining four (4) individuals were not tested at PPMC; while the results of their SARS-CoV-2 test were relayed to the researchers, the sample specimen and test type remain unknown.

2.3. Patient Demographics

All patients were asked to verbally disclose the following information to the researchers: age, race/ethnicity, sex at birth, symptomology, vaccination status, and chief health complaint (what brought them to the hospital). This information can be found in the Appendix A in Table A1.

2.4. Preparation of Collection Materials

2.4.1. Vial Cleaning Procedure

The utilized 10 mL vials were (1) sonicated in a mild, soapy cleaning solution, followed by scrubbing and rinsing under tap water, (2) sonicated in a water bath, (3) sonicated in a bath of deionized water, (4) rinsed with acetone, and (5) baked at 105 °C for a minimum of an hour prior to being used to store collection materials. The sonicator was set to a 30 min cycle and held at 30 °C at each instance of use.

2.4.2. Pre-Treatment of Cotton Gauze

Sterile cotton gauze pads (100% cotton) served as the sorbent medium for the collection of Early-2022 and Mid-2022 human hand odor samples. Gauze pads of 2'' by 2'' and eight-ply density (Dukal Corporation, Syosset, Oyster Bay, NY, USA) were acquired and treated to a cleaning procedure prior to their use in sampling. The undergone pre-treatment procedure entails spiking 1 mL of HPLC-grade methanol (Fisher Chemical, Bridgewater, NJ, USA) onto the sterile gauze and baking the substrate for a minimum of one hour at 105 °C.

2.4.3. Storage & Containment

Pre-treated cotton gauze squares were stored in the cleaned 10 mL vials. The vials were sealed and secured with parafilm around the screw cap opening. All vials were labeled to collect information regarding: (a) Sample #, (b) Date, (c) Sex at Birth, and (d) SARS-CoV-2: Positive or Negative.

2.5. HS-SPME-GC-MS Method

The samples were placed in a digital heating bath set at 50 °C and left to equilibrate for 24 h. After this period, a clean 50/30 µm divinylbenzene/ carboxen/ polydimethylsiloxane (DVB/CAR/PDMS) SPME fiber was exposed to the headspace of the hand odor samples at a 1-inch fiber exposure setting. After 15 h, the SPME fibers were unexposed and removed from the sample headspace.

Analytes on the SPME fibers were thermally desorbed at 270 °C for 5 min (2-inch fiber height) into the heated inlet of the GC (Agilent 8890; Agilent Technologies, Santa Clara, CA, USA). A splitless injection method with a 1 mL/min column flow was implemented on an HP5-MS UI capillary column (15 m × 0.250 mm × 0.25 µm I.D.; Agilent Technologies). UHP Helium was used as the carrier gas. The oven temperature parameters started at 40 °C

(1.25 min hold), increased to 165 °C (5 °C/min rate), and concluded at 270 °C (30 °C/min rate). The total method runtime was 29.75 min. A mass spectrometer (MS) (Agilent 5977B MSD; Agilent Technologies) with an electron impact ionization (EI) source and quadrupole mass analyzer was used, with the following parameters: the MS source was maintained at 230 °C, the MS Quad at 150 °C, the transfer line at 280 °C and the EI source at 70 eV. Samples collected between June 2021 and May 2022 were analyzed using a scan range of m/z 50–550. Samples analyzed between July 2022 and October 2022 were analyzed using a 45–400 m/z scan range.

2.6. Data Pre-Processing

Collected human hand odor samples were analyzed using the described HS-SPME-GC-MS method (Section 2.5). The resulting datafiles were retention time-aligned, and peak matching was performed across the dataset. Following this procedure, the files were background-subtracted using the associated control sample. The background-subtracted samples were separated into four subgroups: (1) Late-2021 SARS-CoV-2-positive samples, (2) Early-2022 SARS-CoV-2-positive samples, (3) Mid-2022 SARS-CoV-2-positive samples, and (4) SARS-CoV-2-negative samples.

Within the SARS-CoV-2-positive subgroupings, the variance of each peak of interest was determined. Peaks of interest that demonstrated zero variance were removed from consideration from the sample set as a whole. Additionally, peaks corresponding to background interferents present in the method blank samples or identified as a recurring non-target such as methylene chloride or siloxanes peaks (present in the column phase and SPME fiber used) were removed from the sample set. The peaks of interest were able to be filtered down to include compounds eluting prior to the 21 min mark. The total ion chromatogram (TIC) peak areas were log₁₀-transformed prior to modeling to reduce the skewedness in the data; all values of 0 were set to 1 before applying the transformation.

2.7. Statistical Analysis

Sparse partial least squares discriminant analysis was used to form predictive models indicative of a patient's SARS-CoV-2 infection status. There were four models developed using the log₁₀-transformed TIC peak areas of the 40 features of interest. The first model demonstrates the outcome of informing a model using multiple sample collection timeframes (2021–2022), and models 2–4 demonstrate predictive models informed by a singular SARS-CoV-2 sample collection timeframe, which is believed to relate to a single variant's dominance in the population.

sPLS-DA modeling was performed using the “mixOmics” packages in R (Version 3.6.1, Vienna, Austria) [29]. In all cases, the sPLS-DA models were informed by an equal number of positive and negative samples. The lesser of the two was used as the defined class size, and the larger group was randomly sampled to provide an equal number of samples from each class. The models were cross-validated using a five-fold cross-validation, repeated 200 times. This resulted in a random division of the samples into 80% training set and 20% test set.

3. Results

3.1. sPLS-DA Modeling for All Timeframes

Following the pre-processing procedure detailed in Section 2.6, the sample set contained 40 features of interest. This sample set was log₁₀-transformed and used to conduct modeling of the SARS-CoV-2-positive and -negative hand odor samples. The sPLS-DA model of the hand odor samples demonstrates the class grouping of the positive and negative samples (Figure 1). The loading contributions of each feature of interest can be found in Figures A1 and A2.

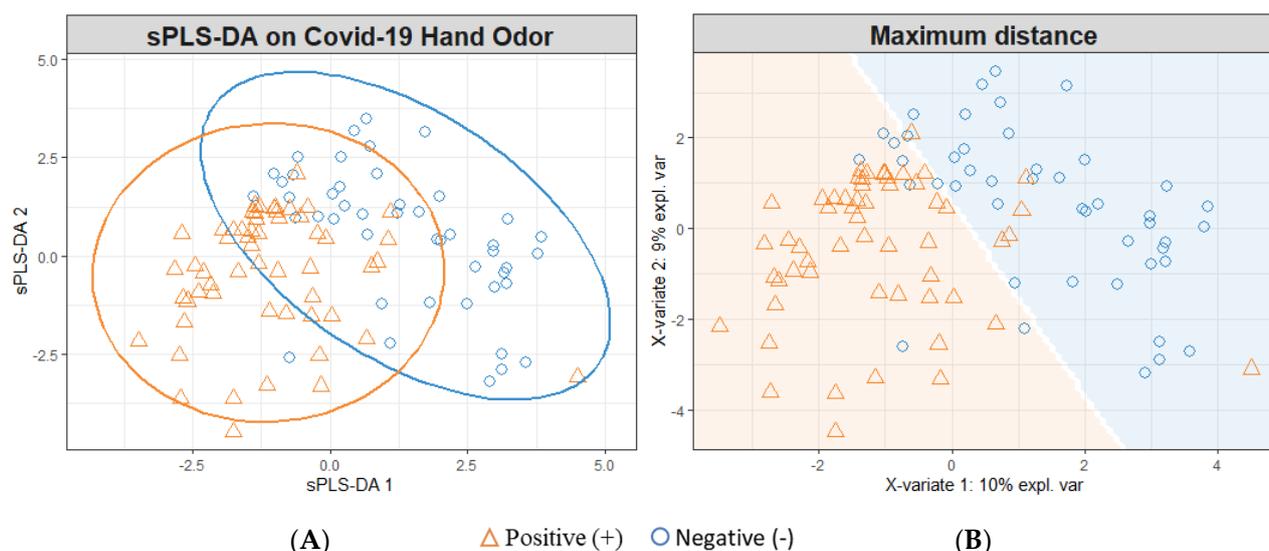


Figure 1. All-Timeframes SARS-CoV-2 sPLS-DA Model; (+) SARS-CoV-2 Samples vs. (−) SARS-CoV-2 Samples. (A) Individual sample plot; ellipses = 95% CI. (B) Individual sample plot; background delineates class assignment.

The sPLS-DA model shown in Figure 1 depicts the individual human hand odor samples collected from SARS-CoV-2-positive ($n = 56$) and SARS-CoV-2-negative ($n = 46$) individuals. These samples were collected across a timespan that included the dominant presence of both the Delta and Omicron variants of the SARS-CoV-2 virus. The model depicted (Figure 1) was cross-validated using a five-fold cross-validation, which was repeated 200 times; this procedure randomly sampled the positive samples to construct models informed by $n = 46$ positive and negative samples. Following this cross-validation, the model was seen to yield an accuracy of 75.8% (± 0.4); the model correctly predicted the SARS-CoV-2 infection status of a sample in 75.8% (± 0.4) of attempts (95% CI). This performance breaks down to a sensitivity/true positive rate (TPR) = 81.8% (± 0.5), specificity/true negative rate (TNR) = 69.7% (± 0.6), false positive rate (FPR) = 30.3% (± 0.6), false negative rate (FNR) = 18.2% (± 0.5), positive predictive value (PPV) = 73.0% (± 0.4), and negative predictive value (NPV) = 79.4% (± 0.5).

3.2. sPLS-DA Modeling for Individual Timeframes

Figure 2 demonstrates the modeling of Late-2021 positive samples ($n = 20$) and negative SARS-CoV-2 infection samples ($n = 46$). The model depicted (Figure 2) was cross-validated using a five-fold cross-validation, which was repeated 200 times; this procedure randomly sampled the negative samples to construct models informed by $n = 20$ positive and negative samples. Following this cross-validation, the model yielded an 86.7% (± 0.6) accuracy rate (95% CI). This performance breaks down to a sensitivity/TPR = 84.2% (± 0.6), specificity/TNR = 89.2% (± 0.9), FPR = 10.8% (± 0.9), FNR = 15.8% (± 0.6), PPV = 89.0% (± 0.9), and NPV = 85.0% (± 0.5).

Figure 3 demonstrates the modeling of Early-2022 positive ($n = 13$) and negative SARS-CoV-2 infection samples ($n = 46$). The model depicted (Figure 3) was cross-validated using a five-fold cross-validation, which was repeated 200 times. The negative samples were randomly sampled to construct models informed by $n = 13$ positive and negative samples. The cross-validated model yielded a 64.4% (± 1.0) accuracy rate. This performance breaks down to a sensitivity/TPR = 73.7% (± 0.7), specificity/TNR = 55.0% (± 1.7), FPR = 45.0% (± 1.7), FNR = 26.3% (± 0.7), PPV = 62.7% (± 1.0), and NPV = 67.1% (± 1.0).

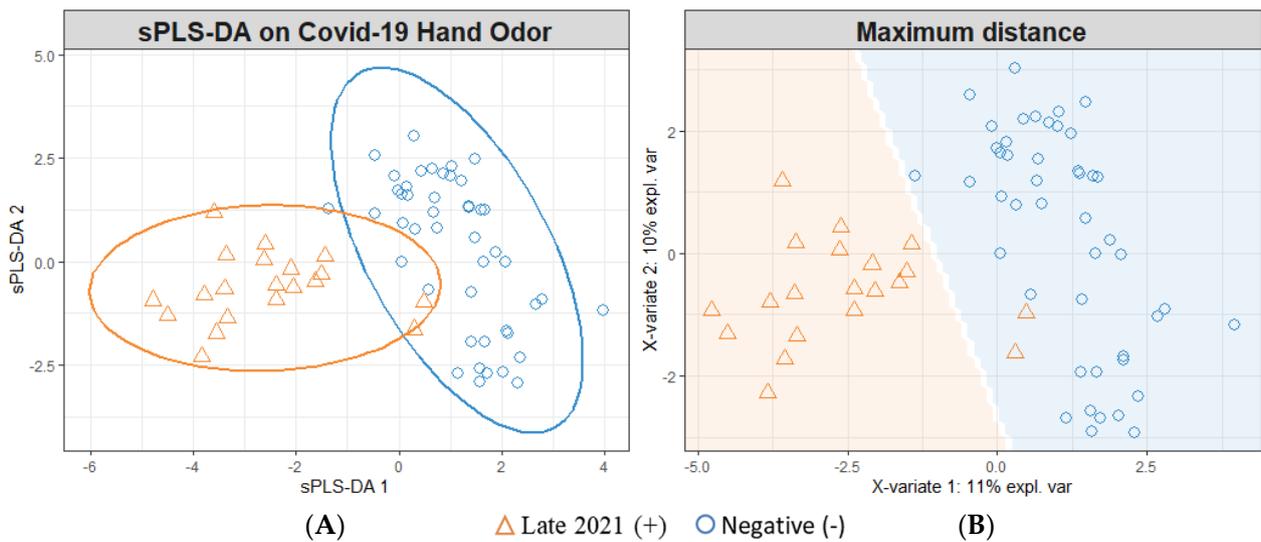


Figure 2. Late-2021 sPLS-DA Model; (+) Late-2021 Hand Odor Samples vs. (−) SARS-CoV-2 Samples. (A) Individual sample plot; ellipses = 95% CI. (B) Individual sample plot; background delineates class assignment.

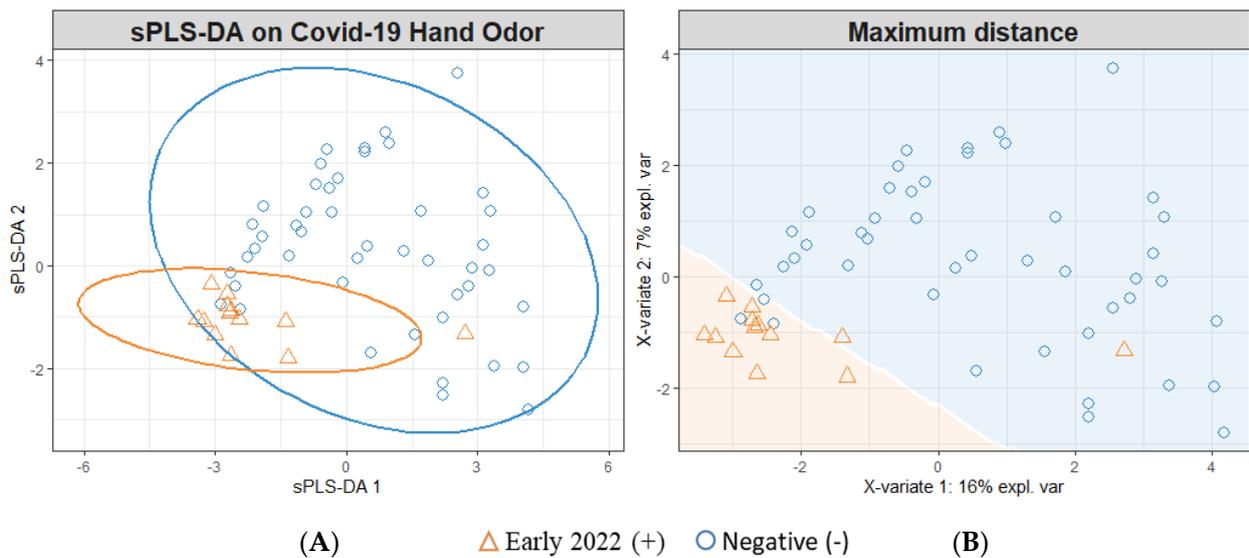


Figure 3. Early-2022 sPLS-DA Model; (+) Early-2022 Hand Odor Samples vs. (−) SARS-CoV-2 Samples. (A) Individual sample plot; ellipses = 95% CI. (B) Individual sample plot; background delineates class assignment.

Figure 4 demonstrates the modeling of Mid-2022 positive ($n = 23$) and negative SARS-CoV-2 infection samples ($n = 46$). The model depicted (Figure 4) was cross-validated using a five-fold cross-validation, which was repeated 200 times; this procedure randomly sampled the negative samples to construct models informed by $n = 23$ positive and negative samples. Following this cross-validation, the model yielded an 83.6% (± 0.8) accuracy rate. This performance breaks down to a sensitivity/TPR = 90.5% (± 0.9), specificity/TNR = 76.7% (± 1.0), FPR = 23.3% (± 1.0), FNR = 9.5% (± 0.9), PPV = 79.7% (± 0.8), and NPV = 89.3% (± 1.0).

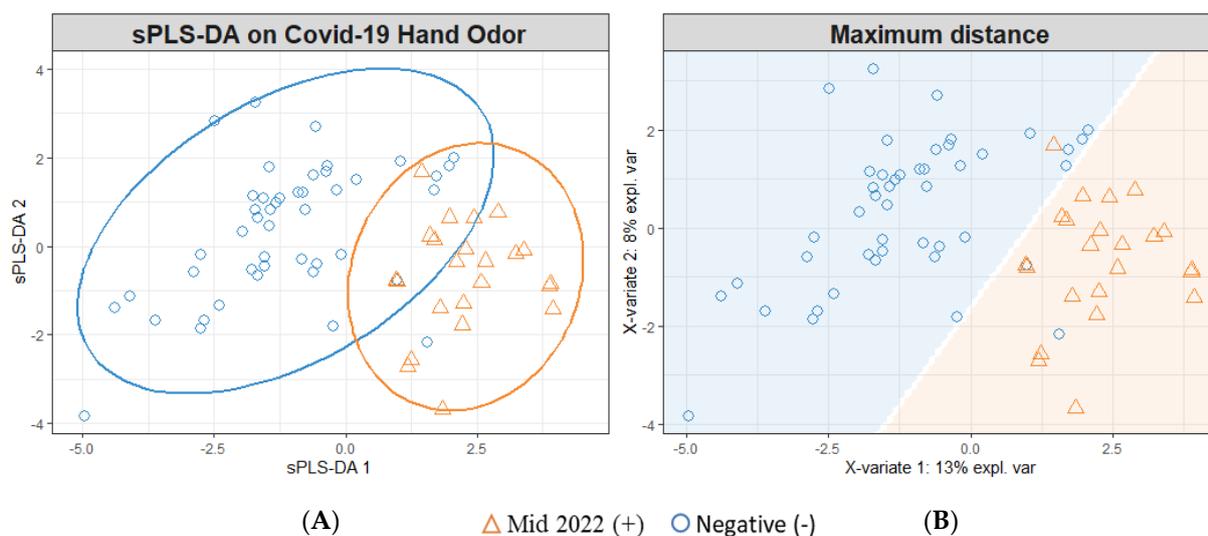


Figure 4. Mid-2022 sPLS-DA Model; (+) Mid-2022 Hand Odor Samples vs. (−) SARS-CoV-2 Samples. (A) Individual sample plot; ellipses = 95% CI. (B) Individual sample plot; background delineates class assignment.

3.3. Identification of Features of Interest

Out of the 40 features of interest used to inform the sPLS-DA models, the corresponding compound identifications of 14 were determined. These compounds (listed in Table 2) were identified through manual confirmation using externally run reference standards.

Table 2. Compounds contributing to the discrimination of SARS-CoV-2 infection.

Retention Time (15 M HP5-MS)	CAS#	Compound of Interest	Reported Presence in Human Skin Emanations
3.109	00123-42-2	Diacetone alcohol	-
4.033	00100-42-5	Styrene	[16,30]
6.577	03777-69-3	2-Pentylfuran	[16,31]
7.897	00122-78-1	Phenylacetaldehyde	[32]
9.552	01120-21-4	Undecane	[16,24–26,31,33,34]
10.366	00111-11-5	Methyl caprylate	[24–26,31,34,35]
11.235	18829-56-6	trans-2-Nonenal	[16,24,25,31,35]
11.679	00143-08-7	1-Nonanol	[24,31,34,35]
12.561	00112-31-2	Decanal	[16,24–26,30,31,33,35]
12.836	00122-99-6	2-Phenoxyethanol	-
17.798	00112-54-9	Dodecanal	[24,26,30,31,35]
18.841	00689-67-8	6,10-Dimethyl-5,9-undecadien-2-one-(E)	[16,24–26,30,31,33–35]
19.366	00112-53-8	1-Dodecanol	[36,37]
21.668	00143-07-7	Dodecanoic acid	[24,30,31,35,38,39]

4. Discussion

4.1. SARS-CoV-2 Infection Diagnostic Model Performance

SARS-CoV-2-infected human odor expression was investigated through the collection and analysis of hand odor samples. Hand odor samples were collected from 102 individuals; of these participants, 56 were SARS-CoV-2-positive and 46 were negative for SARS-CoV-2 at the time of donation. PCR tests were used as the primary confirmatory method in diagnosing SARS-CoV-2 infection status. The samples were collected using sterile gauze and analyzed using HS-SPME-GC-MS. The resulting human hand odor profile was analyzed using the previously described data pre-processing methods, resulting in a key list of 40 features of interest. Class separation (SARS-CoV-2-positive vs. -negative) was

modeled using the log₁₀-transformed total ion chromatogram peak areas of the features of interest. The resulting sPLS-DA modeling of all of the collected SARS-CoV-2-positive and -negative samples (Figure 1) demonstrated class grouping within the negative and positive samples. Although a defined separation of classes was not observed, the model reflected a moderate performance ability with 75.8% accuracy when compared to other screening and diagnostic techniques such as antigen tests and detection canines (as summarized in a previous publication [12]). This rate reflected a sensitivity (TPR) of 81.8% and a specificity (TNR) of 69.7% for predicting the SARS-CoV-2 status. The performance of each model based on a repeated five-fold cross-validation ($n = 200$) can be seen in Table 3 below.

Table 3. Overall and Variant-Specific sPLS-DA Performance.

sPLS-DA Model Correct Prediction Rates *				
Time Range	2021–2022	Late 2021	Early 2022	Mid 2022
Suspected Variant	All	Delta	Omicron BA.2	Omicron BA.5
Accuracy	75.8% (± 0.4)	86.7% (± 0.6)	64.4% (± 1.0)	84.4% (± 0.8)
Sensitivity (TPR)	81.8% (± 0.5)	84.2% (± 0.6)	73.7% (± 0.7)	90.5% (± 0.9)
Specificity (TNR)	69.7% (± 0.6)	89.2% (± 0.9)	55.0% (± 1.7)	76.7% (± 1.0)
FNR	18.2% (± 0.5)	15.8% (± 0.6)	26.3% (± 0.7)	9.5% (± 0.9)
FPR	30.3% (± 0.6)	10.8% (± 0.9)	45.0% (± 1.7)	23.3% (± 1.0)
PPV	73.0% (± 0.4)	89.0% (± 0.9)	62.7% (± 1.0)	79.7% (± 0.8)
NPV	79.4% (± 0.5)	85.0% (± 0.5)	67.1% (± 1.0)	89.3% (± 1.0)

* Error Rate = 100% (Correct Prediction Rate).

The SARS-CoV-2-positive samples were further separated into Late-2021, Early-2022, and Mid-2022 samples and are suspected to correspond to Delta, Omicron BA.2, and Omicron BA.5 variants, respectively. The variant type was not confirmed; it was assigned based upon the variant that was dominant in the collection area (Philadelphia, Pennsylvania) at the time of sample donation. This assignment was conducted using the Centers for Disease Control and Prevention's variant proportional COVID data reporting platform [40]. The Late-2021 SARS-CoV-2-positive samples were modeled against the SARS-CoV-2-negative samples, revealing an increased separability between the classes (Figure 2) compared to that demonstrated in Figure 1. This observed separation translated into an improved performance in correctly identifying SARS-CoV-2-infected sample status (86.7% accuracy). Similarly, the modeling of Mid-2022 SARS-CoV-2 (Figure 4)-positive samples reflected an increase in the rate of accuracy to 84.4%. However, the individual modeling of Early-2022 SARS-CoV-2-positive samples vs. SARS-CoV-2-negative samples (Figure 3) resulted in a reduction in discriminatory power, with a reduced rate of accuracy at 64.4%.

The individual modeling of samples by variant indicates that sPLS-DA models trained on samples collected within the dominance of the Delta or Omicron BA.5 variants yielded a greater predictive power than models trained using samples collected during the height of Omicron BA.2 or informed by multiple variant timeframes. The change in model performance in relation to SARS-CoV-2 variant dominance is an interesting phenomenon which suggests a difference in the underlying VOC expressions between SARS-CoV-2 variants. This variant-dependent expression of SARS-CoV-2 via volatile emanations is a topic of future study and will be further investigated by the researchers. Despite the suggestion of variant dependency in SARS-CoV-2 expression, the predictive power of the sPLS-DA model informed by all samples collected between June 2021 and October 2022 suggests that there is an underpinning commonality in the human odor expression associated with SARS-CoV-2 infection.

This work relied upon the sPLS-DA models which were informed by 40 features of interest. The researchers suggest that these compounds be considered potential biomarkers for the expression of SARS-CoV-2 via human odor. Of the 40 features of interest utilized in the machine learning models, 14 compounds have been identified (Table 2). Many of these compounds consist of aldehydes, alcohols, ketones, and other functional groups that are

common in human scent [26]. The authors were able to identify instances of the published presentation of these compounds in human scent samples for 12/14 compounds (Table 2). The previous reporting of two compounds (diacetone alcohol and 2-phenoxyethanol) in skin emanations was not determined. However, it is noted that 2-phenoxyethanol has been detected in human feces and saliva samples, while diacetone alcohol has been identified as a secondary metabolite and is believed to play a role in signaling processes; both have been identified in human samples, but to the authors' knowledge, they have not been reported in human scent samples.

4.2. Limitations and Future Perspective

Due to the wide performance range of the antigen tests currently on the market, there remains a fair possibility that the three individuals tested by an antigen test alone were not accurately accounted for (positive or negative). While the PCR test is the current gold standard for SARS-CoV-2 detection and diagnostics, it is an imperfect tool as well. As with most diagnostic tools, it is important to acknowledge the possibility of false positive and false negative results produced by PCR tests. Additionally, the tests are bound to their own detection limitations. Participants who tested as "negative" using a PCR test may have attained this result due to a low viral load or stage of infection as opposed to a "truly negative" SARS-CoV-2 infection status.

In addition to concerns over ground truth SARS-CoV-2-infected patient status, there are concerns over previous contractions of SARS-CoV-2 and the possibility that patients were afflicted by long-COVID at the time of sample collection. Recent work in the field of biomedical scent detection has suggested that canines trained to detect SARS-CoV-2-infected positive samples may also be able to detect long-COVID, which is the persistent effects of COVID-19 that remain beyond PCR test positivity (post-COVID affliction) [41]. These new findings suggest that there is more similarity between long-COVID odor presentation and SARS-CoV-2-positive odor presentation than there is between long-COVID and SARS-CoV-2-negative presentations. Due to the difficulty in properly determining previous SARS-CoV-2 infection (individuals may have been asymptomatic or have gone untested during a period of illness), the researchers did not actively seek to identify and remove individuals who may have been suffering from long-COVID. This consideration is viewed as a reasonable approximation of the real-world application of this work, as patients would be subject to SARS-CoV-2 screening regardless of long-COVID status.

The authors intend to continue their investigation of the VOC expression of SARS-CoV-2 infection as it relates to human hand odor. This work served as an initial investigation of the probative value of human hand odor as a medium for SARS-CoV-2 infection status determination. The authors intend to continue this discussion in a secondary manuscript wherein the influence of the suspected SARS-CoV-2 variant is evaluated as a factor driving class separation. Additionally, the separability of SARS-CoV-2 variants will be examined.

The presented methodology of using human hand odor as a potential diagnostic medium is a promising prospect for an adaptive field-deployable technology. While the authors note that the infrastructure does not currently exist for the wide application of this technology, they believe that human hand odor analysis is a viable diagnostic medium that has the capacity to be expanded upon and adapted into a fieldable methodology. The greatest benefit of this technology would reside in mass-screening venues such as airports or other readily visited areas with elevated concerns of transmission. This begins with the adoption of portable GC-MS instrumentation, allowing for lower-cost installations. This methodology could also open a pathway for the continual improvement of the technology through the periodic introduction of new SARS-CoV-2-positive and -negative samples (verified by PCR testing). As status is confirmed, more samples can be added to the model, allowing for continual training; as SARS-CoV-2 variants evolve, so will the model. The adoption of a machine learning-backed approach provides a route to teaching and improving diagnostic models at a rate that is competitive regarding the continual evolution of our targeted virus.

5. Conclusions

This work demonstrated that HS-SPME-GC-MS could be used jointly with a developed sPLS-DA model to predict the SARS-CoV-2 infection status of a patient. Furthermore, the developed method shows that using odor samples from the hands of subjects is sufficient to serve as a potential diagnostic medium. The developed model displayed a 75.8% accuracy. The performance rate was seen to be similar for the sensitivity rate (81.8%) and specificity rate (69.7% correct) for correctly predicting SARS-CoV-2 infection status, implying that this technology has the capacity to be competitive with antigen testing in terms of performance metrics. This model displays an improved overall performance ability when compared to the model informed by a timespan-specific subset corresponding to the dominant Omicron BA.2 and reflects decreased performance when compared to the models trained on the Delta and Omicron BA.5 correlated sample subsets. The computational model was informed by 40 features of interest; at this time, the compound identities of 14 out of 40 features have been identified and confirmed using certified reference standards. This noninvasive, low-preparatory-analysis scheme has the potential to be continually informed to evolve alongside the SARS-CoV-2 virus. The further optimization of this diagnostic analysis could facilitate the rapid screening and mitigation of disease spread and could potentially be adapted to other infectious diseases with expressed human odor profile variations.

Author Contributions: Conceptualization, H.K.H.; methodology, J.C.-C. and V.A.G.; software, V.A.G.; validation, J.C.-C. and V.A.G.; formal analysis, V.A.G. and J.C.-C.; data curation, V.A.G. and J.C.-C.; investigation, J.C.-C., V.A.G. and A.R.T.; resources, L.F., B.S.A. and K.G.F.; writing—original draft preparation, V.A.G. and J.C.-C.; writing—review and editing, V.A.G. and J.C.-C.; visualization, V.A.G.; supervision, H.K.H.; project administration, K.G.F. and A.T.C.J.; funding acquisition, K.G.F., R.P. and A.T.C.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the NIH National Center for Advancing Translational Sciences 1-U18-Tr-003775-01.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the University of Pennsylvania (IRB#843534; approved on 26 June 2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The dataset associated with this study has been submitted to the National Institute of Health's Scientific Data Sharing platform. Access to the dataset can be requested through <https://sharing.nih.gov/>, accessed on 16 January 2023.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Demographic Breakdown of Patients.

Race	Sex at Birth		SARS-CoV-2 Infection Status				Age			
	Male	Female	Delta (+)	Omicron BA.2 (+)	Omicron BA.5 (+)	(−)	18–30	31–45	46–60	60+
Asian	3	0	0	0	0	3	1	2	0	0
Black	29	46	18	4	17	36	15	19	19	22
Hispanic/Latinx *	0	1	0	1	0	0	1	0	0	0
Unknown	1	0	1	0	0	0	0	0	1	0
White	7	15	1	8	6	7	8	3	6	5

* Self-identified as a white Hispanic and/or Latinx person.

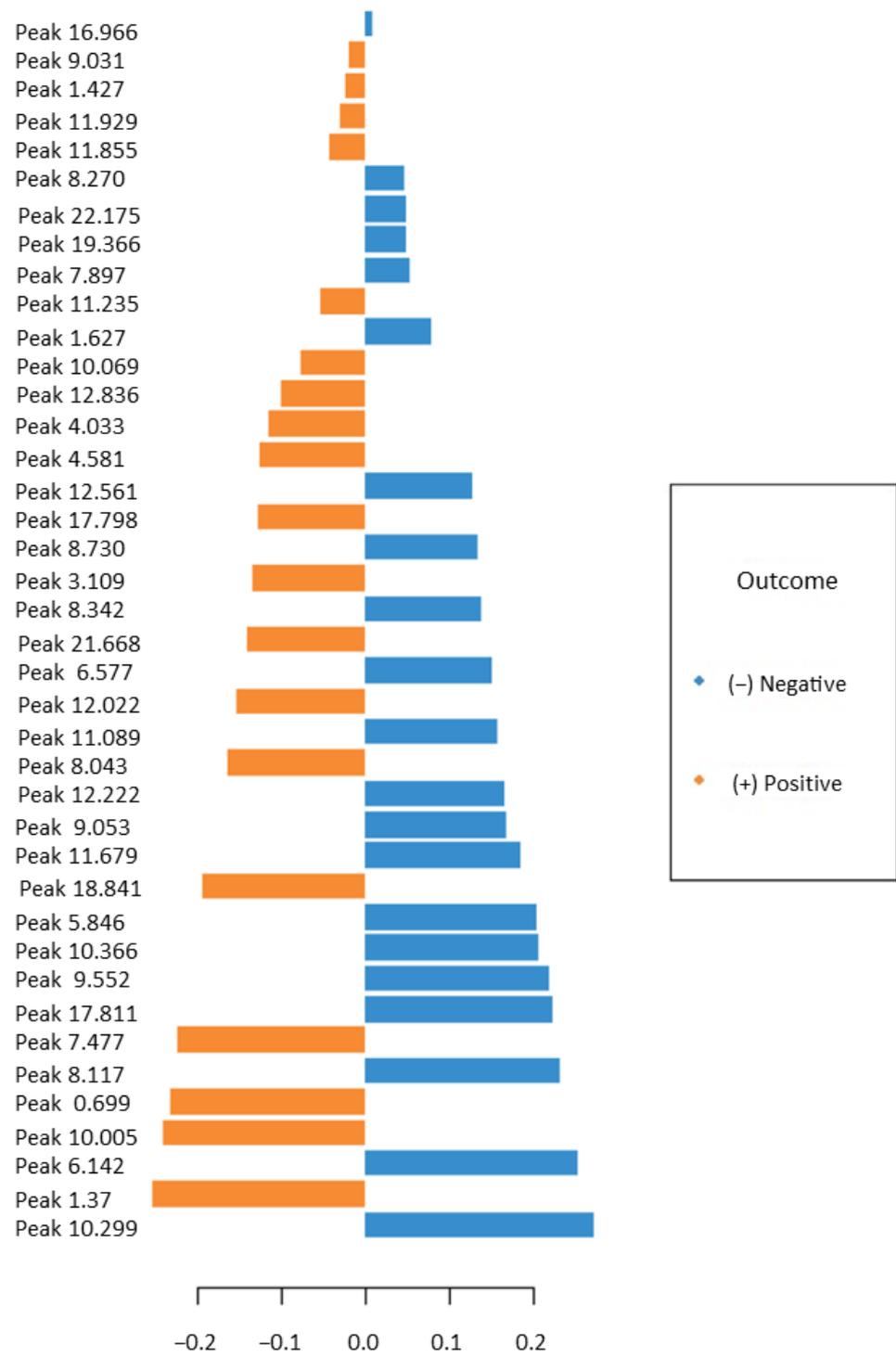


Figure A1. Importance and Feature Contribution of 40 Features of Interest, corresponding to component 1 in Figure 1.

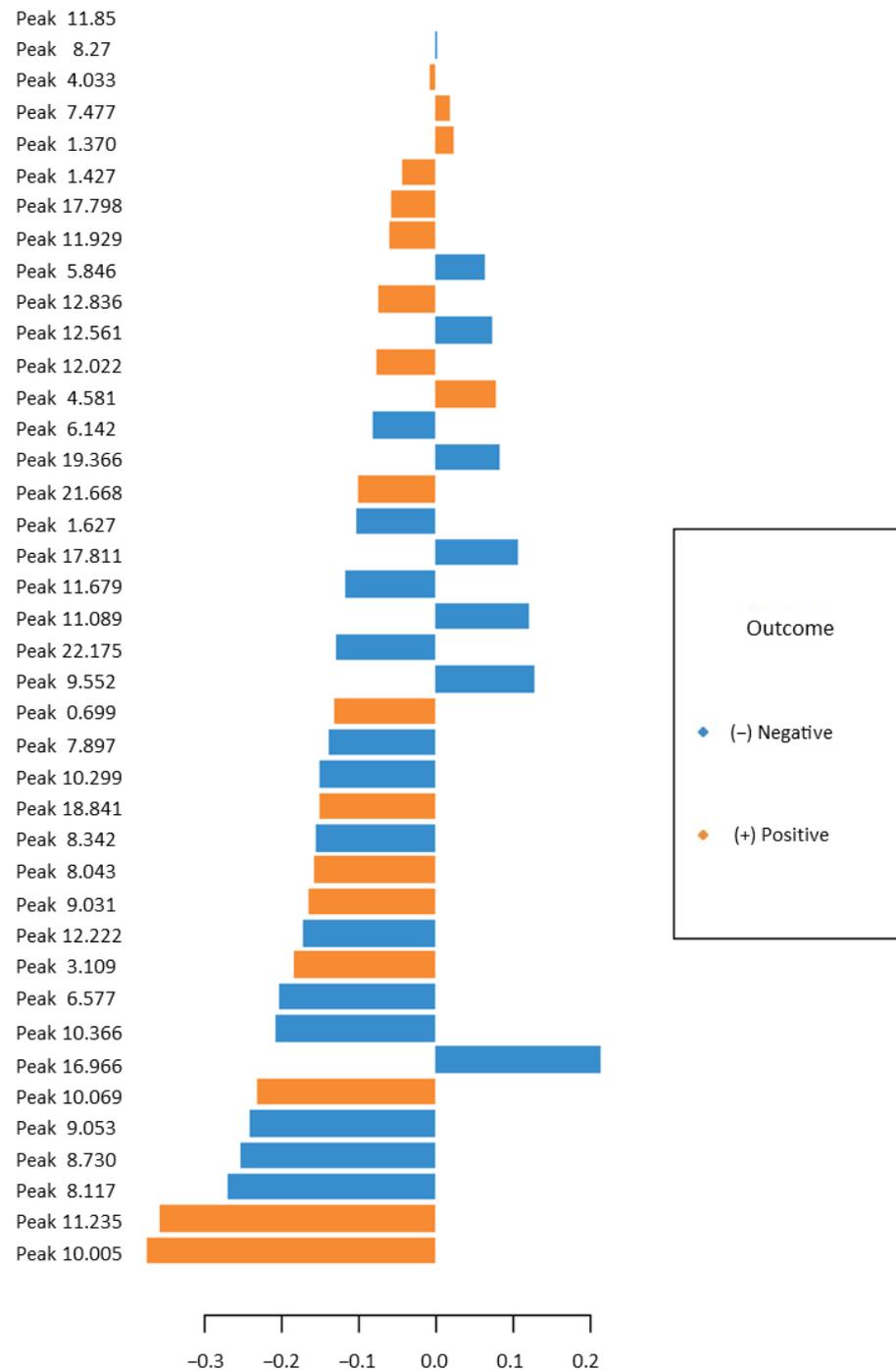


Figure A2. Importance and Feature Contribution of 40 Features of Interest, corresponding to component 2 in Figure 1.

References

1. World Health Organization WHO Coronavirus (COVID-19) Dashboard. Available online: <https://covid19.who.int/> (accessed on 13 August 2022).
2. Fernandes, Q.; Inchakalody, V.P.; Merhi, M.; Mestiri, S.; Taib, N.; Moustafa Abo El-Ella, D.; Bedhiafi, T.; Raza, A.; Al-Zaidan, L.; Mohsen, M.O.; et al. Emerging COVID-19 variants and their impact on SARS-CoV-2 diagnosis, therapeutics and vaccines. *Ann. Med.* **2022**, *54*, 524–540. [CrossRef] [PubMed]
3. Abumeeiz, M.; Elliott, L.; Olla, P. Use of Breath Analysis for Diagnosing COVID-19: Opportunities, Challenges, and Considerations for Future Pandemic Responses. *Disaster Med. Public Health Prep.* **2021**, *16*, 2137–2140. [CrossRef] [PubMed]

4. Mendel, J.; Frank, K.; Edlin, L.; Hall, K.; Webb, D.; Mills, J.; Holness, H.K.; Furton, K.G.; Mills, D.E. Preliminary accuracy of COVID-19 odor detection by canines and HS-SPME-GC-MS using exhaled breath samples. *Forensic Sci. Int. Synerg.* **2021**, *3*, 100155. [[CrossRef](#)] [[PubMed](#)]
5. Sharun, K.; Jose, B.; Tiwari, R.; Natesan, S.; Dhama, K. Biodetection dogs for COVID-19: An alternative diagnostic screening strategy. *Public Health* **2021**, *197*, e10. [[CrossRef](#)]
6. Kritikos, A.; Caruana, G.; Brouillet, R.; Miroz, J.P.; Abed-Maillard, S.; Stieger, G.; Opota, O.; Croxatto, A.; Vollenweider, P.; Bart, P.A.; et al. Sensitivity of rapid antigen testing and rt-pcr performed on nasopharyngeal swabs versus saliva samples in COVID-19 hospitalized patients: Results of a prospective comparative trial (restart). *Microorganisms* **2021**, *9*, 1910. [[CrossRef](#)]
7. Dinnes, J.; Deeks, J.J.; Berhane, S.; Taylor, M.; Adriano, A.; Davenport, C.; Dittrich, S.; Emperador, D.; Takwoingi, Y.; Cunningham, J.; et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst. Rev.* **2021**, *3*, 1–129. [[CrossRef](#)]
8. Alcoba-Florez, J.; Gil-Campesino, H.; de Artola, D.G.M.; González-Montelongo, R.; Valenzuela-Fernández, A.; Ciuffreda, L.; Flores, C. Sensitivity of different RT-qPCR solutions for SARS-CoV-2 detection. *Int. J. Infect. Dis.* **2020**, *99*, 190. [[CrossRef](#)]
9. Brihn, A.; Chang, J.; OYong, K.; Balter, S.; Terashita, D.; Rubin, Z.; Yeganeh, N. Diagnostic Performance of an Antigen Test with RT-PCR for the Detection of SARS-CoV-2 in a Hospital Setting—Los Angeles County, California, June–August 2020. *MMWR. Morb. Mortal. Wkly. Rep.* **2021**, *70*, 702–706. [[CrossRef](#)]
10. Moulou, D.S.; Gourgoulianis, K.I. False-positive and false-negative COVID-19 cases: Respiratory prevention and management strategies, vaccination, and further perspectives. *Expert Rev. Respir. Med.* **2021**, *15*, 993–1002. [[CrossRef](#)]
11. Marra, P.; Colacurcio, V.; Bisogno, A.; De Luca, P.; Calvanese, M.; Petrosino, M.; De Bonis, E.; Troisi, D.; Cassandro, C.; Cavaliere, M.; et al. Evaluation of Discomfort in Nasopharyngeal Swab Specimen Collection for SARS-CoV-2 Diagnosis. *Clin. Ter.* **2021**, *172*, 448–452. [[CrossRef](#)]
12. Gokool, V.A.; Crespo-Cajigas, J.; Mallikarjun, A.; Collins, A.; Kane, S.A.; Plymouth, V.; Nguyen, E.; Abella, B.S.; Holness, H.K.; Furton, K.G.; et al. The Use of Biological Sensors and Instrumental Analysis to Discriminate COVID-19 Odor Signatures. *Biosensors* **2022**, *12*, 1003. [[CrossRef](#)]
13. Bordbar, M.M.; Samadinia, H.; Hajian, A.; Sheini, A.; Safaei, E.; Aboonajmi, J.; Arduini, F.; Sharghi, H.; Hashemi, P.; Khoshsafar, H.; et al. Mask assistance to colorimetric sniffers for detection of COVID-19 disease using exhaled breath metabolites. *Sens. Actuators B Chem.* **2022**, *369*, 132379. [[CrossRef](#)]
14. Eskandari, E.; Ahmadi Marzaleh, M.; Roudgari, H.; Hamidi Farahani, R.; Nezami-Asl, A.; Laripour, R.; Aliyazdi, H.; Dabbagh Moghaddam, A.; Zibaseresht, R.; Akbarialiabadi, H.; et al. Sniffer dogs as a screening/diagnostic tool for COVID-19: A proof of concept study. *BMC Infect. Dis.* **2021**, *21*, 243. [[CrossRef](#)]
15. Curran, A.M.; Rabin, S.I.; Furton, K.G. Analysis of the Uniqueness and Persistence of Human Scent. *Forensic Sci. Commun.* **2005**, *7*, 1–12.
16. Kusano, M.; Mendez, E.; Furton, K.G. Comparison of the Volatile Organic Compounds from Different Biological Specimens for Profiling Potential. *J. Forensic Sci.* **2013**, *58*, 29–39. [[CrossRef](#)]
17. Shirasu, M.; Touhara, K. The scent of disease: Volatile organic compounds of the human body related to disease and disorder. *J. Biochem.* **2011**, *150*, 257–266. [[CrossRef](#)]
18. Wilson, A.D. Noninvasive early disease diagnosis by electronic-nose and related VOC-detection devices. *Biosensors* **2020**, *10*, 73. [[CrossRef](#)]
19. Buljubasic, F.; Buchbauer, G. The scent of human diseases: A review on specific volatile organic compounds as diagnostic biomarkers. *Flavour Fragr. J.* **2015**, *30*, 5–25. [[CrossRef](#)]
20. Chen, H.; Qi, X.; Zhang, L.; Li, X.; Ma, J.; Zhang, C.; Feng, H.; Yao, M. COVID-19 screening using breath-borne volatile organic compounds. *J. Breath Res.* **2021**, *15*, 047104. [[CrossRef](#)]
21. Lamote, K.; Janssens, E.; Schillebeeckx, E.; Lapperre, T.S.; De Winter, B.Y.; Van Meerbeeck, J.P. The scent of COVID-19: Viral (semi-)volatiles as fast diagnostic biomarkers? *J. Breath Res.* **2020**, *14*, 042001. [[CrossRef](#)]
22. Ruszkiewicz, D.M.; Sanders, D.; O'Brien, R.; Hempel, F.; Reed, M.J.; Riepe, A.C.; Bailie, K.; Brodrick, E.; Darnley, K.; Ellerkmann, R.; et al. Diagnosis of COVID-19 by analysis of breath with gas chromatography-ion mobility spectrometry—A feasibility study. *EClinicalMedicine* **2020**, *29*, 100609. [[CrossRef](#)]
23. Curran, A.M.; Rabin, S.I.; Prada, P.A.; Furton, K.G. Comparison of the volatile organic compounds present in human odor using SPME-GC/MS. *J. Chem. Ecol.* **2005**, *31*, 1607–1619. [[CrossRef](#)] [[PubMed](#)]
24. Curran, A.M.; Ramirez, C.F.; Schoon, A.A.; Furton, K.G. The frequency of occurrence and discriminatory power of compounds found in human scent across a population determined by SPME-GC/MS. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2007**, *846*, 86–97. [[CrossRef](#)] [[PubMed](#)]
25. Curran, A.M.; Prada, P.A.; Furton, K.G. The differentiation of the volatile organic signatures of individuals through SPME-GC/ms of characteristic human scent compounds. *J. Forensic Sci.* **2010**, *55*, 50–57. [[CrossRef](#)] [[PubMed](#)]
26. Food and Drug Administration. Cobas SARS-CoV-2—Instructions For Use. Available online: <https://www.fda.gov/media/136049/download> (accessed on 15 January 2023).
27. Food and Drug Administration. BINAXNOW COVID-19 AG CARD (PN 195-000)—Instructions for Use. Available online: <https://www.fda.gov/media/141570/download> (accessed on 15 January 2023).

28. Gokool, V.A.; Holness, H.K.; Furton, K.G. The influence of intra-personal variations in human hand odor on the determination of sample donor. *Forensic Sci. Int.* **2022**, *334*, 111235. [[CrossRef](#)]
29. Le Cao, K.-A.; Rohart, F.; Gonzalez, I.; Dejean, S.; Gautier, B.; Bartolo, F.; Monget, P.; Coquery, J.; Yao, F.; Lique, B. mixOmics: Omics Data Integration Project. R Package Version 6.1.1. Available online: <https://cran.r-project.org/package=mixOmics> (accessed on 13 August 2022).
30. Bernier, U.R.; Kline, D.L.; Barnard, D.R.; Schreck, C.E.; Yost, R.A. Analysis of human skin emanations by gas chromatography/mass spectrometry. 2. Identification of volatile compounds that are candidate attractants for the yellow fever mosquito (*Aedes aegypti*). *Anal. Chem.* **2000**, *72*, 747–756. [[CrossRef](#)]
31. Brown, J.S.; Prada, P.A.; Curran, A.M.; Furton, K.G. Applicability of emanating volatile organic compounds from various forensic specimens for individual differentiation. *Forensic Sci. Int.* **2013**, *226*, 173–182. [[CrossRef](#)]
32. Tavares, D.d.S.; Salgado, V.R.; Miranda, J.C.; Mesquita, P.R.R.; Rodrigues, F.d.M.; Barral-Netto, M.; de Andrade, J.B.; Barral, A. Attraction of phlebotomine sandflies to volatiles from skin odors of individuals residing in an endemic area of tegumentary leishmaniasis. *PLoS ONE* **2018**, *13*, e0203989. [[CrossRef](#)]
33. Zhang, Z.-M.; Cai, J.-J.; Ruan, G.-H.; Li, G.-K. The study of fingerprint characteristics of the emanations from human arm skin using the original sampling system by SPME-GC/MS. *J. Chromatogr. B* **2005**, *822*, 244–252. [[CrossRef](#)]
34. Caraballo, N.I.; Mendel, J.; Holness, H.; La Salvia, J.; Moroosse, T.; Eckenrode, B.; Stockham, R.; Furton, K.; Mills, D.E. An investigation into the concurrent collection of human scent and epithelial skin cells using a non-contact sampling device. *Forensic Sci. Int.* **2016**, *266*, 148–159. [[CrossRef](#)]
35. Prada, P.A.; Curran, A.M.; Furton, K.G. The Evaluation of Human Hand Odor Volatiles on Various Textiles: A Comparison Between Contact and Noncontact Sampling Methods. *J. Forensic Sci.* **2011**, *56*, 866–881. [[CrossRef](#)]
36. Meijerink, J.; Braks, M.A.H.; Brack, A.A.; Adam, W.; Dekker, T.; Posthumus, M.A.; Beek, T.A.; Van Loon, J.J.A. Identification of olfactory stimulants for anopheles gambiae from human scent samples. *J. Chem. Ecol.* **2000**, *26*, 1367–1382. [[CrossRef](#)]
37. Vaglio, S.; Minicozzi, P.; Bonometti, E.; Mello, G.; Chiarelli, B. Volatile signals during pregnancy: A possible chemical basis for mother-infant recognition. *J. Chem. Ecol.* **2009**, *35*, 131–139. [[CrossRef](#)]
38. Colón-Crespo, L.J.; Herrera-Hernández, D.; Holness, H.; Furton, K.G. Determination of VOC marker combinations for the classification of individuals by gender and race/ethnicity. *Forensic Sci. Int.* **2017**, *270*, 193–199. [[CrossRef](#)]
39. Curran, A.M.; Prada, P.A.; Furton, K.G. Canine human scent identifications with post-blast debris collected from improvised explosive devices. *Forensic Sci. Int.* **2010**, *199*, 103–108. [[CrossRef](#)]
40. Centers for Disease Control and Prevention. CDC COVID Data Tracker: Variant Proportions. Available online: <https://covid.cdc.gov/covid-data-tracker/#variant-proportions> (accessed on 13 August 2022).
41. Twele, F.; ten Hagen, N.A.; Meller, S.; Schulz, C.; Osterhaus, A.; Jendry, P.; Ebbers, H.; Pink, I.; Drick, N.; Welte, T.; et al. Detection of Post-COVID-19 Patients Using Medical Scent Detection Dogs—A Pilot Study. *Front. Med.* **2022**, *9*, 877259. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.