Breath biomarkers of pediatric SARS-CoV-2 infection: a pilot study

Amalia Z. Berna[1], Elikplim H. Akaho[1], Rebecca M. Harris[1,2], Morgan Congdon[1,2], Emilie Korn[1,2], Samuel Neher[1,2], Mirna M'Farrej[1,2], Julianne Burns[1,2], Audrey R. Odom John[1,2*]

[1]Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

[2]Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

* Correspondence should be addressed to A.R.O.J. (JOHNA3@email.chop.edu)

Abstract

COVID-19 control efforts have been hampered by transmission from pre-symptomatic individuals infected with SARS-CoV2. Prolonged asymptomatic respiratory viral shedding in children has been described and may be another important reservoir for ongoing transmission. The primary diagnostic approach to identify SARS-CoV2 infection relies on qPCR of specific viral sequences from respiratory samples, which is expensive, uncomfortable, relatively slow, and susceptible to false-negative results. A rapid non-invasive method to detect mild or asymptomatic infection would have a major impact on public health campaigns to control COVID-19. We hypothesize that candidate biomarkers characterize the exhaled breath of children with SARS-CoV2 infection. To test this hypothesis, we enrolled SARS-CoV-2-infected and -uninfected children admitted to a major pediatric academic medical center and analyzed their breath volatile composition. Targeted volatiles analysis revealed that six volatile organic compounds increased significantly in SARS-CoV-2-infected children. Three aldehydes (octanal, nonanal, and heptanal) drew special attention as candidate biomarkers, because viral infections have previously been shown to induce aldehyde production. Together, these biomarkers demonstrated 100% sensitivity and 66.6% specificity. Our work provides a solid framework upon which to build a future “breathalyzer” test for SARS-CoV-2 infection in children.

Introduction

Novel diagnostic strategies are urgently needed to control the current COVID-19 pandemic, caused by infection with the novel coronavirus SARS-CoV-2. The most common strategy to identify acute SARS-CoV-2 infection relies on viral specific RT-PCR from samples collected from either the upper respiratory tract [e.g. nasopharyngeal (NP) swabs] or lower respiratory tract (e.g. bronchoalveolar lavage). Such testing is uncomfortable and relatively expensive, requires multiple reagents for which supplies are limited (swabs, viral transport media, and RNA extraction kits), relies on specialized laboratory equipment and trained personnel, and has a relatively slow turnaround time (hours to days). Moreover, high false-negative rates have also been reported (1). Simple, rapid diagnostic testing for SARS-CoV-2 infection would have a major impact on control of COVID-19 and on clinical care of symptomatic or exposed individuals, especially in resource-limited settings. In addition, a rapid diagnostic suitable for large-scale screening of children for coronavirus infection may facilitate safe school re-openings.

In children, SARS-CoV-2 infection is less likely to result in hospitalization, severe disease, or death than in adults (2). In addition, evidence suggests that children exhibit a distinct immunological response during coronavirus infection (3, 4). However, any novel diagnostic test for SARS-CoV-2 will require excellent performance characteristics in pediatric populations. Symptomatic or mildly symptomatic pediatric cases can lead to SARS-CoV-2 transmission in the household and community (5). While adults are inconvenienced by social distancing measures to control viral transmission, the educational and social development of children may be irreparably harmed (6). Finally, until global control of COVID-19 is achieved, children will require a disproportionately high frequency of testing, due to the burden of clinically

indistinguishable non-SARS-CoV-2 upper respiratory infections in childhood (up to 12 per year) (7) .

Metabolic changes induced by respiratory viral infections can lead to changes in host odor profiles, suggesting that infection-associated volatile organic compounds (VOCs) may be used to develop noninvasive diagnostics through sensor arrays (e.g. “breathalyzers”) or electronic noses. Compelling evidence from canine biosensors suggests that this may be a promising approach for SARS-CoV-2 diagnosis. Trained dogs recognize SARS-CoV-2 infection in both saliva/tracheal samples (8) and sweat samples (9). Breath metabolite profiling is particularly attractive for SARS-CoV-2 biomarker discovery, as a number of upper and lower respiratory tract infections are known to alter the volatile organic compounds (VOCs) present in breath exhalate. A wide variety of respiratory infections leads to characteristic alterations in breath metabolite profiles, including *Mycobacterium tuberculosis* (10), *Aspergillus* spp., and ventilator-associated pneumonia (11). Viral respiratory pathogens likewise impact volatile production *in vitro* and *in vivo*. For example, infection with either rhinovirus (12) or influenza in cell culture (13) results in reproducible VOC changes in vitro. Similarly, animal studies in an influenza model likewise demonstrate an increase in breath concentrations of acetaldehyde, propanal, and n-propyl acetate (14).

To evaluate whether changes in breath metabolites characterize the exhaled breath of pediatric patients with SARS-CoV2 infection, we analyzed breath metabolite profiles from children with and without SARS-CoV-2 admitted to a major pediatric academic medical center. Through targeted GCxGC-mass spectrometric analysis of 84 breath volatiles, we established the reproducible changes in pediatric breath biomarkers in response to SARS-CoV-2 infection.

Results and Discussion:

In this pilot study, we performed breath metabolic profiling on pediatric patients (n=26) from the Children's Hospital of Philadelphia (CHOP), 11 of whom were positive and 15 of whom were negative for SARS-CoV-2 by nasopharyngeal (NP) RT-PCR. One SARS-CoV-2-infected subject was excluded due to poor quality of breath sampling.

Demographic and clinical characteristics in the SARS-CoV-2-infected versus -uninfected subjects are shown in Table 1. SARS-CoV-2-infected and -uninfected patients were broadly similar with respect to age, sex, and racial/ethnic characteristics. As expected, individuals infected with SARS-CoV-2 were more likely to exhibit either fever (50% vs. 0.0%, p=0.004) or cough (40% vs. 0.0%, p=0.016), compared to subjects uninfected with SARS-CoV-2. Two SARS-CoV-2-positive subjects (25%) lacked symptoms consistent with acute infection (specifically fever, sore throat, cough, or GI symptoms). Two SARS-CoV-2-positive subjects were diagnosed with multisystem inflammatory syndrome in children (MIS-C), thought to be a post-infectious complication of SARS-CoV-2.

For each patient, breath volatiles were captured onto sorbent material and subsequently released by thermal desorption for analysis by two-dimensional gas chromatography and time-of-flight mass spectrometry (GCxGC ToF-MS). Isoprene is one of the most common and abundant human breath VOCs. To establish the quality of breath VOC collection, the abundance of isoprene was compared to the abundance of isoprene in ambient air, which was collected in the same room and at the same time as breath collection. For each subject (SARS-CoV-2-infected and -uninfected), we find that the abundance of isoprene was markedly higher than ambient levels, confirming successful breath VOC collection (Figure S3).

For our targeted metabolite analysis, we selected 84 VOCs that have previously been identified as common human odorants (15) or that have previously been found to be associated with host response to viral infection (12-14, 16, 17). Volcano plots (Figure 1A) were used to identify and visualize breath metabolic features that distinguished SARS-CoV-2-infected from -uninfected individuals, using $p<0.05$ as a threshold for statistical significance. Six candidate breath biomarkers were significantly elevated in the breath of children with SARS-CoV-2 infection: three aldehydes [octanal, nonanal, and heptanal (Figure 2B and Figure S4)], as well as decane, tridecane, and 2-penthyl furan (Figure 2B). All compound identities were confirmed by comparison to pure commercial standards. Analytical characteristics of candidate breath biomarkers can be found in Table S1.

Heat map visualization indicates an overall increase in candidate volatile biomarkers in the breath of children with SARS-CoV-2 infection (Figure 1B and Figure S5), suggesting that SARS-CoV-2 infection alters the overall profile of breath VOCs. Because fever alone can alter metabolic profiles, we evaluated whether any candidate biomarkers correlated with fever. Fever was not associated with significant changes in abundance of any biomarker (Figure S6).

To visualize the discriminatory power of the six candidate SARS-CoV-2 biomarkers, principal components analysis (PCA) and hierarchical clustering on principal components were performed (Figures. 2A and 2B). Both techniques indicate substantial differences in breath volatile composition in SARS-CoV-2-infected children compared to uninfected controls (Figure 2B). For hierarchical clustering, the largest cluster encompasses all SARS-CoV-2-positive samples plus five uninfected controls, and the second cluster is formed by 10 uninfected subjects. No SARS-CoV-2-positive sample was misclassified as uninfected. Overall, these data

suggest that breath-based diagnostics are highly sensitive for identification of SARS-CoV-2 but may lack specificity.

Three candidate biomarkers—octanal, nonanal, and heptanal (Figure 1B)—drew special attention, because viral infections have previously been shown to induce aldehyde production. For example, in a study of influenza A-induced breath volatiles, the breath abundance of acetaldehyde, propanal, and n-propyl acetate increased during infection and decreased by the end of infection (14). More recently, *in vitro* studies have demonstrated a significant increase in the levels of fatty acids in coronavirus-infected cells, using HCoV-229E as a model virus (18). Our studies are also in agreement with recent studies by Ruszkiewicz et al (19), which found that the breath abundance of two aldehydes (octanal and heptanal) were increased in patients diagnosed with COVID-19 compared to those with other acute respiratory illnesses (e.g. COPD and pneumonia). Viral infections outside the respiratory tract may also trigger similar metabolic changes, as fecal volatiles from children with rotavirus gastroenteritis were enriched in 2,3-butanedione, octanal, nonanal, and 2-heptenal, when compared to samples from children without rotavirus infection (20). During acute viral infections, aldehyde production is attributed to cellular oxidative stress leading to accumulation of by-products of oxidation of unsaturated fatty acids (21, 22). As a result, further studies will be required to establish the specificity of these SARS-CoV-2-assoicated biomarkers (23).

This study demonstrates the promise of breath testing for the diagnosis of SARS-CoV-2 in children. We found robust global differences in the breath VOC composition based on infection status (Figures 3A and 3B). Moreover, the cumulative abundance of as a few as 6 biomarkers yields a sensitivity of 100% and specificity of 66.6% (Figure 3B). Untargeted biomarker

discovery studies in a larger population may identify additional combinations of breath compounds with diagnostic utility.

Our study indicates that SARS-CoV-2 leads to distinct and characteristic volatile organic compound changes in children. This finding is additionally supported by canine biodetection studies performed in parallel by Cynthia Otto and colleagues [see co-submitted manuscript by Essler et al], in which scent-trained dogs accurately discriminate urine from children and adults infected with SARS-CoV-2, compared to uninfected subjects. After training on urine samples obtained from same cohort of children reported in this breath analysis study, the overall accuracy rate for correct canine detection of novel samples was 96%. These biosensor data provide additional strong support for the development of odor-based screening diagnostics for SARS-CoV-2.

Given the cost, discomfort, and risks of false-negative results with RT-PCR-based tests (24), breath analysis may provide an inexpensive, non-invasive, rapid, and highly sensitive alternative for population-based screening of children for SARS-CoV-2 infection. In this work, we provide the first report of candidate diagnostic breath biomarkers for pediatric SARS-CoV-2 infection. The majority of SARS-CoV-2-infected subjects enrolled demonstrated mild symptoms of infection and were only incidentally found to be infected due to routine pre-admission screening at our institution. Although these studies will require extensive validation, our work provides a solid framework upon which to build a future diagnostic “breathalzyer” for SARS-CoV-2.

Materials and Methods

Study Approval and Enrollment

Prior to enrollment, the study was approved by the CHOP Human Research Ethics Committee (IRB 20-017503) and by the CHOP Institutional Biosafety Committee (IBC 19-000145), for handling of human samples potentially containing SARS-CoV-2. Breath samples were collected from children (4-18 years of age) hospitalized in the Special Isolation Unit at the Children's Hospital of Philadelphia (CHOP), who had been diagnosed as SARS-CoV-2 positive by nasopharyngeal swab RT-PCR on admission (n=11). Samples from uninfected individuals were obtained from nasopharyngeal RT-PCR negative subjects, enrolled from the Emergency Department Extended Care Unit of the Children's Hospital of Philadelphia (n=15). The viral load of patient nasopharyngeal swab samples was estimated by cycle threshold value (Ct-value) of the N gene, with lower Ct-values indicating a higher viral load. Samples were considered positive if the Ct-value was ≤40, and Ct values of positive test results were obtained (Table 1).

Exclusion criteria for control subjects included current rhinorrhea, cough, or diarrhea, in order to exclude individuals that may have false negative SARS-CoV-2 testing. In addition, subjects were excluded if they required oxygen supplementation within preceding 3 hours of breath sample collection. Samples were not screened for common circulating non-SARS-CoV-2 human coronaviruses.

Breath sample collection

Breath collection was performed as previously described (25, 26). In brief, SARS-CoV-2-infected and -uninfected subjects exhaled through a disposable cardboard mouthpiece connected to a chamber. The chamber was then attached using tubing to a 3-L SamplePro

FlexFilm sample bag (SKC Inc, Pennsylvania) (Figure S1). The volunteers were asked to take a few deep breaths, place the cardboard tube between the lips, and exhale completely. Neither a nose clip nor VOC filter were used. Breath from the bags was transferred to a sorbent tube as previously described (25, 27). Briefly, 1 L of the breath sample was transferred to sorbent tube at 200 mL min^{-1} using an electric pump; so that all tubes had consistently the same volume. Three-bed Universal sorbent tubes containing Tenax, Carbograph, and Carboxen were used (Markes International Limited, UK). For each participant, ambient air samples and breath samples were collected from the same room. Samples were stored at 4°C until the time of analysis. Detailed analytical conditions of breath sample analysis as well as chemical standards used for identification of biomarkers and quality control can be found in Supplemental Materials.

Data processing and Statistical Analyses

Data was acquired and processed using ChromSpace (SepSolve Analytical, UK). All statistical analyses were performed using RStudio v1.3.1073 (PBC, Boston, MA) and GraphPad Prism V.8.4.3 (GraphPad Software, San Diego, CA). The workflow for data processing and statistical analysis is shown in Figure S2. Background from the raw BenchTOF data file was removed using ChromSpace, and the Dynamic Background Compensate (DBC) of 0.2s peak width and noise factor 6.9 for typical GCxGC data was applied. DBC files were then integrated using the following parameters: peak detection deconvolution algorithm with a minimum ion count of 2000, absolute minimum peak area was set at 15,000 counts, absolute minimum peak high was set at 10,000 counts, and no relative threshold was set for either mass height or absolute area. Compounds were given annotations using the NIST v.17 reference library. Deconvoluted peaks were exported into .xls format file. The data were then processed using RStudio to generate integrated signal for every isolated feature. Isolated features included 84 targeted volatiles, as described below.

We targeted volatiles that have been previously associated with respiratory viruses from cell culture, from analysis of *in vitro* airway cells infected with human rhinovirus, *in vivo* breath profile in swine during Influenza A infection, known human body volatiles (12-14, 16, 17, 28), and authors' own unpublished breath VOC library (Table S2). The data included three internal standards (Table S2). Chromatographic data was first normalized using internal standard (1,4-difluorobenzene), and a volatile was retained if it was present in more than 50% of the samples in either group (i.e. infected or uninfected). In total, 50 VOCs were retained and used for further statistical analysis. Unpaired t-test was used to identify metabolites that were significantly different between control groups and SARS-CoV-2 groups, with a p-value of 0.05 established as the threshold for statistical significance. Of note, multiple comparison corrections of metabolomic data can increase type II errors, because metabolites are typically highly correlated and not independent features (29, 30).

Six promising breath biomarkers of SARS-CoV-2 infection were identified through volcano plot of p-value versus fold change abundance (SARS-CoV-2-infected/uninfected).

Hierarchical Clustering on Principal Components (HCPC) was applied to classify samples using 6 COVID-19 biomarkers. The algorithm of the HCPC method, implemented in the FactoMineR package, can be summarized as follow: 1) Compute principal component method; at this step the number of dimensions to be retained in the output is chosen (ncp= 3). 2) Compute hierarchical clustering: it is performed using Ward's criterion on the selected principal components. Ward criterion is used in the hierarchical clustering because it is based on the multidimensional variance (as per principal component analysis). 3) Choose the number of clusters based on the hierarchical tree (n=2).

Author contributions

Amalia Z. Berna. Project coordinator, assisted with ethical approval, recruitment, sample collection, instrumental analysis, data interpretation and writing the manuscript.

Elikplim H. Akaho. Assisted with ethical approval, recruitment, sample collection and reviewing the manuscript

Morgan Congdon. Sample collection, revision of manuscript.

Emilie Korn. Sample collection, revision of manuscript

Samuel Neher. Sample collection, revision of manuscript

Mirna M'Farrej. Assisted with subject recruitment, revision of manuscript

Julianne Burns. Sample collection, clinical coordination with the Special Isolation Unit, revision of manuscript and ethical approval

Audrey R. Odom John. Study conception, study design, assisted with ethical approval, data interpretation, and manuscript writing.

Acknowledgments

A.O.J. is supported by NIH/NIAID R01-AI103280, R21-AI123808, and R21-AI130584, and AOJ is an Investigator in the Pathogenesis of Infectious Diseases (PATH) of the Burroughs Wellcome Fund. We express our gratitude to the children and families of the Children's Hospital of Philadelphia for all their support and participation. We acknowledge the unique efforts of the front-line clinical staff during the breath collection.

References:

1. Wang Y, Kang H, Liu X, and Tong Z. Combination of RT-qPCR testing and clinical features for diagnosis of COVID-19 facilitates management of SARS-CoV-2 outbreak. *Journal of Medical Virology.* 2020;92(6):538-9.
2. Li X, Xu W, Dozier M, He Y, Kirolos A, Theodoratou E, et al. The role of children in transmission of SARS-CoV-2: A rapid review. *J Glob Health.* 2020;10(1):011101.
3. Diorio C, Henrickson SE, Vella LA, McNerney KO, Chase J, Burudpakdee C, et al. Multisystem inflammatory syndrome in children and COVID-19 are distinct presentations of SARS-CoV-2. *The Journal of clinical investigation.* 2020;130(11):5967-75.
4. Weisberg SP, Connors T, Zhu Y, Baldwin M, Lin WH, Wontakal S, et al. Antibody responses to SARS-CoV2 are distinct in children with MIS-C compared to adults with COVID-19. *medRxiv.* 2020.
5. NG S, AC M, A M, and al. e. *MMWR Morb Mortal Wkly Report*. June–July 2020:1457–9.
6. Masonbrink AR, and Hurley E. Advocating for Children During the COVID-19 School Closures. *Pediatrics.* 2020;146(3).
7. Nokso-Koivisto J, Pitkaranta A, Blomqvist S, Jokinen J, Kleemola M, Takala A, et al. Viral etiology of frequently recurring respiratory tract infections in children. *Clin Infect Dis.* 2002;35(5):540-6.
8. Jendrny P, Schulz C, Twele F, Meller S, von Kockritz-Blickwede M, Osterhaus A, et al. Scent dog identification of samples from COVID-19 patients - a pilot study. *Bmc Infect Dis.* 2020;20(1):536.
9. Grandjean D, Sarkis R, Tourtier J-P, Julien-Lecocq C, Benard A, Roger V, et al. Detection dogs as a help in the detection of COVID-19 Can the dog alert on COVID-19 positive persons by sniffing axillary sweat samples ? Proof-of-concept study. *bioRxiv.* 2020:2020.06.03.132134.
10. Phillips M, Basa-Dalay V, Blais J, Bothamley G, Chaturvedi A, Modi KD, et al. Point-of-care breath test for biomarkers of active pulmonary tuberculosis. *Tuberculosis.* 2012;92(4):314-20.
11. Filipiak W, Beer R, Sponring A, Filipiak A, Ager C, Schiefecker A, et al. Breath analysis for in vivo detection of pathogens related to ventilator-associated pneumonia in intensive care patients: a prospective pilot study. *J Breath Res.* 2015;9(1).
12. Schivo M, Aksenov AA, Linderholm AL, McCartney MM, Simmons J, Harper RW, et al. Volatile emanations from in vitro airway cells infected with human rhinovirus. *J Breath Res.* 2014;8(3):037110.
13. Purcaro G, Rees CA, Wieland-Alter WF, Schneider MJ, Wang X, Stefanuto PH, et al. Volatile fingerprinting of human respiratory viruses from cell culture. *J Breath Res.* 2018;12(2):026015.
14. Traxler S, Bischoff AC, Sass R, Trefz P, Gierschner P, Brock B, et al. VOC breath profile in spontaneously breathing awake swine during Influenza A infection. *Sci Rep.* 2018;8(1):14857.
15. de Lacy Costello BD, Amann A, Al-Kateb H, Flynn C, Filipiak W, Khalid T, et al. A review of the volatiles from the healthy human body. *J Breath Res.* 2014;8(1).
16. Aksenov AA, Sandrock CE, Zhao W, Sankaran S, Schivo M, Harper R, et al. Cellular scent of influenza virus infection. *Chembiochem.* 2014;15(7):1040-8.
17. Traxler S, Barkowsky G, Sass R, Klemenz AC, Patenge N, Kreikemeyer B, et al. Volatile scents of influenza A and S. pyogenes (co-)infected cells. *Sci Rep-Uk.* 2019;9.
18. Yan B, Chu H, Yang D, Sze KH, Lai PM, Yuan S, et al. Characterization of the Lipidomic Profile of Human Coronavirus-Infected Cells: Implications for Lipid Metabolism Remodeling upon Coronavirus Replication. *Viruses.* 2019;11(1).
19. Ruszkiewicz DM, Sanders D, O'Brien R, Hempel F, Reed MJ, Riepe AC, et al. Diagnosis of COVID-19 by analysis of breath with gas chromatography-ion mobility spectrometry - a feasibility study. *EClinicalMedicine.*
20. Al-Kateb H. 2012.

21. Schwarz KB. Oxidative stress during viral infection: a review. *Free Radic Biol Med.* 1996;21(5):641-9.
22. Ratcliffe N, Wieczorek T, Drabinska N, Gould O, Osborne A, and De Lacy Costello B. A mechanistic study and review of volatile products from peroxidation of unsaturated fatty acids: an aid to understanding the origins of volatile organic compounds from the human body. *J Breath Res.* 2020;14(3):034001.
23. Ayala A, Munoz MF, and Arguelles S. Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. *Oxid Med Cell Longev.* 2014;2014:360438.
24. Ravi N, Cortade DL, Ng E, and Wang SX. Diagnostics for SARS-CoV-2 detection: A comprehensive review of the FDA-EUA COVID-19 testing landscape. *Biosens Bioelectron.* 2020;165:112454.
25. Berna AZ, Debosch B, Stoll J, and John ARO. Breath Collection from Children for Disease Biomarker Discovery. *Jove-J Vis Exp.* 2019(144).
26. Berna AZ, McCarthy JS, Wang XR, Michie M, Bravo FG, Cassells J, et al. Diurnal variation in expired breath volatiles in malaria-infected and healthy volunteers. *J Breath Res.* 2018;12(4).
27. Berna AZ, McCarthy JS, Wang RX, Saliba KJ, Bravo FG, Cassells J, et al. Analysis of breath specimens for biomarkers of *Plasmodium falciparum* infection. *J Infect Dis.* 2015;212(7):1120-8.
28. de Lacy Costello B, Amann A, Al-Kateb H, Flynn C, Filipiak W, Khalid T, et al. A review of the volatiles from the healthy human body. *J Breath Res.* 2014;8(1):014001.
29. McDonald JH, and Delaware Uo. *Handbook of Biological Statistics.* Sparky House Publishing; 2009.
30. Xi B, Gu H, Baniasadi H, and Raftery D. Statistical analysis and modeling of mass spectrometry-based metabolomics data. *Methods Mol Biol.* 2014;1198:333-53.

Table 1. Demographics. Patient demographics and clinical characteristics

Variables	**SARS-CoV-2 negative (*n* = 15)**	**SARS-CoV-2 positive (*n* = 10)**	***p* value**
Demographic characteristics			
Age (years), median (IQR)	15 (12-16)	11 (8.2-17)	0.15
Female, *n* (%)	9 (60)	6 (60)	>0.99[1]
Black or African-American, *n* (%)	6 (40)	5 (50)	0.69[1]
BMI/Age percentile, median (IQR)	86 (62.5-98)[2]	61.5 (45-84)	0.11
Reported Symptoms, n (%)			
Fever (>38.0°C)	0 (0)	5 (50)	0.004[1]
Cough (new onset or worsening of chronic cough)	0 (0)	4 (40)	0.016[1]
Sore throat	0 (0)	1 (10)	0.40[1]
Headache	0 (0)	1 (10)	0.40[1]
Laboratory			
Cycle threshold values (SARS-CoV-2 RT-PCR), median (IQR)	>40 (negative)	36.74 (31.78-37.63)	---

Data represent median value (interquartile range) or number of patients (%).

[1] Fisher's exact test used for contingency table analysis

[2] Data unavailable for two patients

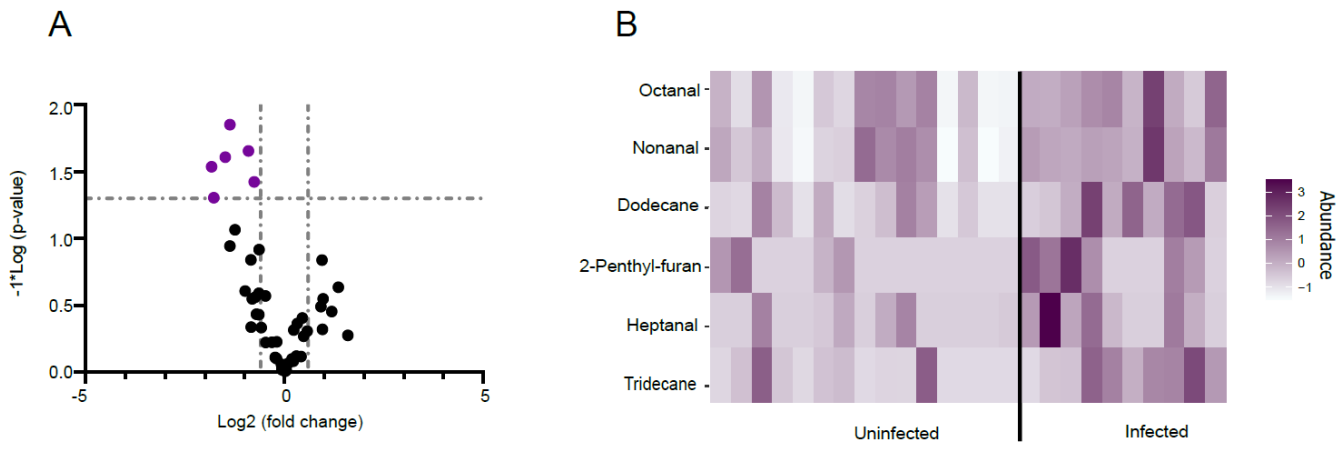

Figure 1: Biomarkers of SARS-CoV-2 in breath. (A) Volcano plot shows − log10 (p-value) on the y-axis versus $\log_2$ (fold change) on the x-axis. Each point represents an individual breath metabolite. In purple, breath metabolites significantly different in SARS-CoV-2-infected subjects compared to uninfected. Identity of these metabolites are shown in heat map. (B) Heat map showing levels of breath biomarkers in uninfected and SARS-CoV-2 positive patients. Abundance is presented as z-scores for visualization purposes.

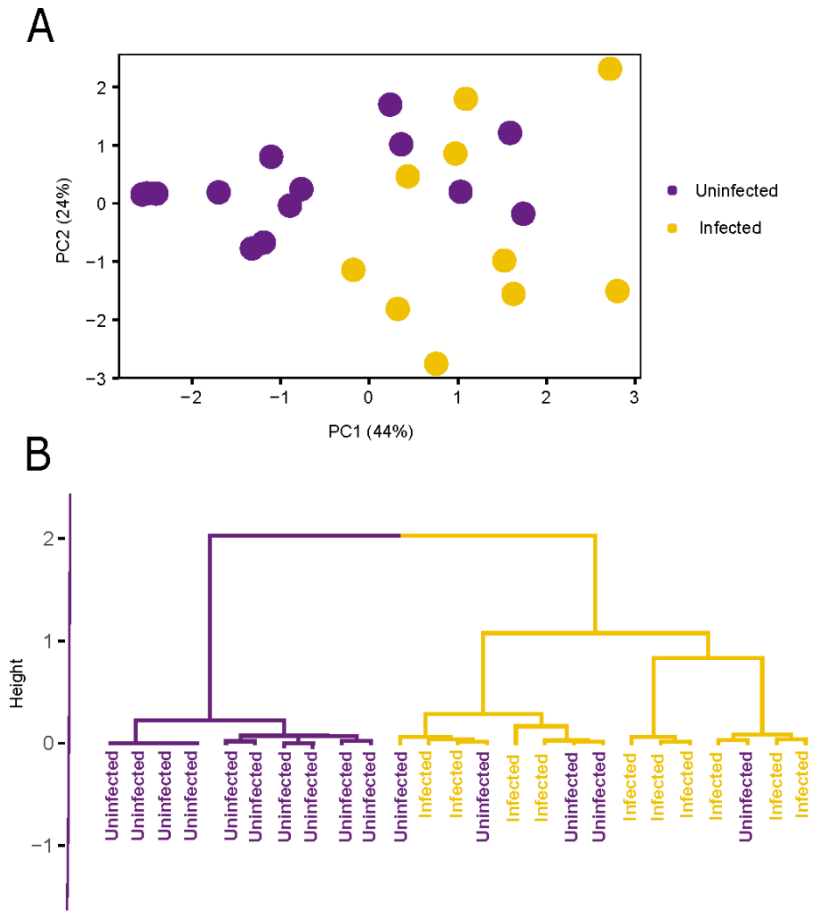

Figure 2. Discrimination power of SARS-CoV-2 breath biomarkers. Principal component analysis (A) performed with six biomarkers from Figure 1. Hierarchical Clustering on Principal Components (B) showing sample identities that comprised the two groups. No SARS-CoV-2 positive samples were misclassified as negative.

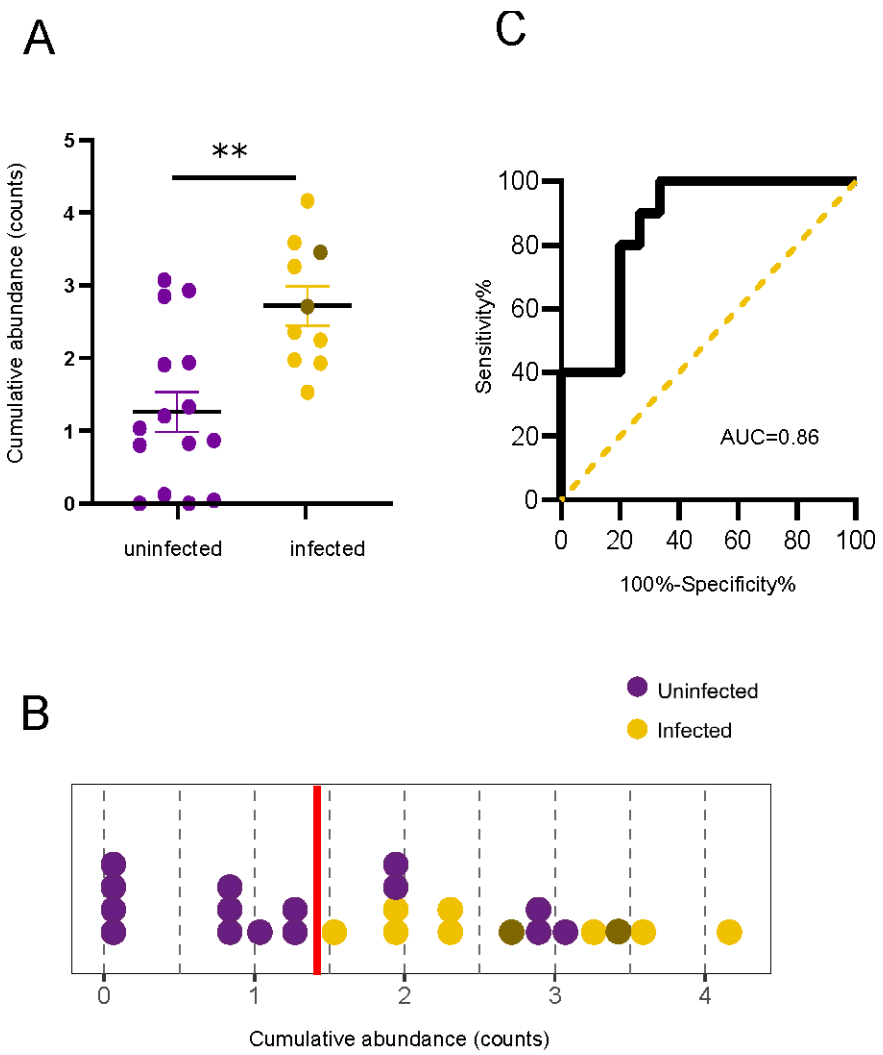

Figure 3. Performance characteristics of cumulative abundance metric for SARS-CoV-2 diagnosis. (A) Cumulative abundance (internal standard normalized) of six candidate biomarkers readily distinguishes breath profiles from children with and without SARS-CoV-2 infection (t-test, p=0.0015)**.** Within the infected group, brown denotes subjects diagnosed with MIS-C. (B) Distribution of cumulative abundance of biomarkers from SARS-CoV-2- infected and uninfected children. Red line, threshold of discrimination between infected and uninfected. Brown, subjects diagnosed with MIS-C. (C) Receiver operator characteristics (ROC) curve for the cumulative abundance of 6 biomarkers. Dotted line indicates expected results if predictive power is no better than random chance. Using threshold, this cumulative abundance metric yields 100% sensitivity and 66.6% specificity.