Supporting information:

Multiplexed Nanomaterial-Based Sensor Array for Detection of COVID-19 in Exhaled Breath

Benjie Shan,†‡ Yoav Y. Broza,§ Wenjuan Li,∥ Yong Wang,∥ Sihan Wu,† Zhengzheng Liu,† Jiong Wang,∥ Shuyu Gui,† Lin Wang,∥ Zhihong Zhang,∥ Wei Liu,∥ Shoubing Zhou,† Wei Jin,† Qianyu Zhang,† Dandan Hu,† Lin Lin,∥‡ Qiujun Zhang,† Wenyu Li,† Jinquan Wang,∥ Hu Liu,∥∗ Yueyin Pan,†∥‡∗ and Hossam Haick,∥∗

† Department of Tumor Biotherapy (5th Ward of the Department of Oncology), Anhui Provincial Cancer Hospital, West District of The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, 230031, Hefei, Anhui Province, People's Republic of China

‡ Department of Oncology, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, 230001, Hefei, Anhui Province, People's Republic of China

§ Department of Chemical Engineering and Russell Berrie Nanotechnology Institute, Technion – Israel Institute of Technology, 3200003, Haifa, Israel

∥ Department of Respiratory Disease, The First Affiliated Hospital of Anhui Medical University, 230022, Hefei, Anhui Province, People's Republic of China

∥ Department of Hematology, The First Affiliated Hospital of Anhui Medical University, 230022, Hefei, Anhui Province, People's Republic of China

∥ Department of Respiratory Disease, Anhui Provincial Cancer Hospital, West District of The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, 230031, Hefei, Anhui Province, People's Republic of China

∥ Department of Respiratory Disease, Anhui Provincial Chest Hospital, 230000, Hefei, Anhui Province, People's Republic of China

∥ Department of Geriatrics, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, 230001, Hefei, Anhui Province, People's Republic of China.

∗ Correspondence to: Hu Liu, e-mail: drliuhu@yahoo.com; Yueyin Pan, e-mail: panyueyin@ustc.edu.cn; Hossam Haick, +972-77-8873087, e-mail: hhossam@technion.ac.il
# Table of Contents

I. Clinical definition ...............................................................................................................................3  
   A. Coronavirus Disease 2019 (COVID-19) ........................................................................................3  
      1. Epidemiological history ...............................................................................................................3  
      2. Clinical manifestations ................................................................................................................3  
   B. Nucleic Acid Detection ...................................................................................................................3  
   C. Serum Antibody ..............................................................................................................................3  
   D. T Lymphocyte Subsets ....................................................................................................................4  
   E. Lung Infections ...............................................................................................................................4  
II. Supplemental Methods ...................................................................................................................4  
   A. Instrument Description and Breath Sampling ..................................................................................4  
   B. Figure S1: schematics of the sensor chip design ...........................................................................4  
   C. Cleaning Step ..................................................................................................................................5  
III. Supplementary Result ....................................................................................................................5  
   A. Table S1 - Demographics, baseline characteristics and laboratory testing of study population* ..............................................................................................................................................5  
   B. Figure S2: Diagnosis of COVID-19 patients based on cumulative breath sample response for all sub groups. .................................................................................................................................................7
I. Clinical definition
   A. Coronavirus Disease 2019 (COVID-19)
      is a respiratory tract infection caused by the recent emergent coronavirus, SARS-CoV-2, that was first
      recognized in Wuhan, China, in December 2019.

      Suspected infection:
      1. Epidemiological history
         a. Within 14 days before the onset of the disease, the patient has a travel or residence history
            in high-risk regions or countries;
         b. Within 14 days before the onset of the disease, the patient has a history of contact with
            those infected with SARS-CoV-2;
         c. Within 14 days before the onset of the disease, the patient had direct contact with patients
            with fever or respiratory symptoms in high-risk regions or countries;
         d. Disease clustering (two or more cases with fever and/or respiratory symptoms occur in such
            places as homes, offices, school classrooms, etc. within two weeks).

      2. Clinical manifestations
         a. The patient has fever and/or respiratory symptoms;
         b. The patient has the following CT imaging features of COVID-19: multiple patchy shadows
            and interstitial changes occur early, particularly at the lung periphery. These develop later
            into multiple ground-glass opacities and infiltrates in both lungs. In severe cases the patient
            may have lung consolidation and rare pleural effusion;
         c. The white blood cell count in the early stage of the disease is normal or decreased, or the
            lymphocyte count decreases over time.

      The patient meets one epidemiological history and two clinical manifestations. The patient has no
      epidemiological history and meets three clinical manifestations.

      Confirmed infection: A person with laboratory confirmation of COVID-19 infection, irrespective
      of clinical signs and symptoms.

   B. Nucleic Acid Detection
      Nucleic acid testing is the preferred method for diagnosing SARS-CoV-2 infection. The testing
      process according to the kit instructions is as follows: Specimens are preprocessed, and the virus is
      lysed to extract nucleic acids. The three specific genes of SARS-CoV-2, namely the Open Reading
      Frame 1a/b (ORF1a/b), nucleocapsid protein (N), and envelope protein (E) genes, are then amplified
      by real-time quantitative PCR. The amplified genes are detected by fluorescence intensity. The criteria
      for positive nucleic acid results are: ORF1a/b gene is positive, and/or N gene/E gene are both positive.
      The combined detection of nucleic acids from multiple types of specimen can improve diagnostic
      accuracy. Among patients with confirmed positive nucleic acid in the respiratory tract, about 30-40%
      have detectable viral nucleic acid in the blood and about 50-60% have detectable viral nucleic acid in
      feces. However, the rate of positive nucleic acid testing in urine samples is quite low. Combined
      testing with specimens from the respiratory tract, feces, blood and other sources is helpful for
      improving diagnostic sensitivity in suspected cases, monitoring treatment efficacy and managing post-
      discharge isolation measures.

   C. Serum Antibody
Specific antibodies are produced after SARS-CoV-2 infection. Methods for determining them include colloidal gold immunochromatography, ELISA, chemiluminescence immunoassay, etc. Positive serum-specific IgM, or specific IgG antibody titer during the recovery phase four or more times higher than during the acute phase, can be used as diagnostic criteria for suspected patients with negative nucleic acid detection. During follow-up monitoring, IgM is detectable 10 days, and IgG 12 days, after symptom onset. The viral load gradually decreases with the increase of serum antibody levels. IgG is the main antibody produced by the second immune response, suggesting a previous infection; IgM is the earliest antibody in the initial humoral immune response, suggesting new infection, and can be used for early diagnosis.

D. **T Lymphocyte Subsets**
These are an important indicator for detecting the cellular immune function, and provide an auxiliary diagnosis for certain diseases (autoimmune diseases, immunodeficiency diseases, malignant tumors, hematological diseases, allergic diseases, etc.), analyzing pathogenesis, observing the efficacy of treatment and monitoring the patient. The prognosis is important. T lymphocyte subsets include CD3, CD4, CD8, etc. CD3 molecules can be expressed on the surface of mature T lymphocytes, while CD4 and CD8 cannot be expressed simultaneously on such surfaces, so mature T lymphocytes can be divided into two subgroups: CD4+ and CD8+ T.

E. **Lung Infections**
Lung diseases caused by various pathogens result in congestion, edema, inflammatory cell infiltration, and exudative lesions. The pathogens are usually bacteria or viruses. High-risk factors for infection include other lung diseases such as cystic fibrosis, COPD, and asthma, diabetes, heart failure, a history of smoking, or a weak immune system. The main clinical manifestations are cough, sputum, fever, and dyspnea. The diagnosis mainly depends on the symptoms, physical examination, chest X-ray, blood tests, and sputum culture. Treatment depends on the underlying cause. Lung infections caused by bacteria are mainly treated with antibiotics and systemic support.

II. **Supplemental Methods**

A. **Instrument Description and Breath Sampling**
The handheld breathalyzer device consists of a number of units; sensors chamber, microfluidics, pumps, valves, GNP gas sensor array (containing eight sensors), breath detection sensors (Nanovation, Israel), auxiliary sensors (humidity, temperature), electronic unit, etc., all integrated into a small case (Figure 2). Communications are via a USB. Measurement involves two main steps: (1) baseline determination and (2) breath sampling. In the first step, the sensors are exposed to the ambient air for 5 s to provide a reference measurement or baseline. In the breath sampling step, the user exhales for at least 4 s directly at the device’s inlet with the device positioned approximately 1 cm from the subject’s mouth. The humidity sensors in the device detect the exhalation and trigger a process by which the end-tidal part of the exhalation is taken. An orifice on the front panel of the device allows air to flow to the sensors without direct contact, and therefore without need for a mouthpiece. The breath sample is held for 10 s, enclosed by two valves. During this time, the sensors are exposed to the sample until they reach stability. Finally, the valves are opened and the breath sample is discarded.

B. **Figure S1: schematics of the sensor chip design**
The device includes eight 1 mm diameter sensors. The electrodes width is 10 microns with a gap of 10 microns. It external size is 150 mm x 153 mm, in addition the device includes two resistors; the first resistor is used for the temperature control and the second one is for temperature measurement.
C. Cleaning Step
To prepare for sampling after each measurement, the device is flushed with air to the clean chamber. After each analysis, the inlet and the surrounding instrument surface are cleaned with a cloth soaked in 70% ethanol. The next measurement is performed after ensuring the ethanol has evaporated.

III. Supplementary Result
A. Table S1 - Demographics, baseline characteristics and laboratory testing of study population*

| Characteristics                     | COVID-19 group (N = 49) | Control group (N = 58) | Lung infections/disease group (N = 33) |
|-------------------------------------|-------------------------|------------------------|---------------------------------------|
| Median age (IQR) – yr               | 58 (51.5-64.5)          | 53 (45-59)             | 64 (53-73)                            |
| Female sex – no. (%)                | 28 (57)                 | 26 (45)                | 15 (45)                               |
| Active Smoking – no. (%)            | 13 (27)                 | 19 (33)                | 9 (27)                                |

Coexisting conditions – no. (%)

|                         | COVID-19 group          | Control group          | Lung infections/disease group        |
|-------------------------|-------------------------|------------------------|--------------------------------------|
| Hypertension            | 18 (37)                 | 8 (12)                 | 8 (24)                               |
| Diabetes                | 5 (10)                  | 2 (3)                  | 4 (12)                               |
| Heart disease           | 1 (2)                   | -                      | 2 (6)                                |
| Malignancy              | -                       | -                      | 8 (24)                               |
| Bacterial/fungal infection | -                      | -                      | 6 (18)                               |
| Other                   | 1 (2)                   | -                      | -                                    |
| Means of positive determination COVID-19 – no. (%) |
| Test Description                                      | Value       |
|------------------------------------------------------|-------------|
| CT angiogram                                         | 33 (67)     |
| RT-PCR, positive (throat swab) – no. (%)             |             |
| nCovORFlab                                           | 29 (59)     |
| 2019nCOV-N                                           | 28 (57)     |
| 2019nCov antibody tests, median (IQR) – AU/ml        |             |
| 2019nCov IgM                                         | 42.85 (13.65-69.24) |
| 2019nCov IgG                                         | 164.68 (133.93-182.55) |
| Count of irregularities in immune system component levels – no. (%) |             |
| CD3 (58.17-84.22%)                                   |             |
| Count of high                                       | 7 (14)      |
| Count of low                                       | 3 (6)       |
| CD4 (25.34-51.37%)                                   |             |
| Count of high                                       | 13 (27)     |
| Count of low                                       | 1 (2)       |
| CD8 (14.23-38.95%)                                   |             |
| Count of high                                       | 4 (8)       |
| Count of low                                       | 4 (8)       |
| CD4/CD8 (0.41-2.72)                                  |             |
| Count of high                                       | 12 (25)     |
| Count of low                                       | -           |

* Patients positive for COVID-19 were identified prior to breath sampling. Only the COVID-19 group was accessed for molecular and immunological characteristics. IQR - interquartile range, CT - computed tomography, RT-PCR - real-time reverse transcriptase polymerase chain reaction.
B. **Figure S2: Diagnosis of COVID-19 patients based on cumulative breath sample response for all sub groups.**

Data classification of cumulative sensor responses to breath samples as represented by the canonical variable of the discriminant analysis, based on the main DFA model of patients vs controls. Box plots of the first canonical score of the data set of the COVID-19 (41 participants), healthy controls (57 participants), other lung infection/condition controls (32 participants), second sampling COVID-19 cured (18 participants), and second sampling COVID-19 uncured (3 participants). Statistically significant P values are for the comparisons of each set of the subgroups marked by connecting lines and asterisk. The horizontal line in the boxes represents the median, the cross represents the mean, and the bottom and top of the boxes represent the 25th and 75th percentiles, respectively. The I bars represent the upper 90 and the lower 10 percentiles, and the square-dots outliers. *P<.001