ChemComm

COMMUNICATION

View Article Online
View Journal

Cite this: DOI: 10.1039/d1cc06384g

Received 12th November 2021,
Accepted 24th March 2022

DOI: 10.1039/d1cc06384g

rsc.li/chemcomm

Breath odor-based individual authentication by an artificial olfactory sensor system and machine learning†

Chaiyanut Jirayupat,[ab] Kazuki Nagashima, *[ac] Takuro Hosomi, [ac] Tsunaki Takahashi, [ac] Benjarong Samransuksamer,[a] Yosuke Hanai,[d] Atsuo Nakao,[d] Masaya Nakatani,[d] Jiangyang Liu,[a] Guozhu Zhang, [a] Wataru Tanaka,[a] Masaki Kanai,[e] Takao Yasui, [cf] Yoshinobu Baba[f] and Takeshi Yanagida*[abe]

Breath odor sensing-based individual authentication was conducted for the first time using an artificial olfactory sensor system. Using a 16-channel chemiresistive sensor array and machine learning, a mean accuracy of >97% was successfully achieved. The impact of the number of sensors on the accuracy and reproducibility was also demonstrated.

Biometric authentication is a convenient and secure individual authentication method in the information technology (IT) field. Its application range covers not only immigration control at airports but also access control of banking, personal computers (PCs)/mobile phones and emerging intelligent vehicles (IVs).[1] To date, various techniques have been developed for biometric authentication, including fingerprint/palmprint verification,[2] iris/retina recognition,[3] facial recognition,[4] hand and finger geometry,[5] voice biometry,[6] finger vein recognition[7] and ear acoustic authentication.[8] All these techniques solely rely on physical information and thus have the risks of being unusable by information alteration due to injury or being compromised by malicious information theft.

Human scent analysis/sensing is a new class of biometric authentication techniques using chemical information.[9–15] Since human scents such as exhaled breath and percutaneous gas have a strong genetic basis,[11,16,17] their chemical composition profiles are inherently different among individuals and therefore can potentially be utilized for individual authentication with low risks of information alteration/theft. Previously, human scent analysis/sensing-based biometric authentication has been conceptualized and attempted mainly *via* percutaneous gas.[9–15] For example, Penn *et al.* analyzed the chemical component profiles of sweat odor from 197 adults using gas chromatography-mass spectrometry (GC-MS) and identified 44 individual specific volatile organic compounds (VOCs).[10] Zheng *et al.* performed skin odor sensing using an artificial olfactory sensor system termed as an electronic nose (e-nose) and classified the sensing data with 91.67% accuracy by machine learning.[13] Despite these achievements, percutaneous gas sensing-based individual authentication has limitations in its performance because the concentrations of VOCs in percutaneous gas are usually lower (ppt to several tens of ppb; ppt: parts per trillion, ppb: parts per billion) than the detection limit of conventional chemical sensors and therefore the detectable number of VOC species is restricted.[18] Exhaled breath is known to have thousands of VOCs and their concentrations are about three orders of magnitude higher than those of percutaneous gas (ppb to several ppm; ppm: parts per million).[18] In this regard, breath odor sensing has great potential to detect a larger number of human-related VOC species and achieve a higher performance in individual authentication compared with that of percutaneous gas sensing. However, breath odor sensing has been mainly directed for pathology/disease diagnosis (*e.g.* cancer, diabetes, COVID-19),[19] and to the best of our knowledge, the feasibility of breath odor sensing-based individual authentication has not been demonstrated so far.

In this study, we demonstrate a primary study for breath odor sensing-based individual authentication using an artificial olfactory sensor system (the workflow is shown in Fig. 1 and the experimental details are shown in Table S1, ESI†). In order to investigate the potential usage of breath odor for individual authentication, we first performed GC-MS

[a] *Department of Applied Chemistry, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. E-mail: kazu-n@g.ecc.u-tokyo.ac.jp, yanagida@g.ecc.u-tokyo.ac.jp*
[b] *Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, 6-1 Kasuga-Koen, Kasuga, Fukuoka, 816-8580, Japan*
[c] *PRESTO, Japan Science and Technology Agency, 4-1-8, Honcho, Kawaguchi-Shi, Saitama 332-0012, Japan*
[d] *Panasonic Corporation, Industry Company, Sensing Solutions Development Center, Kadoma 1006, Kadoma, Osaka 571-8506, Japan*
[e] *Institute for Materials Chemistry and Engineering, Kyushu University, 6-1 Kasuga-Koen, Kasuga, Fukuoka 816-8580, Japan*
[f] *Department of Biomolecular Engineering, Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan*

† Electronic supplementary information (ESI) available: Detailed methods and data, Fig. S1–S8 and Tables S1–S5. See https://doi.org/10.1039/d1cc06384g

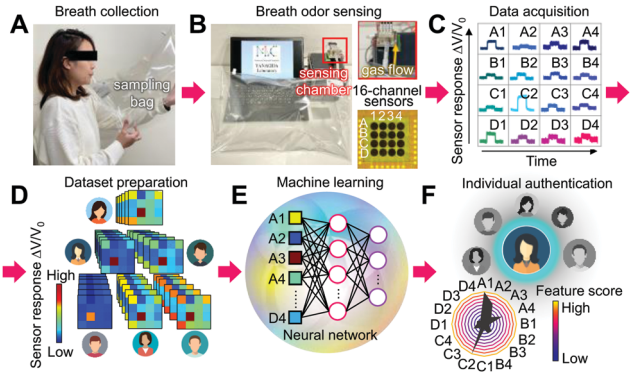

Fig. 1 Graphical workflow of breath odor sensing-based individual authentication. (A) Breath odor sample collection using a gas sampling bag. (B) Breath odor sensing measurements using a 16-channel sensor array. (C) Acquisition of sensor responses. (D) Dataset preparation for machine learning. (E) Machine learning with a neural network algorithm. (F) Individual authentication and evaluation of the feature profiles of the sensors.

measurements and analyzed the individual-specific molecular fragments. For the analysis, two-dimensional (2D) MS maps (*m*/*z* *vs.* retention time) were created and processed using the recently developed data analysis program NPFimg,[20] which combines image processing and machine learning. Fig. 2A–C show the 2D MS maps of the breath odor samples collected from 3 persons (3 males). For visibility, the 2D MS maps are shown in the restricted range (full-range 2D MS maps are shown in Fig. S1, ESI†). Numerous molecular fragment signals are seen in the maps and many of them were common among the 3 tested persons. By learning the datasets of the 2D MS maps, we succeeded in the individual authentication of the 3 persons with 100% accuracy. Fig. 2D–F show the 2D feature score maps of the molecular fragments that contributed to discriminate the individuals from the other two persons. Contrary to the 2D MS maps, the feature score maps were significantly different between the 3 tested persons. Note that the influence of exogenous compounds originating from the diets and the tested environment was negligible because the breath odor samples were collected in the same environment from the persons who fasted for 6 h. We identified the individual-specific marker compounds, *e.g.* benzophenone, decanal, octane, tetradecane, undecane, which were consistently seen in the previous study of sweat odor-based individual authentication (details in Table S2, ESI†).[10,12,15] Thus, these results imply that each person has a unique breath print derived from endogenous compounds as well as indicate the potential feasibility of breath odor-based individual authentication.

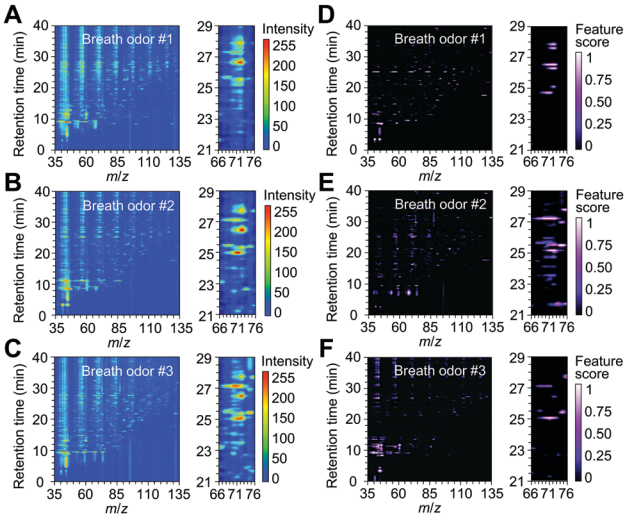

Fig. 2 (A–C) 2D MS maps and (D–F) 2D feature score maps of 3 tested persons (3 males) in the wide-range view (left) and narrow-range view (right). The 2D feature maps were obtained in comparison with the other two breath odor samples.

We then examined individual authentication *via* breath odor sensing. The breath odor samples were first collected using a gas sampling bag (Fig. 1A). The collected breath odor sample was then passed through a sensing chamber with a 16-channel chemiresistive sensor array and breath odor sensing was performed (Fig. 1B). The sensing materials used for the 16-channel sensor array, which were developed for this study, are listed in Table S3 (ESI†). The sensor responses were acquired from the sensing curves of the 16-channel sensors (Fig. 1C) and used as the dataset for machine learning (Fig. 1D). We employed a neural network algorithm for machine learning (Fig. 1E) and demonstrated individual authentication together with feature profile evaluation of the used sensors (Fig. 1F). We tested 6 persons (3 males, 3 females, ages 23–40) of various nationalities (Thai, Chinese and Japanese), as summarized in Table 1. Fig. 3A shows the five successive sensing curves obtained from the 16-channel sensor array in the breath odor sensing of subject $V^{\#1}$. The sensing characteristics such as the maximum sensor response, the initial sensing curve and the recovery curve were different between the sensors. These tendencies were also seen for the other subjects ($V^{\#2}$–$V^{\#6}$, Fig. S2–S6, ESI†), while the sensing characteristics of each sensor strongly depended on the tested person. Fig. 3B shows the heatmaps of the sensor responses of the 16-channel sensor array for the 6 tested persons. The heatmaps are clearly different between the subjects. These results are consistent with those of the GC-MS measurements and therefore anticipate the feasibility of breath odor sensing-based individual authentication.

Fig. 4A shows the box-and-whisker plot of the accuracy of individual authentication for 6 persons calculated by machine learning. The data are displayed as a function of the number of used sensors and the used sensors are arranged in the

Table 1 The details of the tested subjects for breath odor sensing-based individual authentication

Subjects	Nationality	Age	Sex
$V^{\#1}$	Thai	23	Female
$V^{\#2}$	Thai	25	Male
$V^{\#3}$	Chinese	26	Female
$V^{\#4}$	Japanese	28	Male
$V^{\#5}$	Japanese	35	Male
$V^{\#6}$	Japanese	40	Female

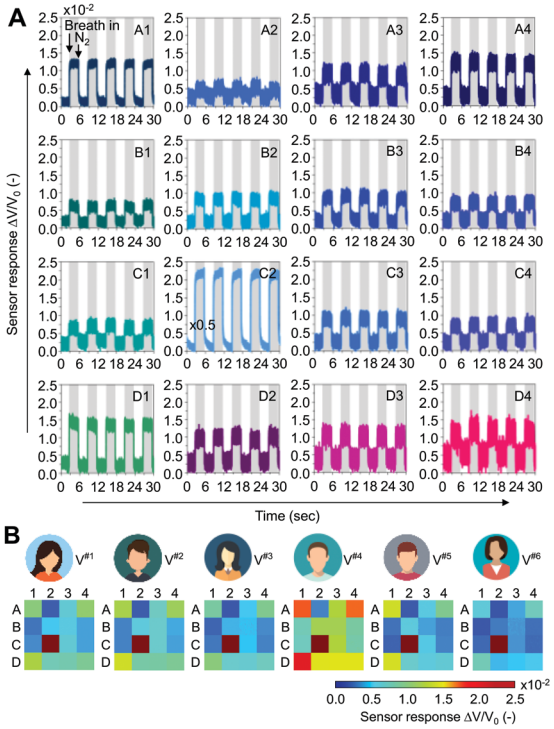

Fig. 3 (A) Sensing curves of the 16-channel sensor array for the breath odor sensing of subject $V^{\#1}$ after the baseline corrections. (B) Heatmaps of the sensor responses of the 16-channel sensor array for the breath odor sensing of each tested person (subject $V^{\#1}$–$V^{\#6}$).

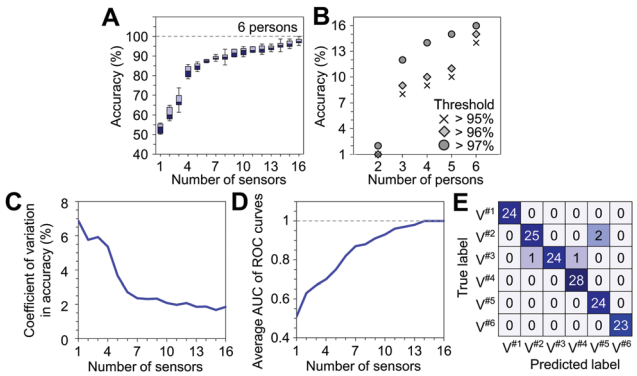

Fig. 4 (A) Accuracy of breath odor sensing-based individual authentication for 6 persons as a function of the number of used sensors. (B) Relationship between the number of persons and the number of required sensors with various thresholds in accuracy (>95%, >96% and >97%). (C) Coefficient of variation in accuracy as a function of the number of used sensors. (D) Averaged AUC of ROC curves as a function of the number of used sensors. (E) Confusion matrix for the breath odor sensing-based individual authentication for 6 persons.

descending order of the amplitude of sensor response to maximize the performance of the data analysis (Table S4, ESI†). The mean accuracies when using a single sensor were 96.9% 84.1%, 80.1%, 68.5% and 54.3% for the individual authentication of 2 persons, 3 persons, 4 persons, 5 persons and 6 persons, respectively. The results indicate that individual authentication tends to be difficult when the number of tested subjects increases. On the other hand, the accuracy of individual authentication was significantly improved when increasing the number of used sensors. The mean accuracy for discriminating 6 persons successfully reached 97.8% using 16 sensors. The relationship between the number of subjects and the number of required sensors for individual authentication is displayed in Fig. 4B. The results indicate that a larger number of sensors are needed to discriminate complex odors, which is consistent with the claim in a recent review paper reported by Lee *et al.*[21] In other words, further discrimination of breath odor would be possible by increasing the number of used sensors. We next evaluated the reliability of the above breath odor sensing results. Fig. 4C and D show the coefficient of variation (CV) values for the accuracy of individual authentication and the averaged area under the curve (AUC) of the receiver operating characteristic (ROC) curve for the classifiers, which are presented as a function of the number of used sensors. The CV values in the accuracy significantly decreased and the averaged AUC of the ROC curves increased as the number of used sensors increased. This shows that both the reproducibility of individual authentication and the reliability of classifiers can be improved by using a larger number of sensors. Furthermore, we found that our sensor was capable of electrically detecting the marker compound at the concentration range found in breath odor (Fig. S7, ESI†). All the above results highlight the feasibility of breath odor sensing-based individual authentication and the impact of the number of integrated sensors on the performance of individual authentication.

Herein, we discuss what critically determined the performance of the breath odor sensing-based individual authentication presented above. Fig. 4E shows the confusion matrix for the individual authentication of 6 persons. While slight false identifications occurred, the errors were randomly distributed, and their pattern was different in the analytical batch. This result indicates that gender, age and nationality did not significantly affect the observed false identifications. Fig. 5 shows the feature score profiles of the used sensors for each tested subject. The data indicate that all sensors contributed to individual authentication, and the profiles were significantly different between the 6 tested persons. These results reasonably

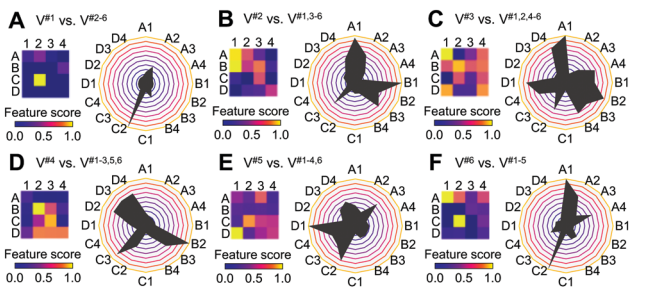

Fig. 5 Feature score patterns of 16-channel sensor array (heatmaps and radar charts) for (A) $V^{\#1}$ *vs.* $V^{\#2-\#6}$, (B) $V^{\#2}$ *vs.* $V^{\#1,\#3-\#6}$, (C) $V^{\#3}$ *vs.* $V^{\#1,\#2,\#4-\#6}$, (D) $V^{\#4}$ *vs.* $V^{\#1-\#3,\#5,\#6}$, (E) $V^{\#5}$ *vs.* $V^{\#1-\#4,\#6}$, (F) $V^{\#6}$ *vs.* $V^{\#1-\#5}$, respectively.

explain why individual authentication was successfully performed. We found that the accuracy was not degraded even after increasing the number of subjects to 20 persons, as shown in Fig. S8 and Table S5 (ESI†). This suggests that the false identification in our study might be caused by the fluctuation/instability of sensor responses and the performance of individual authentication would be better by improving the robustness of the sensing system/material.[22]

In conclusion, we have demonstrated a primary study of breath odor sensing-based individual authentication using an artificial olfactory sensor system. The breath odor samples were tested by a 16-channel chemiresistive sensor array and the acquired sensor responses were analyzed by machine learning with a neural network algorithm. A mean accuracy of >97% was successfully achieved for the individual authentication of up to 20 persons. We found that the accuracy and reproducibility significantly improved by increasing the number of used sensors. While breath odor sensing-based individual authentication was demonstrated for the fasted subjects in this study, it still remains a challenging issue to demonstrate its feasibility under the interference of disease-related metabolites and exogenous compounds originating from the diets and the tested environment towards practical application.[23] The barrier must be overcome by utilizing a larger number of sensors and extracting a larger number of features from the sensing curves. We believe that our findings in this study provide an important foundation for breath odor sensing-based biometrics.

This work was supported by KAKENHI (No. JP18KK0112, JP18H05243, JP20H02208, JP20F20048) and PRESTO Program (No. JPMJPR19J7) of Japan Science and Technology Corporation (JST). K. N. acknowledges JACI. T. Y. was supported by the JST CREST Program (No. JPMJCR19I2) and the Cooperative Research Programs of "Dynamic Alliance for Open Innovation Bridging Human, Environment and Materials in Network Joint Research Center for Materials and Devices", "Network Joint Research Center for Materials and Devices". This study was carried out with the approval of the Research Ethics Committee in Nagoya University (No. 18–32).

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 A. K. Jain, *Nature*, 2007, **449**, 38–40; W. Yang, S. Wang, N. M. Sahri, N. M. Karie, M. Ahmed and C. Valli, *Sensors*, 2021, **21**, 6163; J. J. Lozoya-Santos, V. Sepúlveda-Arróniz, J. C. Tudon-Martinez and R. A. Ramirez-Mendoza, *Cogn. Syst. Res.*, 2019, **55**, 175–191; A. A. Elngar and M. Kayed, *Open Comput. Sci. J.*, 2020, **10**, 17–29.
2 L. Hong, Y. Wan and A. Jain, *IEEE Trans. PAMI*, 1998, **20**, 777–789; W. Jia, D.-S. Huang and D. Zhang, *Pattern Recognit.*, 2008, **41**, 1504–1513.
3 J. Daugman, *IEEE Trans. CSVT*, 2004, **14**, 21–30; H. Borgen, P. Bours and S. D. Wolthusen, *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2008, pp. 1056–1062.
4 A. Mian, M. Bennamoun and R. Owens, *IEEE Trans. PAMI*, 2007, **29**, 1927–1943.
5 A. Kumar and Y. Zhou, *IEEE Trans. Image Process.*, 2012, **21**, 2228–2244.
6 J. Yang, J. Chen, Y. Su, Q. Jing, Z. Li, F. Yi, X. Wen, Z. Wang and Z. L. Wang, *Adv. Mater.*, 2015, **27**, 1316–1326.
7 N. Miura, A. Nagasaka and T. Miyatake, *Mach. Vis. Appl.*, 2004, **15**, 194–203.
8 T. Arakawa, T. Koshinaka, S. Yano, H. Irisawa, R. Miyahara and H. Imaoka, *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2016, pp. 1–4.
9 C. Wongchoosuk, M. Lutz and T. Kerdcharoen, *Sensors*, 2009, **9**, 7234–7249; I. Rodriguez-Lujan, G. Bailador, C. Sanchez-Avila, A. Herrero and G. Vidal-de-Miguel, *Knowl. Based Syst.*, 2013, **52**, 279–289; P. Inbavalli and G. Nandhini, *Int. J. Comput. Sci. Inf. Technol. Adv. Res.*, 2014, **5**, 6270–6274; S. K. Jha, *Rev. Anal. Chem.*, 2017, **36**, 20160028; M. C. Prieto-Blanco, S. P. Barba, Y. Moliner-Martínez and P. Campíns-Falcó, *J. Chromatogr. A*, 2019, **1596**, 241–249.
10 D. J. Penn, E. Oberzaucher, K. Grammer, G. Fischer, H. A. Soini, D. Wiesler, M. V. Novotny, S. J. Dixon, Y. Xu and R. G. Brereton, *J. R. Soc., Interface*, 2007, **4**, 331–340.
11 A. B. Holbert, H. P. Whitelam, L. J. Sooter, L. A. Hornak and J. M. Dawson, *Netw. Model. Anal. Health Inform. Bioinform.*, 2015, **4**, 22.
12 S. K. Jha and K. Hayashi, *Int. J. Mass Spectrom.*, 2017, **415**, 92–102.
13 Y. Zheng, H. Li, W. Shen and J. Jian, *Sens. Actuators, A*, 2019, **285**, 395–405.
14 I. A. Sabilla, Z. A. Cahyaningtyas, R. Sarno, A. Al Fauzi, D. R. Wijaya and R. Gunawan, *2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, 2021, pp. 109–115.
15 P. Kanakam and A. S. N. Chakravarthy, *Soft Comput.*, 2021, **25**, 13015–13025.
16 P. Lenochova and J. Havlicek, Human Body Odour Individuality, in *Chemical Signals in Vertebrates 11*, ed. J. L. Hurst, R. J. Beynon, S. C. Roberts and T. D. Wyatt, Springer, New York, NY, 2008.
17 A. Pol, G. H. Renkema, A. Tangerman, E. G. Winkel, U. F. Engelke, A. P. M. de Brouwer, K. C. Lloyd, R. S. Araiza, L. van der Heuvel, H. Omran, H. Olbrich, M. O. Elberink, C. Gilissen, R. J. Rodenburg, J. O. Sass, K. O. Schwab, H. Schäfer, H. Venselaar, J. S. Sequeira, H. J. M. Op den Camp and R. A. Wevers, *Nat. Genet.*, 2018, **50**, 120–129.
18 K. Toma, S. Suzuki, T. Arakawa, Y. Iwasaki and K. Mitsubayashi, *Sci. Rep.*, 2021, **11**, 10415.
19 J.-W. Yoon and J.-H. Lee, *Lab Chip*, 2017, **17**, 3537–3557; K. Inada, H. Kojima, Y. Cho-Isoda, R. Tamura, G. Imamura, K. Minami, T. Nemoto and G. Yoshikawa, *Sensors*, 2021, **21**, 4742; M. Leja, J. M. Kortelainen, I. Polaka, E. Turppa, J. Mitrovics, M. Padilla, P. Mochalski, G. Shuster, R. Pohle, D. Kashanin, R. Klemm, V. Ikonen, L. Mezmale, Y. Y. Broza, G. Shani and H. Haick, *Cancer*, 2021, **127**, 1286–1292; G. Konvalina and H. Haick, *Acc. Chem. Res.*, 2014, **47**, 66–76; J. Shin, S.-J. Choi, I. Lee, D.-Y. Youn, C. O. Park, J.-H. Lee, H. L. Tuller and I.-D. Kim, *Adv. Funct. Mater.*, 2013, **23**, 2357–2367; R. Xing, L. Xu, J. Song, C. Zhou, Q. Li, D. Liu and H. W. Song, *Sci. Rep.*, 2015, **5**, 10717; B. Shan, Y. Y. Broza, W. Li, Y. Wang, S. Wu, Z. Liu, J. Wang, S. Gui, L. Wang, Z. Zhang, W. Liu, S. Zhou, W. Jin, Q. Zhang, D. Hu, L. Lin, Q. Zhang, W. Li, J. Wang, H. Liu, Y. Pan and H. Haick, *ACS Nano*, 2020, **14**(9), 12125–12132.
20 C. Jirayupat, K. Nagashima, T. Hosomi, T. Takahashi, W. Tanaka, B. Samransuksamer, G. Zhang, J. Liu, M. Kanai and T. Yanagida, *Anal. Chem.*, 2021, **93**, 14708–14715.
21 S.-Y. Jeong, J.-S. Kim and J.-H. Lee, *Adv. Mater.*, 2020, **32**, 2002075.
22 W. Li, K. Nagashima, T. Hosomi, C. Wang, Y. Hanai, A. Nakao, A. Shunori, J. Liu, G. Zhang, T. Takahashi, W. Tanaka, M. Kanai and T. Yanagida, *ACS Sens.*, 2021, **7**, 151–158.
23 J. E. Belizário, J. Faintuch and M. G. Malpartida, *Front. Cell. Infect. Microbiol.*, 2021, **10**, 564194; J. D. Pleil, M. A. Stiegel and T. H. Risby, *J. Breath Res.*, 2013, **7**, 017107.

