Article

Sniff Species: SURMOF-Based Sensor Array Discriminates Aromatic Plants beyond the Genus Level

Salih Okur [1],* , **Chun Li [1]**, **Zejun Zhang [1]**, **Sahi Vaidurya Pratap [2]**, **Mohammed Sarheed [2]**, **Adnan Kanbar [2]**, **Leonard Franke [3]**, **Felix Geislhöringer [3]**, **Lars Heinke [1]**, **Uli Lemmer [3,4]**, **Peter Nick [2]** and **Christof Wöll [1]**

[1] Institute of Functional Interfaces (IFG), Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344 Karlsruhe, Germany; chun.li@partner.kit.edu (C.L.); zejun.zhang@partner.kit.edu (Z.Z.); lars.heinke@kit.edu (L.H.); christof.woell@kit.edu (C.W.)

[2] Molecular Cell Biology, Botanical Institute, Karlsruhe Institute of Technology, Fritz-Haber-Weg, 76131 Karlsruhe, Germany; sahi.vaidurya@kit.edu (S.V.P.); mu.sarheed@gmail.com (M.S.); adnan.kanbar@kit.edu (A.K.); peter.nick@kit.edu (P.N.)

[3] Light Technology Institute, Karlsruhe Institute of Technology, Engesserstraße 13, 76131 Karlsruhe, Germany; leonard.franke@kit.edu (L.F.); felix.geislhoeringer@kit.edu (F.G.); uli.lemmer@kit.edu (U.L.)

[4] Institute of Microstructure Technology, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344 Karlsruhe, Germany

* Correspondence: salih.okur2@kit.edu; Tel.: +49-721-608-26078

Abstract: Lamiaceae belong to the species-richest family of flowering plants and harbor many species that are used as herbs or in medicinal applications such as basils or mints. The evolution of this group has been driven by chemical speciation, mainly volatile organic compounds (VOCs). The commercial use of these plants is characterized by adulteration and surrogation to a large extent. Authenticating and discerning this species is thus relevant for consumer safety but usually requires cumbersome analytics, such as gas chromatography, often coupled with mass spectroscopy. Here, we demonstrate that quartz-crystal microbalance (QCM)-based electronic noses provide a very cost-efficient alternative, allowing for fast, automated discrimination of scents emitted from the leaves of different plants. To explore the range of this strategy, we used leaf material from four genera of Lamiaceae along with lemongrass, which is similarly scented but from an unrelated outgroup. To differentiate the scents from different plants unambiguously, the output of the six different SURMOF/QCM sensors was analyzed using machine learning (ML) methods together with a thorough statistical analysis. The exposure and purging of data sets (four cycles) obtained from a QCM-based, low-cost homemade portable e-Nose were analyzed using a linear discriminant analysis (LDA) classification model. Prediction accuracy with repeated test measurements reached values of up to 0%. We show that it is possible not only to discern and identify plants at the genus level but also to discriminate closely related sister clades within a genus (basil), demonstrating that an e-Nose is a powerful device that can safeguard consumer safety against dangers posed by globalized trade.

Keywords: basil; mint; plant volatiles; electronic nose; principal component analysis; linear discriminant analysis; k-nearest-neighbors analysis

1. Introduction

Plants have developed subtle mechanisms to defend themselves against biotic and abiotic stress factors. One of the ways that plants have evolved to protect themselves is by producing volatile organic compounds (VOCs) [1]. These essential oils, often with monoterpenes as primary components, accumulate in different organs such as leaves, bark, wood, roots, flowers, fruit, specialized glands, or in lysogenic or schizogenic oil ducts [2]. These VOCs are the basis for the human use of aromatic plants, both as spices and for medicinal applications. The multitude of VOC profiles has shaped entire cultures, cuisines, and medical traditions.

One of the most prodigious plant families in this context are the Lamiaceae. With more than 7000 species belonging to more than 200 genera, they are taxonomically the most challenging and diverse group of flowering plants [3]. They secrete complex bouquets of VOCs from their glandular hairs and scales that are often specific for a given species. They also interact with pollinator insects, and this might have been one of the drivers for the immense complexity of this family. Sometimes different chemotypes exist even within a species. Commercially relevant plants, such as mints or basils, belong to this group and are often part of novel food trends fueled by their reputation as so-called superfoods [4].

Because of interest in holistic approaches to health, there is a trend toward the use of supplements and plant-based products that have their roots in traditional medical systems such as Ayurveda or traditional Chinese medicine. With the growing popularity of Ayurveda in Europe, products containing *Ocimum tenuiflorum* L. are readily available in supermarkets. *O. tenuiflorum* (holy basil or Tulsi) has been used for treating ailments such as joint pain, headache, cold, fever, and insect bites [5–9]. In addition, holy basil has been recommended to relieve stress [10] and reduce the effects of diabetes mellitus [11]. Due to the benefits attributed to holy basil, its market in the West is increasing [12], which accentuates the problem of authentication and identification of commercial products that are declared to contain Tulsi [13]. The genus *Ocimum* is composed of many species, several of which are commonly traded. However, each species is endowed with a unique chemical profile that is mostly genetically determined [12]. To authenticate *O. tenuiflorum* by microscopic diagnosis is possible [5] although limited, especially in commercial products that are often processed [14]. In addition, it is possible to discriminate true *O. tenuiflorum* from other basils on the basis of DNA barcoding [12], a rather expensive and time-consuming process. An alternative would be to detect different odorous content since the spectrum of VOCs emitted by *Ocimum* species is unique [15]. In fact, a trained human nose can distinguish *O. tenuiflorum* from other basils due to the emission of specific patterns of volatile phenylpropanoids [16].

The unique VOC profile of a particular plant species provides a means to identify it in commercial products. However, the chemical analysis necessary to identify such gaseous compounds is time consuming and costly, requiring gas chromatography (GC). Since a human nose can discriminate among species, sensor arrays (also referred to as electronic noses) might offer a cost-efficient, convenient, and fast alternative [17,18]. For several decades, e-Noses with different sensing materials have been successfully used: peptides as biosensors [19], molecularly imprinted polystyrene (MIP) membranes for bio-mimicry of terpenes [20], electrochemical sensor array for food quality assessment [21], and metal oxide semiconductor (MOS) sensors to discriminate among medicinal plants based on emissions of their VOCs [22–27]. Quartz crystal microbalance (QCM)-based sensor arrays have also been used to differentiate among plants from the Lamiaceae family [20,28]. For instance, a multichannel QCM (MQCM) with molecularly imprinted polystyrene membranes has been used to discriminate terpenes emanating from freshly dried Lamiaceae species, such as rosemary (*Rosmarinus officinalis* L.), sweet basil (*Ocimum basilicum*), and common Sage (*Salvia officinalis*) [20].

A crucial point when fabricating QCM-based sensors is the detector material used for coating the QCM substrate. In this context, reticular compounds such as metal-organic frameworks (MOFs) carry a huge potential. These porous materials can be easily modified to yield different responses to VOCs, thus allowing us to fabricate sensor arrays with each component showing different sensitivities.

In our previous work, we used an e-Nose to differentiate among different species of mints or VOCs isolated from them [29]. In the present study, we focused on QCM sensors coated with MOF thin films. Six different MOFs were investigated, including HKUST-1, Cu(BDC), Cu(BPDC), $Cu_2(DCam)_2(dabco)$, $Cu_2(DCam)_2(BiPy)$, and $Cu_2(DCam)_2(BiPyB)$ [30]. MOF thin films were deposited using layer-by-layer methods, yielding co-called SURMOFs (surface-anchored MOFs) [31]. To validate the performance of these SURMOF-based QCM arrays, we challenged them by testing the ability of these sensors to discriminate different

chemotypes of closely related species (the two mints *Mentha aquatica* and *Mentha suaveolens*, and the closely related Korean mint *Agastache rugosa*), against the more distant lemon balm (*Melissa officinalis*) and the unrelated but similarly scented lemongrass (*Cymbopogon citratus*). In addition, we used three accessions of basil (*Ocimum campechianum*, Amazonian basil, versus two accessions of *O. tenuiflorum*, holy basil and Tulsi). In contrast to sweet basil, which had previously been addressed by e-Noses [28], we wanted to test to what extent it would be possible to discern true *O. tenuiflorum* from closely related sister species, since *O. campechianum* is a member of the closest haplotype known for the genus *Ocimum* [12]. All accessions were selected from the authenticated reference plant collection at the Botanic Garden of the Karlsruhe Institute of Technology. The response times of the QCM sensors upon exposure to and removal of a particular scent were determined using nonlinear least-square (NLS) fits to an exponential rise (or fall) function and were found to amount to less than 1 min [32,33]. The exposure and purging data sets (four cycles) obtained from a low-cost custom-made portable e-Nose were analyzed using machine learning approaches, employing three different classification methods: principle component analysis (PCA), linear discrimination analysis (LDA), and nearest neighbors (k-NN) [28]. The first and second cycles of the datasets were used for training, and the repeated cycles following were used as unknown data for prediction. A statistical analysis revealed that more than 90% classification accuracy could be achieved within eight different scent classes from three different plant leaves in a very short time (less than 6 min). The prediction accuracies with repeated test measurements reached 90% for LDA and k-NN from unknown data sets.

2. Materials and Methods

2.1. Plant Material

The present study included freshly collected samples: 3 different species of basil, 4 different species of mints, and a control sample of lemongrass grown at the Botanical Institute of Karlsruhe Institute of Technology (KIT), Germany (Table 1). The scents were collected from 3 g weights of fresh leaves from *Ocimum campechianum Mill., Cymbopogon citratus, Ocimum tenuiflorum L., Melissa officinalis L, Mentha aquatica, Agastache rugosa,* and *Mentha suaveolens.* The abbreviations used throughout the text are defined in Table 1.

Table 1. Accessions used in this study. The voucher number gives the code, under which the plants are available in the botanical garden of the KIT. The abbreviations used in the text are also given.

Plant	Common Name	Abbreviation	KIT Voucher
Ocimumcampechianum Mill.	Amazon Basil	Bas7564	7564
Ocimumtenuiflorum L.	Tulsi	Bas5751	5751
Ocimumtenuiflorum L.	Krishna Tulsi	Bas8257	8257
Cymbopogon citratus	Lemongrass	LemGra	5722
Melissa officinalis L	Common Balm	MeliOfL	4643
Mentha aquatica L.	Water Mint	MintAQ	8680
Agastache rugosa	Korean Mint	MintAR	7576
Mentha suaveolens	Apple Mint	MintSU	3638

2.2. Chemicals and Fabrication of QCM Sensors

Commercially available AT-cut HC-49U type 200 nm silver coated 10 MHz quartz crystals with 5 mm electrode radius (J. Walter Thompson, Shanghai, China) were used as QCM electrodes. Six QCM sensors were coated with six different SURMOFs—HKUST-1, Cu(BDC), Cu(BPDC), $Cu_2(DCam)_2(dabco)$, $Cu_2(DCam)_2(BiPy)$, and $Cu_2(DCam)_2(BiPyB)$—to establish the sensor array comprising the e-Nose system, as shown in Figure 1.
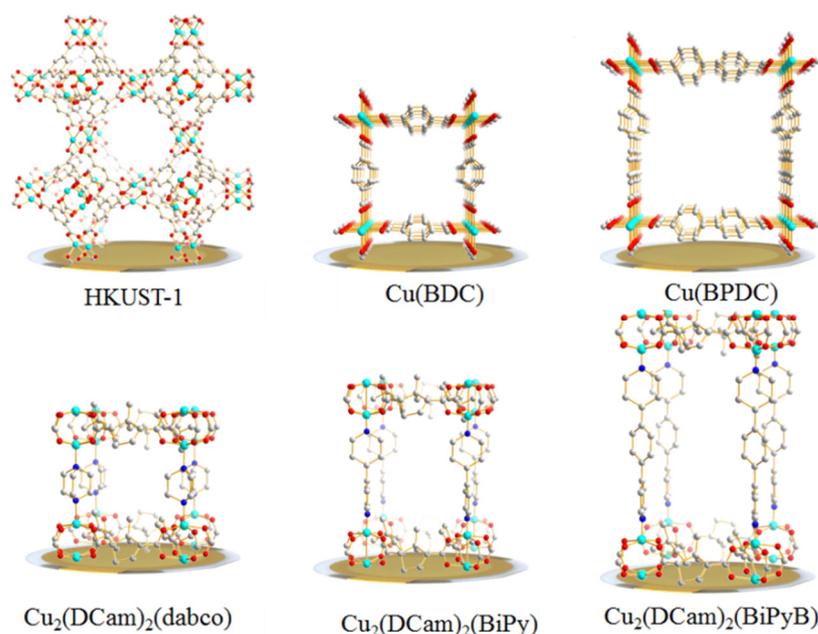
Figure 1. The six different SURMOF structures of the sensor array used in the e-Nose system, namely HKUST-1, Cu(BDC), Cu(BPDC), $Cu_2(DCam)_2(dabco)$, $Cu_2(DCam)_2(BiPy)$, and $Cu_2(DCam)_2(BiPyB)$.

DCam is a layer linker that produces pillared-layer MOF structures. The pillar linkers are diazabicyclo[2.2.2]octane (dabco), 4,4′-bipyridyl (BiPy), and 1,4-bis(4-pyridyl)benzene (BiPyB). BDC stands for benzene-1,4-dicarboxylate and BPDC stands for biphenyl-4,4′-dicarboxylate in the Cu(BDC) and Cu(BPDC) MOF structures [29,30].

Prior to SURMOF deposition, the QCM substrates were functionalized by an O_2 plasma treatment for 30 min. All films were prepared using 30 synthesis cycles. The SURMOF synthesis details are provided in the Supplementary Materials of our previous work [30]. X-ray diffraction (XRD) was used to characterize the SURMOFs thin films grown on the QCM sensors, and the diffractograms are shown in Figure S1. The XRD data reveal the presence of crystalline, oriented MOF thin films with the targeted structure.

2.3. Data Acquisition with the e-Nose

Figure 2 shows a schematic view of the working principle of the six-channel low-cost homemade portable e-Nose system used for discrimination of scents from basil/mint leaves. The sensor array and a humidity/temperature sensor were placed inside a 3D-printed head space in a cylindrical form. For the QCM data acquisition, 5 V/16 MHz ATMega32U4 microcontrollers and open-source Pierce oscillator circuits designed by openQCM were used [34] to read the frequency change. Temperature and humidity were measured with an Adafruit HTU21D-F temperature and humidity sensor breakout board. The temperature of the chamber was kept constant at 25 ± 0.5 °C. The software package MATLAB was used to record and analyze the data.

Three grams of freshly collected leaves from each species of basil and mint were inserted separately into a 100 mL glass vial. The emanating VOCs emitted by the fresh plant leaves inside the bottle were circulated through the sensor array with a 3W small diaphragm pump with a small flow rate 0.1 L/m, while valves 1 and 2 were rotated so that the N_2 gas line was closed. The surface of the sensing thin films inside the head space was activated by purging with N_2. This process led to the removal of residual compounds within the SURMOF pores. For each basil/mint scent accession, the change in resonance frequency was recorded for 6 min for each cycle with 2 min exposure for adsorption and, subsequently, 4 min of purging during cleaning with dry N_2 gas. The exposure and purging cycles were repeated four times.
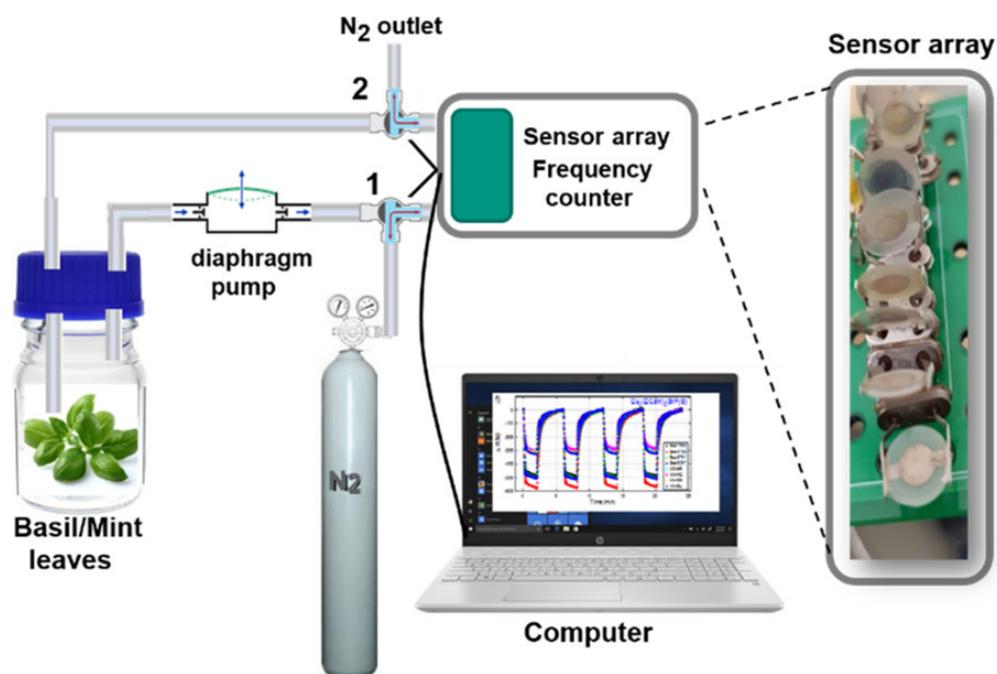
Figure 2. A schematic view of the working principle of the six-channel low-cost homemade portable e-Nose system used for discrimination of scents from basil/mint leaves.

MOFs are highly porous with huge specific surfaces [35]. SURMOFs coated on a QCM adsorb the VOCs on the outer surface as well as inside the pores (see Figure 3) [30]. Of course, for the latter, the pores and channels inside the MOF have to be sufficiently large to accommodate diffusion of the VOC into the pore system. A quantitative determination of the total amount of a particular VOC loaded into a MOF thin film can be carried out using a QCM. In the present case, the scent emitted from a plant consists of a large variety of different compounds, with their number typically exceeding 20 [36,37].
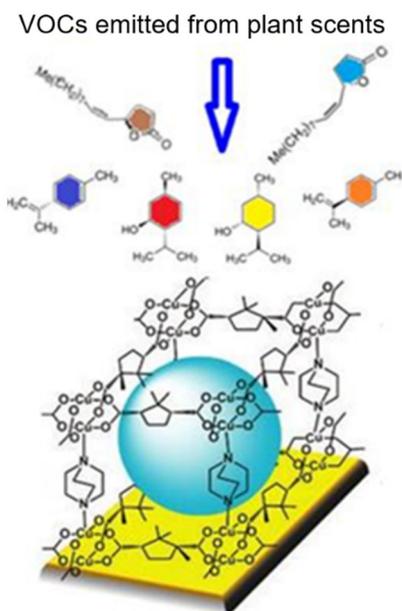
Figure 3. MOFs. Highly porous SURMOFs with huge specific surfaces coated on a QCM adsorb VOCs on the outer surface as well as inside the pores. Part of the Figure 3 was reproduced from Ref. [38] with permission from the Royal Society of Chemistry.

2.4. Data Analysis and Classification

The QCM response after exposure to the plant scent and after purging with dry nitrogen is shown in Figure 4. It was found that single-component rise-and-fall functions well described the QCM data for times up to 60 s after the start of exposure/purging. At later times, there was a linear behavior, indicating diffusion into and out of the pores [39–41].
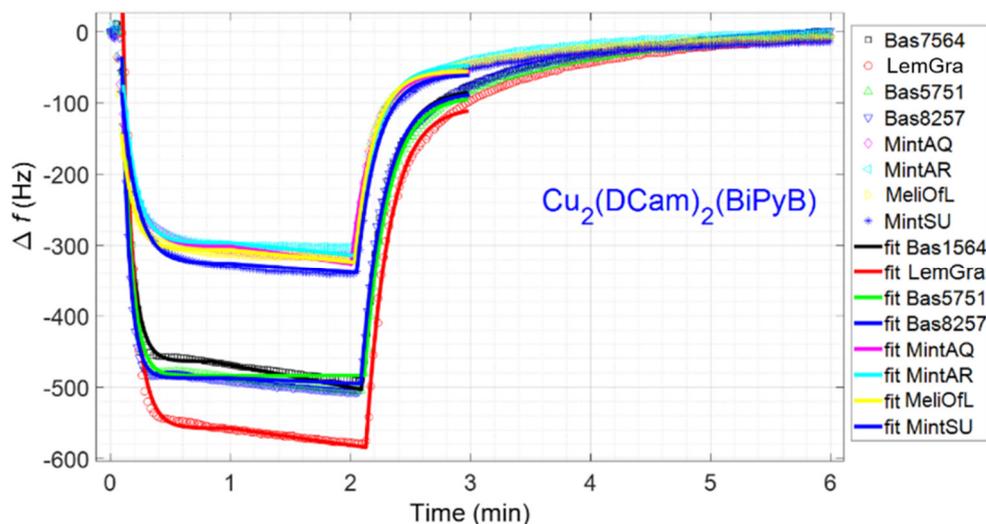
Figure 4. An example illustration of a one-cycle response of the sensor array, the change in resonance frequency of the sensing film $Cu_2(DCam)_2(BiPyB)$ due to adsorption with 2 min exposure, and subsequent 4 min purging during cleaning with dry N_2 gas.

The frequency shift of the QCM sensors is directly proportional to the change in the absorbed mass according to the Sauerbrey relation [42]:

$$\Delta F(t) = -C\Delta m(t) \tag{1}$$

where C is the QCM mass sensitivity constant, which is related to the structural and physical properties of the piezo electrical quartz sensor material. The frequency response times were calculated from nonlinear least-square (NLS) fits of the QCM response to an exponential rise function [32,33] in the time interval between 5 and 60 s.

The QCM signal drop observed after removing a particular scent was determined by an NLS fit to an exponential decay function in the time interval between 125 and 180 s using the following expression:

$$\Delta F(t) = \Delta F_{max} \left(1 - e^{t/\tau_{ads}}\right) \text{(Adsorption),} \tag{2}$$

$$\Delta F(t) = \Delta F_{max} \, e^{-t/\tau_{des}} \text{(Desorption),} \tag{3}$$

where τ_{ads} and τ_{des} are the relaxation times related to the association constant of the adsorption and desorption processes, respectively.

During the discrimination analysis of the scents, the first cycle of the loading/purging curve was used for training while the other three repeated cycles were used to test and predict the eight different classes of scents emitted from the plant leaves as a source. The exposure data with the highest responses between 1 and 2 min just before beginning the purging were cut to be used as the training data set for the discrimination accuracy calculations. Similarly, for prediction tests, one minute of exposure data with the highest responses were cut for the other cycles, e.g., the data between 7 and 8 min for the second cycle, between 13 and 14 min for the third cycle, and between 19 and 20 min for the fourth cycle. Three different classification algorithms were tested: PCA, LDA, and k-NN using scripts written in MATLAB.

PCA is an unsupervised machine-learning method that uses dimension reduction and data visualization [43,44]. This algorithm transforms the original data set into a new set of so-called Principal Components (PC). Hence, a large number of data sets is effectively compressed in a smaller set of PC variables. The 3D-PCA image classification can be obtained by projecting principal component scores in the x, y, and z axes. This makes it possible to visualize the separation of classes or clusters.

On the other hand, LDA is a supervised machine-learning method that maximizes discrimination among known categories by creating a new linear axis and by projecting data points on that axis. LDA and PCA are similar classification techniques [45]. They both compute linear combinations of variables that best explain the data. LDA gives a difference model between the classes of data. PCA, in contrast, does not consider any difference in class. LDA implements data with independent variables of continuous categorical observation. The objective of LDA is to find the projection hyperplane that minimizes the interclass variance and maximizes the separation distance between the projected classes. LDA has also been used in the literature due to its relatively fast model computation. Therefore, in this paper, we evaluate the effectiveness of the model in classifying the scents of two different plant species with eight different classes.

In k-NN discrimination analysis, k nearest neighbors is a simple algorithm that classifies new cases by scanning the distances of the classified elements of the nearest neighbors by comparing all stored known cases. It has been used for statistical estimation and pattern recognition. k is a parameter that defines the number of nearest neighbors before rendering a classification decision. In this paper, we also evaluate the effect of the number of nearest neighbors on classification and prediction accuracy of the scents of two different plant genera with eight different classes.

3. Results and Discussions

3.1. Sensor Array Responses

Figure 4 shows the response of a QCM sensor coated with a SURMOF of the type $Cu_2(DCam)_2(BiPyB)$ after exposure to scents emitted from different plants and after purging with nitrogen gas. For all scents, the sensor reached a saturation frequency (99.3% of $\Delta Fmax$) on average within 29 ± 8 s after the start of exposure and, after purging, recovered (0.7% of $\Delta Fmax$) on average within 54 ± 4 s, as shown in Figure S3 and Table S1 in the Supplementary Materials. Frequency response times were calculated using nonlinear least-square fits to an exponential rise-or-fall function. Among all SURMOF-based QCM sensors, the frequency response time was the fastest for HKUST-1, 27.5 ± 10.8 s for adsorption and 44.0 ± 1.1 s for desorption. The longest frequency response time was observed for $Cu_2(Dcam)_2(BiPy)$, with a value of 69.3 ± 7.3 s for adsorption and 66.7 ± 11.0 s for desorption. In general, adsorption is faster than desorption, showing a strong affinity between the sensing MOFs and the scent molecules emitted from both basil and mint leaves.

The resonance frequency shifts of the sensor array consisting of seven QCM sensors coated with all seven different sensing materials (see Table 1 for the abbreviations) during four cycles of exposure to the individual basil/mint leaves are shown in Figure S2 in Supplementary Materials.

For a comparison of the effect of SURMOF modification on bare Ag-coated QCM sensors, the maximum scale of the plots was kept constant at -600 Hz. For all MOF materials, the QCM sensors showed the highest response to LemGra and the lowest response to MintAQ. Interestingly, the responses of all sensors to both basil and mint species can be separated into two categories. The red circle with the highest response belongs to the scent of the control sample LemGra.

A radar plot (Figure 5) of the maximum frequency shift response for the sensor arrays shows that each sensor responds differently. In the case of the latter, for all scents, the response was very small, less than -10 Hz, as expected. The maximum frequency shift response values of the different SURMOF-based QCM sensors for the different scents shown in the radar plot are also listed in Table 2. The highest response of around -600 Hz comes

from the sensor coated with a $Cu_2(Dcam)_2$ (BiPyB) SURMOF thin film. The lowest response was obtained from the sensor coated with Cu2(Dcam)2(dabco). The large difference in response between different MOFs for the same scent results from the different chemical structure of the various SURMOFs. In addition to the chemical structure, the different pore sizes can also have an influence. Since each scent contains many different VOCs, precise identification of the underlying mechanisms is beyond the scope of this article.
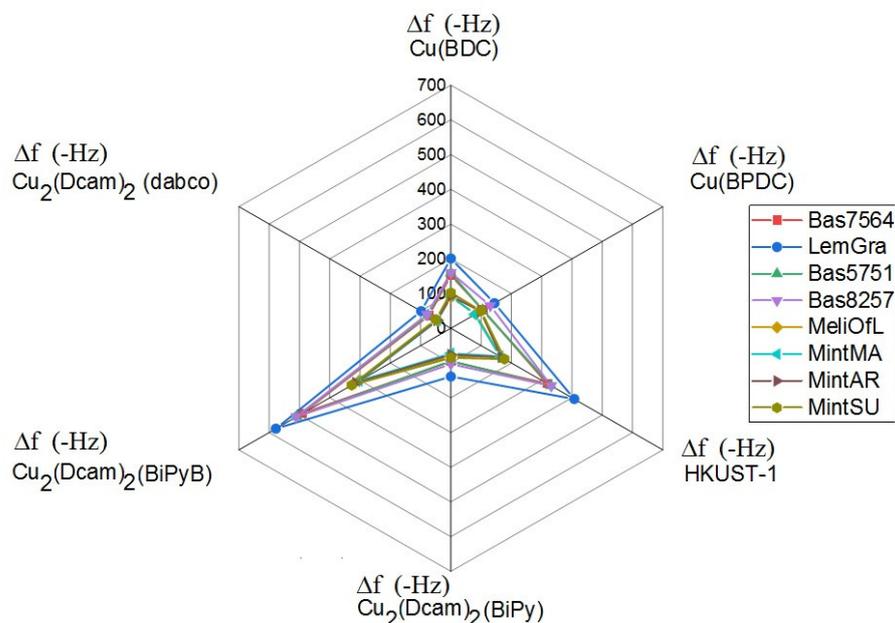
**Figure 5.** Radar plot of the maximum frequency shift responses of the sensor array.

Table 2. The maximum frequency shift response values of the sensor arrays shown in the radar plot in Figure 5.

	Bas7564	LemGra	Bas5751	Bas8257	MeliOfL	MintMA	MintAR	MintSU
	$\Delta f(-Hz)$	$\Delta f(-Hz)$	$\Delta f(-Hz)$	$\Delta f(-Hz)$	$\Delta f(-Hz)$	$\Delta f(-Hz)$	$\Delta f(-Hz)$	$\Delta f(-Hz)$
Cu(BDC)	153.1	200.9	159.4	160.0	89.5	94.4	93.1	102.4
Cu(BPDC)	105.4	144.0	104.8	130.1	82.6	78.9	98.2	100.1
HKUST-1	319.1	407.6	330.6	331.3	163.3	162.1	166.7	177.3
$Cu_2(Dcam)_2$ (BiPy)	95.9	139.0	95.0	103.0	80.4	72.5	76.1	84.8
$Cu_2(Dcam)_2$ (BiPyB)	492.6	578.3	509.0	510.9	302.8	303.4	314.4	327.6
$Cu_2(Dcam)_2$ (dabco)	72.9	98.2	75.4	79.0	44.7	44.5	45.9	49.3

3.2. Principal Component Analysis (PCA)

Figure 6 shows a 3D projection of the principal component scores in the *x*, *y*, and *z* axes calculated using principal component analysis from 451 measurements for the eight different scents. These components are grouped into clearly separated clusters. Interestingly, the two accessions from Mentha clustered with *Agastache rugosa* (Korean mint, AR, belonging to a neighboring clade) and with the more distantly related *Melissa officinalis* (MeliOfL). The three Basil scents were clearly separated: here, the two accessions for *O. tenuiflorum* (Tulsi) were close to each other, but unequivocally resolved from the closely related *O. campechianum*. This is astonishing because the latter species belongs to the sister clade closest to *O. tenuiflorum* within the entire genus. The other surprise comes from the complete separation of lemongrass (LemGra) from *Melissa officinalis* (MeliOfL), since both species have a very similar lemon-like scent and are often used for mutual surrogation in commercial samples. The clear separation indicates that the e-Nose can pick up even subtle differences in the VOC profile that go unnoticed by most human noses.

The sum of the three scores of the total variance explained by each principal component in the 3D plot given in Figure 6 is equal to 96.1%. By introducing the fourth and fifth PCA components, the visual PCA discrimination accuracy reaches 99.8%.
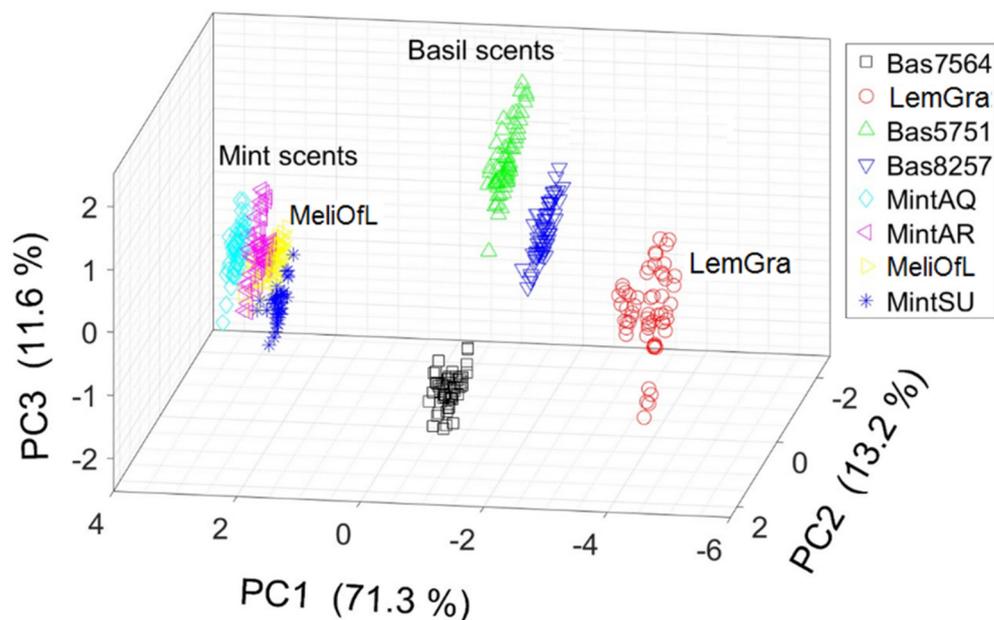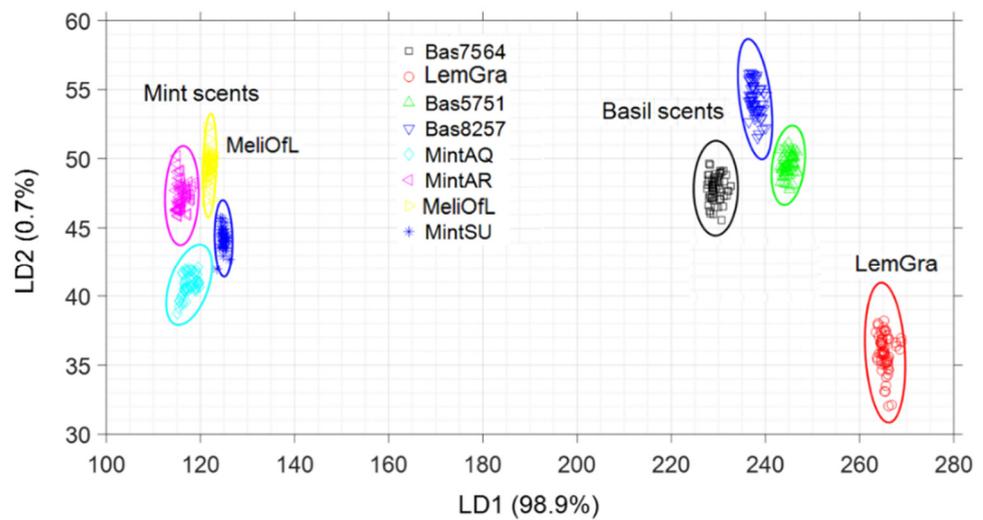

Figure 6. Three-dimensional plot of the principal component coefficients from 451 observations.

3.3. Linear Discrimination Analysis (LDA)

The 2D plot of the linear discriminant analysis for the eight different species with a 95% confidence ellipse is presented in Figure 7a. The so-called confusion matrix was calculated from a 10-fold LDA cross-validation partition using 451 observations with 406 training sizes and 45 test sizes obtained from the first cycle of the e-Nose measurements, as shown in Figure 7b. The LDA plot again shows an obvious clustering. The sum of the first two LDA vector components is 99.6%, and the LDA discrimination accuracy reached 100%. A calculated confusion matrix chart given in Figure 7 also confirms that the categorized (raw) labels match 100% with the true labels (columns) given during the training. The diagonal cells show correctly classified observations, while the off-diagonal values show the percentage of misclassification.

Figure 7 show the linear discriminant analysis of eight basil/mint species including a control sample. The 2D plot of the 10-fold linear discriminant analysis was obtained from the training data sets shown with the colored symbols (first-cycle e-Nose measurements) in Figure 7a and the prediction confusion matrix for the unknown data sets from the second cycle of the e-Nose measurements shown in Figure 7b. As clearly seen from Figure 8, the Mentha group of plants form a cluster separated from the basil group of plants, and lemongrass, which was used as an outlier, forms a completely separate cluster. The prediction matrix (Figure 8b) shows 9.8% overlap in the case of Bas8257 (Krishna Tulsi) and Bas5751 (Tulsi). This could be attributed to the fact that both samples belong to plants of the same species.
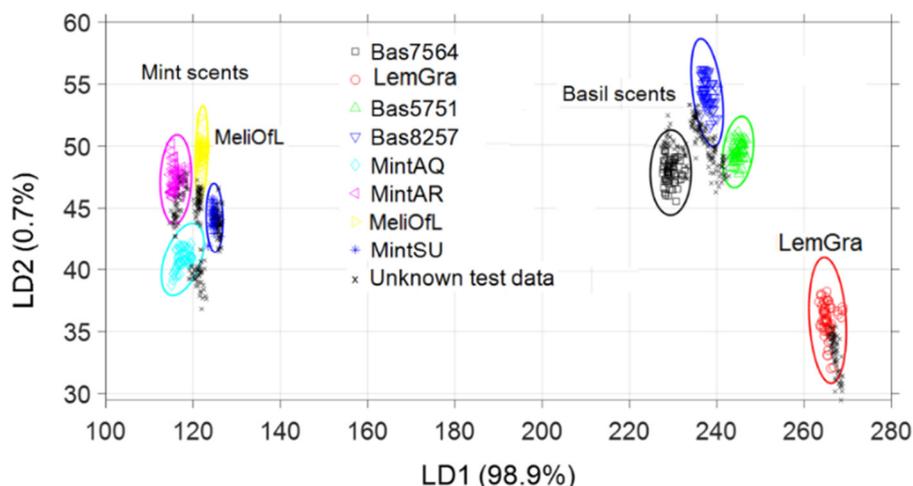
(**a**)

(**b**)

Figure 7. Linear discriminant analysis of eight species: (**a**) 2D plot of the linear discriminant analysis with 95% confidence ellipse and (**b**) confusion matrix obtained from 10-fold LDA calculations using the training dataset obtained from the first cycle of the e-Nose measurements.

Figure 8. Linear discriminant analysis of six mint species: (**a**) 2D plot of the 10-fold linear discriminant analysis obtained from the training data sets shown with the colored symbols (first-cycle e-Nose measurements) and from the prediction data sets shown with the black symbols (second cycle of the e-Nose measurements), and (**b**) a prediction confusion matrix for the unknown data sets from the second cycle of the e-Nose measurements.

Table 3 shows a summary of the LDA prediction results for unknown data sets obtained from different cycles of measurement after training with data sets from the first, second, and third cycles. The discrimination accuracy for each cycle is 100%. Nevertheless, the crosscheck prediction accuracies are obtained between 73.5 and 90.2% with an average of 79.2%. The prediction accuracies in Table 3 show similar overlap between Bas8257

(Krishna Tulsi) and Bas5751 (Tulsi), confirming that both samples originate from the same plant species.

Table 3. LDA prediction results for unknown data sets obtained from different cycles of measurement after training with data sets from the first cycle, second cycle, and third cycle.

Training Data	Unknown Test Data	Discrimination Accuracy	Prediction Accuracy
first cycle	second cycle	100	90.2%
first cycle	third cycle	100	75.3%
first cycle	fourth cycle	100	77.7%
second cycle	third cycle	100	79.3%
second cycle	fourth cycle	100	79.0%
third cycle	fourth cycle	100	73.5%

3.4. Nearest Neighbor Analysis (k-NN)

As a third nonparametric classification scheme, we applied the k-NN analysis with a 10-fold (k = 10) calculation of the unknown data sets from the second cycle compared to the true assignment from the training data set collected during the first cycle. The data sets from the second cycle of the e-Nose measurement were used for the k-NN calculation to determine the prediction accuracy for unknown observations. The k-NN discrimination accuracy was obtained as 94.2% with 5.8% misclassification (see Figure 9a). The overall prediction accuracy for the unknown data sets was smaller than in the case of LDA with 82.3% corresponding to 17.7% misclassification.

The change in the k-NN discrimination and prediction accuracies with an increasing number of nearest neighbors between 2 and 50 is given in Figure S4 in the Supplementary Materials. The k-NN discrimination accuracy drops from 100 to 90.2% with increasing nearest neighbor due to the overlap of classified data. Similarly, the k-NN prediction accuracy drops from 85.1 to 72.3% with the increasing number of nearest neighbors.
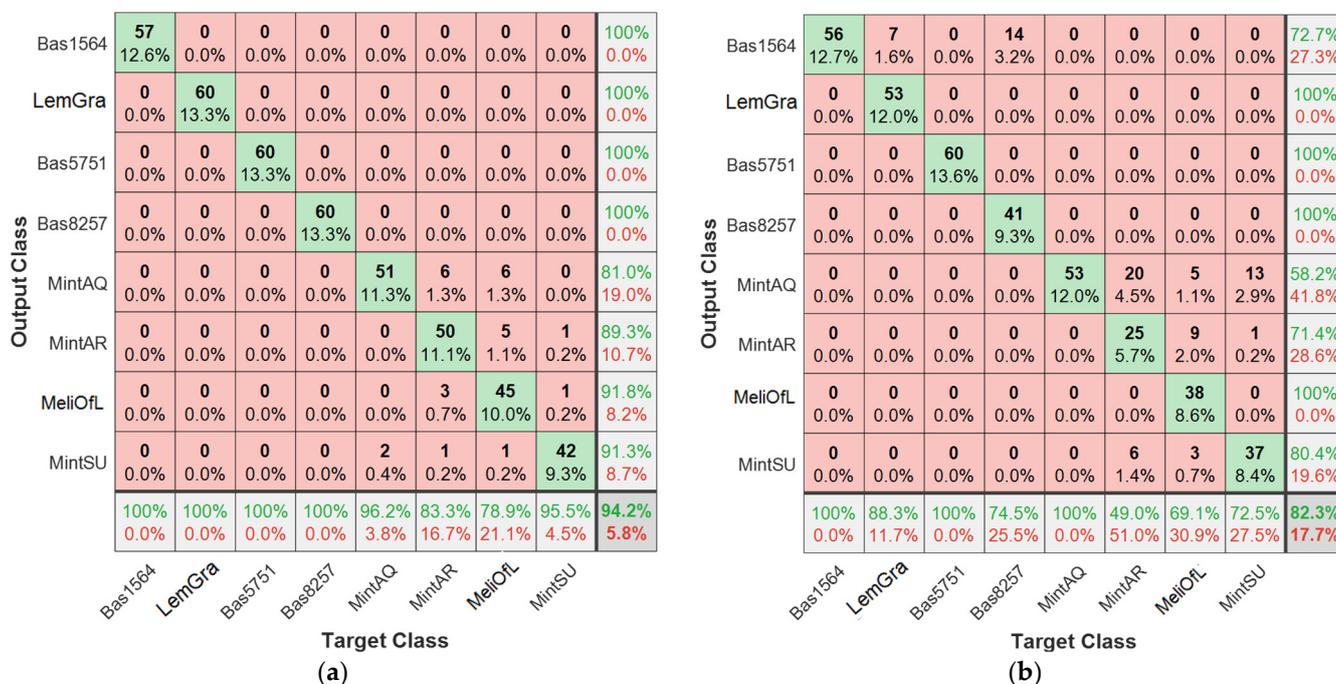
Figure 9. The discrimination (**a**) and prediction (**b**) confusion matrixes obtained from the k-NN analysis with 10-fold (k = 10) calculation using the unknown data sets from the second cycle to compare with the training data set (true labels) obtained from the first cycle.

A comparative GC–MS study performed by Chalchat et al. on *Ocimum basilicum* L. found 58.26% estragole, 19.4% limonene, and 2.40% p-cymene in the essential oil [46]. Another GC–MS analysis showed that most of the essential oil was composed of three main terpenes: linalool, 1,8-cineol, and eugenol [47]. Sarheed at al. [36] also showed that Mentha-type plants can constitute more than 20 different VOC molecules. Therefore, the most abundant constituent molecule is a dominant factor in the absorption signal on the SURMOF sensor array. A quantitative determination of the total amount of a particular VOC loaded into a SURMOF thin film can be carried out by different analytical methods such as GC-MS [46,47] and MIP-QCM [20,48]. It has been shown in our previous work that a SURMOF-based QCM sensor array can be used for detection and discrimination of plant oil scents and their mixtures [29]. In the present case, the scent emitted from a plant consists of a large variety of different compounds [36]. This reveals that a QCM-type e-Nose has high potential advantages for quickly analyzing the sample constituents of a complex mixture. A miniaturized portable multichannel QCM-based e-Nose is an economical artificial receptor option compared to the costly and time-consuming GC–MS.

4. Conclusions

In this work, sensor arrays based on six different SURMOFs were used successfully to discriminate eight aromatic plants, seven of which belonged to the taxonomically challenging family Lamiaceae. The exposure and purging data sets (four cycles) obtained from a low-cost custom-made portable e-Nose were analyzed using a linear discriminant analysis (LDA) classification model. The first and second cycles of the datasets were used for training, and the repeated cycles following were used as unknown data for prediction. More than a 90% classification accuracy was obtained within eight different scent classes. The prediction accuracies with repeated test measurements reached up to 90% for LDA from unknown data sets. We can show that it is possible not only to discern and identify plants on the genus level (*Mentha*, *Agastache*, and *Melissa*, all belonging to the Mentheae tribe within the Lamiaceae family) but also to discriminate closely related sister clades within a genus (basil). In addition, we were able to separate lemongrass (*Cymbopogon citratus*) unequivocally from common balm (*Melissa officinalis* L.) although these species share an intense lemon-like scent and are often used for mutual surrogation and adulteration, demonstrating that the e-Nose exceeds the performance of most human noses, which can be easily tricked by these two species. This study paves the way for the potential use of sensors in the detection of food adulterants. The portability and quick response of the sensor arrays demonstrate a huge potential for future fabrication of cheap monitoring devices for use in the food industry and in food surveillance.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/chemosensors9070171/s1, Figure S1: The X-ray diffractograms of the SURMOFs thin films of the sensor array used in the e-Nose system., e.g. Cu2(DCam)2(dabco), Cu2(DCam)2 (BiPy), Cu2(DCam)2 (BiPyB), HKUST-1, Cu(BDC), Cu(BPDC). The data indicates crystalline, oriented growth of the MOF films with the targeted structure, Figure S2: Resonance frequency shifts of the sensor array with 7 different sensing materials (see Table 1 for abbreviations) during 4 cycles of exposure to the individual Basel/Mint/Lemon Grass/Melissa O.L. leaves, Figure S3: Nonlinear Least square fit to an exponential rise function describing (Adsorption process) and an exponential drop function as (desorption process) to find response time values of the sensor array, Figure S4: The change in the k-NN discrimination and prediction accuracies with increasing number of nearest neighbor between 2 and 50, Table S1: The response times calculated from Nonlinear Least square fit to an exponential rise function describing (Adsorption process) and an exponential drop function as (desorption process).

Author Contributions: Conceptualization, S.O., M.S., A.K., S.V.P., L.F., F.G., P.N., and U.L.; methodology, S.O., M.S., A.K., P.N., C.L., Z.Z., and U.L.; software, S.O., L.F., and F.G.; validation, S.O., M.S., A.K., S.V.P., P.N., and U.L.; formal analysis, S.O., A.K., L.H., S.V.P., P.N., and U.L.; investigation, S.O.; resources, S.O., C.W., P.N., and U.L.; data curation, S.O. and L.H.; writing—original draft preparation, S.O., M.S., C.L., Z.Z., S.V.P., and P.N.; writing—review and editing, A.K., C.L., Z.Z., L.H., C.W., S.V.P.,

L.F., F.G., P.N., and U.L.; visualization, S.O.; supervision, C.W., P.N., and U.L.; project administration, S.O., C.W., P.N., and U.L.; funding acquisition, C.W., P.N., and U.L. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable for this study since it is not involving humans or animals.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available for now.

Conflicts of Interest: No conflict of interest.

References

1. Choesin, D.N.; Boerner, R.E. Allyl isothiocyanate release and the allelopathic potential of Brassica napus (Brassicaceae). *Am. J. Bot.* **1991**, *78*, 1083–1090. [CrossRef]
2. Bruneton, J. *Pharmacognosy, Phytochemistry, Medicinal Plants*; Lavoisier Publishing: Paris, France, 1995.
3. Kubitzki, K.; Rohwer, J.; Bittrich, V. *The Families and Genera of Vascular Plants*; Springer: Berlin/Heidelberg, Germany, 1990; Volume 1.
4. Butterworth, M.; Davis, G.; Bishop, K.; Reyna, L.; Rhodes, A. What Is a Superfood Anyway? Six Key Ingredients for Making a Food "Super". *Gastronomica* **2020**, *20*, 46–58. [CrossRef]
5. Zoller, A.; Nordwig, H. *Heilpflanzen der Ayurvedischen Medizin*; Narayana Verlag: Kandern, Germany, 2017.
6. Murthy, K.S.R.; Reddy, M.C.; Rani, S.S.; Pullaiah, T. Bioactive principles and biological properties of essential oils of Burseraceae: A review. *J. Pharmacogn. Phytochem.* **2016**, *5*, 247.
7. Malladi, S.; Ratnakaram, V.N.; Pullaiah, T. Phytochemical screening of Caralluma lasiantha: Isolation of C21 pregnane steroid. *Orient. J. Chem.* **2017**, *33*, 963–967. [CrossRef]
8. Mediratta, P.; Sharma, K. Effect of essential oil of the leaves and fixed oil of the seeds of Ocimum sanctum on immune responses. *J. Med. Aromat. Plant Sci.* **2000**, *22*, 694–700.
9. Jürges, G.; Beyerle, K.; Tossenberger, M.; Häser, A.; Nick, P. Development and validation of microscopical diagnostics for 'Tulsi'(*Ocimum tenuiflorum* L.) in ayurvedic preparations. *Eur. Food Res. Technol.* **2009**, *229*, 99–106. [CrossRef]
10. Cohen, Y.; Ben Naim, Y.; Falach, L.; Rubin, A.E. Epidemiology of basil downy mildew. *Phytopathology* **2017**, *107*, 1149–1160. [CrossRef]
11. Agrawal, P.; Rai, V.; Singh, R. Randomized placebo-controlled, single blind trial of holy basil leaves in patients with noninsulin-dependent diabetes mellitus. *Int. J. Clin. Pharmacol. Ther.* **1996**, *34*, 406.
12. Jürges, G.; Sahi, V.; Rios Rodriguez, D.; Reich, E.; Bhamra, S.; Howard, C.; Slater, A.; Nick, P. Product authenticity versus globalisation—The Tulsi case. *PLoS ONE* **2018**, *13*, e0207763. [CrossRef]
13. Riemer, B.; Weber, B. *Lebensmittelbedarfsgegenstände und Verpackungsrecht*; Behr's Verlag: Hamburg, Germany, 2014.
14. Ichim, M.C.; Häser, A.; Nick, P. Microscopic authentication of commercial herbal products in the globalized market: Potential and limitations. *Front. Pharmacol.* **2020**, *11*, 876. [CrossRef]
15. Zaheer, S.H.; Prasad, B.; Chopra, R.N.; Santapau, H.; Krishnan, M.S.; Deshaprabhu, S.B. *Wealth of India–A Dictionary of Raw Materials and Industrial Products*; CSIR: New Delhi, India, 1966; Volume VII, p. 10.
16. Vani, S.R.; Cheng, S.; Chuah, C. Comparative study of volatile compounds from genus Ocimum. *Am. J. Appl. Sci.* **2009**, *6*, 523. [CrossRef]
17. Ganie, S.H.; Upadhyay, P.; Das, S.; Sharma, M.P. Authentication of medicinal plants by DNA markers. *Plant Gene* **2015**, *4*, 83–99. [CrossRef]
18. Zou, H.-Q.; Lu, G.; Liu, Y.; Bauer, R.; Tao, O.; Gong, J.-T.; Zhao, L.-Y.; Li, J.-H.; Ren, Z.-Y.; Yan, Y.-H. Is it possible to rapidly and noninvasively identify different plants from Asteraceae using electronic nose with multiple mathematical algorithms? *J. Food Drug Anal.* **2015**, *23*, 788–794. [CrossRef]

19. Wasilewski, T.; Szulczyński, B.; Wojciechowski, M.; Kamysz, W.; Gębicki, J. A Highly Selective Biosensor Based on Peptide Directly Derived from the HarmOBP7 Aldehyde Binding Site. *Sensors* **2019**, *19*, 4284. [CrossRef] [PubMed]

20. Iqbal, N.; Mustafa, G.; Rehman, A.; Biedermann, A.; Najafi, B.; Lieberzeit, P.A.; Dickert, F.L. QCM-arrays for sensing terpenes in fresh and dried herbs via bio-mimetic MIP layers. *Sensors* **2010**, *10*, 6361–6376. [CrossRef] [PubMed]

21. Wojnowski, W.; Majchrzak, T.; Dymerski, T.; Gębicki, J.; Namieśnik, J. Portable electronic nose based on electrochemical sensors for food quality assessment. *Sensors* **2017**, *17*, 2715. [CrossRef] [PubMed]

22. Kiani, S.; van Ruth, S.M.; Minaei, S.; Ghasemi-Varnamkhasti, M. Hyperspectral imaging, a non-destructive technique in medicinal and aromatic plant products industry: Current status and potential future applications. *Comput. Electron. Agric.* **2018**, *152*, 9–18. [CrossRef]

23. Kiani, S.; Minaei, S.; Ghasemi-Varnamkhasti, M. Real-time aroma monitoring of mint (*Mentha spicata* L.) leaves during the drying process using electronic nose system. *Measurement* **2018**, *124*, 447–452. [CrossRef]

24. Rasekh, M.; Karami, H.; Wilson, A.D.; Gancarz, M. Classification and Identification of Essential Oils from Herbs and Fruits Based on a MOS Electronic-Nose Technology. *Chemosensors* **2021**, *9*, 142. [CrossRef]

25. Dong, W.; Tan, L.; Zhao, J.; Hu, R.; Lu, M. Characterization of fatty acid, amino acid and volatile compound compositions and bioactive components of seven coffee (*Coffea robusta*) cultivars grown in Hainan Province, China. *Molecules* **2015**, *20*, 16687–16708. [CrossRef] [PubMed]

26. Gancarz, M.; Malaga-Toboła, U.; Oniszczuk, A.; Tabor, S.; Oniszczuk, T.; Gawrysiak-Witulska, M.; Rusinek, R. Detection and measurement of aroma compounds with the electronic nose and a novel method for MOS sensor signal analysis during the wheat bread making process. *Food Bioprod. Process.* **2021**, *127*, 90–98. [CrossRef]

27. Slimani, S.; Bultel, E.; Cubizolle, T.; Herrier, C.; Rousselle, T.; Livache, T. Opto-Electronic Nose Coupled to a Silicon Micro Pre-Concentrator Device for Selective Sensing of Flavored Waters. *Chemosensors* **2020**, *8*, 60. [CrossRef]

28. Okur, S.; Sarheed, M.; Huber, R.; Zhang, Z.; Heinke, L.; Kanbar, A.; Wöll, C.; Nick, P.; Lemmer, U. Identification of Mint Scents Using a QCM Based E-Nose. *Chemosensors* **2021**, *9*, 31. [CrossRef]

29. Okur, S.; Zhang, Z.; Sarheed, M.; Nick, P.; Lemmer, U.; Heinke, L. Towards a MOF e-Nose: A SURMOF sensor array for detection and discrimination of plant oil scents and their mixtures. *Sens. Actuators B Chem.* **2020**, *306*, 127502. [CrossRef]

30. Okur, S.; Qin, P.; Chandresh, A.; Li, C.; Zhang, Z.; Lemmer, U.; Heinke, L. An enantioselective e-nose: An array of nanoporous homochiral MOF films for stereospecific sensing of chiral odors. *Angew. Chem. Int. Ed.* **2020**, *132*, 1–7.

31. Shekhah, O.; Wang, H.; Kowarik, S.; Schreiber, F.; Paulus, M.; Tolan, M.; Sternemann, C.; Evers, F.; Zacher, D.; Fischer, R.A. Step-by-step route for the synthesis of metal−organic frameworks. *J. Am. Chem. Soc.* **2007**, *129*, 15118–15119. [CrossRef] [PubMed]

32. Okur, S.; Kuş, M.; Özel, F.; Yılmaz, M. Humidity adsorption kinetics of water soluble calix [4] arene derivatives measured using QCM technique. *Sens. Actuators B Chem.* **2010**, *145*, 93–97. [CrossRef]

33. Horzum, N.; Taşçıoglu, D.; Okur, S.; Demir, M.M. Humidity sensing properties of ZnO-based fibers by electrospinning. *Talanta* **2011**, *85*, 1105–1111. [CrossRef]

34. Muckley, E.S.; Anazagasty, C.; Jacobs, C.B.; Hianik, T.; Ivanov, I.N. Low-cost scalable quartz crystal microbalance array for environmental sensing. In Proceedings of the Organic Sensors and Bioelectronics IX, San Diego, CA, USA, 28 August–1 September 2016; p. 99440Y.

35. Heinke, L.; Wöll, C. Surface-Mounted Metal–Organic Frameworks: Crystalline and Porous Molecular Assemblies for Fundamental Insights and Advanced Applications. *Adv. Mater.* **2019**, *31*, 1806324. [CrossRef]

36. Sarheed, M.M. Allelopathic Compounds from Mint Target the Cytoskeleton from Cell Biology Towards Application as Bioherbicides. Ph.D. Thesis, Karlsruher Institut für Technologie (KIT), Karlsruhe, Germany, 2019.

37. Sarheed, M.M.; Rajabi, F.; Kunert, M.; Boland, W.; Wetters, S.; Miadowitz, K.; Kaźmierczak, A.; Sahi, V.P.; Nick, P. Cellular base of mint allelopathy: Menthone affects plant microtubules. *Front. Plant Sci.* **2020**, *11*, 1320. [CrossRef]

38. Gu, Z.-G.; Grosjean, S.; Bräse, S.; Wöll, C.; Heinke, L. Enantioselective adsorption in homochiral metal–organic frameworks: The pore size influence. *Chem. Commun.* **2015**, *51*, 8998–9001. [CrossRef]

39. Heinke, L.; Gu, Z.; Wöll, C. The surface barrier phenomenon at the loading of metal-organic frameworks. *Nat. Commun.* **2014**, *5*, 1–6. [CrossRef] [PubMed]

40. Zhou, W.; Wöll, C.; Heinke, L. Liquid-and gas-phase diffusion of ferrocene in thin films of metal-organic frameworks. *Materials* **2015**, *8*, 3767–3775. [CrossRef]

41. Heinke, L. Diffusion and photoswitching in nanoporous thin films of metal-organic frameworks. *J. Phys. D Appl. Phys.* **2017**, *50*, 193004. [CrossRef]

42. Sauerbrey, G.; Jung, G. Vibrational Modes of Planoconvex Quartz Plates. *Z. Angew. Phys.* **1968**, *24*, 100–106.

43. Kambhatla, N.; Leen, T.K. Dimension reduction by local principal component analysis. *Neural Comput.* **1997**, *9*, 1493–1516. [CrossRef]

44. Sharifzadeh, S.; Ghodsi, A.; Clemmensen, L.H.; Ersbøll, B.K. Sparse supervised principal component analysis (SSPCA) for dimension reduction and variable selection. *Eng. Appl. Artif. Intell.* **2017**, *65*, 168–177. [CrossRef]

45. Shaoning, P.; Ozawa, S.; Kasabov, N. Incremental linear discriminant analysis for classification of data streams. *IEEE Trans. Syst. ManCybern. Part B (Cybern.)* **2005**, *35*, 905–914.

46. Chalchat, J.-C.; Özcan, M.M. Comparative essential oil composition of flowers, leavesand stems of basil (*Ocimum basilicum* L.) used as herb. *Food Chem.* **2008**, *110*, 501–503. [CrossRef]

47. Bilal, A.; Jahan, N.; Ahmed, A.; Bilal, S.N.; Habib, S.; Hajra, S. Phytochemical and pharmacological studies on Ocimum basilicum Linn-A review. *Int. J. Curr. Res. Rev.* **2012**, *4*, 73–83.
48. Iqbal, N. QCM Sensor Arrays for Monitoring Volatile Plant Emanations via Molecularly Imprinted Polymers. Ph.D. Thesis, uniwien, Wien, Austria, 2011.