



Research Article

<https://doi.org/10.1631/jzus.B2600550>

Novel olfactory epithelium-based bioelectronic nose for cigarette aroma recognition

Yongsheng LI^{1*}, Xinrui TANG^{2*}, Qingxiang LI¹, Chengxu ZHENG¹, Zhenjie ZHAO¹, Yingying XUE^{2,3}, Tan Shuan GIN², Xuliang FU², Jianguo WU², Deming JIANG², Hao WAN^{2,4}, Liujing ZHUANG^{2✉}, Ping WANG^{2,3,4✉}, Nan JIANG^{2✉}

¹China Tobacco Zhejiang Industrial Co., Ltd., Hangzhou 310009, China

²Biosensor National Special Laboratory, Key Laboratory for Biomedical Engineering of Education Ministry, Zhejiang Key Laboratory of Intelligent Sensing Technology and Advanced Medical Instrument, Department of Biomedical Engineering, Zhejiang University, Hangzhou 310027, China

³State Key Laboratory of Transducer Technology, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China

⁴Innovation Center for Smart Medical Technologies & Devices, Binjiang Institute of Zhejiang University, Hangzhou 310053, China

Abstract: The objective evaluation of complex odors, such as cigarette aromas, remains a significant challenge due to the limitations of traditional analytical methods in reflecting holistic sensory perception. In this study, we developed a bioelectronic nose by coupling the olfactory epithelium (OE) with a 64-channel microelectrode array (MEA) to emulate mammalian odor perception. The OE was stimulated using representative cigarette additives representing four different cigarette aromas and smoke samples from five commercial cigarettes. By recording the response signals at multiple sites within the OE, different spatiotemporal response patterns to different stimuli could be obtained. On the basis of these electrophysiological responses, spectral features were extracted and evaluated using multiple conventional machine learning classifiers. Among them, the Random Forest (RF) classifier achieved the best performance, reaching an accuracy of 96.83% for aroma recognition using only 10 highly informative spectral features. Furthermore, commercial cigarette recognition achieved an accuracy of 94.36% using a 50-feature subset. These results demonstrate that our new OE-based bioelectronic nose provides a sensitive and objective platform for complex odor analysis, showing great potential for tobacco flavor profiling, objective quality assessment, and quality control.

Key words: Microelectrode array; Olfactory epithelium; Bioelectronic nose; Electrophysiology recording; Machine learning

1 Introduction

Olfaction is an important chemical sense for mammals to perceive external environmental information. Within the nasal cavity, volatile odor molecules bind to specific olfactory receptors (ORs) on the membranes of olfactory sensory neurons (OSNs). This process triggers neural electrical signals that are transmitted to the olfactory bulb (OB) and the brain to form olfactory perception. (Mombaerts et al., 1996; Firestein, 2001) Using a complex neural network of OSNs and a combinatorial coding mechanism, the human olfactory system can

✉ Nan JIANG, 21915031@zju.edu.cn

Ping WANG, cnpwang@zju.edu.cn

Liujing ZHUANG, zhuangliujing@zju.edu.cn

* The two authors contributed equally to this work.

Ping WANG, <https://orcid.org/0000-0001-6474-2722>

Liujing ZHUANG, <https://orcid.org/0000-0001-9607-1302>

distinguish over a trillion distinct scents.(Guo et al., 2023) Inspired by this natural mechanism, artificial olfactory systems, commonly known as bioelectronic noses, have been developed to detect diverse odors. These systems typically integrate biological recognition elements, such as OR proteins or cultured cells, with secondary physical transducers to convert chemical interactions into measurable electrical signals. While conventional bioelectronic noses demonstrate excellent sensitivity for detecting simple, single-odor molecules(Dung et al., 2018), they face significant limitations when dealing with complex odor mixtures in real-world environments. This is mainly because most current systems rely on isolated receptor proteins or two-dimensional (2D) cultured cells as sensing elements.(Qin et al., 2023) These simplified elements lack the specific three-dimensional (3D) synaptic connections and complex neural networks unique to the real olfactory system, making it difficult to reproduce high-dimensional combinatorial coding.

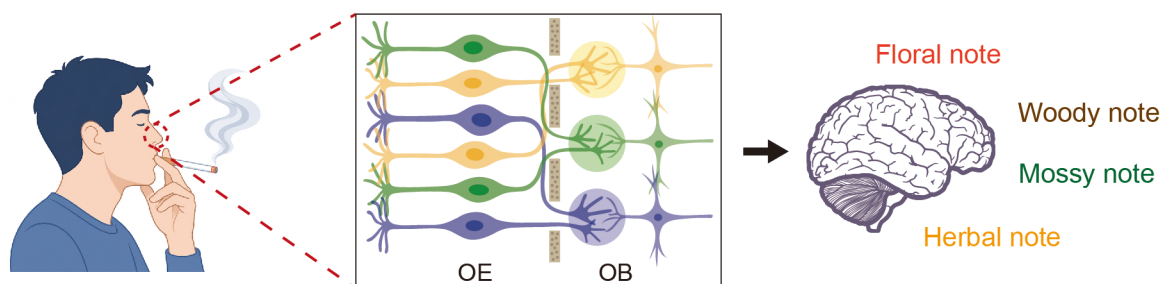
This challenge is particularly evident in the tobacco industry, where analyzing complex odor mixtures is a critical task. Cigarette smoke is a highly complex mixture, and conventional artificial noses struggle to holistically and synergistically perceive these multi-component aromas; the assessment of cigarette quality still relies on human sensory evaluation. However, this subjective approach is time-consuming and prone to sensory fatigue, highlighting the need for a more reliable analytical alternative.

To overcome the limitations of conventional biosensors and subjective human evaluation, intact olfactory epithelium (OE) tissue can be used as a biological sensing element. Since this tissue preserves the native extracellular environment, the natural spatial distribution of OSNs, and the complex cross-talk between neurons, coupling acutely isolated OE with microelectrode arrays (MEAs) provides a non-invasive and high-throughput recording platform. This system can record the spatiotemporal local field potentials (LFPs) evoked by complex mixtures, mimicking the natural biological sense of smell far better than the conventional bioelectronic noses.

In addition to detecting odor molecules, accurately decoding and recognizing the electrical signals is also highly important. Because odor-evoked electrophysiological signals exhibit highly dynamic temporal and spatial patterns, traditional analytical methods cannot easily extract the underlying coding rules. Recent studies have demonstrated the potential of machine learning (ML)-based pattern recognition for decoding olfactory information from biological electrophysiological signals. Commonly used ML algorithms include random forests (RFs), k-nearest neighbors (KNN), support vector machines (SVMs), and neural networks. Dimensionality reduction techniques, such as principal component analysis (PCA), can further simplify high-dimensional electrophysiological data and facilitate classification. For example, PCA-SVM frameworks have achieved over 80% accuracy in classifying six odors from multichannel mouse olfactory bulb recordings, while RF classifiers have been used to distinguish individual odorants and odor mixtures based on locust electroantennogram signals.(Shor et al., 2022; Neta et al., 2023; Zhu et al., 2025) These findings highlight the feasibility of applying ML-based approaches to extract and classify odor-related information from biological olfactory electrophysiological signals.

This study establishes an objective evaluation system that can assist professionals in distinguishing complex cigarette aromas. A new bioelectronic nose was developed based on intact mouse OE coupled with a MEA chip (Fig. 1), and this platform was evaluated using two distinct groups of stimuli. The first group consisted of four representative cigarette additives corresponding to fundamental flavor notes (floral, woody, mossy, and herbal), while the second group comprised five complex commercial cigarette smoke samples. By combining spectral feature extraction with multiple ML classifiers, we constructed an analytical framework for decoding odor-evoked electrophysiological responses and evaluating the odor discrimination capability of the OE-based bioelectronic nose. The results demonstrated that the system can successfully capture the distinct neural response patterns elicited by different stimuli, achieving high accuracy in classifying both the fundamental flavor additives and the specific commercial cigarette categories.

Mammalian olfactory system



OE-based bioelectronic nose

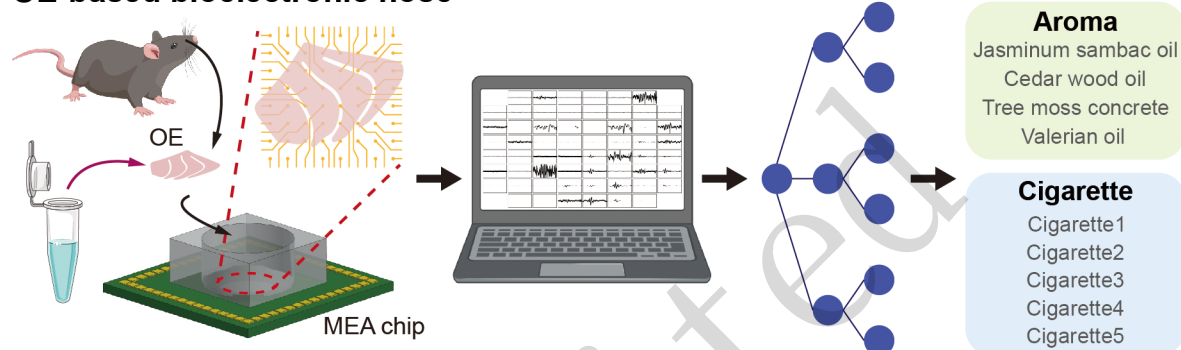


Fig. 1 Schematic illustration of the mammalian olfactory system and the proposed OE-based bioelectronic nose. The natural olfactory pathway involves volatile molecules being detected by specialized neurons in the OE and transmitted through the OB to the brain for the perception of specific flavor notes (Top). In the biomimetic sensing framework, an acutely isolated mouse OE is coupled onto a planar MEA chip to capture odor-evoked spatiotemporal electrophysiological signals. These signals are subsequently decoded by machine learning algorithms to classify four representative cigarette additives (aromas) and five commercial cigarettes (Bottom).

2 Materials and methods

2.1 Chemicals

Four commonly used cigarette additives (jasminum sambac oil, cedar wood oil, tree moss concrete, and valerian oil) were selected as stimuli to represent floral, woody, mossy, and herbal flavor notes, respectively. These additives were initially dissolved in ethanol to prepare stock solutions. Before each experiment, the stock solutions were diluted 1,000 times with phosphate-buffered saline (PBS) for use as a stimulus. Furthermore, five commercial cigarette products from the Liqun brand (Modern, Xinban, Ruanchangzui, Yangguang, and Xiuxian, China Tobacco Zhejiang Industrial Co., Ltd., Hangzhou, China) were evaluated. For brevity, these five cigarette types are hereafter referred to as C1, C2, C3, C4, and C5, respectively. The cigarette smoke was generated using a homemade smoking device and collected into 10 mL of PBS, which was used as the stock solution for subsequent stimulation experiments. The smoke was aspirated at a flow rate of 17.5 mL/s, with each aspiration lasting 2 s and an interval of 10 s between successive aspirations. This procedure was repeated until five cigarettes had been completely smoked. After collection, the sample was thoroughly mixed to ensure the uniform distribution of smoke-derived components in the PBS.

2.2 OE preparation

Three-week-old C57BL/6 mice were used for in vitro electrophysiological recordings of the OE. The mice

were euthanized via an overdose of urethane, followed by decapitation. The heads were sagittally bisected to isolate the olfactory mucosa, which was immediately incubated at 28°C in Ringer's solution saturated with 95% O₂/5% CO₂ (pH 7.4, 315 mOsm).

2.3 Electrophysiological measurements

In vitro electrophysiological recordings of OE were conducted using a multi-channel MEA system (MEA2100, Multichannel Systems) and 64-channel MEA chips. After incubation, the mucosa was positioned over the working electrode area of the MEA (60 working electrodes, each 30 μ m in diameter, with a spacing of 200 μ m) and secured with a slice anchor. The tissue was then allowed to stabilize in oxygenated Ringer's solution for 10 min before recording commenced. During recording, the Ringer's solution was removed until a thin layer remained, and the chemical stimulus (100 μ L) was applied to the OE using a pipette. The OE was washed at least three times with oxygenated Ringer's solution when transitioning between stimuli. Signals were amplified, filtered (0.1-3500 Hz) and digitized at 1 kHz.

2.4 LFP signal analysis

The initial processing of the recorded MEA signals was performed using custom-written MATLAB scripts (MATLAB R2022b, The Mathworks, Natick, MA, USA). Raw LFPs were first band-pass filtered (0.1-300 Hz) using a zero-phase digital filter 'filtfilt' to prevent phase distortion. The power spectrum was estimated as the magnitude squared of the fast Fourier transform (function 'fft') of the time domain signal. To create a spectrogram, the LFP was convolved by complex Morlet wavelets, with the central frequency ranging from 0.1 to 5 Hz in 0.1 Hz steps. (Tallon-Baudry et al., 1997; Zhuang et al., 2019) Power analysis was performed for periods of interest, and signals were cut into non-overlapping 10-s-long segments. The band power of LFPs was calculated using the 'bandpower' function. The results were averaged across all segments.

2.5 Machine learning-based classification of MEA responses

Based on the electrophysiological responses recorded by the MEA, two classification tasks were performed, including aroma recognition and commercial cigarette recognition. Each classification task was evaluated using three independent recording sets, and an independent machine learning model was constructed for each recording set. Both tasks followed the same analytical workflow, including signal segmentation, spectral feature extraction, feature selection, and model training and evaluation (Fig. 2).

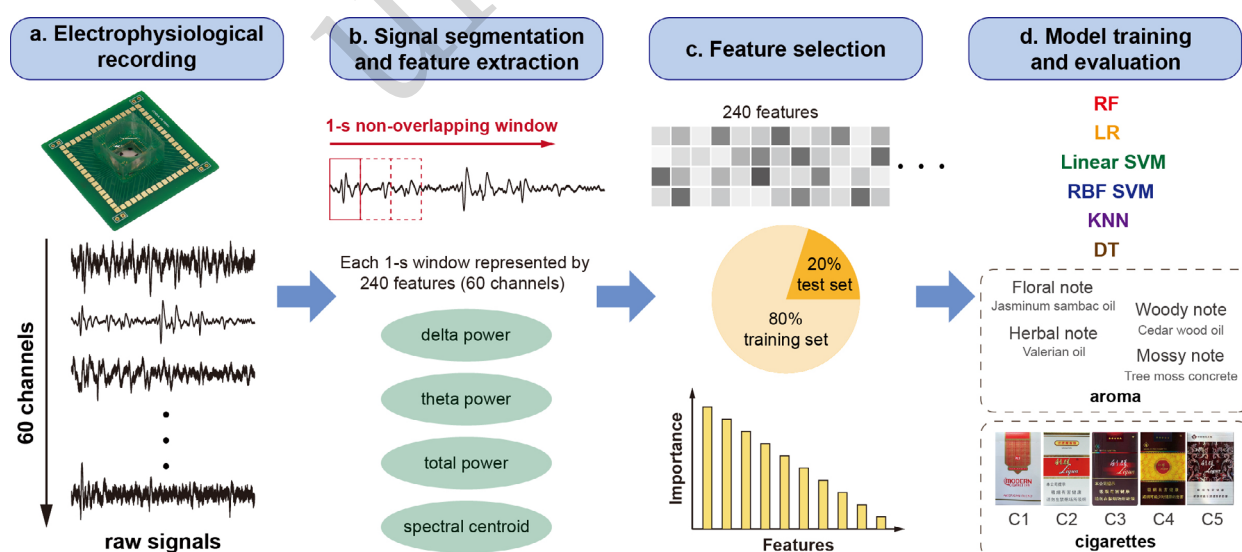


Fig. 2 Machine learning workflow for MEA signal processing and pattern recognition. (a) Acquisition of raw electro-

physiological signals using the 64-channel MEA system. (b) Signal segmentation utilizing 1-second non-overlapping windows, followed by the extraction of 240 spectral features from the 60 recording channels, including delta power, theta power, total spectral power, and spectral centroid. (c) Dataset partitioning into training (80%) and test (20%) sets, coupled with feature selection based on calculated importance scores. (d) Model training and evaluation utilizing multiple classifiers for aroma recognition and commercial cigarette recognition. (RF: Random Forest, LR: Logistic Regression, Linear SVM: Linear Support Vector Machine, RBF SVM: Radial Basis Function Support Vector Machine, KNN: K-Nearest Neighbors, DT: Decision Tree)

The first step of the analytical workflow comprised signal segmentation and spectral feature extraction. The raw signals were segmented into consecutive, non-overlapping 1-s time windows. For each time window, four spectral features were extracted from each of the 60 recording channels, including delta (0.1-4 Hz) power, theta (4-12 Hz) power, total spectral power, and spectral centroid. These features were selected to characterize complementary aspects of odor-evoked electrophysiological responses: the delta and theta power describe low-frequency components previously implicated in olfactory information processing (Jiang et al., 2017; Iravani et al., 2021), the total spectral power reflects the overall response energy, and the spectral centroid characterizes the distributional center of the power spectrum. Accordingly, each time window was represented by a 240-dimensional spectral feature vector.

The extracted spectral features were used for model construction. For each independent recording set, the 1-s windows of each class were sorted according to their acquisition time. The first 80% of the windows were used for training, whereas the remaining 20% were reserved for testing. Across the three recording sets, the aroma recognition task comprised 779 windows, including 620 training windows and 159 test windows, whereas the commercial cigarette recognition task comprised 2,247 windows, including 1,792 training windows and 455 test windows. Detailed class-specific numbers of windows for each recording set are provided in Tables S1 and S2, respectively. Feature selection, model training and performance evaluation were then performed independently for each recording set using the same analytical procedure. To avoid data leakage, feature selection was performed exclusively on the training set. Specifically, a Random Forest (RF) model was first trained using the training data (Breiman, 2001), and the importance scores of all 240 spectral features were calculated and ranked in a descending order. To investigate the influence of feature dimensionality on classification performance, feature subsets containing the top 10, 20, 30, 50, and 100 ranked features were constructed, and their classification performances were compared.

After the optimal feature subset had been determined for each classification task, it was used as the common input for classifier comparison. Six conventional machine learning classifiers, including RF, Logistic Regression (LR), Linear Support Vector Machine (Linear SVM), Radial Basis Function Support Vector Machine (RBF SVM), KNN, and Decision Tree (DT), were subsequently evaluated (Ma et al., 2023; Zhu, et al., 2025).

Class-balanced weights were applied to all applicable classifiers to account for potential class imbalance. The RF classifier consisted of 500 decision trees; LR employed L2 regularization with a maximum of 5,000 iterations; linear SVM used a linear kernel ($C = 1$); RBF SVM used an RBF kernel ($C = 10$). KNN was implemented with $k = 5$. The random seed was set to 42 for all stochastic models. For LR, SVM and KNN, features were standardized using the training set, and the same transformation parameters were subsequently applied to the corresponding test set.

The classification performance of each model was evaluated using overall accuracy (Accuracy), macro-averaged precision (Macro Precision), macro-averaged recall (Macro Recall), and macro-averaged F1-score (Macro F1-score). These performance metrics were calculated independently for each of the three recording sets and reported as the mean $\pm$ standard error of the mean (SEM). For each classification task, the true and predicted labels of the test windows were concatenated across the three independently trained and evaluated recording sets. A single confusion matrix was then calculated from the pooled window-level predictions. The pooled confusion matrix was normalized row-wise, with each cell divided by the total number of test windows in the corresponding true class and expressed as a percentage. Consequently, each row sums to

100%. The resulting confusion matrix was row-normalized and presented as percentages. Beyond evaluating classification performance, the composition of the selected feature subsets was further analyzed to facilitate the interpretation of the model's decision basis by calculating the proportions of the four spectral feature types (delta power, theta power, total spectral power, and spectral centroid) within the selected feature subsets. Statistical analyses were performed using GraphPad Prism Software (GraphPad Prism v9.0, GraphPad Software, San Diego, CA, USA).

3 Results and discussion

3.1 Construction of OE-based bioelectronic nose

To emulate the complex odor discrimination capabilities of the mammalian olfactory system, an OE-based bioelectronic nose was constructed (Fig. 1). In natural odor perception, volatile molecules are detected by specialized ORs expressed on the cell membranes of OSNs within the OE. These chemical stimuli are converted into electrical signals and transmitted to the OB and olfactory cortex, where they are perceived as specific flavor notes.(Lledo et al., 2005) Bioelectronic noses mimic this physiological process by integrating biological recognition elements (OR proteins, cells or tissues) with secondary transducers to convert biological interactions into measurable electrical or optical signals.(Zhang et al., 2018) Such systems have realized artificial odor perception to a certain extent, and thus have tremendous potential for applications in food safety, environmental monitoring and clinical diagnosis.

Motivated by the high application potential of the bioelectric nose, in this study, a sensing interface was constructed by integrating an acutely isolated mouse OE onto a 64-channel planar MEA chip (Fig. 3a). MEA is one of the most important biosensors and has been commonly used to detect the extracellular electrophysiological activities of cells and tissues in vitro due to its long-term and label-free features.(Thomas et al., 1972) The fabrication process of micro-electro-mechanical systems (MEMS)-based MEA sensor is similar to that described in previous reports.(Wei et al., 2019; Qin et al., 2022) The utilized MEA chip comprises 60 working electrodes arranged in the central area and 4 reference electrodes located at the corners.

The spatiotemporal electrophysiological signals evoked by odorants were acquired in real time and subsequently decoded using machine learning algorithms, which acted as an artificial brain for pattern recognition. To systematically evaluate the capability of the system in detecting and discriminating tobacco aromas, two distinct groups of stimuli were employed: the first comprised four representative cigarette additives (jasminum sambac oil, cedar wood oil, tree moss concrete, and valerian oil), corresponding to floral, woody, mossy, and herbal notes, respectively; the second group consisted of five types of commercial cigarettes.

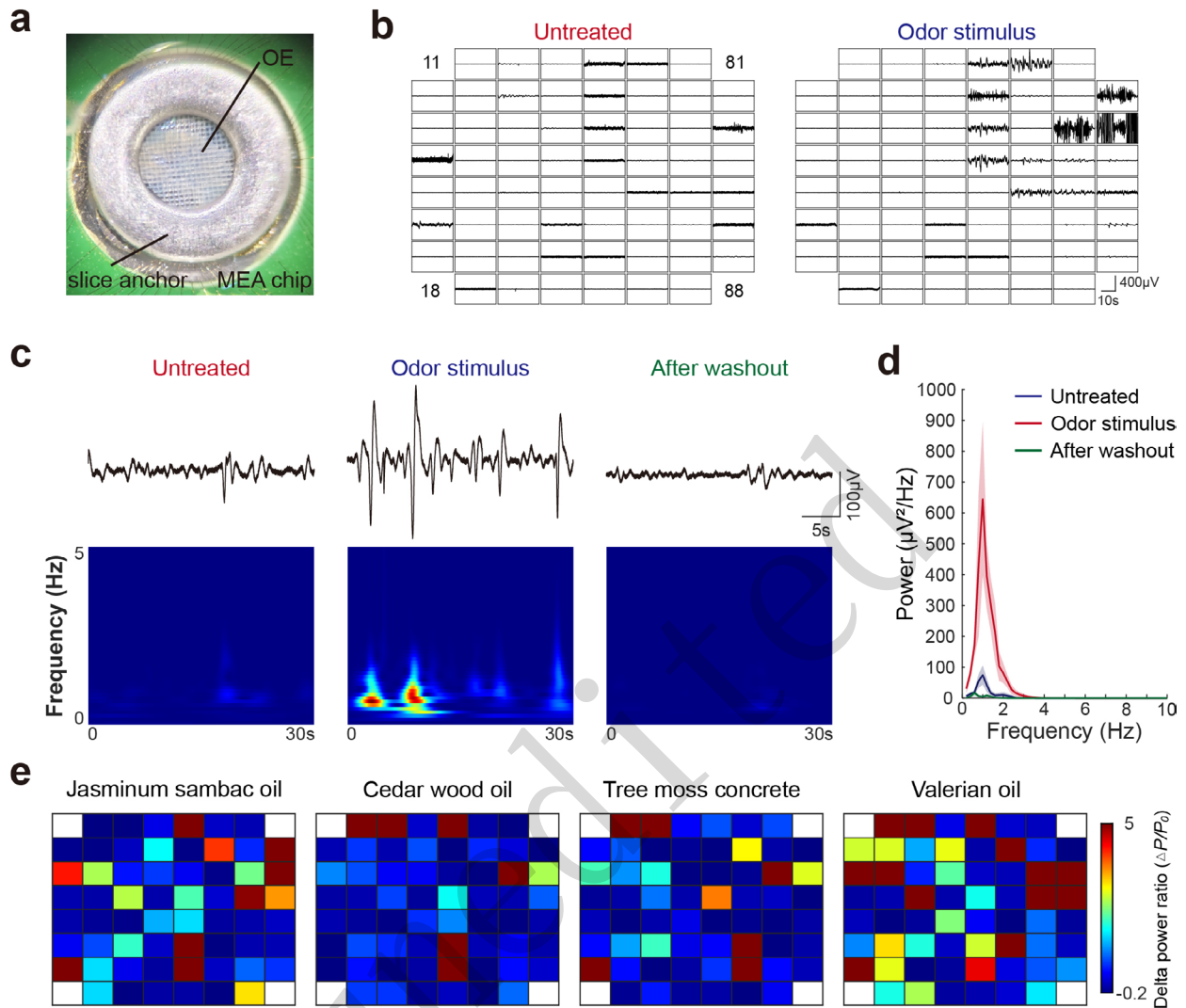


Fig. 3 Performance of the OE-based bioelectronic nose. (a) Image of OE on the planar MEA chip. (b) Raw trace of spontaneous (30 s long, left) and tree moss concrete-evoked (30 s long, right) activities from 60 recording channels. (c) Representative raw traces (0.1-300 Hz, top) and wavelet spectrograms (0-5 Hz, bottom) of LFP signals recorded from a single electrode before, during and after the washout of bath-applied tree moss concrete, representing the baseline, treatment, and washout periods, respectively. (d) Average LFP power spectrum before (baseline) and after odor application showing a dramatic increase in delta power. (e) The acquired odor response of each channel was converted to $\Delta P/P_0$ (delta power ratio: change in delta power/baseline power) and presented in a heatmap. The OE tissue exhibited distinct response patterns to different cigarette additives when recorded using the same MEA chip.

3.2 Performance and spatiotemporal coding of the OE-based bioelectronic nose

To verify the performance of the OE-MEA interface, it was crucial to first confirm whether the isolated OE maintained odor responsiveness and that its synchronous activities could be reliably recorded by the MEA system. Upon mounting the OE tissue onto the MEA chip, both spontaneous and odor-evoked LFPs were recorded. LFP oscillations represented the synchronous rhythmic activity of a large population of neurons in the immediate vicinity of the recording electrodes. As shown in Fig. 3b, odor-evoked oscillations exhibited specific spatiotemporal distribution. While several electrodes detected miniature potential fluctuations prior to stimulation, odor presentation simultaneously triggered high-amplitude oscillations across multiple channels.

Wavelet spectrogram and spectral analysis revealed that odor stimulation induced a significant increase in power predominantly within the delta frequency range, which returned to baseline levels after washout (Figs. 3c and 3d).

To further investigate the effects of different odorants on the spatiotemporal response patterns in OE tissue, four representative cigarette additives (jasminum sambac oil, cedar wood oil, tree moss concrete, and valerian oil) were chosen as stimuli. To visualize odor selectivity, delta power ratio ($\Delta P/P_0$) heatmaps were generated by computing the relative change in delta power between the baseline and odor presentation periods (Fig. 3e). For each of the four odorants, the heatmaps revealed distinct spatial and intensity activation features. These differential spatial patterns can be attributed to the inherent biological coding mechanisms of the mammalian olfactory system. The genes encoding ORs constitute one of the most diverse gene families in mammalian genomes. (Zhang and Firestein, 2002) Each OSN randomly expresses only one OR gene, which is known as the “one-neuron-one-receptor” rule. (Mombaerts, 2004; Tan et al., 2015) Nevertheless, single receptors can bind various odorants and individual odorants can activate multiple receptors. (Ressler et al., 1994) The discrimination of different odors depends on a diverse combination of receptors. When the OE tissue is mounted on the MEA, each electrode contacts a different population of neurons expressing various types of ORs. The oscillation waveforms were determined by the cumulative contributions of every neural source, and the amplitudes were associated with the distance from the source. (Zhuang et al., 2021) Therefore, the change in odor stimulation will influence the response pattern and intensity. The successful capture of these differential spatial distributions demonstrates that the OE-based bioelectronic nose preserves the natural, cross-reactive receptor mapping of the olfactory system, and this biological fidelity provides a robust, multidimensional physiological data foundation for the subsequent machine learning-based classification of tobacco flavors.

3.3 Aroma recognition

To evaluate the ability of the OE-based bioelectronic nose to discriminate representative cigarette flavor notes, a four-class aroma recognition task was first performed using the responses evoked by floral, woody, mossy, and herbal additives. Three independent recording sets were analyzed independently, and the classification results were summarized as the mean $\pm$ SEM across the three recording sets.

The influence of feature dimensionality was first examined by constructing RF classifiers with the Top10, Top20, Top30, Top50, and Top100 ranked spectral features selected from the original 240-dimensional feature set. As shown in Fig. 4a, the classification performance was already saturated when only the Top10 features were used. The Top10 feature subset achieved the highest performance, with an average Accuracy of $96.83 \pm 1.67\%$ and a Macro F1-score of $96.82 \pm 1.67\%$. Increasing the number of selected features beyond Top10 did not further improve classification performance, indicating that the most discriminative information for aroma recognition was concentrated in a small number of spectral features.

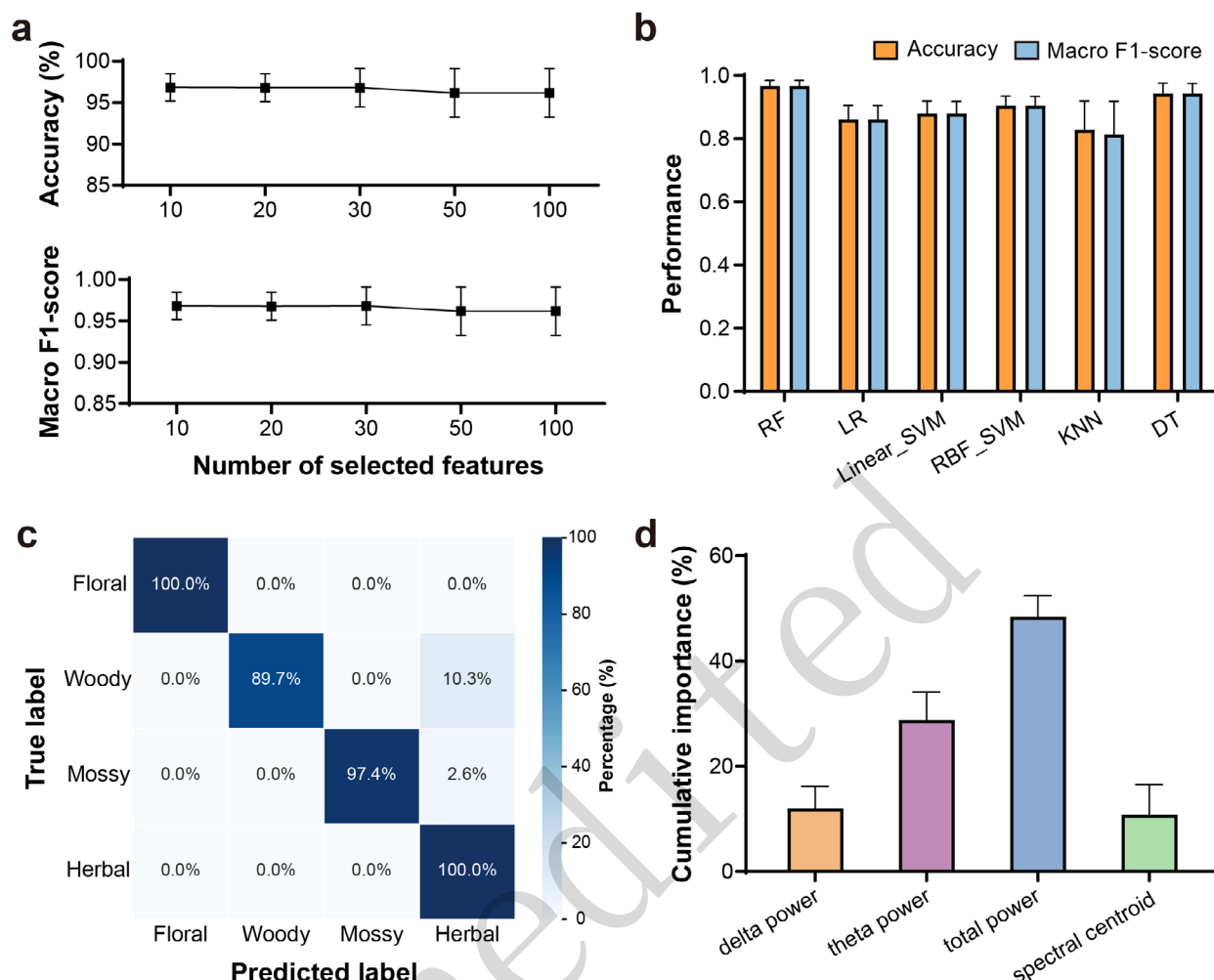


Fig. 4 Machine learning-based aroma recognition using MEA spectral features. (a) Classification performance of the RF classifier across varying quantities of spectral features, with the best performance achieved when using the Top10 feature subset. (b) Comparative evaluation of different machine learning classifiers based on the selected Top10 features. (c) Row-normalized confusion matrix constructed by pooling the test-window predictions from the three independently trained and evaluated RF models, one for each recording set. (d) Cumulative importance of the four spectral feature types within the Top10 feature subset. Error bars represent the SEM derived from three separate recording sets.

Based on the above result, the Top10 feature subset was used for classifier comparison. As shown in Fig. 4b, RF achieved the best overall performance among the tested classifiers, with an Accuracy of $96.83 \pm 1.67\%$ and a Macro F1-score of $96.82 \pm 1.67\%$. The performance of DT was also relatively high, whereas those of LR, Linear SVM, RBF SVM, and KNN were lower. These results indicate that RF is well suited for decoding aroma-evoked spectral response patterns. The confusion matrix further confirmed the high discriminative capability of the system for aroma recognition (Fig. 4c). Floral and Herbal samples were completely correctly classified. Woody samples were mainly correctly classified, with a small proportion misclassified as Herbal, while Mossy samples showed only minor misclassification. Overall, the four aroma categories were clearly separated. To interpret the RF-selected features, the cumulative importance of the Top10 feature subset was analyzed according to the four spectral feature types (Fig. 4d). Total spectral power contributed the largest proportion of cumulative importance ($48.38 \pm 4.06\%$), followed by theta power ($28.77 \pm 5.37\%$). In contrast, delta power and spectral centroid showed lower contributions, accounting for $12.01 \pm 4.17\%$ and $10.84 \pm 5.69\%$,

respectively. These results suggest that aroma recognition was mainly driven by changes in overall spectral energy and theta band-specific activity. Together, these findings demonstrate that the OE-based bioelectronic nose can accurately discriminate representative tobacco flavor notes using a compact set of interpretable spectral features.

3.4 Commercial cigarette recognition

After validating the aroma recognition capability of the OE-based bioelectronic nose using representative flavor additives, we examined its ability to discriminate complex commercial cigarette smoke samples. We first evaluated whether compact subsets of RF-ranked spectral features were sufficient for the reliable classification of these complex stimuli. As shown in Fig. 5a, the Top10 feature subset was insufficient to fully capture the discriminative information required for commercial cigarette recognition. As more RF-ranked spectral features were included, the classification performance increased and reached the highest value with the Top50 feature subset. The Top50 feature subset achieved an average Accuracy of $94.36 \pm 0.85\%$ and a Macro F1-score of $94.40 \pm 0.74\%$, whereas further increasing the feature number to Top100 did not improve the performance. This suggests that, compared with representative aroma additives, commercial cigarette smoke evokes more complex and heterogeneous electrophysiological response patterns, requiring a broader set of spectral features for reliable discrimination.

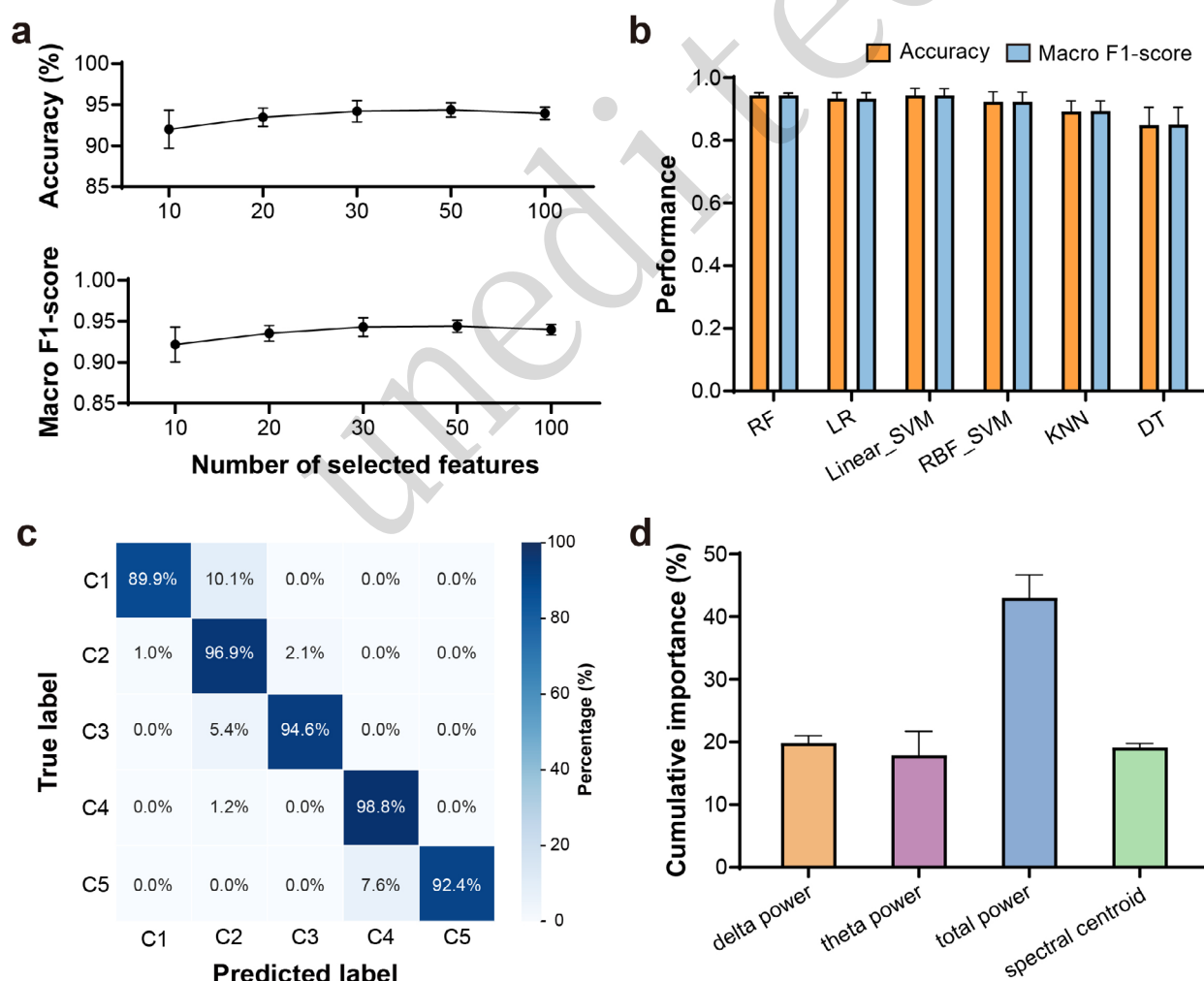


Fig. 5 Machine learning-based commercial cigarette recognition using MEA spectral features. (a) Classification per-

formance of the RF classifier across varying quantities of spectral features, with the best performance achieved when using the Top50 feature subset. (b) Comparative evaluation of different machine learning classifiers based on the selected Top50 features. (c) Row-normalized confusion matrix constructed by pooling the test-window predictions from the three independently trained and evaluated RF models, one for each recording set. (d) Cumulative importance of the four spectral feature types within the Top50 feature subset. Error bars represent the SEM derived from three separate recording sets.

The classifier comparison based on the Top50 feature subset is shown in Fig. 5b. RF and Linear SVM achieved comparable performance in commercial cigarette recognition; it showed a slightly higher average Accuracy, whereas RF achieved a slightly higher Macro F1-score. Considering its consistently high performance across both classification tasks, RF was used for the final confusion matrix analysis. The pooled and row-normalized confusion matrix further confirmed the discriminative capability of the RF model for commercial cigarette recognition (Fig. 5c). Most samples were correctly assigned to their corresponding cigarette categories, with only minor misclassifications mainly occurring between C1 and C2, C2 and C3, and C5 and C4. These results indicate that although several commercial cigarette products evoked partially similar electrophysiological response patterns, the overall decision boundaries among the five categories remained well-defined.

To interpret the basis of RF feature selection, the cumulative importance of the Top50 feature subset was analyzed according to the four spectral feature types (Fig. 5d). Total spectral power contributed the largest proportion of cumulative importance ($43.07 \pm 3.62\%$), followed by delta power ($19.84 \pm 1.18\%$), spectral centroid ($19.15 \pm 0.68\%$), and theta power ($17.95 \pm 3.77\%$). Compared with aroma recognition, the feature importance distribution in commercial cigarette recognition was more balanced across different spectral feature types, suggesting that the discrimination of complex commercial cigarette smoke samples relies on more distributed spectral information rather than a small number of dominant feature types.

4 Conclusions

Although the composition of complex cigarette smoke can be identified using techniques such as gas chromatography–mass spectrometry (GC-MS), the evaluation of its aromas still relies on human description and scoring. However, as human sensory evaluation is subjective and susceptible to various factors, an objective analytical method is needed to distinguish different cigarette categories and their specific flavor notes. In this study, we constructed a bioelectronic nose based on OE tissue coupled with a MEA to detect four representative cigarette additives and five distinct commercial cigarettes from the Liqun brand. The MEA chip successfully recorded olfactory information from multiple sites within the OE tissue as electrical signals. We found that this system could preserve the combinatorial coding mechanism of the biological olfactory system, generating distinct response patterns to different stimuli. Subsequently, we utilized multiple machine learning classifiers to accurately recognize and classify these electrophysiological signals. For odor-type recognition involving four fundamental cigarette additives, the RF classifier achieved a superior accuracy of 96.83% using a highly informative subset of only 10 spectral features, indicating that specific flavor notes elicit distinct and easily separable neural activity patterns. For the more challenging task of commercial cigarette recognition, the system accurately discriminated among five different complex smoke profiles with an accuracy of 94.36%. Notably, this complex task required a broader subset of 50 features, highlighting that the neural encoding of complex aerosols is more distributed and relies on a richer combination of spectral dynamics. In summary, this study proves the feasibility and high performance of combining intact OE tissues with MEA interfaces and ML algorithms for complex odor analysis. The proposed OE-based bioelectronic nose acts as a highly sensitive and objective platform for cigarette aroma recognition, offering broad application prospects in food evaluation and quality control.

Data availability statement

The authors declare that the data supporting the findings of this study are included within this article. If any raw data files should be needed in another format, they are available from the corresponding author upon reasonable request.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2024YFB3212300), Science and Technology Project of China Tobacco Zhejiang Industrial Co., Ltd (ZJZY-2023C019), and Foundation of the State Key Laboratory of Transducer Technology (No. SKT2520).

Author contributions

Yongsheng LI, Hao WAN, Liujiang ZHUANG, Ping WANG, and Nan JIANG conceived and designed the experiments; Yongsheng LI, Xinrui TANG, Tan Shuan GIN, and Xuliang FU performed the experiments and analyzed the data; Qingxiang LI, Chengxu ZHENG, Zhenjie ZHAO, Jianguo WU, and Deming JIANG participated in the experiments; Yongsheng LI, Xinrui TANG, Yingying XUE, and Nan JIANG drafted and revised the manuscript. All authors have read and approved the final manuscript, and therefore, have full access to all the data in the study and take responsibility for the integrity and security of the data.

Compliance with ethics guidelines

Liujiang ZHUANG is a Young Scientist Committee Member for Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) and was not involved in the editorial review or the decision to publish this article. Yongsheng LI, Xinrui TANG, Qingxiang LI, Chengxu ZHENG, Zhenjie ZHAO, Yingying XUE, Xuliang FU, Jianguo WU, Deming JIANG, Hao WAN, Liujiang ZHUANG, Ping WANG, and Nan JIANG declare that they have no conflicts of interest.

All animal studies were approved by the Institutional Animal Care and Use Committee and the Ethics Committee of Zhejiang University, Hangzhou, China (ZJU20240774).

Declaration on the use of generative AI tools

No generative AI tools were used in the preparation of this manuscript.

References

- Breiman L, 2001. Random forests. *Machine Learning*, 45(1):5-32. <https://doi.org/10.1023/A:1010933404324>
- Dung TT, Oh Y, Choi S-J, et al., 2018. Applications and advances in bioelectronic noses for odour sensing. *Sensors*, 18(1):103. <https://doi.org/10.3390/s18010103>
- Firestein S, 2001. How the olfactory system makes sense of scents. *Nature*, 413(6852):211-218. <https://doi.org/10.1038/35093026>
- Guo L, Cheng J, Lian S, et al., 2023. Structural basis of amine odorant perception by a mammal olfactory receptor. *Nature*, 618(7963):193-200. <https://doi.org/10.1038/s41586-023-06106-4>
- Iravani B, Arshamian A, Lundqvist M, et al., 2021. Odor identity can be extracted from the reciprocal connectivity between olfactory bulb and piriform cortex in humans. *NeuroImage*, 237:118130. <https://doi.org/10.1016/j.neuroimage.2021.118130>
- Jiang H, Schuele S, Rosenow J, et al., 2017. Theta oscillations rapidly convey odor-specific content in human piriform cortex. *Neuron*, 94(1):207-219.e204. <https://doi.org/10.1016/j.neuron.2017.03.021>
- Lledo P-M, Gheusi G, Vincent J-D, 2005. Information processing in the mammalian olfactory system. *Physiological Reviews*, 85(1):281-317. <https://doi.org/10.1152/physrev.00008.2004>
- Ma J, Guan Y, Xing F, et al., 2023. Accurate and non-destructive monitoring of mold contamination in foodstuffs based on whole-cell biosensor array coupling with machine-learning prediction models. *Journal of Hazardous Materials*, 449:131030. <https://doi.org/10.1016/j.jhazmat.2023.131030>
- Mombaerts P, Wang F, Dulac C, et al., 1996. Visualizing an olfactory sensory map. *Cell*, 87(4):675-686. [https://doi.org/10.1016/s0092-8674\(00\)81387-2](https://doi.org/10.1016/s0092-8674(00)81387-2)
- Mombaerts P, 2004. Odorant receptor gene choice in olfactory sensory neurons: The one receptor-one neuron hypothesis revisited. *Current Opinion in Neurobiology*, 14(1):31-36. <https://doi.org/10.1016/j.conb.2004.01.014>
- Neta S, Ariel G, Yossi Y, et al., 2023. The locust antenna as an odor discriminator. *Biosensors and Bioelectronics*, 221:114919. <https://doi.org/10.1016/j.bios.2022.114919>
- Qin C, Chen C, Yuan Q, et al., 2022. Biohybrid tongue for evaluation of taste interaction between sweetness and sourness. *Analytical Chemistry*, 94(19):6976-6985. <https://doi.org/10.1021/acs.analchem.1c05384>

- Qin C, Wang Y, Hu J, et al., 2023. Artificial olfactory biohybrid system: An evolving sense of smell. *Advanced Science*, 10(5):2204726. <https://doi.org/https://doi.org/10.1002/advs.202204726>
- Ressler KJ, Sullivan SL, Buck LB, 1994. Information coding in the olfactory system: Evidence for a stereotyped and highly organized epitope map in the olfactory bulb. *Cell*, 79(7):1245-1255. [https://doi.org/https://doi.org/10.1016/0092-8674\(94\)90015-9](https://doi.org/https://doi.org/10.1016/0092-8674(94)90015-9)
- Shor E, Herrero-Vidal P, Dewan A, et al., 2022. Sensitive and robust chemical detection using an olfactory brain-computer interface. *Biosensors and Bioelectronics*, 195:113664. <https://doi.org/https://doi.org/10.1016/j.bios.2021.113664>
- Tallon-Baudry C, Bertrand O, Delpuech C, et al., 1997. Oscillatory gamma-band (30-70 hz) activity induced by a visual search task in humans. *Journal of Neuroscience*, 17(2):722-734. <https://doi.org/10.1523/jneurosci.17-02-00722.1997>
- Tan L, Li Q, Xie XS, 2015. Olfactory sensory neurons transiently express multiple olfactory receptors during development. *Molecular Systems Biology*, 11(12):844. <https://doi.org/https://doi.org/10.15252/msb.20156639>
- Thomas CA, Springer PA, Loeb GE, et al., 1972. A miniature microelectrode array to monitor the bioelectric activity of cultured cells. *Experimental Cell Research*, 74(1):61-66. [https://doi.org/https://doi.org/10.1016/0014-4827\(72\)90481-8](https://doi.org/https://doi.org/10.1016/0014-4827(72)90481-8)
- Wei X, Qin C, Gu C, et al., 2019. A novel bionic *in vitro* bioelectronic tongue based on cardiomyocytes and microelectrode array for bitter and umami detection. *Biosensors and Bioelectronics*, 145:111673. <https://doi.org/https://doi.org/10.1016/j.bios.2019.111673>
- Zhang X, Firestein S, 2002. The olfactory receptor gene superfamily of the mouse. *Nature Neuroscience*, 5(2):124-133. <https://doi.org/10.1038/nn800>
- Zhang X, Cheng J, Wu L, et al., 2018. An overview of an artificial nose system. *Talanta*, 184:93-102. <https://doi.org/https://doi.org/10.1016/j.talanta.2018.02.113>
- Zhu P, Chen Y, Wang M, et al., 2025. An *in vivo* bioelectronic nose utilizing support vector machine techniques for quantitative assessment of the odor detection performance. *Chemical Engineering Journal*, 523:168346. <https://doi.org/https://doi.org/10.1016/j.cej.2025.168346>
- Zhuang L, Zhang B, Qin Z, et al., 2019. Nasal respiration is necessary for the generation of γ oscillation in the olfactory bulb. *Neuroscience*, 398:218-230. <https://doi.org/https://doi.org/10.1016/j.neuroscience.2018.12.011>
- Zhuang L, Wei X, Jiang N, et al., 2021. A biohybrid nose for evaluation of odor masking in the peripheral olfactory system. *Biosensors and Bioelectronics*, 171:112737. <https://doi.org/https://doi.org/10.1016/j.bios.2020.112737>

Supplementary information

Tables S1 and S2