



# Co-optimization of algorithm and sensor array for precise diagnosis of HER2-positive breast cancer and trastuzumab response prediction

Tianyu Zeng<sup>1</sup> · Zeying Li<sup>1</sup> · Yanting Sun<sup>1</sup> · Yincheng Liu<sup>2</sup> · Jialin Xu<sup>1</sup> · Xiang Huang<sup>1</sup> · Yan Liang<sup>1</sup> · Hai Shi<sup>3</sup> · Shuai Wu<sup>4</sup> · Genxi Li<sup>5,6</sup> · Yongmei Yin<sup>1</sup>

Received: 14 October 2025 / Revised: 5 January 2026 / Accepted: 10 January 2026  
© The Author(s) 2026

## Abstract

Accurate and rapid diagnosis of human epidermal growth factor receptor-2 (HER2)-positive breast cancer, coupled with prediction of trastuzumab therapeutic efficacy, is critical for clinical decision-making to the patients with breast cancer. However, there is still no standard to be clinically used without suffering from inherent limitations. In this work, we propose a machine learning-assisted multifunctional biosensing platform utilizing enzyme-embedded hydrogen-bonded organic frameworks (HOFs). In this design, diverse HOFs@enzyme composites with distinct assembly configurations serve as sensitive array elements to interact with breast cancer-derived exosomes. Moreover, these interactions can modulate HOF-enzyme activity, generating diagnostic signal patterns that form unique exosomal molecular “fingerprint” profiles. Simultaneously, coordination with machine learning enables processing of complex sensor array-based data to amplify subtle differences of exosome between different subtypes of breast cancer, thereby enhancing the discriminatory capacity of this platform. By establishing reference fingerprints using exosomes from 96 training-set patients and validating classification accuracy against immunohistochemical in 76 test-set patients, the platform achieved 100% concordance in identifying the HER2-positive subtype, demonstrating exceptional discriminative capacity. Remarkably, the platform can also predict trastuzumab treatment response with 87.5% accuracy through clinical outcome correlation. So, by enabling precise exosome characterization from peripheral blood, this non-invasive liquid biopsy technology offers a transformative approach for precision oncology in HER2-positive breast cancer, overcoming critical limitations of current diagnostic paradigms.

**Keywords** Sensor array · Hydrogen-bonded organic frameworks · HER2-positive breast cancer · Exosome · Machine learning

## 1 Introduction

Human epidermal growth factor receptor 2 (HER2)-positive breast cancer (HER2<sup>+</sup> BC) represents 20–30% of all breast cancer cases and is clinically characterized by aggressive

tumor behavior, rapid disease progression, and poor patient outcomes [1–3]. Drugs targeting HER2, such as trastuzumab, have significantly improved the poor clinical outcomes of HER2<sup>+</sup> BC patients and reshaped the diagnosis and treatment landscape of breast cancer [4–6]. While

✉ Shuai Wu  
shwu@njmu.edu.cn

✉ Genxi Li  
genxililab@nju.edu.cn

✉ Yongmei Yin  
ymyin@njmu.edu.cn

<sup>1</sup> Department of Oncology, The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, P.R. China

<sup>2</sup> Department of Plastic and Reconstructive Surgery, Shanghai Geriatric Medical Center, Shanghai 200444, P.R. China

<sup>3</sup> Department of Biology, School of Basic Medical Sciences, Guizhou Medical University, Guiyang 561113, P.R. China

<sup>4</sup> Clinical Medicine Research Institute, The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, P.R. China

<sup>5</sup> State Key Laboratory of Analytical Chemistry for Life Science, School of Life Sciences, Nanjing University, Nanjing 210023, P.R. China

<sup>6</sup> Center for Molecular Recognition and Biosensing, School of Life Sciences, Shanghai University, Shanghai 200444, P.R. China

the appearance of trastuzumab has revolutionized treatment paradigms, providing significant survival benefits for HER2<sup>+</sup> BC, therapeutic resistance remains a major clinical challenge. Approximately 25–30% of patients exhibit primary resistance, while the majority of initially responsive advanced-stage patients develop acquired resistance within 12 months of treatment initiation. These resistance mechanisms frequently lead to disease recurrence and metastatic dissemination, ultimately contributing to cancer-related mortality [7, 8]. Therefore, there is an urgent need for novel diagnostic approaches that can accurately identify treatment-sensitive populations and guide personalized therapeutic strategies.

Current diagnostic standards for breast cancer rely on immunohistochemical (IHC) analysis of tumor biopsies [9], but the time-consuming, invasive nature, heterogeneity of tumor, and inconsistent test reagents may profoundly frustrate the subjective judgment. Moreover, dynamic molecular changes during treatment creates a major challenge to cancer management [10–14]. In recent years, exosomes have gained prominence as particularly promising circulating biomarkers due to their stable carriage of parent cell molecular signatures, ubiquitous presence in bodily fluids, and demonstrated elevation in breast cancer patients relative to healthy controls [15–22]. However, the multifaceted heterogeneity among distinct breast cancer subtypes cannot be sufficiently characterized by single or even combinatorial exosomes, necessitating the development of novel detection platforms capable of resolving holistic inter-exosomal variations to enhance exosome identification accuracy. Alternatively, mammalian gustatory system-inspired sensor arrays have gained significant research interest [23]. These platforms employ nonspecific interactions with samples to generate characteristic signal fingerprints, permitting both discrimination of multiple analytes with minor difference in structure or properties and comprehensive analysis of complex biological specimens [24, 25].

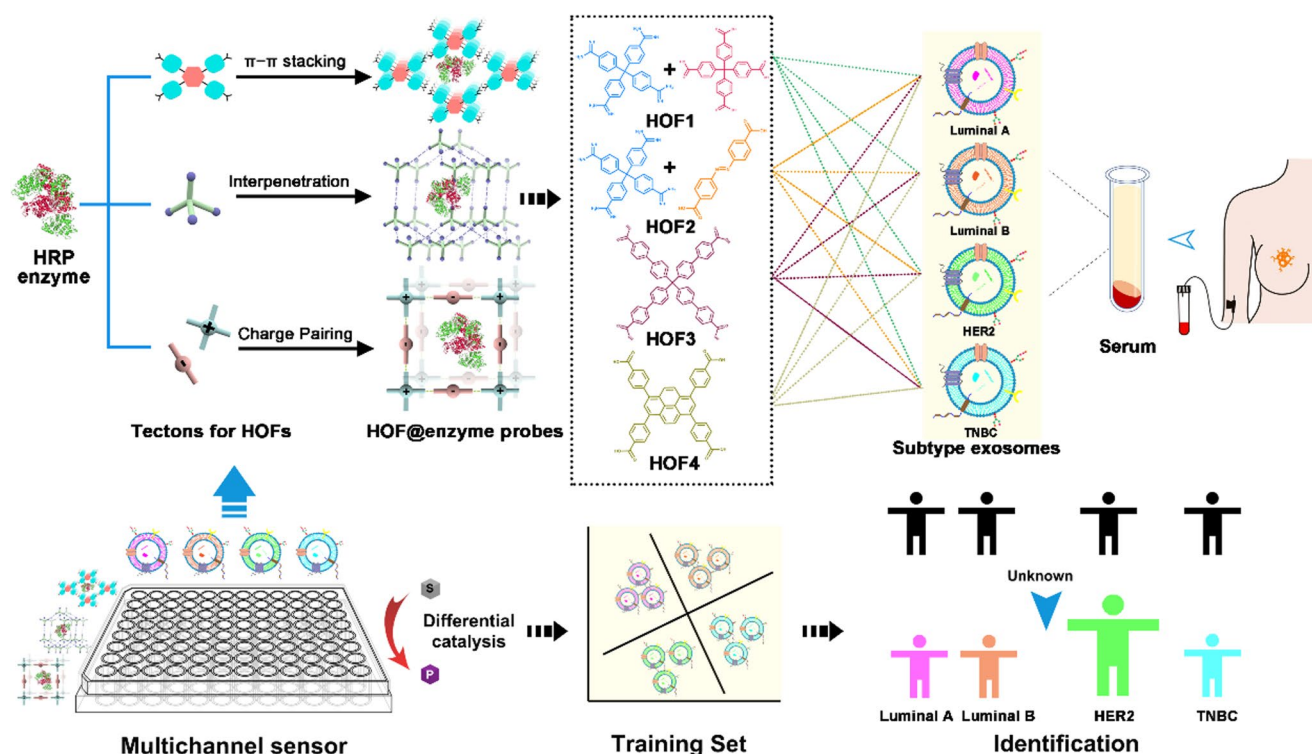
Hydrogen-bonded organic frameworks (HOFs), crystalline porous materials formed through coordinated intermolecular interactions including hydrogen bonding, electrostatic forces, and  $\pi$ - $\pi$  stacking, present an innovative solution to the fabrication of sensor array. Their structural stability, synthetic versatility, and tunable porosity make HOFs ideal for biosensor applications [26–28]. HOFs permit fabrication using benign synthetic conditions and demonstrate superior biocompatibility due to their metal-free composition, significantly enhancing their potential for clinical translation [29]. Furthermore,  $\pi$ -conjugated molecules building blocks endow certain HOFs with intrinsic luminescent properties, enabling their successful deployment as fluorescent probes for detecting diverse analytes ranging from organic molecules to metal

ions [30–32]. Building upon these advantages, we have engineered a novel HOF@enzyme composite in this work that simultaneously functions as both recognition and signal transduction element, enabling direct exosomal fingerprint generation through multidimensional signal output analysis. Moreover, the fabricated sensor array permits concurrent discrimination of structurally similar analytes while maintaining robust performance in intricate clinical specimens. In the meantime, the sensor array eliminates the need for complex probe design by employing four HOF@enzyme sensor units, which generate a  $3 \times 4$  data matrix upon exosome interaction to construct exosomal fingerprints. Besides, capitalizing on exceptional capacity for processing complex datasets and analytical advantages of machine learning (ML), we have further implemented ML to decode the sensor array-derived exosomal fingerprints, enabling sensitive and high-throughput exosome characterization. Our advanced sensor array has demonstrated dual clinical functionality, enabling not only accurate HER2<sup>+</sup> BC diagnosis but also reliable prediction of trastuzumab treatment efficacy, which represents a significant advancement over current diagnostic standards by combining non-invasive sampling with comprehensive molecular profiling and therapeutic response forecasting. The operational simplicity and rapid analysis time also address critical unmet needs in precision oncology, offering real-time treatment monitoring capabilities that could fundamentally transform HER2<sup>+</sup> BC management paradigms.

## 2 Results and discussion

### 2.1 Design principle of the sensor array using HOF@enzyme

The foundational premise of this study stems from the well-established molecular distinctions among the four major breast cancer subtypes. Focusing specifically on exosomes derived from trastuzumab-sensitive, primary resistant, and secondary resistant phenotypes (Fig. S1), our preliminary proteomic and metabolomic analyses have revealed extensive differential expression patterns that cannot be adequately captured through single-target detection methods. Align with the MISEV2018/2023 guidelines [33], exosomes may be difficult to separate from other vesicles in blood samples. Therefore, ExoQuick exosome precipitation solution (EXOQ20A-1) rather than ExoQuick ULTRA EV Isolation Kit for Serum and Plasma (EQUltra-20 A-1) has been chosen to isolate exosome. At the core of this technology lies a four-channel HOF@enzyme-based sensor array designed for HER2<sup>+</sup> BC identification and trastuzumab



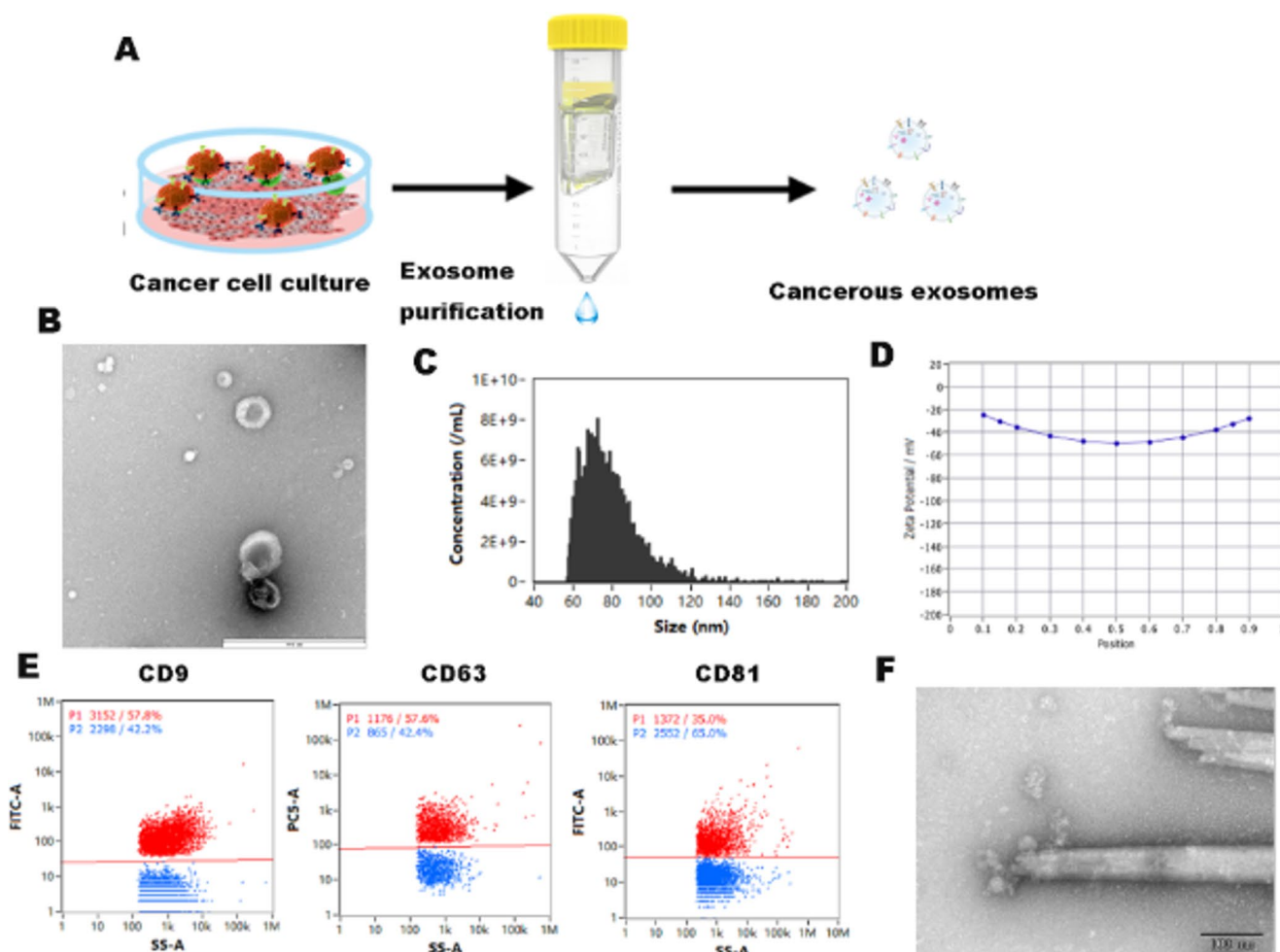
**Scheme 1** Schematic illustration of exosomal fingerprints identification for HER2<sup>+</sup> BC through HOFs sensors and ML

response prediction. As shown in Scheme 1, the sensor array exploits the unique molecular engineering flexibility of HOFs, which can form tunable probe units through precisely controlled co-assembly with catalytic enzymes. This assembly process is governed by three key intermolecular interactions:  $\pi$ - $\pi$  stacking, interpenetration, and charge pairing. Critically, the catalytic activity of these HOF@enzyme complexes demonstrates exquisite sensitivity to subtle variations in these interactions - a property we may harness to create responsive sensor elements. When exosomes interact with the HOF@enzyme, distinct surface markers induce conformational rearrangements that differentially expose catalytic sites. This generates subtype-specific modulation of enzymatic substrate conversion, producing a multidimensional response pattern ( $3 \times 4$  data matrix) that serves as a unique exosomal fingerprint. We have further enhanced this sensor array through integration with machine learning algorithms, which process the multichannel optical signals (measured via microplate reader) to achieve dual functionality: accurate HER2<sup>+</sup> BC diagnosis and quantitative prediction of trastuzumab therapeutic efficacy directly from serum exosomes. This combination of nanoscale engineering and computational analytics represents a significant advancement over current single-analyte detection paradigms, offering both high specificity (> 90% accuracy in validation studies) and operational simplicity suitable for clinical translation.

## 2.2 Characterization of HOFs and exosomes

To systematically investigate the detection efficacy of the array-based platform, exosomes derived from cell lines were utilized as a validation model (Fig. 1A). Firstly, transmission electron microscopy (TEM) characterization confirmed the intact structures of exosomes (Fig. 1B). Nanoparticle tracking analysis (NTA) revealed exosome concentrations averaging  $1.4 \times 10^{10}$  particles/mL with a mean diameter of 82 nm (Fig. 1C). The measured zeta potential was  $-30.86$  mV at physiological pH (Fig. 1D), which also allowed for electrostatic adsorption with HOF@enzyme. Nano-Flow cytometry analysis also demonstrated high CD9, CD63 and CD81 levels in the exosomes (Fig. 1E), which indicated the successful extraction and stability of exosomes.

Subsequently, the nanostructure and morphology of the HOFs@enzyme sensor unit (HOF1-HOF4) were observed by TEM, which present regular structures ranged from 200 to 500 nm (Fig. S2). The crystal structures of four sensor units were further characterized by X-ray diffraction (XRD) (Fig. S3). Fourier transform infrared (FT-IR) spectroscopy displayed the chemical compositions of HOFs@enzyme which further confirmed the successful synthesis and stability of the nanosensors (Fig. S4). After verifying the characterization of exosomes and sensor array, we co-incubated the exosomes with the HOFs@enzyme. It was observed that the exosomes would bind to the surface of the HOFs@enzyme



**Fig. 1** Characterization of exosomes. (A) Scheme of the extraction process of cancerous exosomes. (B) TEM images of exosomes. Scale bars: 200 nm. (C) NTA results of exosomes. (D) Zeta potential of

exosomes at physiological pH. (E) Nano-Flow cytometry analysis of CD9, CD63 and CD81 levels. (F) TEM images of exosomes after the interaction with HOFs. Scale bars: 100 nm

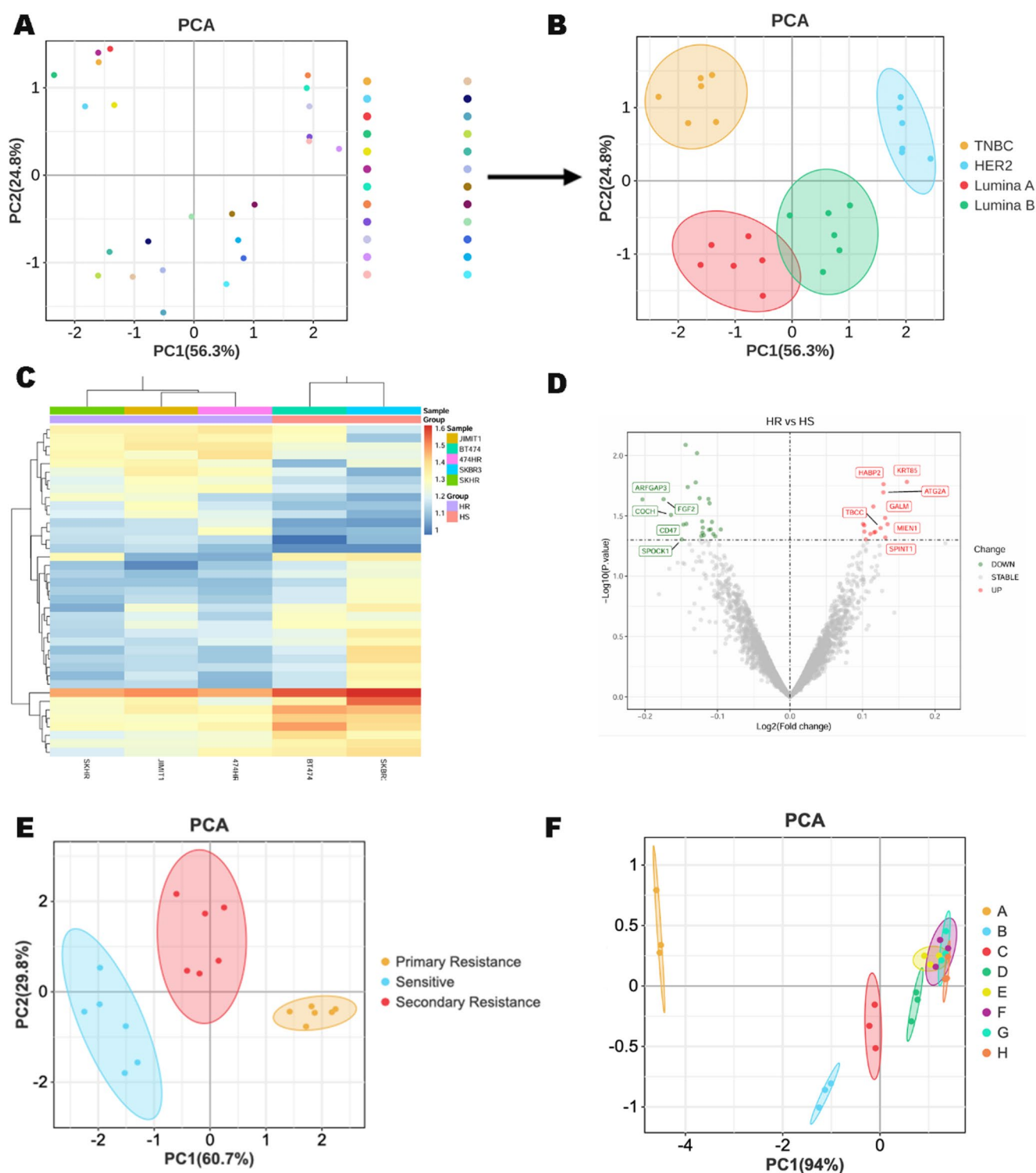
(Fig. 1F). The experiments of structural analysis like CD and Brunauer-Emmett-Teller (BET) surface area analysis were also performed. As shown in Fig. S5, it can be confirmed that the secondary structure of the protein remained unchanged within HOFs. However, after incubation with exosomes, changes were observed in the conformational distribution of HOF@enzyme. Nitrogen adsorption (Fig. S6, Table S1) further demonstrated that exosome interaction would alter the pore structure of HOF@enzyme, thereby affecting its ability to facilitate TMB diffusion and generating different signal intensities.

### 2.3 Optimization of experimental conditions

After verifying the characterization of exosomes and HOFs@enzyme respectively, we preliminarily verified the feasibility of sensor array through cancer cells derived exosomes. Firstly, exosomes secreted by the T47D, MCF-7, ZR-75-1, BT474, MDA-MB-453, SKBR3, HCC1806 and

MDA-MB-231 were extracted (T47D and MCF-7 were luminal A, ZR-75-1 and BT474 were luminal B, MDA-MB-453 and SKBR3 were HER2+, HCC1806 and MDA-MB-231 were triple negative breast cancer (TNBC)). These exosomes were added to the sensor array to produce a multidimensional dataset comprising four distinct probes, six replicates, and four cellular subtypes. Modern breast cancer classification leverages multiple machine learning algorithms. Principal component analysis (PCA) was employed for preliminary data exploration and feature extraction. As shown in Fig. 2A and B, PCA converted the original scattered data into a discrimination map. It can be seen that our sensor array can well classify the exosomes derived from the 8 cell lines into four categories, which exactly correspond to the molecular typing of the corresponding cells. Considering that the expressions of Luminal A and Luminal B type molecules are closer, it is expected that the distances of these two clusters are close and slightly overlapping, which indicates the potential of the future application of the





**Fig. 2** Analysis of cell-derived exosomes. **(A)** All exosomal fingerprint distributions. **(B)** PCA of exosomes secreted by four molecular subtypes of breast cancer cells demonstrates distinct clustering patterns. **(C)** Heatmap visualization of proteomic profiles reveals differentially expressed proteins in exosomes derived from cells with varying trastuzumab response phenotypes. HR: trastuzumab-resistant cells; HS: trastuzumab-sensitive cells. **(D)** Volcano plot analysis identifies statistically significant protein expression changes ( $p < 0.05$ ) between

trastuzumab-sensitive and resistant exosome groups. **(E)** PCA score plot of exosomes stratified by trastuzumab treatment efficacy shows separation between sensitive, primary resistance and secondary resistance populations. **(F)** Concentration-dependent response curves of the sensor array. (A:  $1 \times 10^8$  particles/mL, B:  $1 \times 10^7$  particles/mL, C:  $1 \times 10^6$  particles/mL, D:  $1 \times 10^5$  particles/mL, E:  $1 \times 10^4$  particles/mL, F:  $1 \times 10^3$  particles/mL, G:  $1 \times 10^2$  particles/mL, H: 10 particles/mL.)

four-channel sensor array. To mitigate the challenge of standardization in exosome isolation, we implemented an ultracentrifugation protocol to isolate exosomes for downstream detection. The consistency of the results could validate the robustness of our detection approach despite inherent isolation variables. Subsequently, further condition optimization will be carried out to improve the accuracy rate of detection.

To balance probe economy with precise identification of nuanced exosomal differences, we optimized the concentration of the HOF@enzyme probes, and the results showed the catalytic efficiency increased gradually with higher concentrations of HOFs@enzyme and prolonged catalytic times, reaching a peak efficiency at the HOF1 concentration of 100  $\mu\text{g/mL}$  (Fig. S7A), HOF2 concentration of 80  $\mu\text{g/mL}$  (Fig. S7B), HOF3 concentration of 20  $\mu\text{g/mL}$  (Fig. S7C) and HOF4 concentration of 40  $\mu\text{g/mL}$  (Fig. S7D). At the same time, the catalytic reaction reaches its peak at 5 min when the under the above optimal conditions, meeting the requirements of rapid detection and avoiding long waiting times (Fig. S8). In addition, the incubation time of exosomes is another key factor affecting the recognition efficiency. It was observed that the incubation period with HOFs@enzyme for achieving maximum interaction efficiency could be reduced to 30 min with continuous gentle rotation (Fig. S9).

For HER2<sup>+</sup> BC, trastuzumab has become an indispensable component of treatment across all disease stages - whether as neoadjuvant/adjuvant therapy for early-stage disease or salvage therapy for advanced cases [34–37]. While it has significantly improved survival outcomes for most patients, the emergence of primary or acquired resistance in some individuals remains a major obstacle to precision medicine in HER2<sup>+</sup> BC. To address this challenge, we attempted to draw the fingerprint patterns of exosomes with different efficacy of trastuzumab. We first isolated exosomes secreted from: (1) trastuzumab-sensitive cells (BT474 and SKBR3), (2) primary trastuzumab-resistant cells (JIMT1), and (3) induced trastuzumab-resistant cells (BT474-HR and SKBR3-HR) to establish comprehensive models of drug sensitivity, primary resistance, and acquired resistance. Drug resistance in HER2<sup>+</sup> breast cancer often arises from a combination of genetic, epigenetic, microenvironmental, and metabolic adaptations. Integrated proteomic and metabolomic analyses revealed significant differential expression of multiple proteins (Fig. 2C and D) and metabolites (Fig. S1) among these exosome subgroups ( $p < 0.05$ ). Isolating resistance-specific signals from exosomal profiles remains challenging due to the overlapping signals and cellular heterogeneity. To systematically characterize these variations at a global level, we incorporated all three exosome types into our array-based sensing platform. PCA revealed clear segregation among sample clusters, enabling clear

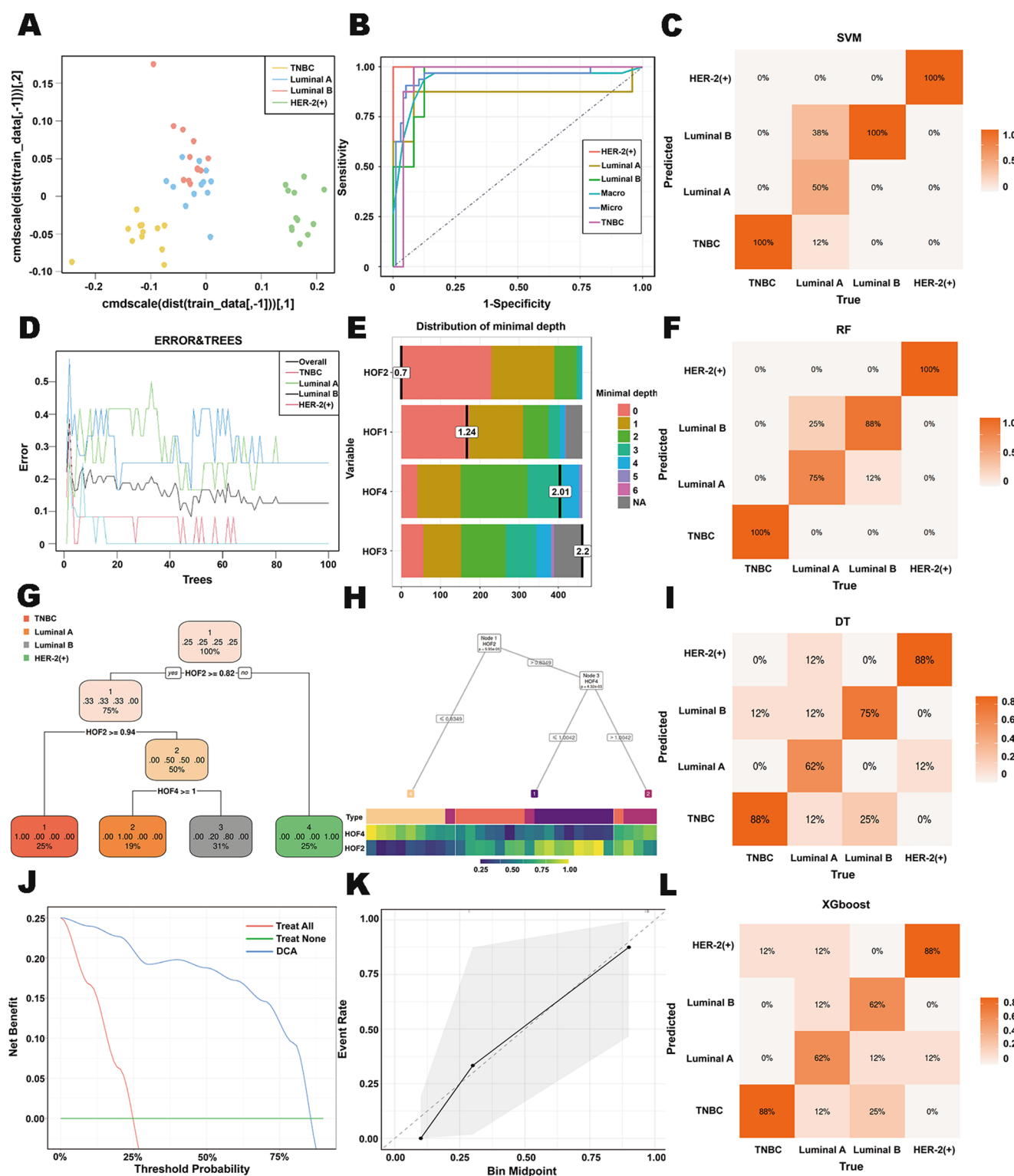
discrimination among exosomes from different resistance phenotypes (Fig. 2E). These findings strongly suggest the potential utility of this sensing array for predicting trastuzumab treatment outcomes.

A concentration-response relationship was evident in the 2D plot (Fig. 2F), with distinct clusters corresponding to different exosomal levels, demonstrating quantitative analysis capability of the platform. Remarkably, the array exhibited detection sensitivity down to  $1 \times 10^4$  particles/mL. Comparatively, the detection limit of an aptamer-based metasurfaces for label-free plasmonic biosensing of breast tumor-derived exosomes is as low as  $1.04 \times 10^4$  particle/mL [38]. Another an colorimetric Detection of HER2-over-expressing-cancer-derived exosomes is as low as  $8.5 \times 10^8$  particle/mL [39]. Considering that the exosomes extracted from 250  $\mu\text{L}$  of serum can reach more than  $10^{11}$  particles/mL in clinical practice, the sensitivity of our array and the prospect of clinical application are verified (Table S2).

## 2.4 HER2 phenotype identification among BC patients

We employed this array sensor to identify HER2<sup>+</sup> BC based on the exosomal fingerprints. All exosome samples were purified from patients with confirmed clinical diagnoses, with their identity confirmed by IHC prior to analysis. The data sets consisted of UV-vis absorption spectra of HOFs originating from their interactions with different exosomes. The absorbance at 650 nm was quantified for each HOFs component, generating a response matrix that captured the interaction profiles with individual exosome samples. Random stratified splitting was employed to create training and testing datasets for subsequent applications.

To strengthen the adaptability to variable input patterns, we systematically evaluated the ML algorithms and ultimately selected support vector machines (SVM), random forest (RF) analysis, decision tree (DT) and XGBoost for exosomal fingerprint analysis based on their complementary analytical strengths [40, 41]. All the models were tuned exclusively on the training set, with the test set kept completely isolated until the final evaluation, thereby preventing any data leakage. The SVM algorithm offers multiple kernel functions, including linear and polynomial kernels. We initially selected the optimal kernel function based on data characteristics, then fine-tuned model parameters through cross-validation to accommodate small-scale clinical samples. Operating on the maximum margin principle, SVM effectively minimizes interference from outliers while maintaining robust classification performance even with limited sample sizes. The test dataset was processed through the trained model, generating randomized planar point distribution as illustrated in Fig. 3A. The visualization demonstrates



**Fig. 3** ML-assisted HER2<sup>+</sup> exosomal fingerprint profiling. **(A)** Two-dimensional scatter plot of single-exosome distribution patterns. **(B)** ROC curve demonstrating the diagnostic performance of SVM-based exosomal fingerprint identification. **(C)** Confusion matrix evaluating SVM classification accuracy for exosomal molecular subtyping (overall accuracy=87.5%). **(D)** Representative decision tree from RF classification, generated through interaction testing. **(E)** Minimal depth distribution analysis quantifying the relative importance of each array

sensing element. **(F)** Classification accuracy of RF-based molecular subtyping (overall accuracy=90.625%). **(G)** Visualization of the DT process for the training set. **(H)** Visualization of the DT process for the validation set. **(I)** Performance metrics of DT model in molecular subtyping (overall accuracy=78.12%). **(J)** The decision curve of HER2<sup>+</sup> identification. **(K)** The calibration curve of HER2<sup>+</sup> identification. **(L)** Predictive accuracy of XGboost-based classification for molecular subtyping (overall accuracy=75.00%)

clear spatial segregation, with HER2<sup>+</sup> exosomes predominantly clustered in the right-sided region, exhibiting distinct separation from other exosomal subpopulations. The model achieved an overall prediction accuracy of 87.5% on the test set (Fig. 3B and C), with particularly outstanding performance in HER2<sup>+</sup> BC classification (100% sensitivity).

To mitigate bias from imbalanced data, we employed an RF model comprising multiple decision trees to enhance recognition capability. The RF algorithm facilitates feature selection by evaluating importance metrics. We quantified each feature's predictive contribution using partitioning indicators such as information gain and Gini coefficient, subsequently ranking these features. Visualization of the decision-making process further enhanced our understanding of critical data features. We initially investigated the relationship between the number of decision trees and the out-of-bag (OOB) error rate to minimize misclassification in the random forest algorithm (Fig. 3D). Subsequently, we employed minimal depth distribution visualization to analyze the depth characteristics of each HOF sensing unit, identifying HOF2 as the unit with the highest average minimal depth value (Fig. 3E). The optimized model demonstrated robust performance on the test dataset, achieving 90.625% of overall prediction accuracy and 100% of HER2<sup>+</sup> BC classification sensitivity (Fig. 3F). As shown in the confusion matrix and further supported by confidence interval analysis, the 95% CI of the RF model is (0.7498, 0.9802), indicating robust model performance. Statistically, the high specificity (96.875%) and F1 score (0.9059) further validate the reliability of HER2<sup>+</sup> identification. The 5-fold cross-validation was performed on the RF model using the training set. The results demonstrated consistent validation accuracy above 0.8 across all folds, with no strong evidence of overfitting observed. (Fig. S10). The calibration curve for class 4 (HER2<sup>+</sup>) confirms the credibility of the model output (Fig. S11). While slight deviations exist for other classes (Fig. S11A-C), the overall calibration trend remains satisfactory. SHAP analysis also confirms that HOF2 is the key sensor for HER2<sup>+</sup> recognition (Figs. S12 and S13).

Regarding nonlinear effects and model interaction capabilities, we have systematically evaluated more complex algorithms to examine whether they could capture higher-order feature interactions that might be missed by RF. Specifically, in addition to the original RF model, we introduced XGBoost and DT models for comparison. To further investigate the relationships among the selected features, we employed decision tree visualization on the training and validation sets to render the decision tree process interpretable, thereby making every decision path clearly traceable (Fig. 3G, H). The results were consistent with prior observations, showing partial overlap between Luminal A and Luminal B exosomal fingerprints while demonstrating

clear separation from other subtypes. Final evaluation using a confusion matrix yielded an overall accuracy of 78.12% (Fig. 3I). To further evaluate predictive performance, XGBoost was also implemented for HER2<sup>+</sup> breast cancer classification. Decision curve analysis and calibration curves were generated (Fig. 3J, K), yielding an overall prediction accuracy of 75.00% (Fig. 3L). This performance was comparatively modest when benchmarked against the RF model, which additionally showed a more robust profile characterized by a narrower confidence interval and higher specificity (Table S3).

The fingerprint profiles generated from the test set were input into the pre-established RF model. Each exosome generated an exclusive fingerprint. The fingerprint heatmap visualized the differential interactions between each HOFs variant and exosome sample. Comparative analysis demonstrated high concordance (100% sensitivity) between our machine learning-based assessment and conventional IHC results for HER2<sup>+</sup> BC classification (Fig. S14). Notably, our diagnostic platform achieved 100% accuracy in identifying TNBC, while demonstrating slightly reduced but still clinically acceptable performance in distinguishing Luminal A and Luminal B subtypes. This marginal difference in diagnostic efficiency between hormone receptor-positive subtypes carries limited clinical implications, as both subtypes typically receive endocrine therapy as standard treatment. This limitation should be framed not as a dismissal of the clinical question but as a current technical boundary of our platform. More efforts are needed to integrate additional molecular features into the sensor or model. Such integration would likely improve classification accuracy for these subtypes, thereby enhancing the tool's general applicability across diverse clinical contexts. Future refinements of the platform should focus on enhancing subtype discrimination accuracy to enable more precise molecular characterization, which could further optimize therapeutic decision-making for patients with luminal-type breast cancer. This finding robustly validates the good performance of exosomal fingerprint profiling in determining HER2<sup>+</sup> BC.

## 2.5 Trastuzumab efficacy prediction in clinical samples

While our previous work established a rapid and accurate diagnostic platform for HER2<sup>+</sup> BC, the predictive capacity for trastuzumab-based therapeutic outcomes remained unexplored. Given the demonstrated ability of our array-based sensing system to precisely distinguish exosomes derived from cells with differential treatment responses, we subsequently extended its application to predict clinical outcomes in trastuzumab-treated patients. Following the same experimental paradigm, we randomly selected 48 patients



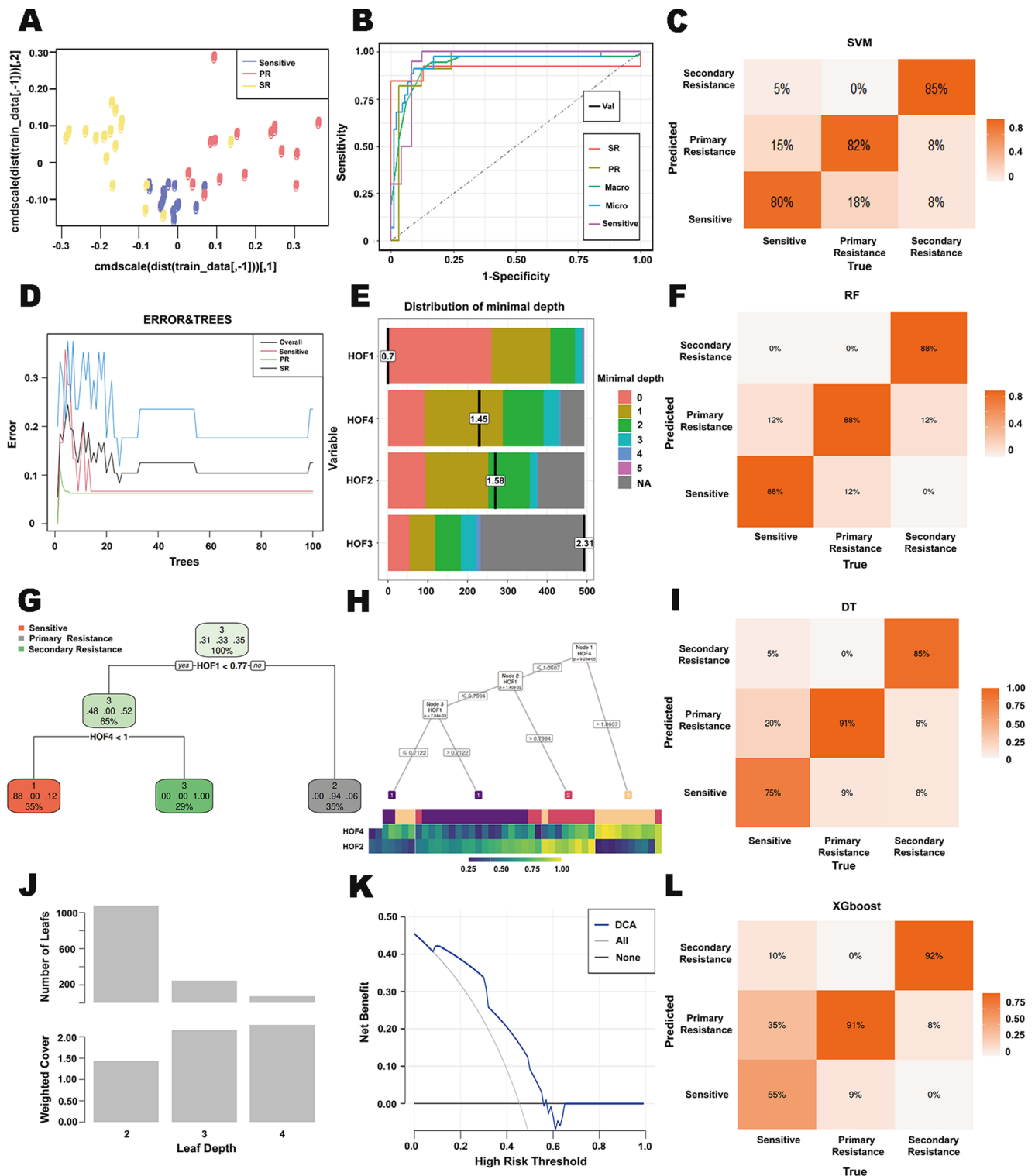
from each of three clinical subgroups (that is 15 for sensitive, 16 for primary resistance and 17 for secondary resistance) to form the training cohort. Their isolated exosomes were incubated with four distinct sensing units to generate array-based profiling data. The developed ML algorithm was subsequently applied to analyze testing cohort exosomal fingerprints for predicting efficacy of trastuzumab ( $n = 24$ ; that is, 8 for sensitive, 8 for primary resistance and 8 for secondary resistance). Additionally, 20 patients from the Affiliated Hospital of Xuzhou Medical University were included as an independent external validation cohort. The resulting fingerprints served as input features for SVM, RF, DT and XGBoost trained on the reference dataset. The SVM model demonstrated effective stratification of the three patient cohorts with an overall prediction accuracy of 81.82% (Fig. 4A, B). Specifically: 80.00% sensitivity for trastuzumab-sensitive, 82.00% for primary resistant groups and 85.00% prediction efficiency for secondary resistant cases (Fig. 4C). RF analysis was subsequently implemented to minimize misclassification and evaluate feature importance among HOFs for trastuzumab response prediction (Fig. 4D, E). HOF2 emerged as the most discriminative framework, though the aggregate prediction accuracy reached 87.50% of the model (Fig. 4F). Comparative analysis using DT and XGboost algorithm revealed reduced performance. Visualization of the decision tree process on both the training and validation sets enables us to trace each decision path, thereby elucidating the underlying decision logic of the model (Fig. 4G, H). The overall prediction accuracy reached 81.82% (Fig. 4I). Subsequently, the XGBoost algorithm was also applied to predict trastuzumab treatment efficacy, only achieving an overall prediction accuracy of 75.00% (Fig. 4J-L). Validation studies confirmed the ML platform's clinical utility through successful analysis of test-set exosomal signatures. The current IHC methods require binding specific antibodies to HER2 protein in tumor tissues to determine the susceptibility to trastuzumab [42, 43].

While existing diagnostic methods suffer from prolonged processing times and limited predictive capacity, this platform permits timely therapeutic evaluation following brief exosome incubation (Fig. 5A). The entire detection process, from the addition of exosomes to the HOF sensing unit to the final signal output and fingerprint prediction matching, only takes 40 min in total. This significantly reduces the waiting time for patients. Compared with other detection methods, it also has certain advantages [15, 44]. Meanwhile, compared with other existing liquid biopsy approaches, our approach is notably more straightforward, requiring only three key steps: exosome incubation, catalytic color development, and signal readout. This simplicity reduces the technical demand on operators and facilitates broader adoption and reproducibility across different laboratories [45,

46]. A total of 92 patients were enrolled in this study, exosomes from each patient generated an exclusive fingerprint. Of these, 72 patients from Jiangsu Province Hospital were split into a training set ( $n = 48$ ) and an internal validation set ( $n = 24$ ). Additionally, 20 patients from the Affiliated Hospital of Xuzhou Medical University were included as an independent external validation cohort. The RF-based prediction model achieved an accuracy of 87.50% in internal validation and maintained an accuracy of 75.00% in the external cohort, demonstrating its generalizability (Fig. 5B). The observed difference in performance is reasonable and commonly seen in real-world multicenter studies, reflecting heterogeneity in sample processing protocols and baseline patient characteristics across different hospital settings. Although further optimization is required to improve the detection efficiency for primary resistance, the current system already represents a robust complementary tool for clinical decision-making, particularly in stratifying patients who may benefit from trastuzumab-based regimens versus those requiring alternative therapeutic strategies. This finding robustly validates the good performance of exosomal fingerprint profiling in determining breast cancer molecular subtypes and predicting trastuzumab efficacy. When compared with recently published platforms that utilize machine learning for exosomal proteomic [46] or non-coding RNA analyses [47], our label-free phenotypic analysis platform shows somewhat lower performance in terms of sample size, sensitivity, and specificity. Moreover, our approach offers distinct practical advantages. It is significantly more cost-effective, requires shorter detection time, and operates without reliance on large-scale instrumentation. These features make it particularly suitable for rapid initial screening in broader populations, thereby helping to reduce both waiting times and testing costs for patients. Therefore, this rapid, non-invasive diagnostic approach enables accurate identification of HER2<sup>+</sup> BC, thereby facilitating timely implementation of personalized anti-HER2 therapeutic regimens.

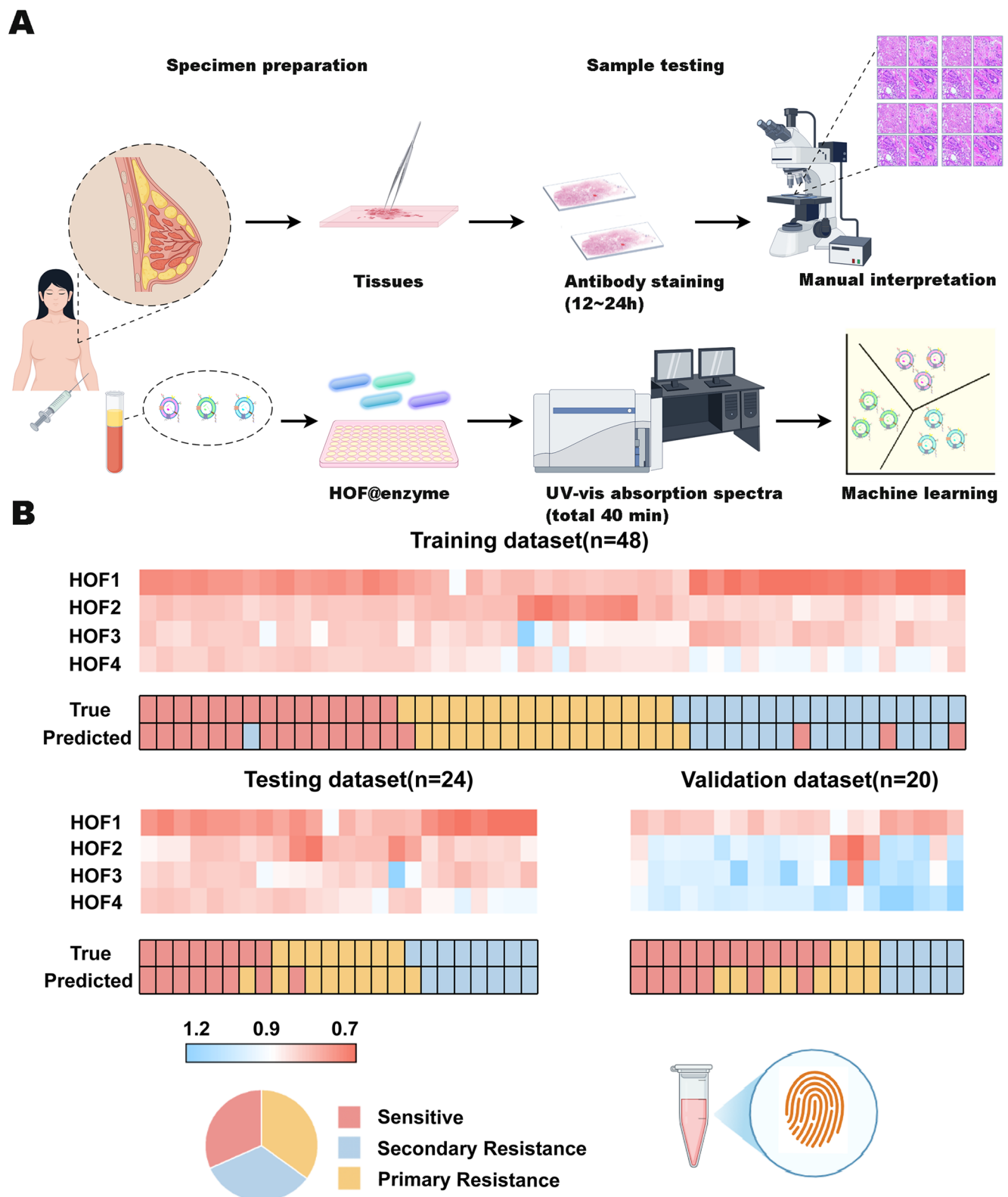
### 3 Conclusion

Collectively, this study exploits co-optimization of algorithm and sensor array for clinical HER2<sup>+</sup> BC detection and trastuzumab treatment outcome prediction through serum exosomal fingerprint analysis. By leveraging the comprehensive molecular profile of serum exosomes, this platform provided superior diagnostic strategies compared to approaches using a single exosomal biomarker or restricted biomarker panels that risk missing critical pathological information. The HOFs@ enzyme-based sensor array pre-trained with UV-Vis absorption spectra of exosomes derived from four molecularly distinct subtypes achieves high



**Fig. 4** ML-based exosomal fingerprint analysis for trastuzumab therapeutic response prediction. **(A)** Two-dimensional scatter plot illustrating the distribution patterns of individual exosomes. **(B)** ROC curve demonstrating the predictive performance of SVM classification using exosomal fingerprints. **(C)** Confusion matrix evaluating SVM prediction accuracy for trastuzumab response (overall accuracy = 81.82%). **(D)** Representative decision tree from RF classification, generated through interaction testing. **(E)** Minimal depth distribution analysis assessing the relative importance of each sensing element in the array.

**(F)** Predictive accuracy of RF-based classification for trastuzumab efficacy (overall accuracy = 87.50%). **(G)** Visualization of the DT process for the training set. **(H)** Visualization of the DT process for the validation set. **(I)** Performance metrics of DT model in predicting trastuzumab response (overall accuracy = 81.82%). **(J)** XGboost model complexity. **(K)** The decision curve of trastuzumab response prediction. **(L)** Predictive accuracy of XGboost-based classification for trastuzumab efficacy (overall accuracy = 75.00%).



**Fig. 5** Machine-learning-based identification of exosomal fingerprint derived from breast cancer patients. **(A)** Comparison of IHC and this sensor array-based method for exosomal fingerprint identification. **(B)**

Comparative analysis between RF-based classification (training set:  $n=48$ ; test set:  $n=24$ ) and clinically evaluated therapeutic outcomes in the validation cohort ( $n=20$ )

concordance (100% sensitivity) in identification of patients with HER2<sup>+</sup> BC. Our algorithm-sensor array co-optimization approach has been also demonstrated with outstanding performance in evaluating the trastuzumab efficacy, so the presented sensor array can discriminate serum exosomes from distinct pathological sources through subtle spectral variations, and simultaneously resolve clinical limitations including operational complexity, inconsistent reproducibility and sample heterogeneity. The sensor array can authenticate the HER2<sup>+</sup> BC patients and trastuzumab efficacy with the help of machine learning, which might act as a powerful tool for clinical diagnosis after extensive analytical validation and clinical validation in larger, multi-center cohorts.

## 4 Methods

### 4.1 Materials

All chemical reagents were obtained from commercial suppliers and used without further purification. 4,4',4'',4'''-Methanetetrayltetrabenzimidamide tetrahydrochloride, tetrakis(4-carboxyphenyl)methane, azobenzene-4,4'-dicarboxylic acid, and 1,3,6,8-tetrakis(4-carboxyphenyl)pyrene (H4TBAPy) were purchased from J&K Scientific Ltd. (Beijing, China). N, N-Dimethylformamide (DMF) and horseradish peroxidase (HRP) were acquired from Sigma-Aldrich (Shanghai, China). For exosome isolation, the ExoQuick exosome precipitation kit (EXOTC10A-1 for cell culture media; EXOQ20A-1 for clinical samples) was procured from System Biosciences (Palo Alto, CA, USA). Ammonia monohydrate (AMH) was obtained from local chemical distributors. Anti-CD9, anti-CD63 and anti-CD81 antibodies were purchased from Abcam Inc (USA).

### 4.2 Cell lines

The human breast cancer cell lines, T47D, MCF-7, ZR-75-1, BT474, MDA-MB-453, SKBR3, HCC1806 and MDA-MB-231, were obtained from American Type Culture Collection (Manassas, VA, USA).

### 4.3 Synthesis of HOF@ enzyme probes

An aqueous stock solution of horseradish peroxidase (HRP) was prepared at a concentration of 20 mg/mL for subsequent use. Four distinct hydrogen-bonded organic frameworks (HOFs) were synthesized as follows:

For HOF1, 5 mg of 4,4',4'',4'''-methanetetrayltetrabenzimidamide tetrahydrochloride was dissolved in 1.8 mL of deionized water, while 6 mg of tetrakis(4-carboxyphenyl)

methane was first dissolved in 0.02 mL of ammonia monohydrate (AMH) and then diluted to 2 mL with water. These two solutions were subsequently mixed with 0.2 mL of the prepared HRP solution.

HOF2 was synthesized by dissolving 5 mg of 4,4',4'',4'''-methanetetrayltetrabenzimidamide tetrahydrochloride in 1.8 mL deionized water, and separately dissolving 3 mg of azobenzene-4,4'-dicarboxylic acid in 0.02 mL AMH before adjusting the volume to 2 mL with water. The resulting solutions were combined with 0.2 mL HRP solution.

The preparation of HOF3 involved dissolving 3.7 mg of tetrakis(4-carboxyphenyl) methane in 3.8 mL deionized water, followed by the addition of another 3.8 mL water.

For HOF4, 2 mg of 1,3,6,8-tetrakis(4-carboxyphenyl) pyrene (H4TBAPy) was completely dissolved in 0.2 mL dimethylformamide (DMF) under ultrasonication, after which 0.2 mL HRP solution and 3.6 mL deionized water were added.

All reaction mixtures were stirred at 1200 rpm for 5 min at room temperature. The products were then collected by centrifugation at 12,000 rpm for 10 min, washed three times with double-distilled water, and finally dried at 45 °C under light-protected conditions for storage.

### 4.4 Characterization of HOF@ enzyme probes

HOF@ enzyme probes were characterized including the TEM, XRD patterns and FTIR spectra. The morphology of HOF@ enzyme was characterized by TEM (FEI Tecnai G2 spirit, USA). The XRD patterns were measured by a Burker D8 Advance (Germany). FTIR spectroscopy was carried out on a Nicolet IS10 FTIR spectrometer (USA).

### 4.5 Exosome nanoscale flow cytometry (NanoFCM) analysis

A 10 µL aliquot of exosomes was diluted to 30 µL with PBS, followed by incubation with 0.5 µL of primary antibodies (CD9, CD63, CD81) at 37 °C for 30 min. The mixture was diluted with 1 mL of ice-cold PBS and pelleted via ultracentrifugation (110,000 × g, 4 °C, 70 min) using a fixed-angle rotor. The supernatant was carefully removed. The pellet was resuspended in 100 µL PBS and incubated with 0.5 µL fluorescent secondary antibody (with a secondary-only control) at 37 °C for 30 min. The labeled exosomes were washed with 1 mL PBS and centrifuged again (110,000 × g, 4 °C, 70 min). The pellet was resuspended in 1 mL PBS and subjected to a third ultracentrifugation (110,000 × g, 4 °C, 70 min) to ensure purity. The final exosome pellet was resuspended in 50 µL ice-cold PBS for NanoFCM analysis. Samples were analyzed using NanoFCM, and protein



marker expression levels (CD9, CD63, CD81) were quantified based on fluorescence signals.

#### 4.6 Generation of the exosomal fingerprint

The four types of HOFs were individually diluted with phosphate-buffered saline (PBS, pH 7.4) to prepare working solutions (99  $\mu\text{L}$  per tube). Subsequently, 1  $\mu\text{L}$  of isolated exosomes was added to each HOF solution, thoroughly mixed, and transferred to a 96-well microplate. After 30 min of incubation at room temperature ( $25 \pm 1$  °C), 100  $\mu\text{L}$  of 3,3',5,5'-tetramethylbenzidine (TMB) chromogenic substrate was added to each well. The exosomal fingerprint was generated from the absorbance at 650 nm obtained from the multimode microplate reader. Each experimental condition was performed in sextuplicate ( $n=6$ ) to ensure statistical reliability. For a single isolation, 4 absorbance and 6 repeated values were obtained from nanosensors to generate the exosomal fingerprint. The matrix was collected as the data set for further analysis.

#### 4.7 Machine-learning analysis

Machine learning pipeline integrated multiple analytical approaches in R 4.0 to characterize exosome-distinctive fingerprints. Supervised classification models, including SVM, RF, XGboost, and DT analyses were based on the classification criteria of the sample by “caret”, “randomForest”, “XGboost”, “rpart”.etc. For SVM, we selected the radial basis function (RBF) kernel and employed a grid search approach with a step size of 0.5 to determine the optimal cost (C) and gamma ( $\gamma$ ) parameters. RF models were optimized by minimizing the out-of-bag (OOB) error, while DTs underwent cost-complexity pruning to prevent overfitting. Model performance was rigorously evaluated using multiple metrics: accuracy, F1 score, specificity, positive predictive value (PPV), and negative predictive value (NPV). All supervised analyses were conducted on a labeled dataset, with model development restricted to the training set. In addition, five-fold cross validations were conducted in the derivation cohort for the validation of the prediction model. To address the “black-box” nature of complex models and enhance clinical interpretability, we applied SHapley Additive exPlanations (SHAP) to provide both global feature importance rankings and local explanations for predictions, elucidating the relationship between input features (e.g., HOF sensor responses) and model outputs. The heatmap was finally constructed based on the absorbance at 650 nm and visualized by package “heatmap”.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s42114-026-01638-5>.

**Acknowledgements** Thank you to Li Ben from the Affiliated Hospital of Xuzhou Medical University for providing the external validation.

**Author contributions** S. W., T. Z., G. L. and Y. Y. conceived the project. S. W., T. Z., Z.L. and Y.S. designed and performed most of the experiments. S. W. and H. S. synthesized HOF. T. Z., Y. L., J. X. and Y. S. performed cell culture and experiments. X.H., and Y. L. assessed efficacy of patients. H. S., T. Z. and S. W. wrote the paper. T. Z., Z. L. and Y. S. made the same contribution. All authors reviewed and modified the manuscript.

**Funding** This work was supported by National Natural Science Foundation of China (82403696, 82202294, 82272667, 81972484), Jiangsu Funding Program for Excellent Postdoctoral, the Major Basic Research Fund of Jiangsu Province Hospital (QY202406), Clinical Diagnosis and Treatment Technology Innovation “Open Bidding and Leader Selecting” Project of Jiangsu Province Hospital (CZ1420240207), The D-Class Talent Cultivation Program of Jiangsu Province Hospital.

**Data availability** The main data supporting the findings of this study are available within the article and its Supplementary Information. All raw data generated during the study are available from the corresponding authors on reasonable request.

#### Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

#### References

1. Loibl S, Gianni L (2017) HER2-positive breast cancer. *Lancet* 389(10087):2415–2429. [https://doi.org/10.1016/S0140-6736\(16\)32417-5](https://doi.org/10.1016/S0140-6736(16)32417-5)
2. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235(4785):177–182. <https://doi.org/10.1126/science.3798106>
3. Hayes DF (2019) HER2 and breast cancer - a phenomenal success story. *N Engl J Med* 381(13):1284–1286. <https://doi.org/10.1056/NEJMcibr1909386>
4. Pernas S, Tolaney SM (2019) HER2-positive breast cancer: new therapeutic frontiers and overcoming resistance. *Ther Adv Med Oncol* 11:1758835919833519. <https://doi.org/10.1177/1758835919833519>

5. Oh DY, Bang YJ (2020) HER2-targeted therapies - a role beyond breast cancer. *Nat Rev Clin Oncol* 17(1):33–48. <https://doi.org/10.1038/s41571-019-0268-3>
6. Sharma P (2020) Major strides in HER2 Blockade for metastatic breast cancer. *N Engl J Med* 382(7):669–671. <https://doi.org/10.1056/NEJMe1916310>
7. Subbiah IM, Gonzalez-Angulo AM (2014) Advances and future directions in the targeting of HER2-positive breast cancer: implications for the future. *Curr Treat Options Oncol* 15(1):41–54. <https://doi.org/10.1007/s11864-013-0262-4>
8. Swain SM, Shastry M, Hamilton E (2023) Targeting HER2-positive breast cancer: advances and future directions. *Nat Rev Drug Discov* 22(2):101–126. <https://doi.org/10.1038/s41573-022-00579-0>
9. Wolff AC, Somerfield MR, Dowsett M, Hammond MEH, Hayes DF, McShane LM, Saphner TJ, Spears PA, Allison KH (2023) Human epidermal growth factor receptor 2 testing in breast cancer: ASCO-College of American pathologists guideline update. *J Clin Oncol* 41(22):3867–3872. <https://doi.org/10.1200/JCO.22.02864>
10. Press MF, Sauter G, Bernstein L, Villalobos IE, Mirlacher M, Zhou JY, Wardeh R, Li YT, Guzman R, Ma Y, Sullivan-Halley J, Santiago A, Park JM, Riva A, Slamon DJ (2005) Diagnostic evaluation of HER-2 as a molecular target: an assessment of accuracy and reproducibility of laboratory testing in large, prospective, randomized clinical trials. *Clin Cancer Res* 11(18):6598–6607. <https://doi.org/10.1158/1078-0432.CCR-05-0636>
11. Siddiqui S, Rimm DL (2010) Pre-analytic variables and phospho-specific antibodies: the Achilles heel of immunohistochemistry. *Breast Cancer Res* 12(6):113. <https://doi.org/10.1186/bcr2782>
12. Tozbikian G, Krishnamurthy S, Bui MM, Feldman M, Hicks DG, Jaffer S, Khoury T, Wei S, Wen H, Pohlmann P (2024) Emerging landscape of targeted therapy of breast cancers with low human epidermal growth factor receptor 2 protein expression. *Arch Pathol Lab Med* 148(2):242–255. <https://doi.org/10.5858/arpa.2022-0335-RA>
13. Perez EA, Cortes J, Gonzalez-Angulo AM, Bartlett JM (2014) HER2 testing: current status and future directions. *Cancer Treat Rev* 40(2):276–84. <https://doi.org/10.1016/j.ctrv.2013.09.001>
14. Sun H, Kang EY, Chen H, Sweeney KJ, Suchko M, Wu Y, Wen J, Krishnamurthy S, Albarracin CT, Ding QQ, Foo WC, Sahin AA (2024) Immunohistochemical assessment of HER2 low breast cancer: interobserver reproducibility and correlation with digital image analysis. *Breast Cancer Res Treat* 205(2):403–411. <https://doi.org/10.1007/s10549-024-07256-3>
15. Shi H, Zeng T, Liu Y, Liang Q, Yang F, Liang Y, Fu Z, Li W, Li G, Yin Y (2023) Assembly of nanocatalyst as a noninvasive tool for breast cancer diagnosis and immunotherapy surveillance designed by visually tracking tumor-derived exosomal PD-L1. *ACS Mater Lett* 5(2):282–290. <https://doi.org/10.1021/acsmater.1c01036>
16. van Niel G, D'Angelo G, Raposo G (2018) Shedding light on the cell biology of extracellular vesicles. *Nat Rev Mol Cell Biol* 19(4):213–228. <https://doi.org/10.1038/nrm.2017.125>
17. Kalluri R, LeBleu VS (2020) The biology, function, and biomedical applications of exosomes. *Science* 367(6478). <https://doi.org/10.1126/science.aau6977>
18. Jiang L, Gu Y, Du Y, Liu J (2019) Exosomes: diagnostic biomarkers and therapeutic delivery vehicles for cancer. *Mol Pharm* 16(8):3333–3349. <https://doi.org/10.1021/acs.molpharmaceut.9b00409>
19. Li S, Yi M, Dong B, Tan X, Luo S, Wu K (2021) The role of exosomes in liquid biopsy for cancer diagnosis and prognosis prediction. *Int J Cancer* 148(11):2640–2651. <https://doi.org/10.1002/ijc.33386>
20. Xu G, Huang R, Wumaier R, Lyu J, Huang M, Zhang Y, Chen Q, Liu W, Tao M, Li J, Tao Z, Yu B, Xu E, Wang L, Yu G, Gires O, Zhou L, Zhu W, Ding C, Wang H (2024) Proteomic profiling of serum extracellular vesicles identifies diagnostic signatures and therapeutic targets in breast cancer. *Cancer Res* 84(19):3267–3285. <https://doi.org/10.1158/0008-5472.CAN-23-3998>
21. Wang Y, Gao W, Sun M, Feng B, Shen H, Zhu J, Chen X, Yu S (2023) A filter-electrochemical microfluidic chip for multiple surface protein analysis of exosomes to detect and classify breast cancer. *Biosens Bioelectron* 239:115590. <https://doi.org/10.1016/j.bios.2023.115590>
22. Cao Y, Yu X, Zeng T, Fu Z, Zhao Y, Nie B, Zhao J, Yin Y, Li G (2022) Molecular characterization of exosomes for subtype-based diagnosis of breast cancer. *J Am Chem Soc* 144(30):13475–13486. <https://doi.org/10.1021/jacs.2c00119>
23. Chen J, Hickey BL, Wang L, Lee J, Gill AD, Favero A, Pinalli R, Dalcanele E, Hooley RJ, Zhong W (2021) Selective discrimination and classification of g-quadruplex structures with a host-guest sensing array. *Nat Chem* 13(5):488–495. <https://doi.org/10.1038/s41557-021-00647-9>
24. Chen J, Gill AD, Hickey BL, Gao Z, Cui X, Hooley RJ, Zhong W (2021) Machine learning aids classification and discrimination of noncanonical DNA folding motifs by an arrayed host:guest sensing system. *J Am Chem Soc* 143(32):12791–12799. <https://doi.org/10.1021/jacs.1c06031>
25. Jiang M, Yan X, Wang Y, Pu F, Liu H, Li Y, Yang C, Zhu J, Liu X, Ren J, Qu X (2023) One-component artificial gustatory system based on hydrogen-bond organic framework for discrimination of versatile analytes. *Adv Funct Mater* 33(24):2300091. <https://doi.org/10.1002/adfm.202300091>
26. Yu D, Zhang H, Ren J, Qu X (2023) Hydrogen-bonded organic frameworks: new horizons in biomedical applications. *Chem Soc Rev* 52(21):7504–7523. <https://doi.org/10.1039/d3cs00408b>
27. Lin RB, He Y, Li P, Wang H, Zhou W, Chen B (2019) Multifunctional porous hydrogen-bonded organic framework materials. *Chem Soc Rev* 48(5):1362–1389. <https://doi.org/10.1039/c8cs00155c>
28. Zhang Z, Ye Y, Xiang S, Chen B (2022) Exploring multifunctional hydrogen-bonded organic framework materials. *Acc Chem Res* 55(24):3752–3766. <https://doi.org/10.1021/acs.accounts.2c00686>
29. Wang B, Lin RB, Zhang Z, Xiang S, Chen B (2020) Hydrogen-bonded organic frameworks as a tunable platform for functional materials. *J Am Chem Soc* 142(34):14399–14416. <https://doi.org/10.1021/jacs.0c06473>
30. Yang J, Wang J, Hou B, Huang X, Wang T, Bao Y, Hao H (2020) Porous hydrogen-bonded organic frameworks (HOFs): from design to potential applications. *Chem Eng J* 399:125873. <https://doi.org/10.1016/j.cej.2020.125873>
31. Sun Z, Li Y, Chen L, Jing X, Xie Z (2015) Fluorescent hydrogen-bonded organic framework for sensing of aromatic compounds. *Cryst Growth Des* 15(2):542–545. <https://doi.org/10.1021/cg501652r>
32. Yang J, Zhang X, Chen M, Huang Y, Tian B, Wang N, Hao H (2022) Versatile hydrogen-bonded organic framework (HOF) platform for simultaneous detection and efficient removal of heavy metal ions. *J Environ Chem Eng* 10(6):108983. <https://doi.org/10.1016/j.jece.2022.108983>
33. Welsh JA, Golderdhan DCI, O'Driscoll L, Buzas EI, Blenkinson C, Bussolati B, Cai H, Di Vizio D, Driedonks TAP, Erdbrugger U, Falcon-Perez JM, Fu QL, Hill AF, Lenassi M, Lim SK, Mahoney MG, Mohanty S, Moller A, Nieuwland R, Ochiya T, Sahoo S, Torrecillas AC, Zheng L, Zijlstra A, Abuelreich S, Bagabas R, Bergese P, Bridges EM, Bruciale M, Burger D, Carney RP, Cocucci E, Crescitelli R, Hanser E, Harris AL, Haughey NJ, Hendrix A, Ivanov AR, Jovanovic-Talisman T, Kruh-Garcia NA, Kyburz V, Lasser C,

- Lennon KM, Lotvall J, Maddox AL, Martens-Uzunova ES, Mizenko RR, Newman LA, Ridolfi A, Rohde E, Rojalin T, Rowland A, Saftics A, Sandau US, Saugstad JA, Shekari F, Swift S, Ter-Ovanesyan D, Tosar JP, Useckaite Z, Valle F, Varga Z, van der Pol E, van Herwijnen MJC, Wauben MHM, Wehman AM, Williams S, Zendrini A, Zimmerman AJ, Consortium M, Thery C (2024) K.W. Witwer, Minimal information for studies of extracellular vesicles (MISEV2023): from basic to advanced approaches. *J Extracell Vesicles* 13(2):e12404 <https://doi.org/10.1002/jev2.12404>
34. Tolane SM, Tarantino P, Graham N, Tayob N, Pare L, Villacampa G, Dang CT, Yardley DA, Moy B, Marcom PK, Albain KS, Rugo HS, Ellis MJ, Shapira I, Wolff AC, Carey LA, Barroso-Sousa R, Villagrana P, DeMeo M, DiLullo M, Zanudo JGT, Weiss J, Wagle N, Partridge AH, Waks AG, Hudis CA, Krop IE, Burstein HJ, Prat A, Winer EP (2023) Adjuvant paclitaxel and trastuzumab for node-negative, HER2-positive breast cancer: final 10-year analysis of the open-label, single-arm, phase 2 APT trial. *Lancet Oncol* 24(3):273–285. [https://doi.org/10.1016/S1470-2045\(23\)00051-7](https://doi.org/10.1016/S1470-2045(23)00051-7)
35. Nitz UA, Gluz O, Christgen M, Grischke EM, Augustin D, Kummel S, Braun M, Potenberg J, Kohls A, Krauss K, Stefek A, Schumacher C, Forstbauer H, Reimer T, Fischer H, Liedtke C, Wuerstlein R, Schumacher J, Kates R, Kreipe H, Harbeck N (2017) De-escalation strategies in HER2-positive early breast cancer (EBC): final analysis of the WSG-ADAPT HER2+/HR-phase II trial: efficacy, safety, and predictive markers for 12 weeks of neoadjuvant dual blockade with trastuzumab and pertuzumab +/- weekly paclitaxel. *Ann Oncol* 28(11):2768–2772. <https://doi.org/10.1093/annonc/mdx494>
36. von Minckwitz G, Procter M, de Azambuja E, Zardavas D, Benyunes M, Viale G, Suter T, Arahmani A, Rouchet N, Clark E, Knott A, Lang I, Levy C, Yardley DA, Bines J, Gelber RD, Piccart M, Baselga J, Committee AS (2017) Investigators, adjuvant pertuzumab and trastuzumab in early HER2-positive breast cancer. *N Engl J Med* 377(2):122–131. <https://doi.org/10.1056/NJMoal703643>
37. Swain SM, Miles D, Kim SB, Im YH, Im SA, Semiglazov V, Ciruelos E, Schneeweiss A, Loi S, Monturus E, Clark E, Knott A, Restuccia E, Benyunes MC, Cortes J (2020) S. group, pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA): end-of-study results from a double-blind, randomised, placebo-controlled, phase 3 study. *Lancet Oncol* 21(4):519–530. [https://doi.org/10.1016/S1470-2045\(19\)30863-0](https://doi.org/10.1016/S1470-2045(19)30863-0)
38. Hu R, Lin S, Li F, Shen J, Xie Y, Deng B, Zhang Y, Dong Z, Zhu J (2024) Exploring aptamer-based metasurfaces for label-free plasmonic biosensing of breast tumor-derived exosomes. *Adv Opt Mater* 12(28):2401180. <https://doi.org/10.1002/adom.202401180>
39. Kim R, Mun B, Lim S, Park C, Kim J, Lim J, Jeong H, Son HY, Rho HW, Lim EK, Haam S (2024) Colorimetric detection of HER2-overexpressing-cancer-derived exosomes in mouse urine using magnetic-polydiacetylene nanoparticles. *Small* 20(13):e2307262. <https://doi.org/10.1002/smll.202307262>
40. Yang X, Wang Y, Byrne R, Schneider G, Yang S (2019) Concepts of artificial intelligence for computer-assisted drug discovery. *Chem Rev* 119(18):10520–10594. <https://doi.org/10.1021/acs.chemrev.8b00728>
41. Capman NSS, Zhen XV, Nelson JT, Chaganti V, Finc RC, Lyden MJ, Williams TL, Freking M, Sherwood GJ, Buhlmann P, Hogan CJ, Koester SJ (2022) Machine learning-based rapid detection of volatile organic compounds in a graphene electronic nose. *ACS Nano* 16(11):19567–19583. <https://doi.org/10.1021/acsnano.2c10240>
42. Ferrari A, Vincent-Salomon A, Pivot X, Sertier AS, Thomas E, Tonon L, Boyault S, Mulugeta E, Treilleux I, MacGrogan G, Arnould L, Kielbassa J, Le Texier V, Blanche H, Deleuze JF, Jacquemier J, Mathieu MC, Penault-Llorca F, Bibeau F, Mariani O, Mannina C, Pierga JY, Tredan O, Bachelot T, Bonnefoi H, Romieu G, Fumoleau P, Delaloge S, Rios M, Ferrero JM, Tarpin C, Bouteille C, Calvo F, Gut IG, Gut M, Martin S, Nik-Zainal S, Stratton MR, Pauporte I, Saintigny P, Birnbaum D, Viari A, Thomas G (2016) A whole-genome sequence and transcriptome perspective on HER2-positive breast cancers. *Nat Commun* 7:12222. <https://doi.org/10.1038/ncomms12222>
43. McNamara KL, Caswell-Jin JL, Joshi R, Ma Z, Kotler E, Bean GR, Kriner M, Zhou Z, Hoang M, Beechem J, Zoeller J, Press MF, Slamon DJ, Hurvitz SA, Curtis C (2021) Spatial proteomic characterization of HER2-positive breast tumors through neoadjuvant therapy predicts response. *Nat Cancer* 2(4):400–413. <https://doi.org/10.1038/s43018-021-00190-z>
44. Liu Y, Zeng T, Wu S, Sun Y, Liang Q, Liang Y, Huang X, Li W, Shi H, Li G, Yin Y (2025) A colorimetric biosensor for the analysis of circulating PD-L1 with application to diagnose breast cancer. *Sens Actuators B* 430:137334. <https://doi.org/10.1016/j.snb.2025.137334>
45. Xu W, Lin Y, Huang Z, Li Y, Lu Y, Liu M, Cui S, Zhang T, Shi N, Sheng Y, Hu J (2025) Split proximity circuit initiated CRISPR-Cas12a system profiling exosomal surface proteins for early cancer detection. *Biosens Bioelectron* 295:118280. <https://doi.org/10.1016/j.bios.2025.118280>
46. Qu X, Lu B, Gao C, Zhao W, Zeng Y, Wu S, Ji C, Li G (2025) AI-enabled new sensing technology: colorimetric analysis of exosomes for precise diagnosis of breast cancer. *Chem Sci*. <https://doi.org/10.1039/d5sc06650f>
47. Cai ZR, Zheng YQ, Hu Y, Ma MY, Wu YJ, Liu J, Yang LP, Zheng JB, Tian T, Hu PS, Liu ZX, Zhang L, Xu RH, Ju HQ (2025) Construction of exosome non-coding RNA feature for non-invasive, early detection of gastric cancer patients by machine learning: a multi-cohort study. *Gut* 74(6):884–893. <https://doi.org/10.1136/gutjnl-2024-333522>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.