Supporting Information

Spatiotemporal Identification of Therapeutic Markers via Dual-targeted SERS Bioprobes During Breast Cancer Progression

Xinyu Miao [a,b,c,d,⊥], *Xiawei Xu* [c,d,⊥], *Li Sun* [c,d,⊥], *Xiaoxia Wu* [b *], *Aochi Liu* [c,d], *Pan Fu* [a,b], *Pengtao Chen* [b], *Yujiao Xie* [c,d], *Jiabao Guo* [c,d], *Zhiwei Hou* [c,d], *Dong Xu* [a b *], *Aiguo Wu* [c,d *], *Jie Lin* [c,d *]

[a] Second clinical college, Zhejiang Chinese Medical University, Hang Zhou 310053, China.

[b] Department of Interventional Radiology, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Zhejiang Key Laboratory of Imaging and Interventional Medicine, Hangzhou 310022, China.

[c] Ningbo Key Laboratory of Biomedical Imaging Probe Materials and Technology, Laboratory of Advanced Theranostic Materials and Technology, Chinese Academy of Sciences (CAS) Key Laboratory of Magnetic Materials and Devices, Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, Ningbo 315201, China.

[d] Zhejiang International Cooperation Base of Biomedical Materials Technology and Application, Zhejiang Engineering Research Center for Biomedical Materials, Ningbo Cixi Institute of Biomedical Engineering, Ningbo 315300, China.

⊥ Xinyu Miao, Xiawei Xu and Li Sun contributed equally.

***Corresponding authors. E-mail:**

xiaoxiawu@ucas.ac.cn (*Xiaoxia Wu*)

xudong@zjcc.org.cn (*Dong Xu*)

aiguo@nimte.ac.cn (*Aiguo Wu*)

linjie@nimte.ac.cn (*Jie Lin*)

Characterization equipment. The morphology and elemental mapping were characterized using high-resolution transmission electron microscopy (HRTEM, Talos F200x, Thermo Fisher Scientific). X-ray diffraction (XRD) analysis of α radiation (λ = 1.54056 Å) was conducted with a BRUKER D8 ADVANCE DAVINCI diffractometer equipped with Cu-K radiation. Size measurements were performed using dynamic light scattering (DLS, Malvern). The ultraviolet-visible (UV-Vis) absorption spectrum was obtained with a Lambda T10 CS spectrophotometer (PerkinElmer). The concentrations of iron and gold in the bioprobes were determined by inductively coupled plasma optical emission spectroscopy (ICP-OES, Spectro Arcos, Spectro). Fluorescence spectra were recorded using confocal laser scanning microscopy (CLSM, TCS SP8, Leica). The SERS spectra and images were obtained using a confocal Raman microscope (LabRAM Odyssey, Horiba).

Preparation of IO NPs: according to previous reports,[1, 2] iron oxide (IO) nanoparticles (IO NPs) in the oil phase were synthesized as follows: 1.0 g of sodium hydroxide (NaOH) was dissolved in a mixed solution of 10 mL of oleic acid and 10 mL of ethanol, and the mixture was stirred until uniformly dispersed. Additionally, 0.78 g of ammonium iron (II) sulfate hexahydrate ($Fe_2(NH_4)_2(SO_4)_2 \cdot 6H_2O$) was dissolved in 20 mL of deionized water to form a precursor solution. This aqueous mixture was then slowly added to the oil phase and stirred until the solution turned brown due to the oxidation. Subsequently, the mixture was transferred to a reactor for hydrothermal reaction at 230 °C for 8 hours. After cooling, the precipitate was collected by centrifugation (6000 rpm for 10 minutes) or magnetic separation, washed three times with ethanol

and cyclohexane, and finally dispersed in 20 mL of cyclohexane for storage. The water phase modification of IO NPs involved mixing 2 mL of the oil phase IO NPs with 500 mg of citric acid, 5 mL of dimethylformamide (DMF), and 5 mL of chloroform. This mixture was subjected to ultrasonic treatment for 8 hours to facilitate complete ligand exchange. Subsequently, the unreacted reagents were removed by centrifugation with ethanol, repeated three times. Finally, the citric acid-modified IO NPs were dispersed in a total of 100 mL of deionized water, with 10 mL allocated per portion when divided into ten parts, and stored at 4 °C for future use.

Preparation of Au NPs: as previously reported, gold nanoparticles (Au NPs) were synthesized using the citrate reduction method.[3] This procedure involved adding 1.5 mL of $HAuCl_4 \cdot 4H_2O$ (33.33 mmol/L) to 48.5 mL of ultrapure water, followed by stirring and heating the mixture to boiling. Sodium citrate solution (2.3 mL, 1%) was then added rapidly. After 20 minutes, the solution turned purple; at this point, heating was stopped, and stirring was continued for an additional 40 minutes. Once the Au NPs had cooled to room temperature, they were stored at 4 °C for subsequent experiments.

Cell culture. Three breast cancer cell lines, MDA-MB-231, MCF-7, and SK-BR-3, were cultured under standard conditions (5% CO_2, 37 °C). The cells were grown in DMEM supplemented with 10% fetal bovine serum (Gibco Life Technologies). The SKBR3 cell line was purchased from Procell system. Co., Ltd; catalogue number: CL-0211; Species: Human; Purchase Date: December 15, 2024. The MCF-7 cell line was obtained from Suzhou Haixing Biotechnology Co., Ltd; catalogue number: TCH-C247; Species: Human; Purchase Date: June 11, 2024. The

MDA-MB-231 cell line was obtained from Cellverse Co., Ltd.; catalogue number: iCell-0055a; Species: Human; Purchase Date: August 14, 2024.

In vitro cytotoxicity test. *In vitro* Cytotoxicity of SERS bioprobes, the CCK-8 assay was employed to assess the cytotoxicity of antibodies. Briefly, 100 μL of MCF-7 cells (10^5 cells per well) were incubated in 96-well plates for 24 hours. Various concentrations of IO NPs@AR@PDA@aPD-L1 (0-80 μg/mL) and Au NPs@4-MBA@PDA@aHER2 (0-80 μg/mL) were then added and incubated for an additional 24 hours. Following this, 10 μL of CCK-8 solution was introduced to each well. After a 2-hour incubation period, the optical density (OD) was measured at 450 nm using a microplate reader. Cell viability was calculated using the following formula: Cell viability (%) = (OD of the test group - OD of the blank group) / (OD of the control group - OD of the blank group) × 100%. Reason: Improved clarity and readability by enhancing vocabulary, correcting punctuation, and ensuring consistent formatting.

SERS signal collection of cells. MDA-MB-231, MCF-7, and SK-BR-3 cells (1×10^5 cells/mL) were incubated with 1 mL of IO NPs@AR@PDA@aPD-L1 (20 μg/mL) and 1 mL of Au NPs@4-MBA@PDA@aHER2 (10 μg/mL), respectively. The dual-targeted bioprobes were also included in the incubation for 30 minutes. Subsequently, each type of cell was fixed onto a silicon wafer, and 100 cells from each line were randomly selected using a laser confocal Raman spectrometer to collect 100 spectra and mapping of each cell line was performed.

Statistical analysis of Raman spectroscopy with machine learning. In this study, a Python

3.13 environment utilizing the Scikit-learn 1.4.0 framework was employed to construct classification models for Raman spectra. 100 SERS spectra from each cell line, spanning a range of 800 to 1800 cm-1, were analyzed. The model implemented a systematic data processing and validation approach. During the data preprocessing stage, Z-score normalization was applied to mitigate dimensional discrepancies among spectral features, while label encoding was utilized to facilitate the digital conversion of categories. A cascade dimensionality reduction strategy was employed to optimize the feature space. Initially, principal components were dynamically selected using principal component analysis (PCA), with a user-defined dimension chosen to achieve a cumulative explained variance contribution rate of 80%. This process resulted in the extraction of a low-dimensional representation. Subsequently, linear discriminant analysis (LDA) was applied within the PCA subspace to facilitate supervised dimensionality reduction by maximizing inter-class separability, targeting a dimension of n_components = 2. Firstly, the performance of the test set was evaluated using hierarchical random partitioning (with a training set to test set ratio of 7:3, and fixed random seeds). The classification effectiveness was quantified through the multi-class receiver operating characteristic curve (ROC) curve (calculation of the AUC value) and confusion matrix analysis. The mean classification accuracy and standard deviation were subsequently calculated using hierarchical cross-validation with a user-defined number of iterations (5-fold cross-validation). Random seed control (random_state = 42) was implemented to ensure the experiment's repeatability. The parameterized interface allowed for the customization of key parameters, such as the PCA dimension threshold and the number of cross-validation iterations.

The visualization module generates 2D and 3D PCA projection maps, LDA discriminant spatial distribution maps, and decision boundary maps. It simultaneously produces standardized analysis reports, which include cumulative variance tables, ROC datasets, and confusion matrices. This module offers a comprehensive system solution that is both explanatory and reliable for multi-class classification of spectral data. Support Vector Machine (SVM) with a radial basis function kernel, Random Forest (RF), and Linear Discriminant Analysis (LDA). All are integrated within a scikit-learn Pipeline that couples principal component analysis (PCA) for dimensionality reduction with the respective classifier. Each workflow utilizes StratifiedShuffleSplit for stratified cross-validation and fixes random_state=42 to ensure reproducibility. Hyperparameter tuning is minimal across all approaches: the SVM uses default settings (C=1.0, gamma='scale') without optimization; LDA is applied in its standard form without regularization or solver customization; and although the Random Forest implementation includes a preliminary GridSearchCV step (searching over n_estimators, max_depth, and min_samples_split), this tuning is performed on the full dataset prior to cross-validation and is not incorporated into the evaluation loop, thereby lacking nested validation and risking optimistic bias. Consequently, none of the three pipelines implements rigorous, nested hyperparameter optimization, limiting the robustness of the reported performance metrics.

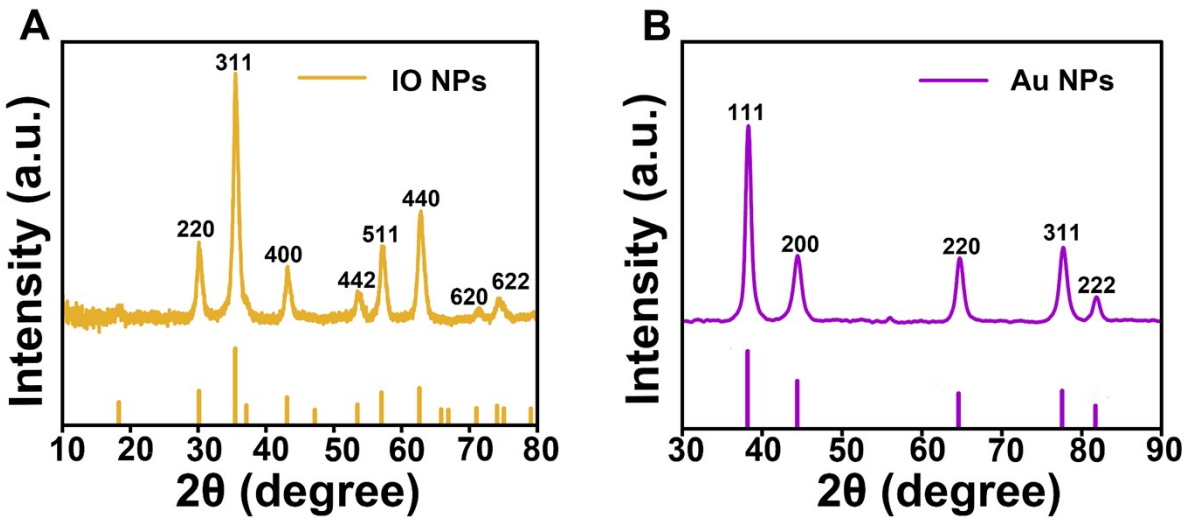

Figure. S1. (A-B) TEM image of IO NPs@PDA (A) and Au NPs@PDA (B). (A, B) XRD spectrum of IO NPs (C) and Au NPs (D).

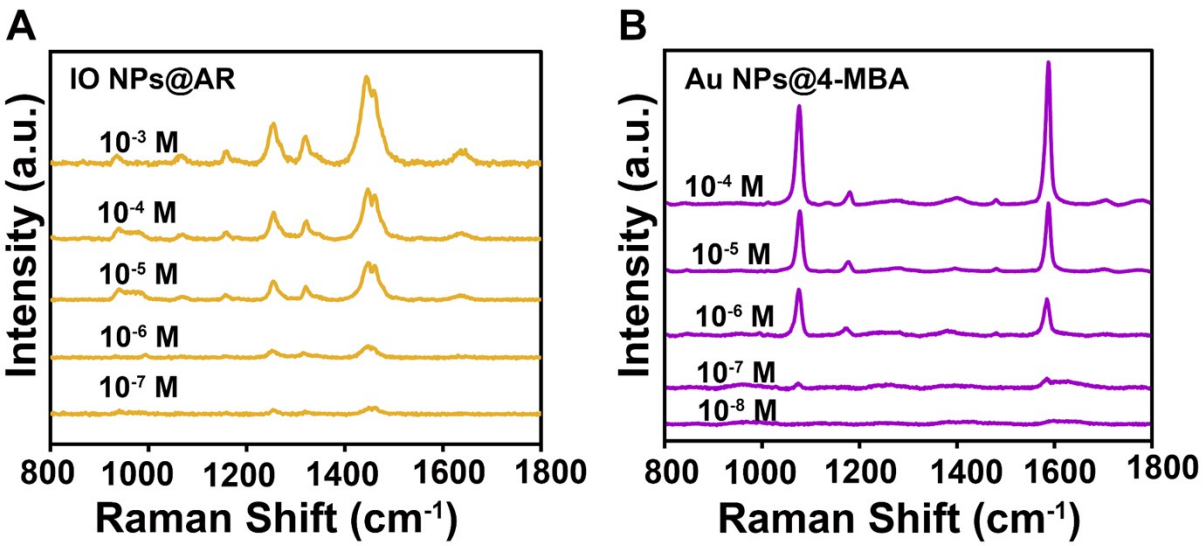

Figure. S2. (A) SERS spectra of IO NPs@AR during the optimization process for AR modification on IO NPs. (B) SERS spectra of Au NPs@4-MBA during the optimization process for 4-MBA modification on Au NPs.

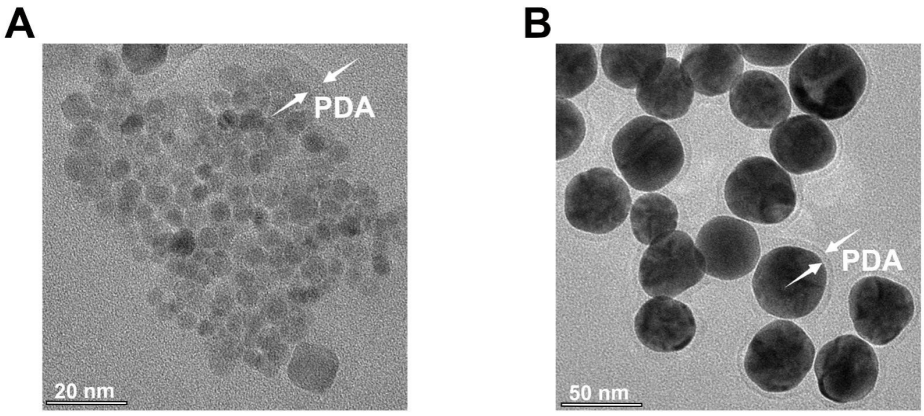

Figure. S3. (A-B) TEM image of IO NPs@PDA (A) and Au NPs@PDA (B).

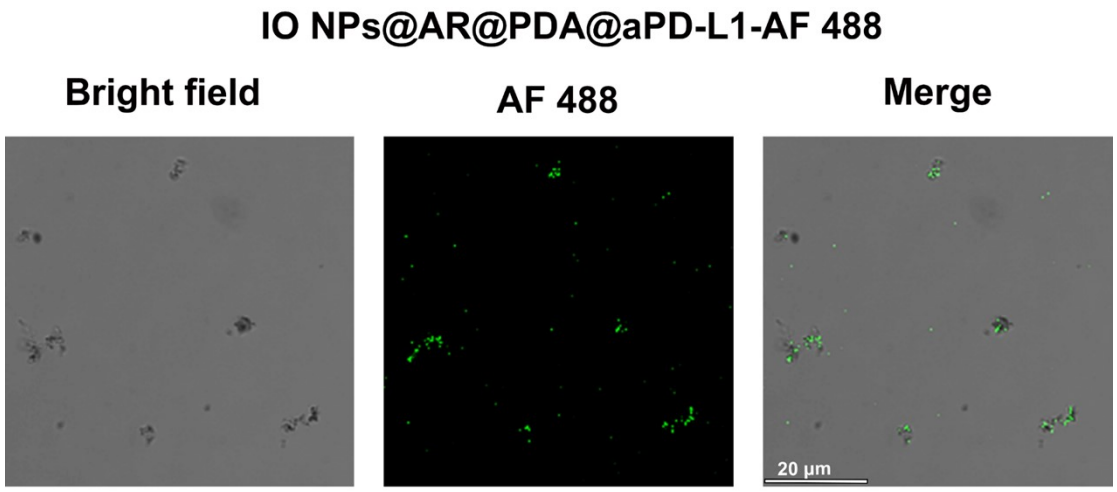

Figure. S4. Fluorescence image of Alexa Fluor 488 labeled PD-L1 antibody after conjugated to IO NPs@AR@PDA.

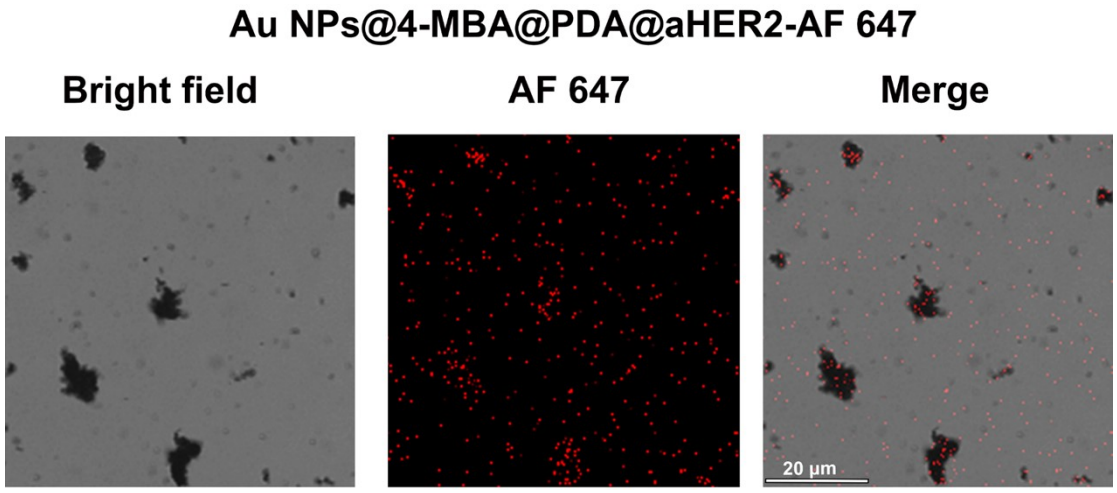

Figure. S5. Fluorescence image of Alexa Fluor 647 labeled HER2 antibody after conjugated to Au NPs@4-MBA@PDA.

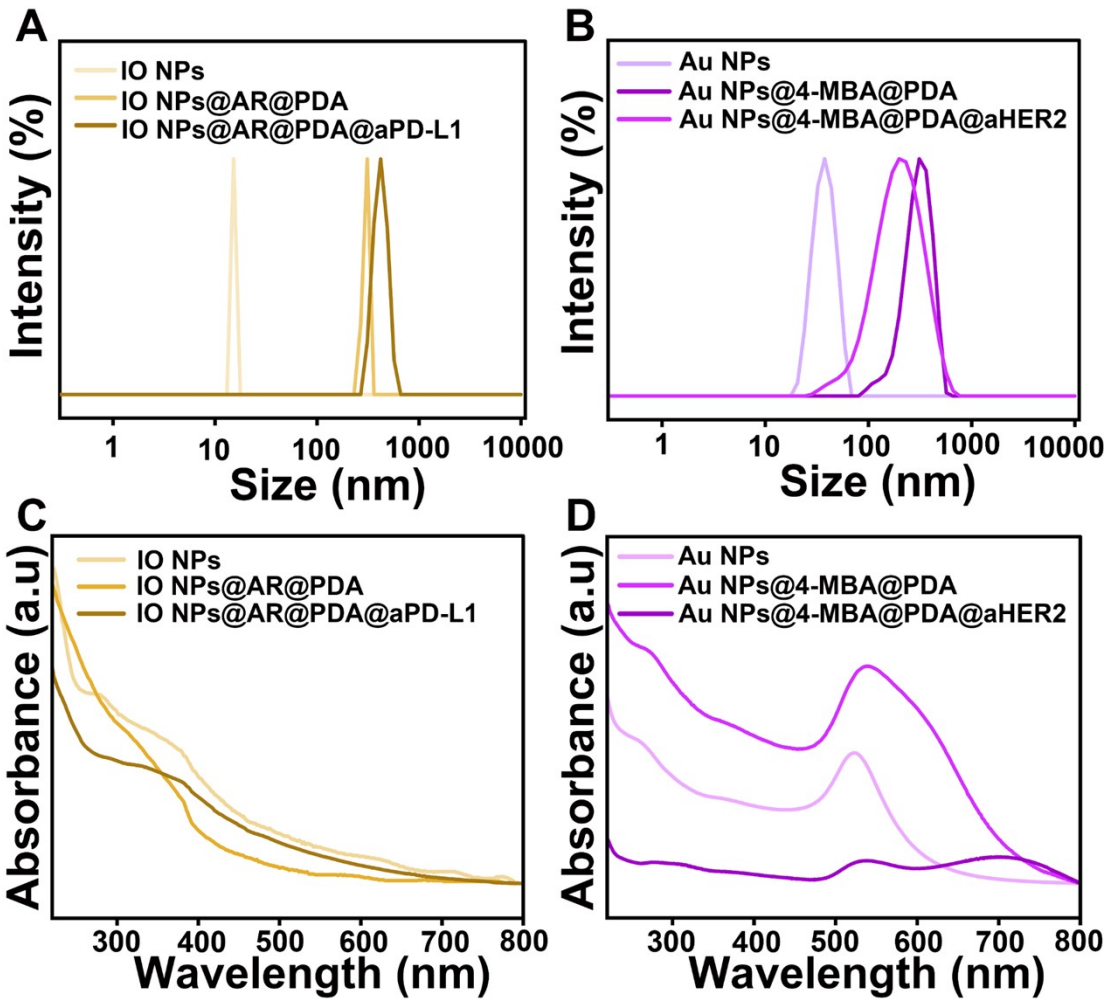

Figure. S6. (A) Size distributions of IO NPs, IO NPs@AR@PDA, IO NPs@AR@PDA@aPD-L1. (B) Size distributions of Au NPs, Au NPs@4-MBA@PDA, Au NPs@4-MBA@PDA@aHER2. (C) UV-Vis spectra of IO NPs, IO NPs@AR@PDA, IO NPs@AR@PDA@aPD-L1. (D) UV-Vis spectra of Au NPs, Au NPs@4-MBA@PDA, Au NPs@4-MBA@PDA@aHER2.

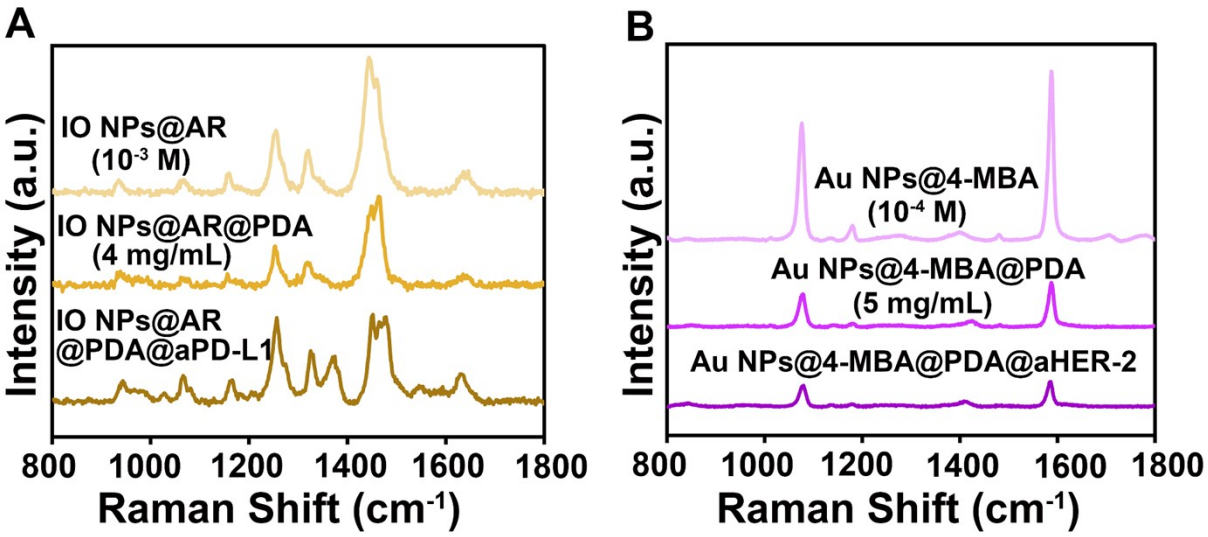

Figure S7. (B) SERS spectra of IO NPs@AR, IO NPs@AR@PDA and IO NPs@AR@PDA@aPD-L1 (AR: 5 × 10-3 mol/L, PDA: 4 mg/mL). (D) SERS spectra of Au NPs@4-MBA, Au NPs@4-MBA@PDA, Au NPs@4-MBA@PDA@aHER2 (4-MBA: 2 × 10-4 mol/L, PDA: 5 mg/mL).

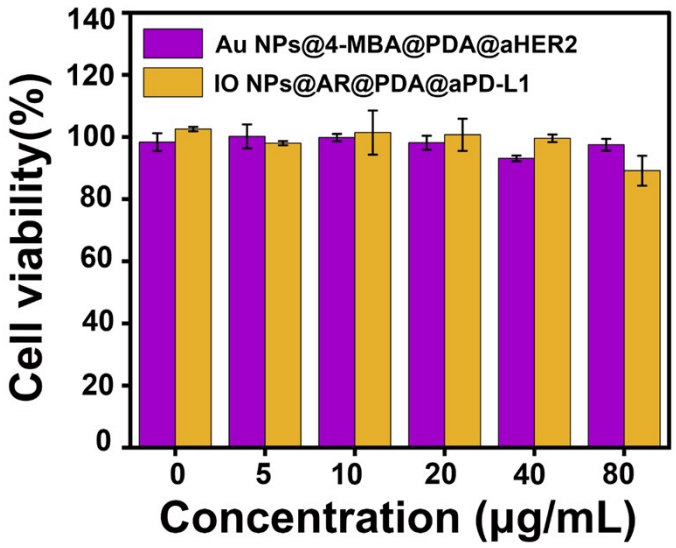

Figure. S8. CCK-8 assay of two bioprobes with different concentrations in MCF-7 cells.

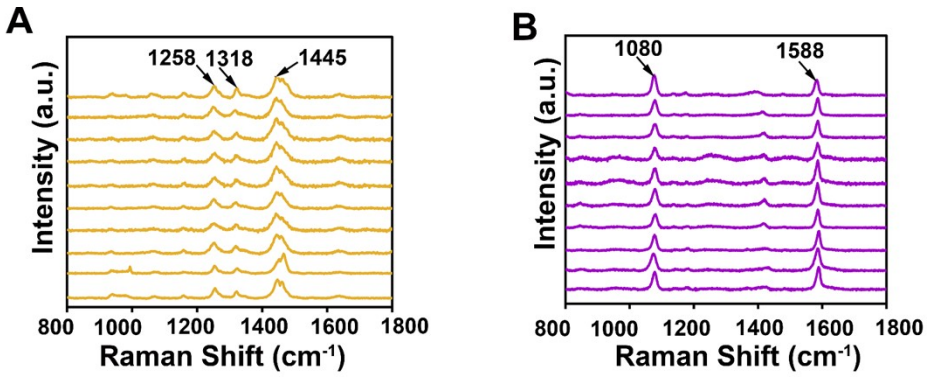

Figure. S9. (A) SERS spectra of 10 points (B) for IO NPs@AR@PDA@aPD-L1. (B) SERS spectra of 10 points (E) for Au NPs@4-MBA@PDA@aHER2. Laser wavelength: 532 nm, laser power: 268 mW, lens: 50× objective for IO NPs@AR@PDA@aPD-L1 bioprobes; laser wavelength: 633 nm; laser power: 104 mW; lens: 50× objective for Au NPs@4-MBA@PDA@aHER2 bioprobes

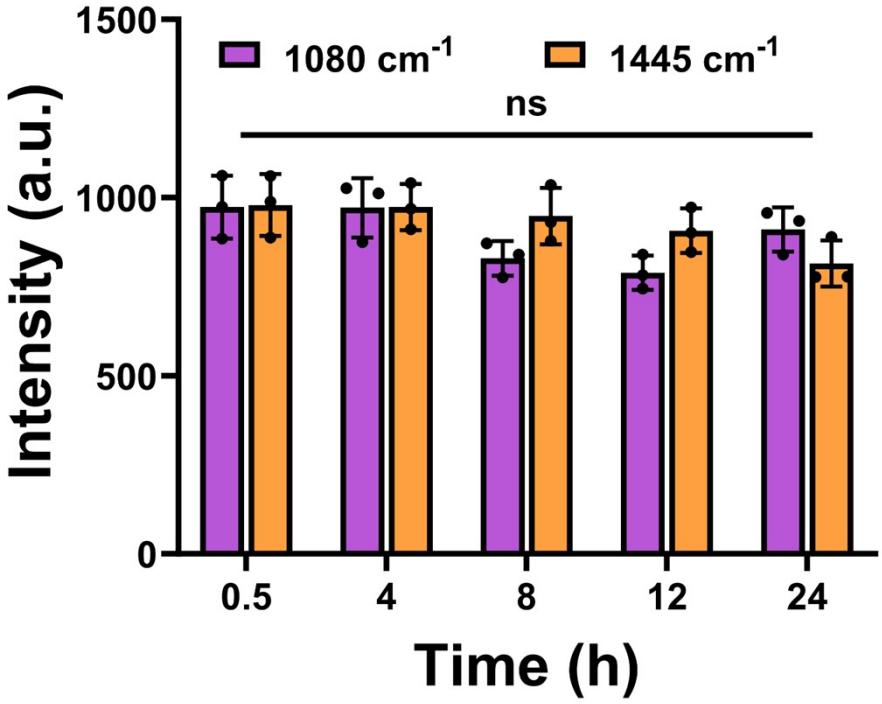

Figure S10. Au NPs@4-MBA@PDA@aHER2 and the stability specific peak signals at 1080 cm-1 and IO NPs@AR@PDA@aPD-L1 and the stability specific peak signals at 1445 cm-1. (n=3, ns means no statistical difference, $p>0.05$)

Table S1. Raman shift 1080 cm-1 and 1445 cm-1 RSD values at different time in a simulated plasma environment.

Time (h)	1080 cm^{-1}	1445 cm^{-1}
0.5	9.06%	8.85%
4	8.60%	6.68%
8	5.88%	8.38%
12	6.09%	6.89%
24	6.82%	7.93%

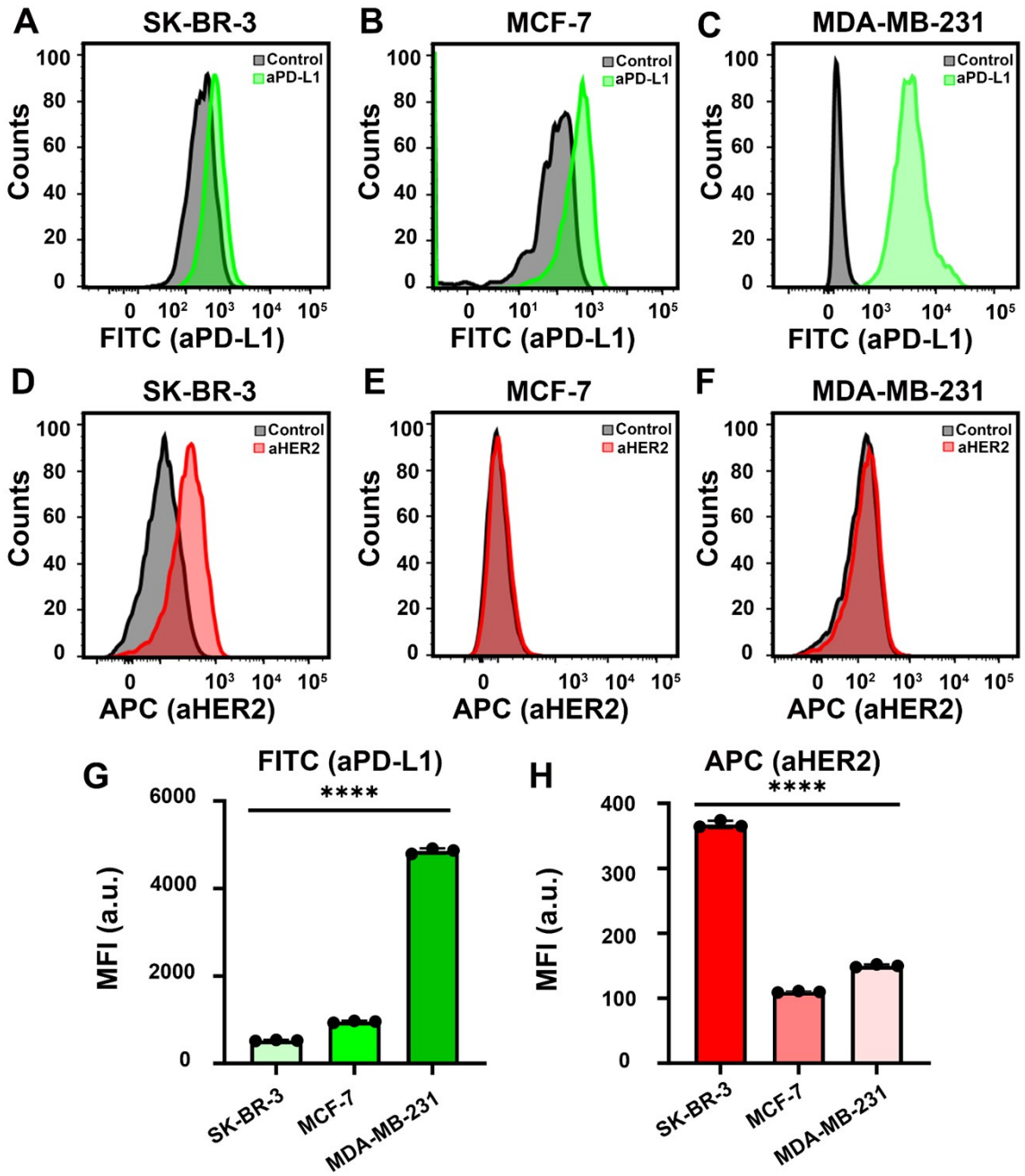

Figure. S11. (A-C) Flow cytometry of PD-L1 expression of SK-BR-3 (A), MCF-7(B), and MDA-MB-231(C). (D-F) Flow cytometry analysis of the function of aHER2 in SK-BR-3 (D), MCF-7(E), and MDA-MB-231(F). (G, H) Mean F.L. intensity of FITC (G) and APC (H) in SK-BR-3, MCF-7, MDA-MB-231 cells. (n=3, ****$p < 0.0001$).

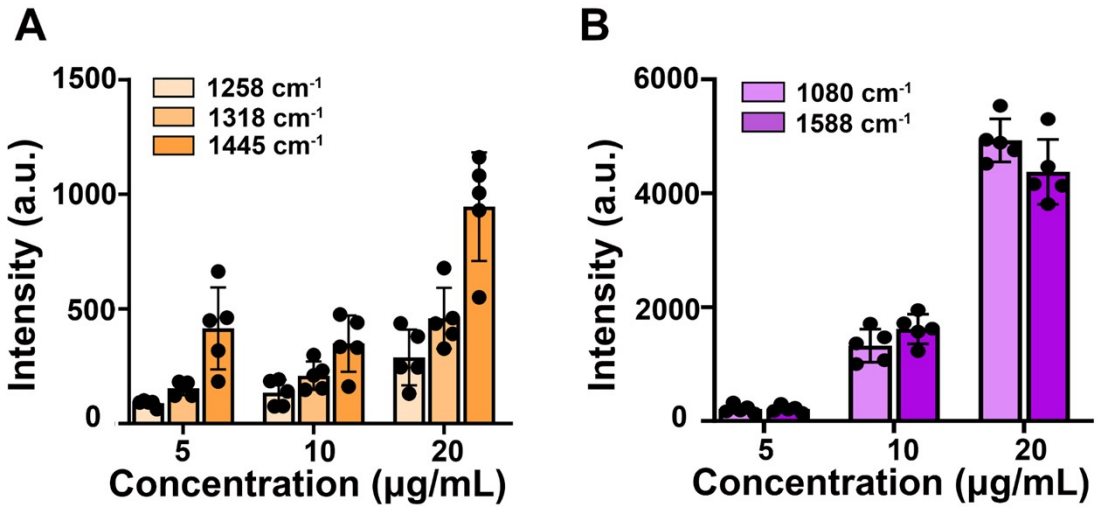

Figure. S12. (A) The average SERS intensity at 1258, 1318 and 1445 cm^{-1} (corresponding to IO NPs@AR@PDA@aPD-L1) and 1080 and 1588 cm^{-1} (corresponding to Au NPs@4-MBA@PDA@aHER2) at different concentrations. (n=5). Laser wavelength: 532 nm, laser power: 670 mW, lens: 50× objective for IO NPs@AR@PDA@aPD-L1 probes; laser wavelength: 633 nm; laser power: 260 mW; lens: 50× objective for Au NPs@4-MBA@PDA@aHER2 probes.

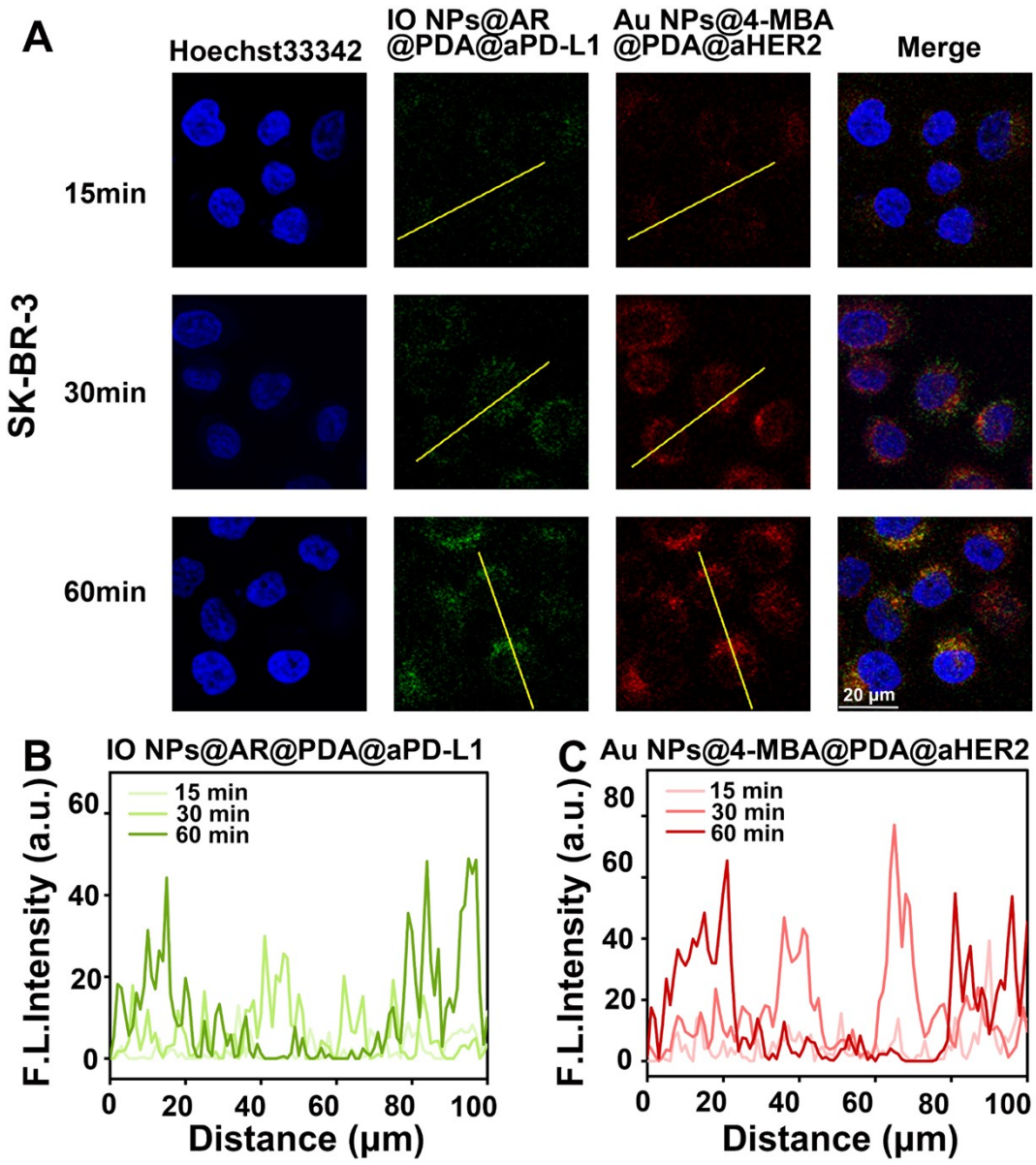

Figure. S13. The colocalization of two SERS probes in SK-BR-3 by fluorescence imaging. (A) The CLSM images of SK-BR-3 cells incubated with two targeted SERS nanoprobes for 15, 30, 60 min, respectively. (B, C) Semiquantitative fluorescence intensity of SK-BR-3 cells after incubation with two targeted SERS probes for different times.

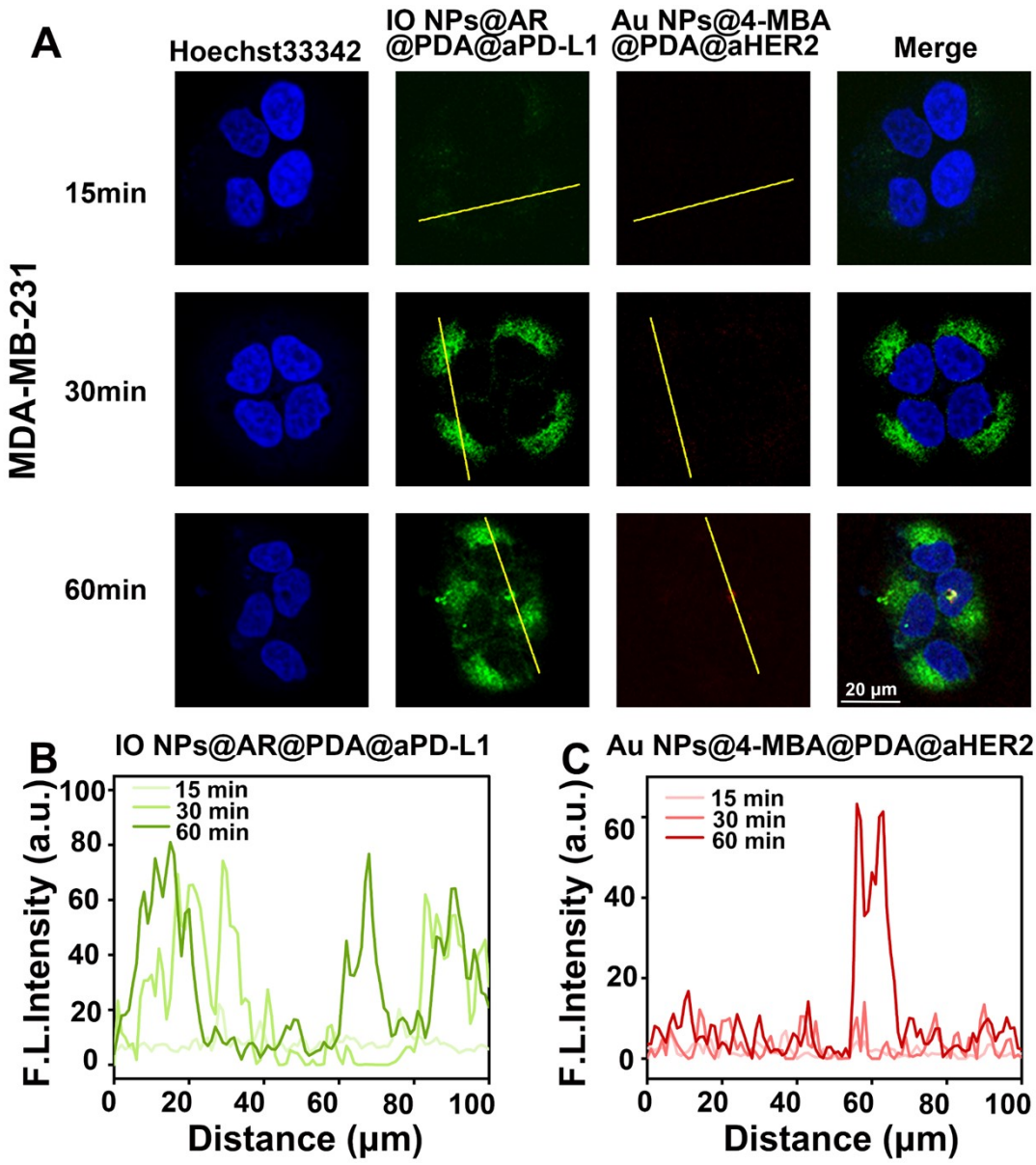

Figure. S14. The colocalization of two SERS probes in MDA-MB-231 cells by fluorescence imaging. (A) The CLSM images of MDA-MB-231 cells incubated with two targeted SERS nanoprobes for 15, 30, 60 min, respectively. (B, C) Semiquantitative fluorescence intensity of MDA-MB-231 cells after incubation with two targeted SERS probes for different times.

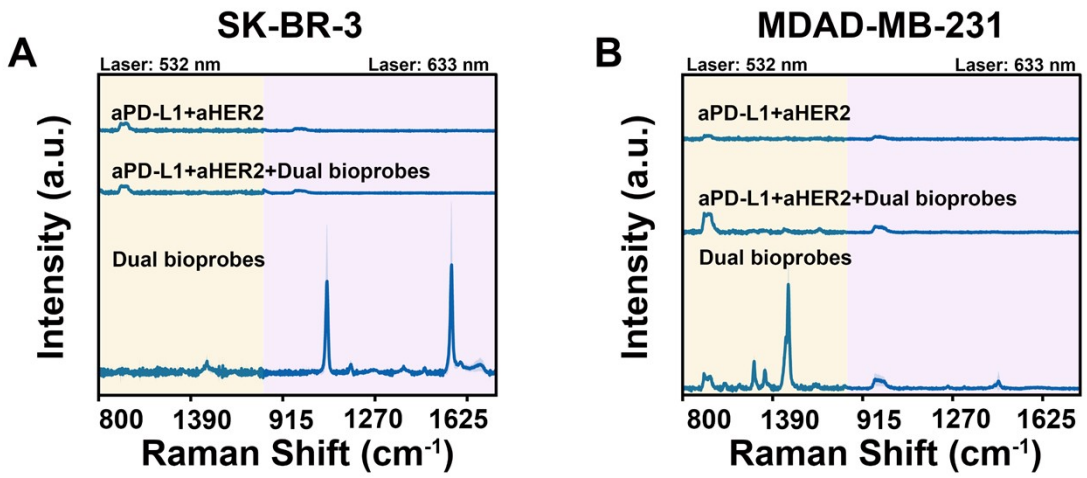

Figure S15. Competitive blocking assays of SK-BR-3 cells (A), MDA-MB-231 (B).

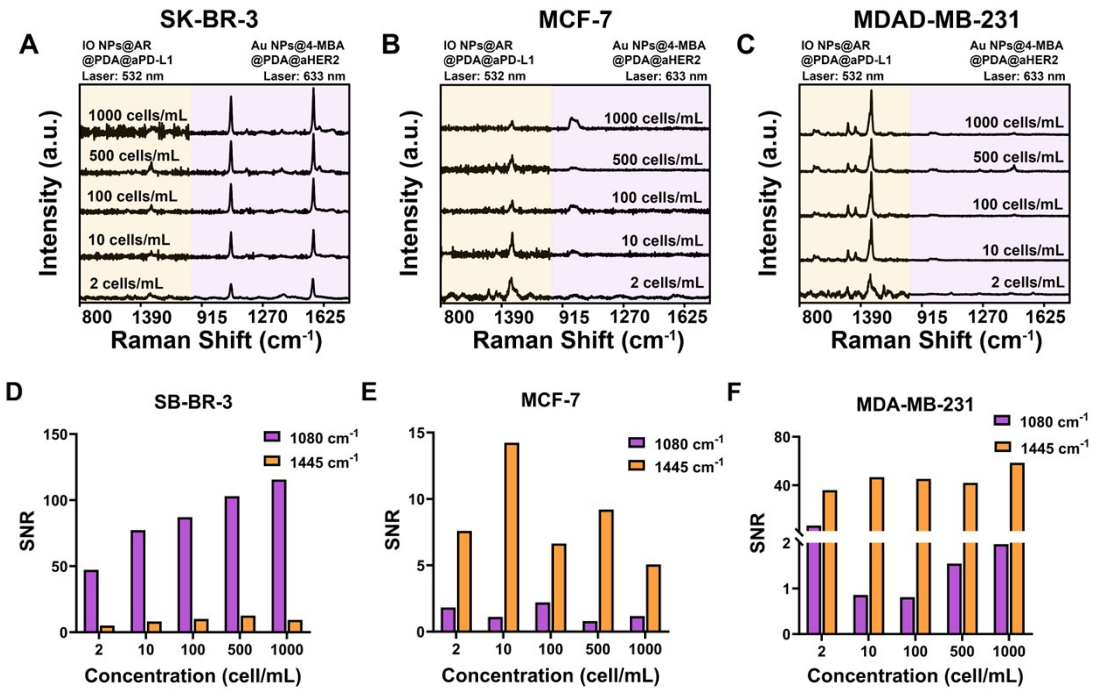

Figure S16. (A-C) The sensitivity of CTC detection by two SERS bioprobes with cell concentration of 2-1000 cells/mL (SK-BR-3, MCF-7, and MDA-MB-231 cells in blood). (D-F) The SNR of 1080 cm-1 and 1445 cm-1 with cell concentration of 2-1000 cells/mL (SK-BR-3, MCF-7, and MDA-MB-231 cells in blood).

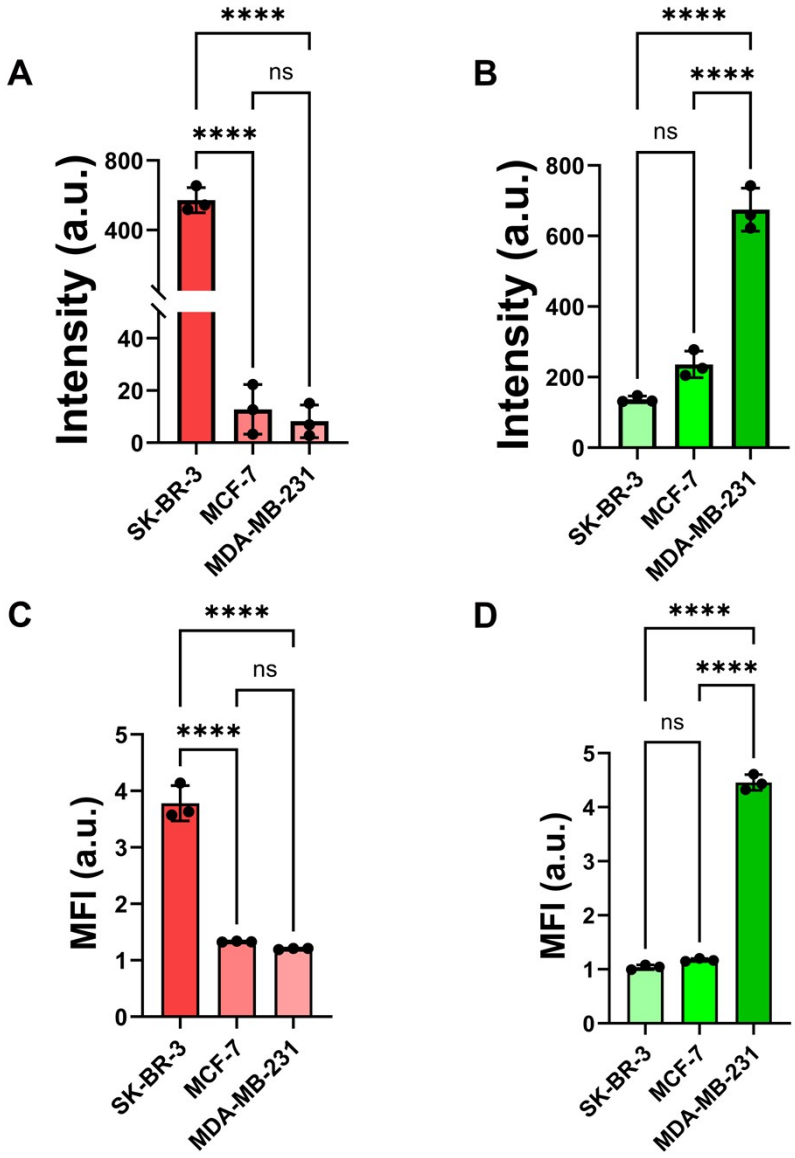

Figure S17. SERS intensity of Au NPs@4-MBA@PDA@aHER2 at 1080 cm^{-1} (A) and IO NPs@AR@PDA@aPD-L1 at 1445cm^{-1} (B). Mean F.L. intensity of Au NPs@4-MBA@PDA@aHER2-AF647 (C) and IO NPs@AR@PDA@aPD-L1-AF488 (D) in SK-BR-3, MCF-7, MDA-MB-231 cells. (n=3, ns means no statistical difference, ****$p < 0.0001$).

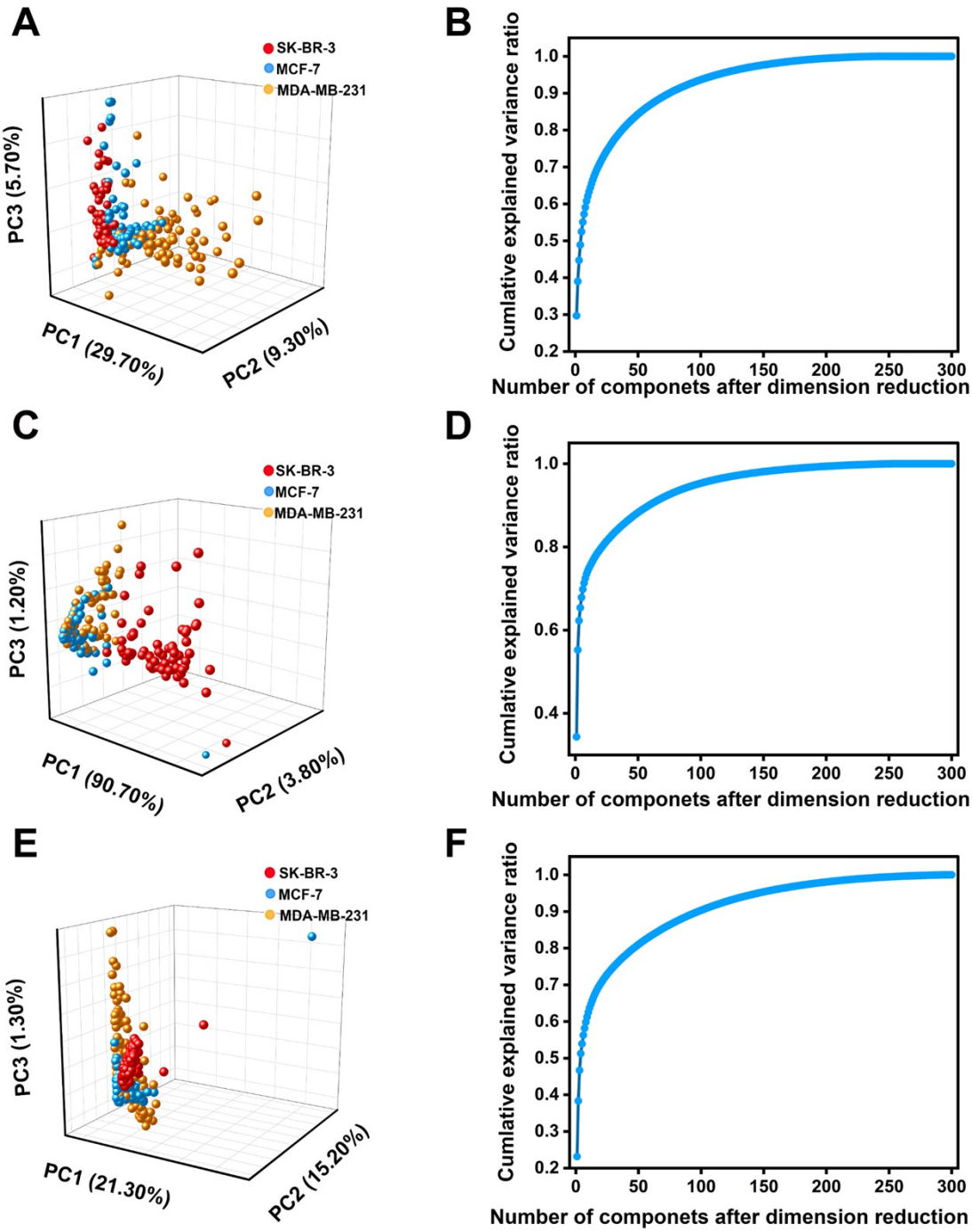

Figure. S18. (A, B) 3D PCA projection (A) and cumulative explained variance curve corresponding (B) of SK-BR-3, MCF-7, and MDA-MB-231 cells classified using the IO@AR@PDA-aPD-L1 probe; (C, D) 3D PCA projection (C) and cumulative explained variance

curve corresponding (D) of SK-BR-3, MCF-7, and MDA-MB-231 cells classified using the Au NPs@4-MBA@PDA@aHER2 probe; (E, F) 3D PCA projection (E) and cumulative explained variance curve corresponding (F) of SK-BR-3, MCF-7, and MDA-MB-231 cells classified using the dual-targeting SERS probe.

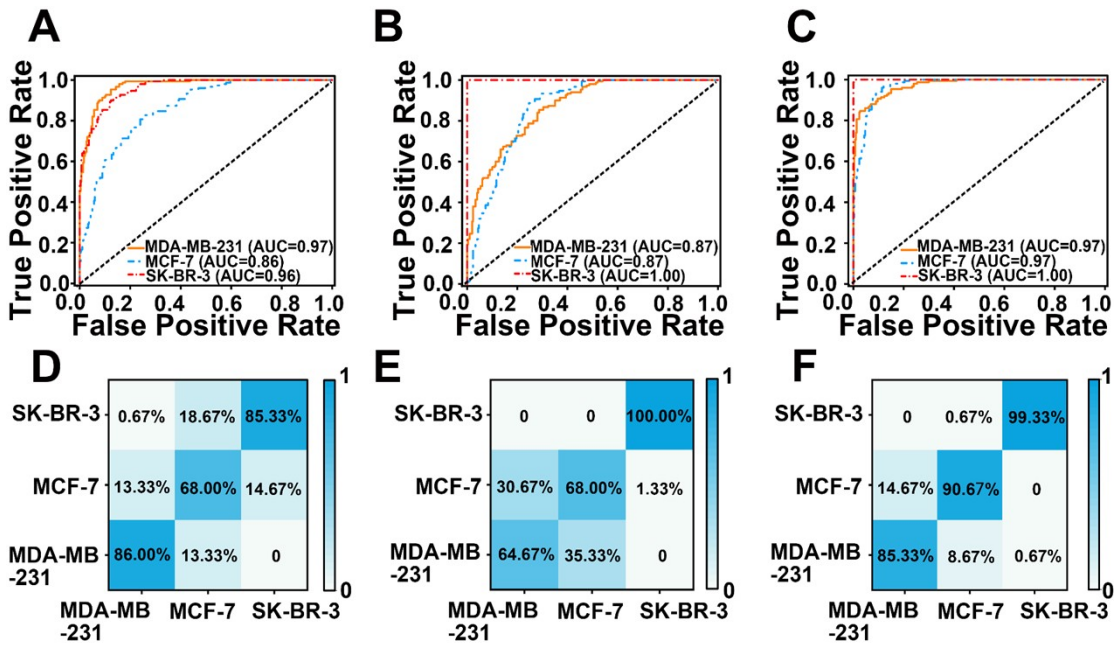

Figure. S19. The machine learning (PCA-Random Forest) analysis of SERS spectra in CTCs detection for the identification of therapeutic targets. (A-C) the ROC curve derived from the validation set results of the PCA-Random Forest mode of SK-BR-3 cells, MCF-7 cells and MDA-MB-231 cells by IO@AR@PDA-aPD-L1 (A) Au NPs@4-MBA@PDA@aHER2 (B) dual-probes (C). (D-F) confusion matrix heatmap of the validation set as predicted by the trained PCA-Random Forest model of SK-BR-3 cells, MCF-7 cells and MDA-MB-231 cells by IO@AR@PDA-aPD-L1 (D) Au NPs@4-MBA@PDA@aHER2 (E) dual-probes (F).

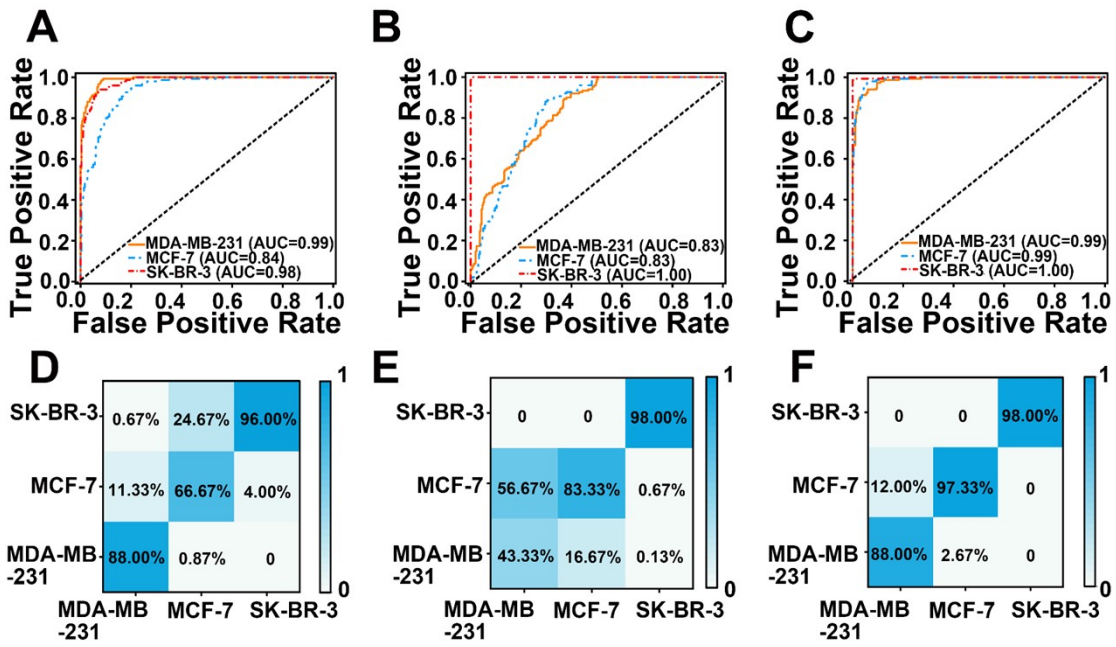

Figure. S20. The machine learning (PCA-SVM) analysis of SERS spectra in CTCs detection for the identification of therapeutic targets. (A-C) the ROC curve derived from the validation set results of the PCA-SVM mode of SK-BR-3 cells, MCF-7 cells and MDA-MB-231 cells by IO@AR@PDA-aPD-L1 (A) Au NPs@4-MBA@PDA@aHER2 (B) dual-probes (C). (D-F) confusion matrix heatmap of the validation set as predicted by the trained PCA-SVM model of SK-BR-3 cells, MCF-7 cells and MDA-MB-231 cells by IO@AR@PDA-aPD-L1 (D) Au NPs@4-MBA@PDA@aHER2 (E) dual-probes (F).

Table S2. Comparison of the discrimination results of three different machine learning models.

Machine learning model	IO NPs@AR@PDA @aPD-L1	Au NPs@4-MBA@PDA @aHER2	Dual-probes
PCA-LDA	79.33%	73.78%	94.22%
PCA- Random Forest	79.78%	77.76%	91.78%
PCA-SVM	83.56%	74.89%	94.44%

Table S3. Comparison of sensitivity and specificity of dual-probe detection of different cell lines assisted by three machine learning models.

Cell line	Sensitivity			Specificity		
	PCA-LDA	PCA-Random Forest	PCA-SVM	PCA-LDA	PCA-Random Forest	PCA-SVM
MDA-MB-231	86.00%	85.33%	88.00%	99.00%	95.33%	97.67%
MCF-7	98.00%	90.67%	97.33%	92.33%	92.67%	94.00%
SK-BR-3	98.67%	99.33%	98.00%	100.00%	99.67%	100.00%

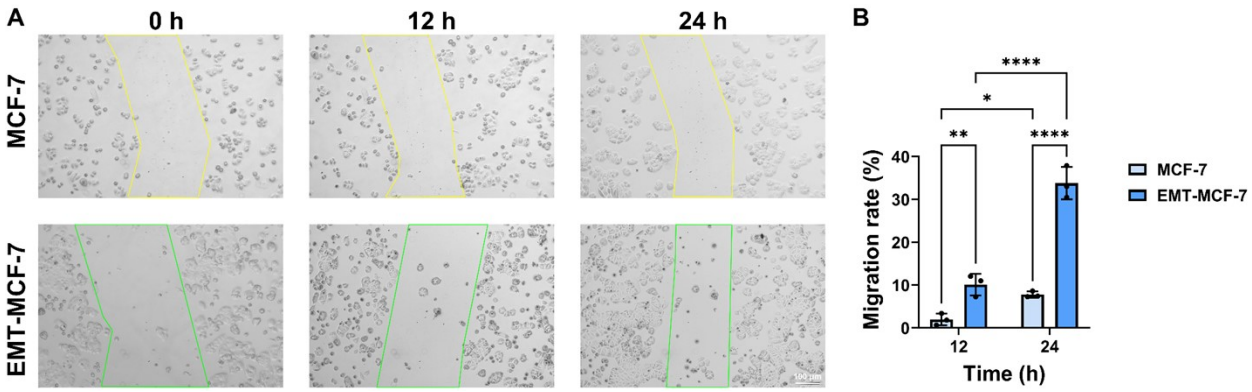

Figure S21. TGF-β–induced EMT enhances the migratory capacity of MCF-7 cells. (A) Representative bright-field images of wound healing assays at 0, 12, and 24 h after scratch in MCF-7 and EMT-MCF-7 cells. (B) Quantification of cell migration rate at 12 h and 24 h (n=3, *p<0.05, **p<0.01, ****p<0.0001).

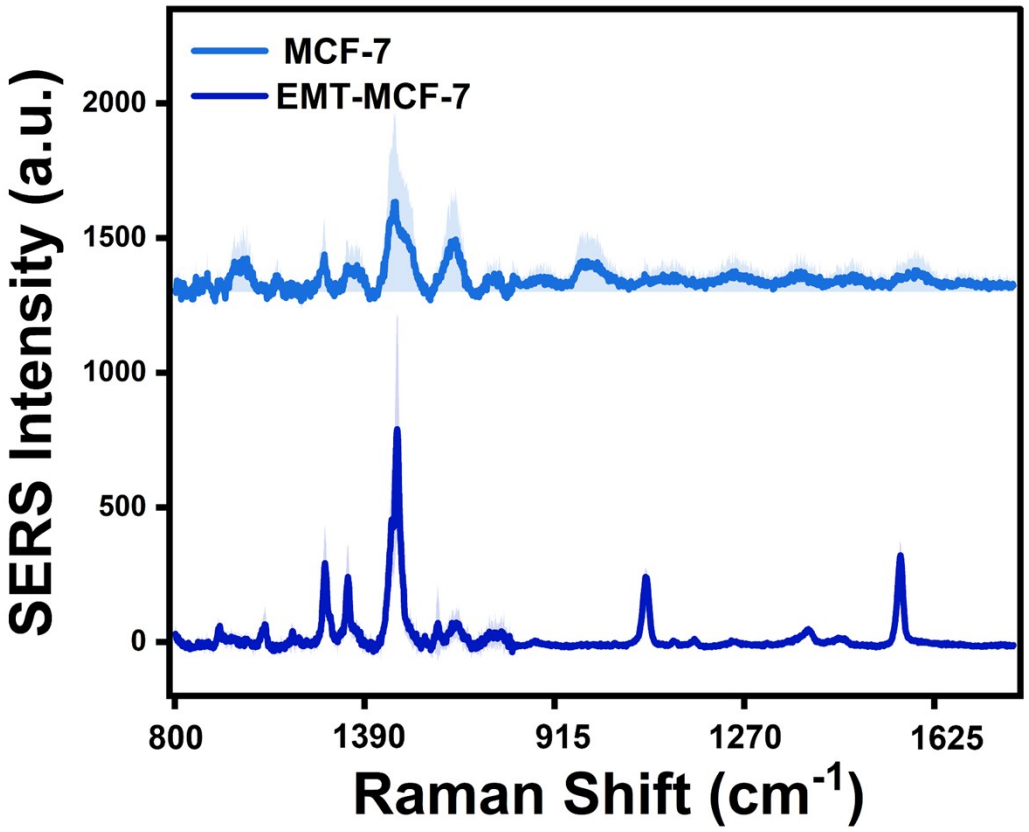

Figure S22. The SERS spectra of MCF-7 and EMT-MCF-7 cells for dual-targeted bioprobes. Laser wavelength: 532 nm, laser power: 268 mW, lens: 50× objective for the IO NPs@AR@PDA@aPD-L1 bioprobe; laser wavelength: 633 nm; laser power: 104 mW; lens: 50× objective for the Au NPs@4-MBA@PDA@aHER2 bioprobe.

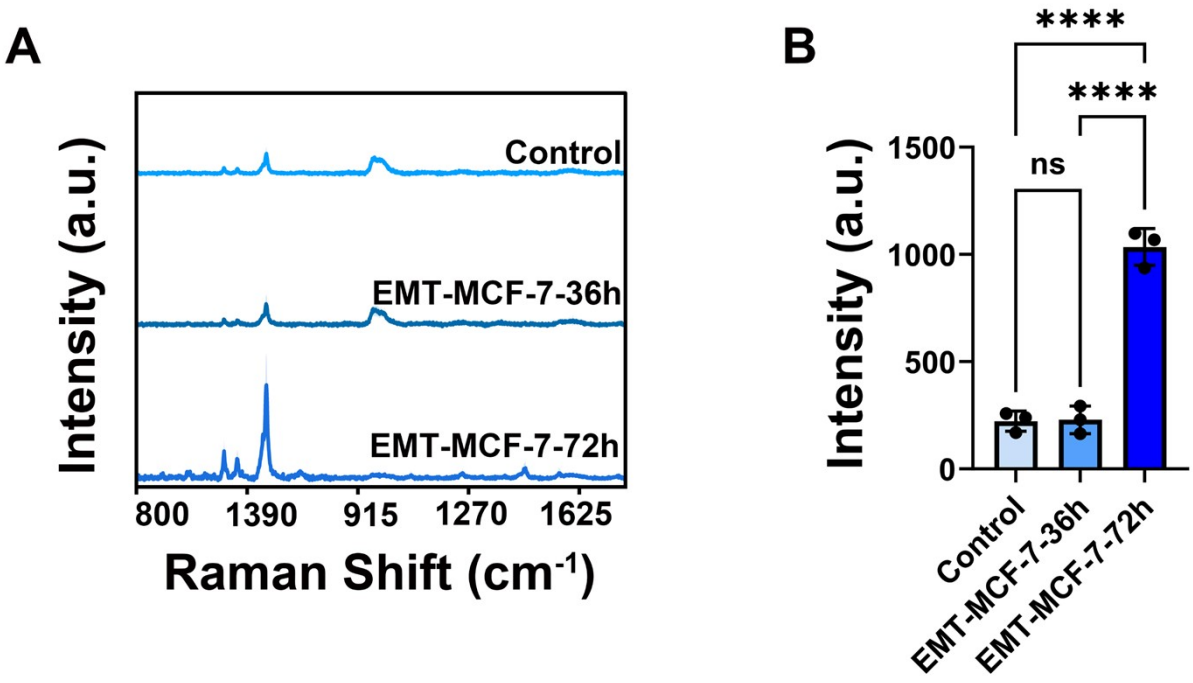

Figure S23. (A) The SERS spectra of MCF-7, EMT-MCF-7-36h and EMT-MCF-7-72h for dual-targeted bioprobes, (B) peak signals at 1445 cm-1 of MCF-7, EMT-MCF-7-36h and EMT-MCF-7-72h. (n=3, ns means no statistical difference, ****$p < 0.0001$)

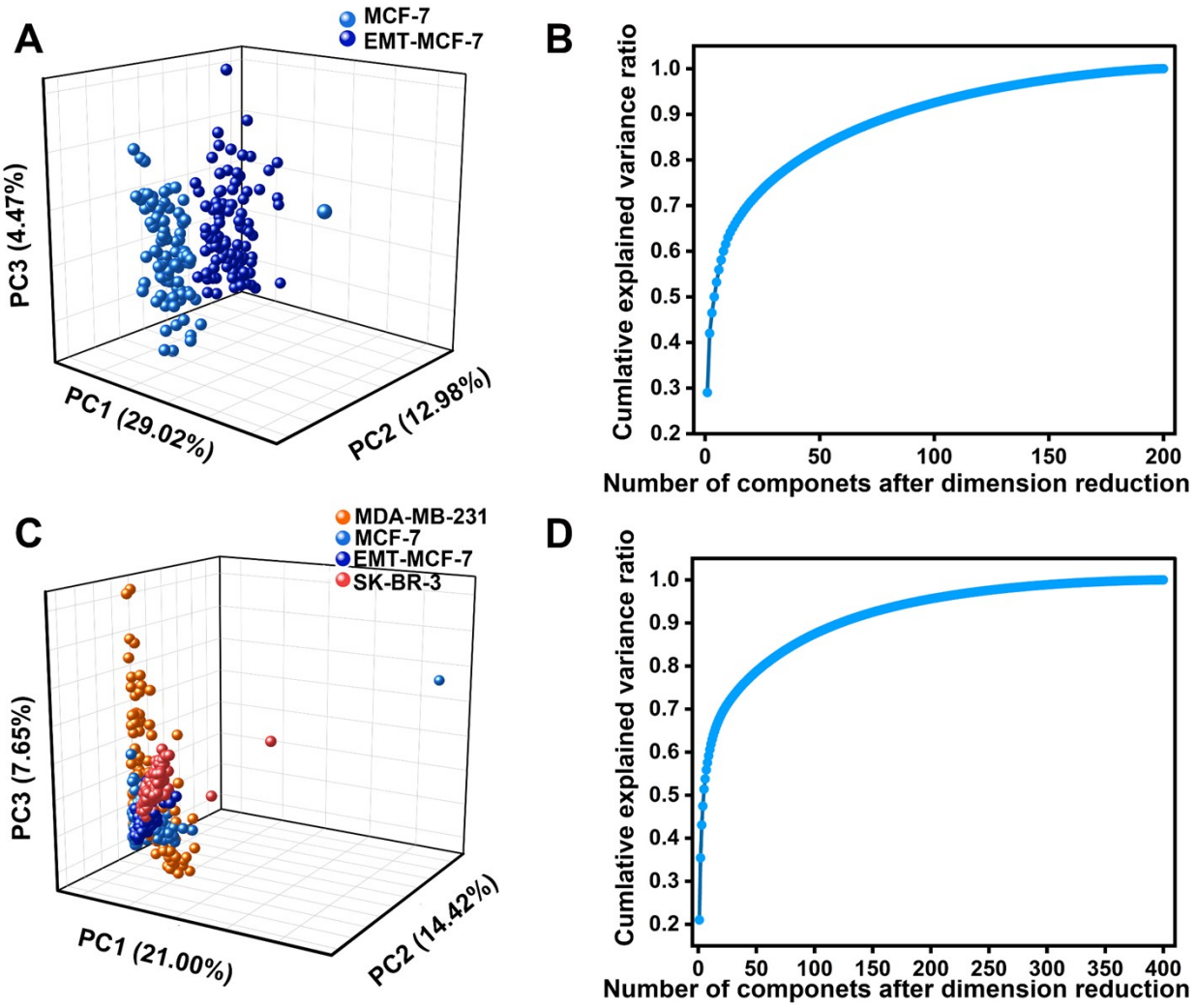

Figure S24. (A, B) 3D PCA projection (A) and cumulative explained variance curve corresponding (B) of MCF-7, and EMT-MCF-7 cells classified using the dual-targeting SERS probes. (C, D) 3D PCA projection (C) and cumulative explained variance curve corresponding (D) of SK-BR-3, MCF-7, EMT-MCF-7 and MDA-MB-231 cells classified using the dual-targeting SERS probes.

REFERENCE

1. T. Pan, D. Zhang, G. You, X. Wu, C. Zhang, X. Miao, W. Ren, Y. He, L. He, Y. Gong, J. Lin, A. Wu and G. Shao, PD-L1 targeted iron oxide SERS bioprobe for accurately detecting circulating tumor cells and delineating tumor boundary, *Chin. Chem. Lett.*, 2025, **36**, 109857.
2. C. Zhang, L. Xu, X. Miao, D. Zhang, Y. Xie, Y. Hu, Z. Zhang, X. Wang, X. Wu, Z. Liu, W. Zang, C. He, Z. Li, W. Ren, T. Chen, C. Xu, Y. Zhang, A. Wu and J. Lin, Machine learning assisted dual-modal SERS detection for circulating tumor cells, *Biosens. Bioelectron.*, 2025, **268**, 116897.
3. X. Wu, D. Zhang, Y. Ding, F. Cao, Y. Li, J. Yao, X. Miao, L. He, J. Luo, J. Li, J. Lin, A. Wu and J. Zheng, Self-assembled co-delivery system of gold nanoparticles and paclitaxel based on in-situ dynamic covalent chemistry for synergistic chemo-photothermal therapy, *Rare Met.*, 2025, **44**, 417-429.