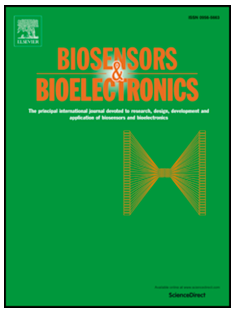


Profiling DNA mutation patterns by SERS fingerprinting for supervised cancer classification

Lei Wu, Alexandra Teixeira, Alejandro Garrido-Maestu, Laura Muinelo-Romay, Luis Lima, Lúcio Lara Santos, Marta Prado, Lorena Diéguez

PII: S0956-5663(20)30386-9

DOI: https://doi.org/10.1016/j.bios.2020.112392

Reference: BIOS 112392

To appear in: *Biosensors and Bioelectronics*

Received Date: 24 March 2020

Revised Date: 4 June 2020

Accepted Date: 13 June 2020

Please cite this article as: Wu, L., Teixeira, A., Garrido-Maestu, A., Muinelo-Romay, L., Lima, L., Santos, Lú.Lara., Prado, M., Diéguez, L., Profiling DNA mutation patterns by SERS fingerprinting for supervised cancer classification, *Biosensors and Bioelectronics* (2020), doi: https://doi.org/10.1016/j.bios.2020.112392.

Credit Author Statement

Lei Wu: *Conceptualization, Methodology, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization.*

Alexandra Teixeira: *Validation, Investigation, Writing - Review & Editing.*

Alejandro Garrido-Maestu: *Validation, Formal analysis, Investigation.*

Laura Muinelo-Romay: *Validation, Investigation, Resources, Writing - Review & Editing.*

Luis Lima: *Resources.*

Lúcio Lara Santos: *Resources.*

Marta Prado: *Resources, Writing - Review & Editing.*

Lorena Diéguez: *Conceptualization, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.*

Profiling DNA Mutation Patterns by SERS Fingerprinting for Supervised Cancer Classification

Lei Wu[1], Alexandra Teixeira[1], Alejandro Garrido-Maestu[2], Laura Muinelo-Romay[3,4], Luis Lima[5,6,7], Lúcio Lara Santos[5,8,9], Marta Prado[2], Lorena Diéguez[1]*

[1] Medical Devices, International Iberian Nanotechnology Laboratory - INL, 4715-330 Braga, Portugal

[2] Food Quality and Safety, International Iberian Nanotechnology Laboratory - INL, 4715-330 Braga, Portugal

[3] Liquid Biopsy Analysis Unit, Oncomet, Health Research Institute of Santiago (IDIS), Complejo Hospitalario de Santiago de Compostela, Trav. Choupana s/n, 15706, Santiago de Compostela, Spain

[4] CIBERONC, Centro de Investigación Biomédica en Red Cáncer Calle de Melchor Fernández Almagro, 3, 28029, Madrid, Spain

[5] Experimental Pathology and Therapeutics Group, Portuguese Institute of Oncology, R. Dr. António Bernardino de Almeida 865, 4200-072 Porto, Portugal;

[6] Glycobiology in Cancer, Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), R. Alfredo Allen 208, 4200-135 Porto, Portugal.

[7] Instituto de Investigação e Inovação em Saúde (I3S), University of Porto, R. Alfredo Allen 208, 4200-135 Porto, Portugal.

[8] Health School of University Fernando Pessoa, Porto, Portugal

[9] Department of Surgical Oncology, Portuguese Institute of Oncology, Porto, Portugal

*E-mail: *lorena.dieguez@inl.int*

Abstract

Profiling DNA mutation patterns for cancer classification plays an essential role in precision and personalized medicine. Conventional PCR-based mutation assay is limited by the extensive labour on target amplification. We herein create an amplification-free surface enhanced Raman spectroscopy (SERS) biochip which enables direct and simultaneous identification of multiple point mutations in tumor cells. Without pre-amplifying the target sequences, the SERS assay reads out the presence of cellular mutations through the interpretation of Raman fingerprints. The SERS sensor is integrated into a microfluidic chip, achieving one-step multiplex analysis within 40 min. Importantly, by combining SERS spectra encoding technique with supervised learning algorithm, a panel of nucleotide mixtures can be well distinguished according to their mutation profiles. We initially demonstrate an excellent levels of classification in samples from colorectal cancer and melanoma cell lines. For final clinical validation, the system performance is verified by classifying cancer patient samples, which shows an accuracy above 90%. Due to the simplicity and rapidness, the SERS biosensor is expected to become a promising tool for clinical point-of-care diagnosis towards precision medicine.

Keywords:

Surface enhanced Raman spectroscopy (SERS); DNA mutation; multiplex detection; spectral encoding; supervised cancer classification

1. Introduction

Cancer is a remarkably heterogeneous disease which requires personalized therapy for each specific patient (Alizadeh et al., 2015; Bedard et al., 2013; Friedman et al., 2015; Ben-David et al., 2019). Increasing evidences have indicated that tumors with similar molecular features could share certain similarities in pathological characteristics and their responses to clinical therapies (van't Veer et al., 2002; Ciriello et al., 2013; Kuijjer et al., 2018). Accordingly, the accurate classification of cancer patients into clinically-meaningful subtypes would help doctors make comprehensive decisions on therapeutic strategies. Cancer is also very dynamic, so molecular analysis needs to be a continuous effort through patient treatment and evolution (Bedard et al., 2013). The attempt to classify cancer patients based on expression profiles of biomarkers, such as surface proteins and mRNAs (Network, 2011; Network, 2012; Collisson et al., 2014), has been successfully applied to stratifying the subtypes of breast cancer and glioblastoma (Koboldt et al., 2012; Sidaway, 2017; Zhang et al., 2018). However, for some other types of cancer, like colorectal and small-cell lung cancer, the expression levels of surface biomarkers do not seem to correlate well with clinical phenotypes, including the survival rate and response to therapies (Muzny et al., 2012; Hofree et al., 2013; Rudin et al., 2019). In this case, the analysis of cancer driver and resistance mutations has proven to be crucial in lung cancer (Mitsudomi et al., 2010; Giaccone, 2005) and in colorectal cancer (De Roock et al., 2010; Misale et al., 2012), respectively. It is widely acknowledged that the accurate identification of particular mutations which affect the efficacy of targeted therapy would help improve drug sensitivity, which further increases the overall survival rate (Jao et al., 2018; Jin et al., 2019). Consequently, to overcome the limitation of expression-based stratification, it is crucial to investigate the mutational spectrum in cancer patients for precision medicine.

Classifying cancer types through mutational screening is challenging, mainly because of the low frequency and rich diversity of mutations in tumor cells. As mutant genes can represent an extremely small proportion in the entire genome network (Hofree et al., 2013; Martincorena et al., 2017), an ultrasensitive and specific approach is required to detect the sparse mutation among thousands of wild-type genes. In addition, the mutation patterns across different patients are remarkably diverse. Even within the same type of cancer, there usually exhibits different mutation profiles. It is common that patients with the same type of cancer share no more than a single mutation (Kuijjer et al., 2018; Hofree et al., 2013; Salk et al., 2010; Zhang et al., 2015). To divide the heterogeneous population of tumors into clinically relevant subtypes, it is required to develop a multiplex sensing tool which provides a comprehensive view of the mutational landscape. As such, molecular diagnosis has become the paradigm of precision medicine but, despite the added value, it is not yet fully implemented in clinical routine practice.

The current gold standard for profiling cellular mutation relies on sequencing (Mamanova et al., 2010) or polymerase chain reaction (PCR) (Thierry et al., 2014) techniques. DNA/RNA sequencing is a powerful

tool to gain insights into the biological and medical mysteries of cancer, but not an ideal choice for the clinical routine due to the long processing time and high cost. Some advanced PCR techniques, including COLD-PCR (Li et al., 2008), PCR clamping (Nagai et al., 2005) and ligase chain reaction (LCR) (Shi et al., 2004), have successfully identified point mutations by selectively amplifying target sequence with a single base variation. However, it requires trained personnel and a number of handling steps, which limits its general application as a point-of-care tool. Additionally, multiplex PCR assays tend to induce unbalanced amplification of different DNA fragments. The primer pairs need to be carefully designed and optimized to achieve specific and even amplification of multiple target sequences (Quick et al., 2017). Therefore, a PCR-free method which simultaneously profiles multiple mutations is urgently needed.

Surface enhanced Raman spectroscopy (SERS) has gained increasing popularity in biosensing, especially due to the ultrasensitivity and multiplexing capability (Wang et al., 2013; Langer et al., 2019; Wang et al., 2019). On the one hand, the strong plasmonic enhancement allows SERS to detect targets at ultralow level down to single-molecule resolution (Chen et al., 2018; Zhang et al., 2019; Fang et al., 2019). On the other hand, the narrow width of vibrational Raman bands endows SERS excellent spectra encoding capacity to simultaneously detect multiple target molecules (Wang et al., 2018; Wang et al., 2020). Several research groups combined SERS technique with PCR amplification to analyse nucleic acids (Gracie et al., 2014; Wu et al., 2020), which have been sucessfully adapted for mutation detection (Wee et al., 2016; Li et al., 2018). Neverthless, there are some other impressive works aiming at amplification-free SERS mutation analysis. For example, Xu *et al*. employed iodide-modified silver nanoparticles for ultrasensitive SERS detection of DNAs with single-base selectivity (Xu et al., 2015). Tian *et al.* synthesized aggregated aluminium nanocrystals as the SERS substrate, which enabled direct quantification of nucleobase content without additional probe molecules (Tian et al., 2017). Recently, Huang *et al.* achieved SERS discrimination of single DNA bases in liquid phase by electro-plasmonic trapping (Huang et al., 2019). Luo *et al.* developed a plasmonic gold nanohole array for label-free SERS detection of DNA methylation (Luo et al., 2019). Kowalczyk *et al.* employed the sputtered gold substrate to identify gene mutation based on the conformational change of alkanethiol DNA linker (Kowalczyk et al., 2019). While these amplification-free SERS sensors can be employed for direct identification of single base variation, they have currently been limited to analyse one kind of mutation at a time. It remains challenging to simultaneously identify multiple point mutations in real biological samples without pre-amplifying the target sequences.

In our previous work, we developed an amplification-free SERS sensor to detect *KRAS* point mutation (Wu et al., 2019). However, analysing one kind of mutation can hardly classify heterogeneous tumors with diverse mutation patterns. Additionally, the analysis of clinical samples through multiplexed SERS spectrum is often limited by the band overlap and matrix interferences (Wang et al., 2017; Panneerselvam et al., 2018). To achieve cancer classification in real clinical samples, we herein created an advanced SERS biochip to profile multiplex mutational patterns in patient DNAs, in which Raman spectral encoding is combined with

supervised learning strategy for computational analysis and calculation. The intelligent SERS classifier was initially validated with the task of analysing cellular DNAs from different cancer cell lines, and finally challenged for analysing tumor DNAs from colorectal cancer patients, to demonstrate its potential for clinical applications.

2. Experimental Section

2.1 Materials and instruments

Information on materials and instruments is detailed in supporting information (SI).

2.2 Preparation of SERS nanosubstrate

The SERS nanosubstrate was fabricated by assembling Au@Ag nanorods (NRs) onto glass slides. Briefly, gold nanorods were synthesized by the typical seed-medicated growth method and then coated with silver shell through redox reaction. The obtained Au@Ag NRs were purified, concentrated and then assembled onto the thiol-modified glass slides through Ag-S bonding. Here, to immobilize the nanoparticles onto the substrate in the multi-stripe pattern, a drilled polydimethylsiloxane (PDMS) piece was employed to confine the nanoparticles in the desired regions. The protocols are detailed in SI.

2.3 Fabrication of the SERS-microfluidic sensor

The prepared multi-stripe SERS substrate was conjugated with molecular beacon probes (sequences are shown in SI) for multiplex analysis. First, all of the nucleotide probes were heated to 95 °C for 10 min, and then slowly cooled down to room temperature at a rate of around 2 °C /min. Second, 100 μL of 0.5 μM probe (in phosphate buffered saline (PBS)) was added to each stripe of the SERS nanosubstrate for 12 h reaction at room temperature. To achieve stable DNA immobilization, the substrate was then treated with 100 μL of 0.1 M sodium chloride (NaCl) (in PBS) for 12 h. Finally, the PDMS piece was peeled off the substrate and the glass slide was washed 3 times with milli-Q water and dried with nitrogen gas. For spatial encoding-based parallel analysis, each strip was conjugated with one kind of molecular beacon probe. For spectral encoding-based multiplex sensing, 3 molecular beacon probes were mixed and then modified onto one stripe of SERS substrate. The concentration ratio among the probes was carefully adjusted to normalize the peak intensity of Raman bands from different Raman reporters (Cy5, FAM and ROX on *KRAS G12V, BRAF V600E* and *PIK3CA P449T* probes, respectively). Typically, the ratio among *KRAS G12V, BRAF V600E* and *PIK3CA P449T* probes was 1:3:3.

The functionalized glass slides were bonded against PDMS microchannels to form the SERS-microfluidic sensor. Here, the PDMS microchannels were fabricated through replica molding as previously reported (Garrido-Maestu et al., 2017). To promote the irreversible bonding, the PDMS channel was treated

with oxygen plasma for 1 min and then pressed onto the functionalized glass slide. To ensure effective bonding, a heavy weight is pressed onto the chip for 1 h.

2.4 Mutation assay on chip

All the mutant and wild-type control samples were prepared by diluting with the reaction buffer solution (0.01 M tris (hydroxymethyl) aminomethane-hydrochloric acid (Tris-HCl) buffer with 0.1 M potassium chloride (KCl) and 1 mM magnesium chloride ($MgCl_2$), pH=8) and then introduced into microtubes using the negative pressure induced by the syringe. Next, 25 μL of samples were loaded into the microchannels at 5 μL/min for 5 min, followed by incubation at 50 □ for 30 min. To remove unbound nucleotides, PBS-tween (PBST) solution was flowed through the microchannels at 5 μL/min for 5 min. Prior to SERS measurement, the chip was sealed in a container to avoid liquid evaporation.

2.5 Cell culture and DNA extraction

All cell lines used in this work were cultured in Dulbecco's Modified Eagle's Medium (DMEM) with 10% fetal bovine serum and 1% penicillin-streptomycin. They were seeded into flasks and cultured under standard cell culture conditions (5% CO_2, 37 □). DNA extraction was carried out using the PureLink Genomic DNA Mini Kit (ThermoFisher Scientific), following manufacturer instructions. In brief, for cancer cell lines (SW480, MDA-MB-435 and HT-29), 5×10^6 cells were detached by trypisinization and resuspended in 200 μL PBS solution. For paraffin-embedded patient tissue samples, each sample was mixed with 1 mL of xylene, purified with pure ethanol, and then resuspended in 200 μL PBS solution. Each of these pre-processed samples was added with 20 μL of proteinase K, 20 μL of RNase A and 200 μL of lysis/binding buffer, followed by incubation at 55 □ for 10 min. Next, the cell lysate was mixed with 200 μL of pure ethanol. Finally, a spin column was used to purify the DNAs. The concentration and quality of the obtained genomic DNAs were measured with NanoDrop Spectrophotometer. All the DNA extractions used for mutation tests were ensured high purity with an absorption ratio (A260/A280) above 1.80. They were stored at -20 □ for later use. Before mutation analysis, they were incubated at 95 □ for 10 min and immediately immersed into iced water for 5 min.

3. Results and discussion

3.1 Principle and overview

The working principle of this SERS-based DNA sensor is depicted in a schematic view in **Figure 1**. Here, the representative detection targets were *KRAS, BRAF* and *PIK3CA* genes, which are commonly altered in cancer cells. The following three cancer cell lines were selected as representatives for these mutations: SW480 colorectal cancer cell line carrying *KRAS G12V* (mutation on exon 2 in codon 12), MDA-MB-435

melanoma cell line carrying *BRAF V600E* (mutation on exon 15 in codon 600), and HT-29 colorectal cancer cell line which simultaneously carries *BRAF V600E* and *PIK3CA P449T* (mutation on exon 12 in codon 449) (Ahmed et al., 2013; Ikediobi et al., 2006). Three kinds of molecular beacon probes, which were designed to specifically hybridize with *KRAS G12V, BRAF V600E* and *PIK3CA P449T* respectively, were covalently bound to the SERS nanosubstrates assembled with gold@silver nanorods (Au@Ag NRs) (Figure 1b). Upon hybridization with the target DNA, the hairpin nucleic acid probe experienced a structure change, in which the Raman reporter labelled at the free end of the probe moved away from the surface of the SERS substrate (Figure 1e), thus leading to a decrease in SERS intensity (Figure 1f). As different probes were labelled with different Raman reporters, their distinct Raman fingerprint can be identified through the spectrum decoding, which enables simultaneous detection of multiple mutations. By clustering the SERS spectra of mutation assay for various cancer cells and patient samples, different types of cancer were classified with the help of supervised training algorithm based on classical least square (CLS)-linear discriminant analysis (LDA) (Figure 1g).

3.2 Specific mutation analysis with array-based SERS sensor

Array-based sensor is a perfect choice for the cross-validation of the specificity (Geng et al., 2019; Wu et al., 2015). Three molecular beacon probes labelled with Cy5, FAM and ROX molecules were designed to specifically detect the *KRAS G12V*, *BRAF V600E* and *PIK3CA P449T* mutations respectively (sequences in supporting information (SI)). To inspect the selectivity of each molecular beacon probe, a PDMS chip with 8 microchannels was bonded against a glass slide patterned with 3 stripes of SERS substrate (**Figure 2a**). As each strip was modified with one specific probe, the target sequences flowing through the microfluidic channels successively hybridized with the corresponding probes in the 3×8 array, which allowed parallel testing of *KRAS G12V, BRAF V600E* and *PIK3CA P449T* in 8 samples on one chip.

The specificity of these probes was evaluated with the synthetic mutant sequences of *KRAS G12V, BRAF V600E* and *PIK3CA P449T*, respectively. For each molecular beacon, the complementary mutant sequence was tested as the specific target while the other two non-complementary mutants were used as negative controls. As shown in Figure S4, the cross reaction of 3 probes with 3 sequences results in 9 SERS spectra in the 3×3 array. The histogram in Figure S4 was plot using the peak area at 1589, 1325 and 1493 cm^{-1} to characterize the SERS intensity of Cy5, FAM and ROX respectively (Figure 2b). Among the several predominant Raman bands, these three were chosen due to their relatively strong intensity and less overlap. It can be observed that the hybridization of specific target DNAs with their corresponding probes results in a lower SERS intensity, while that of negative controls leads to a higher intensity, which demonstrated that the designed molecular probes can specifically detect each of these mutations.

To examine the specificity for testing a mixture of multiple targets, a group of samples containing a combination of *KRAS G12V, BRAF V600E* and *PIK3CA P449T* were tested (Figure 2a, table). As shown in

Figure 2c, a significant decrease in SERS intensity of the probe has been observed for samples carrying the corresponding mutations. For example, the measurement results of Sample F carrying *KRAS G12V* and *PIK3CA P449T* exhibited a significant decrease in SERS intensity for Cy5 and ROX, which confirmed the presence of these two mutants. For all of these 8 samples, the colour map in Figure 2c demonstrated that multiple targets can be identified with excellent specificity using this spatially encoded array sensor.

Moreover, to evaluate the sensitivity of this SERS assay, the dependence of SERS intensity on the concentration of target mutation was investigated. Concentrations of target DNAs ranging from 1 nM to 10 fM were measured together with a blank sample. The peak areas at 1589, 1325 and 1493 cm^{-1} were employed to characterize the intensity of the corresponding *KRAS*, *BRAF*, *PIK3CA* mutation probes respectively. It is reflected in Figure S5 and Figure 2d that the SERS intensity decreased with the increasing of target concentrations over 5 orders of magnitude, achieving a detection limit of 10 fM.

3.3 Multiplex mutation analysis & Cancer type classification

SERS possesses excellent multiplexing ability due to the narrow Raman bands. With SERS spectral encoding, the signatures of different Raman reporters can be clearly identified even in a mixture, making it possible to simultaneously profile multiple targets with one SERS measurement (Li et al., 2018; Kamil Reza et al., 2017; Wang et al., 2018). Different from the parallel multiplex detection (as described in the previous section) which collects a single spectrum on each stripe of the SERS substrate, the SERS spectral encoding technique enables one-channel multiplex analysis with just one spectrum acquisition. By decoding the superimposed Raman fingerprints, simultaneous detection can be achieved with one sensing element and one SERS measurement.

In the experiment, molecular beacon probes targeting at *KRAS G12V*, *BRAF V600E* and *PIK3CA P449T* were immobilized onto the same strip of SERS substrate (**Figure 3a**), resulting in a superimposed Raman signature of Cy5, FAM and ROX molecules. The hybridization of target DNAs with their specific probes would lead to the changes in SERS intensity of the corresponding Raman labels. By monitoring the signal change of each Raman reporter, simultaneous detection of three DNA mutations can be achieved, the results of which were subsequently employed for sample classification.

Figure 3b illustrates the workflow for the spectrum acquisition and analysis. The overall idea was to establish an algorithm to classify unknown samples into one of the classes, through a supervised learning strategy using samples of known classes. First, samples of known classes were tested with the SERS-based multiplex mutation assay, in which the superimposed SERS spectra were obtained. Second, to accurately and automatically identify the contribution of each Raman reporter, CLS algorithm was applied to decompose the superimposed spectra (Chuong et al., 2017), resulting in an estimation of contributions from each Raman reporter, as represented by a set of coefficient: $\{P_{i1}, P_{i2}, P_{i3}\}$ (i refers to the sequence of group). Third, using $\{P_{i1}, P_{i2}, P_{i3}\}$ as the input data, the supervised learning algorithm, LDA, was employed for

classification, as plotted by the clusters using canonical variables $\{C_{i1}, C_{i2}\}$ (i refers to the sequence of group).

To evaluate the multiplexing ability using this platform, 8 groups of samples with a combination of three mutations were tested (the composition of each group is shown in Figure 2a). Applying CLS algorithm to decompose the spectra, the contributions of each Raman reporter were obtained and plotted in Figure 3c, which was used as a primary training dataset. The superimposed SERS spectrum is assumed as a linear combination of the individual Cy5, FAM and ROX spectra, and the calculation gives an estimated contribution of each reporter with minimum errors. Afterwards, LDA was further applied to create an algorithm that can organize the samples into one of the defined classes. The resulting LDA scatter plot in Figure 3d showed well-separated clusters, which indicated the selectivity and discrimination ability of this CLS-LDA classifier. The cross-validation results in Figure 3f shows an accuracy of 98.75% for 80 tests (1 error in 80 tests). In addition, the average SERS spectrum of each group shown in Figure 3e provides a distinct Raman fingerprint for each group. All the above results confirmed that the combination of multiplex SERS DNA assay with supervised CLS-LDA classifier enables accurate and automatic classification of unknown samples into different groups.

To evaluate the ability for cancer classification, the developed assay was employed to test DNA lysates from different cancer cell lines. Here, two colorectal cancer cell lines, SW480 and HT-29, as well as a melanoma cell line, MDA-MB-435, were used as classical representatives to validate the ability of this system to detect *KRAS G12V*, *BRAF V600E* and *PIK3CA P449T*. Multiplex SERS detection was performed with the DNA lysates extracted from these cancer cell lines, in which SW480 contains *KRAS G12V*, HT-29 contains *BRAF V600E* and *PIK3CA P449T*, while MDA-MB-435 contains *BRAF V600E* (**Figure 4d**). The mutational status of cancer cell lines was confirmed with droplet digital PCR (Table S1). Considering that the components of DNA lysates from cancer cell lines are much more complicated than those of the pure synthetic DNAs used in our prior tests, it is necessary to establish a new database for the following validation. Hence, we collected SERS spectra of assays on 3 cancer cell lines to establish a training database, in which 10 SERS assay results for each cell line were used. Figure 4a shows the SERS spectra for different cancer cell lines, which exhibit a distinct fingerprint. As shown in the clusters in Figure 4b and c, SW480, MDA-MB-435 and HT-29 cancer cell lines were well-differentiated. Additionally, the cross-validation and another 15 blind tests (5 tests for each cell line) were performed to validate the whole working network (Figure 4e). It is reflected in the table that the blind samples were correctly allocated to the corresponding groups.

Finally, the platform was challenged with the task of classifying patients diagnosed with colorectal cancer, in which DNAs extracted from 8 patient tissue samples were tested. In parallel to SERS detection, the *KRAS G12V* and *BRAF V600E* mutations were tested with real-time PCR. As shown in **Figure 5d** and Figure S8, 2 of these 8 samples (patients 2 and 3) were positive for KRAS G12V while all of the 8 samples

were negative for BRAF V600E. Consequently, we allocated patient 1, 4, 5, 6, 7 and 8 into Group A and patient 2 and 3 into Group B. In the trials analysing *KRAS G12V* and *BRAF V600E* with the SERS method, the corresponding probes were attached to the substrate for multiplex mutation detection. Figure 5a shows the SERS spectra of 10 tests for each sample, among which the spectra for patients 1-6 were employed as the training database for patient classification. As shown in Figure 5b and c, patients with different mutation patterns were well separated, where the cross validation showed an accuracy of 90% (54 of 60 tests) (Figure 5e). Next, another 20 blind tests (10 tests for each of the patients 7-8) were performed with the established CLS-LDA classifier, where the accuracy reached 95% (19 of 20 tests) (Figure 5e). Additionally, the developed system was also capable of classifying colorectal cancer patients through the mutational profiling of plasma DNA lysates (Figure S9). As patients with same kind of gene mutations could show similar pathological characteristics and responses to the specific therapy, this supervised SERS classifier could become a promising tool for personalized medicine.

4. Conclusion

We have developed a SERS approach to classify cancer types by profiling mutational patterns in tumor DNAs. The chip-based mutation sensor which combines SERS spectral encoding technique with supervised learning strategy enabled intelligent cancer classification according to their mutational landscape. Compared with the conventional PCR-related methods, the proposed SERS classifier possesses the following advantages. First, due to the ultrasensitivity of Au@Ag NRs substrate, amplification-free detection of cellular mutations has been accomplished. Second, only one SERS measurement is needed to simultaneously profile multiple mutations, thanks to the great encoding capacity of Raman spectroscopy. Third, heterogeneous samples carrying different mutations can be intelligently classified into groups through the supervised spectrum analysis based on CLS-LDA algorithm. Our experimental results demonstrated rapid (40 min), straightforward and highly automated clustering of tumor cell lines and colorectal cancer patients according to their mutational profiles. While it is difficult to classify all types of cancers through the expression profiles, this SERS mutational classifier could become an indispensable and complementary tool for clinical precision medicine.

Acknowledgements

This research was supported by the CANCER project (NORTE-01-0145-FEDER-000029) co-funded through the NORTE-45-2015-02 program, the “Innovative Microfluidic Platform for Analysis of myeloid Leukemia blasts (IMPAct-L)” project (030782) co-funded by FCT and the ERDF through COMPETE2020, and the Raman4clinics H2020 COST Action. The authors thank Yizhi Zhang for helping with the development of algorithms. The authors also thank Sara Abalde-Cela and Sandra Carvalho for helping with the TEM imaging and DNA extracting, respectively.

References

Ahmed, D., Eide, P.W., Eilertsen, I.A., Danielsen, S.A., Eknæs, M., Hektoen, M., Lind, G.E., Lothe, R.A., 2013. Oncogenesis 2, e71–e71.

Alizadeh, A.A., Aranda, V., Bardelli, A., Blanpain, C., Bock, C., Borowski, C., Caldas, C., Califano, A., Doherty, M., Elsner, M., Esteller, M., Fitzgerald, R., Korbel, J.O., Lichter, P., Mason, C.E., Navin, N., Pe'er, D., Polyak, K., Roberts, C.W.M., Siu, L., Snyder, A., Stower, H., Swanton, C., Verhaak, R.G.W., Zenklusen, J.C., Zuber, J., Zucman-Rossi, J., 2015. Nat. Med. 21, 846–853.

Bedard, P.L., Hansen, A.R., Ratain, M.J., Siu, L.L., 2013. Nature 501, 355–364.

Ben-David, U., Beroukhim, R., Golub, T.R., 2019. Nat. Rev. Cancer 19, 97–109.

Chen, C., Li, Y., Kerman, S., Neutens, P., Willems, K., Cornelissen, S., Lagae, L., Stakenborg, T., Van Dorpe, P., 2018. Nat. Commun. 9, 1733.

Chuong, T.T., Pallaoro, A., Chaves, C.A., Li, Z., Lee, J., Eisenstein, M., Stucky, G.D., Moskovits, M., Soh, H.T., 2017. Proc. Natl. Acad. Sci. U. S. A. 114, 9056–9061.

Ciriello, G., Miller, M.L., Aksoy, B.A., Senbabaoglu, Y., Schultz, N., Sander, C., 2013. Nat. Genet. 45, 1127-1133.

Collisson, E.A., Campbell, J.D., Brooks, A.N., Berger, A.H., Lee, W., Chmielecki, J., Beer, D.G., Cope, L., Creighton, C.J., Danilova, L., Ding, L., Getz, G., Hammerman, P.S., Hayes, D.N., Hernandez, B., Herman, J.G., Heymach, J. V, Jurisica, I., Kucherlapati, R., Kwiatkowski, D., Ladanyi, M., Robertson, G., Schultz, N., Shen, R., Sinha, R., Sougnez, C., Tsao, M.-S., Travis, W.D., Weinstein, J.N., Wigle, D.A., Wilkerson, M.D., Chu, A., Cherniack, A.D., Hadjipanayis, A., Rosenberg, M., Weisenberger, D.J., Laird, P.W., Radenbaugh, A., Ma, S., Stuart, J.M., Byers, L.A., Baylin, S.B., Govindan, R., Meyerson, M., Rosenberg, M., Gabriel, S.B., Cibulskis, K., Sougnez, C., Kim, J., Stewart, C., Lichtenstein, L., Lander, E.S., Lawrence, M.S., Getz, Kandoth, C., Fulton, R., Fulton, L.L., McLellan, M.D., Wilson, R.K., Ye, K., Fronick, C.C., Maher, C.A., Miller, C.A., Wendl, M.C., Cabanski, C., Ding, L., Mardis, E., Govindan, R., Creighton, C.J., Wheeler, D., Balasundaram, M., Butterfield, Y.S.N., Carlsen, R., Chu, A., Chuah, E., Dhalla, N., Guin, R., Hirst, C., Lee, D., Li, H.I., Mayo, M., Moore, R.A., Mungall, A.J., Schein, J.E., Sipahimalani, P., Tam, A., Varhol, R., Robertson, A.G., Wye, N., Thiessen, N., Holt, R.A., Jones, S.J.M., Marra, M.A., Campbell, J.D., Brooks, A.N., Chmielecki, J., Imielinski, M., Onofrio, R.C., Hodis, E., Zack, T., Sougnez, C., Helman, E., Pedamallu, C.S., Mesirov, J., Cherniack, A.D., Saksena, G., Schumacher, S.E., Carter, S.L., Hernandez, B., Garraway, L.,

Beroukhim, R., Gabriel, S.B., Getz, G., Meyerson, M., Hadjipanayis, A., Lee, S., Mahadeshwar, H.S., Pantazi, A., Protopopov, A., Ren, X., Seth, S., Song, X., Tang, J., Yang, Lixing, Zhang, J., Chen, P.-C., Parfenov, M., Xu, A.W., Santoso, N., Chin, L., Park, P.J., Kucherlapati, R., Hoadley, K.A., Auman, J.T., Meng, S., Shi, Y., Buda, E., Waring, S., Veluvolu, U., Tan, D., Mieczkowski, P.A., Jones, C.D., Simons, J. V, Soloway, M.G., Bodenheimer, T., Jefferys, S.R., Roach, J., Hoyle, A.P., Wu, J., Balu, S., Singh, D., Prins, J.F., Marron, J.S., Parker, J.S., Hayes, D.N., Perou, C.M., Liu, J., Cope, L., Danilova, L., Weisenberger, D.J., Maglinte, D.T., Lai, P.H., Bootwalla, M.S., Van Den Berg, D.J., Triche Jr., T., Baylin, S.B., Laird, P.W., Rosenberg, M., Chin, L., Zhang, J., Cho, J., DiCara, D., Heiman, D., Lin, P., Mallard, W., Voet, D., Zhang, H., Zou, L., Noble, M.S., Lawrence, M.S., Saksena, G., Gehlenborg, N., Thorvaldsdottir, H., Mesirov, J., Nazaire, M.-D., Robinson, J., Getz, G., Lee, W., Aksoy, B.A., Ciriello, G., Taylor, B.S., Dresdner, G., Gao, J., Gross, B., Seshan, V.E., Ladanyi, M., Reva, B., Sinha, R., Sumer, S.O., Weinhold, N., Schultz, N., Shen, R., Sander, C., Ng, S., Ma, S., Zhu, J., Radenbaugh, A., Stuart, J.M., Benz, C.C., Yau, C., Haussler, D., Spellman, P.T., Wilkerson, M.D., Parker, J.S., Hoadley, K.A., Kimes, P.K., Hayes, D.N., Perou, C.M., Broom, B.M., Wang, J., Lu, Y., Ng, P.K.S., Diao, L., Byers, L.A., Liu, W., Heymach, J. V, Amos, C.I., Weinstein, J.N., Akbani, R., Mills, G.B., Curley, E., Paulauskis, J., Lau, K., Morris, S., Shelton, T., Mallery, D., Gardner, J., Penny, R., Saller, C., Tarvin, K., Richards, W.G., Cerfolio, R., Bryant, A., Raymond, D.P., Pennell, N.A., Farver, C., Czerwinski, C., Huelsenbeck-Dill, L., Iacocca, M., Petrelli, N., Rabeno, B., Brown, J., Bauer, T., Dolzhanskiy, O., Potapova, O., Rotin, D., Voronina, O., Nemirovich-Danchenko, E., Fedosenko, K. V, Gal, A., Behera, M., Ramalingam, S.S., Sica, G., Flieder, D., Boyd, J., Weaver, J., Kohl, B., Thinh, D.H.Q., Sandusky, G., Juhl, H., Duhig, E., Illei, P., Gabrielson, E., Shin, J., Lee, B., Rogers, K., Trusty, D., Brock, M. V, Williamson, C., Burks, E., Rieger-Christ, K., Holway, A., Sullivan, T., Wigle, D.A., Asiedu, M.K., Kosari, F., Travis, W.D., Rekhtman, N., Zakowski, M., Rusch, V.W., Zippile, P., Suh, J., Pass, H., Goparaju, C., Owusu-Sarpong, Y., Bartlett, J.M.S., Kodeeswaran, S., Parfitt, J., Sekhon, H., Albert, M., Eckman, J., Myers, J.B., Cheney, R., Morrison, C., Gaudioso, C., Borgia, J.A., Bonomi, P., Pool, M., Liptay, M.J., Moiseenko, F., Zaytseva, I., Dienemann, H., Meister, M., Schnabel, P.A., Muley, T.R., Peifer, M., Gomez-Fernandez, C., Herbert, L., Egea, S., Huang, M., Thorne, L.B., Boice, L., Salazar, A.H., Funkhouser, W.K., Rathmell, W.K., Dhir, R., Yousem, S.A., Dacic, S., Schneider, F., Siegfried, J.M., Hajek, R., Watson, M.A., McDonald, S., Meyers, B., Clarke, B., Yang, I.A., Fong, K.M., Hunter, L., Windsor, M., Bowman, R. V, Peters, S., Letovanec, I., Khan, K.Z., Jensen, M.A., Snyder, E.E., Srinivasan, D., Kahn, A.B., Baboud, J., Pot, D.A., Shaw, K.R.M., Sheth, M., Davidsen, T., Demchok, J.A., Yang, Liming, Wang, Z., Tarnuzzer, R., Zenklusen, J.C., Ozenberger, B.A., Sofia, H.J., Travis, W.D., Cheney, R., Clarke, B., Dacic, S., Duhig, E., Funkhouser, W.K., Illei, P., Farver, C., Rekhtman, N., Sica, G., Suh, J., Tsao, M.-S., Network, C.G.A.R., 2014. Nature 511, 543–550.

De Roock, W., Claes, B., Bernasconi, D., De Schutter, J., Biesmans, B., Fountzilas, G., Kalogeras, K.T.,

Kotoula, V., Papamichael, D., Laurent-Puig, P., Penault-Llorca, F., Rougier, P., Vincenzi, B., Santini, D., Tonini, G., Cappuzzo, F., Frattini, M., Molinari, F., Saletti, P., De Dosso, S., Martini, M., Bardelli, A., Siena, S., Sartore-Bianchi, A., Tabernero, J., Macarulla, T., Di Fiore, F., Gangloff, A.O., Ciardiello, F., Pfeiffer, P., Qvortrup, C., Hansen, T.P., Van Cutsem, E., Piessevaux, H., Lambrechts, D., Delorenzi, M., Tejpar, S., 2010. Lancet Oncol. 11, 753–762.

Fang, W., Jia, S., Chao, J., Wang, Liqian, Duan, X., Liu, H., Li, Q., Zuo, X., Wang, Lihua, Wang, Lianhui, Liu, N., Fan, C., 2019. Sci. Adv. 5, 4506.

Friedman, A.A., Letai, A., Fisher, D.E., Flaherty, K.T., 2015. Nat. Rev. Cancer 15, 747–756.

Garrido-Maestu, A., Azinheiro, S., Carvalho, J., Abalde-Cela, S., Carbó-Argibay, E., Diéguez, L., Piotrowski, M., Kolen'ko, Y. V., Prado, M., 2017. Front. Microbiol. 8, 2159.

Geng, Y., Peveler, W.J., Rotello, V.M., 2019. Angew. Chemie Int. Ed. 58, 5190–5200.

Giaccone, G., 2005. Nat. Clin. Pract. Oncol. 2, 296–297.

Gracie, K., Correa, E., Mabbott, S., Dougan, J.A., Graham, D., Goodacre, R., Faulds, K., 2014. Chem. Sci. 5, 1030–1040.

Hofree, M., Shen, J.P., Carter, H., Gross, A., Ideker, T., 2013. Nat. Methods 10, 1108–1115.

Huang, J.-A., Mousavi, M.Z., Zhao, Y., Hubarevich, A., Omeis, F., Giovannini, G., Schütte, M., Garoli, D., De Angelis, F., 2019. Nat. Commun. 10, 5321.

Ikediobi, O.N., Davies, H., Bignell, G., Edkins, S., Stevens, C., O'Meara, S., Santarius, T., Avis, T., Barthorpe, S., Brackenbury, L., Buck, G., Butler, A., Clements, J., Cole, J., Dicks, E., Forbes, S., Gray, K., Halliday, K., Harrison, R., Hills, K., Hinton, J., Hunter, C., Jenkinson, A., Jones, D., Kosmidou, V., Lugg, R., Menzies, A., Mironenko, T., Parker, A., Perry, J., Raine, K., Richardson, D., Shepherd, R., Small, A., Smith, R., Solomon, H., Stephens, P., Teague, J., Tofts, C., Varian, J., Webb, T., West, S., Widaa, S., Yates, A., Reinhold, W., Weinstein, J.N., Stratton, M.R., Futreal, P.A., Wooster, R., 2006. Mol. Cancer Ther. 5, 2606–2612.

Jao, K., Tomasini, P., Kamel-Reid, S., Korpanty, G.J., Mascaux, C., Sakashita, S., Labbé, C., Leighl, N.B., Liu, G., Feld, R., Bradbury, P.A., Hwang, D.M., Pintilie, M., Tsao, M.-S., Shepherd, F.A., 2018. Lung Cancer 123, 22–29.

Jin, J., Wu, X., Yin, J., Li, M., Shen, J., Li, J., Zhao, Y., Zhao, Q., Wu, J., Wen, Q., Cho, C.H., Yi, T., Xiao, Z., Qu, L., 2019. Front. Oncol. 9, 263.

Kamil Reza, K., Wang, J., Vaidyanathan, R., Dey, S., Wang, Y., Trau, M., 2017. Small 13, 1602902.

Koboldt, D.C., Fulton, R.S., McLellan, M.D., Schmidt, H., Kalicki-Veizer, J., McMichael, J.F., Fulton, L.L., Dooling, D.J., Ding, L., Mardis, E.R., Wilson, R.K., Ally, A., Balasundaram, M., Butterfield, Y.S.N., Carlsen, R., Carter, C., Chu, A., Chuah, E., Chun, H.-J.E., Coope, R.J.N., Dhalla, N., Guin, R., Hirst, C., Hirst, M., Holt, R.A., Lee, D., Li, H.I., Mayo, M., Moore, R.A., Mungall, A.J., Pleasance, E., Gordon Robertson, A., Schein, J.E., Shafiei, A., Sipahimalani, P., Slobodan, J.R., Stoll, D., Tam, A., Thiessen, N., Varhol, R.J., Wye, N., Zeng, T., Zhao, Y., Birol, I., Jones, S.J.M., Marra, M.A., Cherniack, A.D., Saksena, G., Onofrio, R.C., Pho, N.H., Carter, S.L., Schumacher, S.E., Tabak, B., Hernandez, B., Gentry, J., Nguyen, H., Crenshaw, A., Ardlie, K., Beroukhim, R., Winckler, W., Getz, G., Gabriel, S.B., Meyerson, M., Chin, L., Park, P.J., Kucherlapati, R., Hoadley, K.A., Todd Auman, J., Fan, C., Turman, Y.J., Shi, Y., Li, L., Topal, M.D., He, X., Chao, H.-H., Prat, A., Silva, G.O., Iglesia, M.D., Zhao, W., Usary, J., Berg, J.S., Adams, M., Booker, J., Wu, J., Gulabani, A., Bodenheimer, T., Hoyle, A.P., Simons, J. V, Soloway, M.G., Mose, L.E., Jefferys, S.R., Balu, S., Parker, J.S., Neil Hayes, D., Perou, C.M., Malik, S., Mahurkar, S., Shen, H., Weisenberger, D.J., Triche Jr, T., Lai, P.H., Bootwalla, M.S., Maglinte, D.T., Berman, B.P., Van Den Berg, D.J., Baylin, S.B., Laird, P.W., Creighton, C.J., Donehower, L.A., Getz, G., Noble, M., Voet, D., Saksena, G., Gehlenborg, N., DiCara, D., Zhang, J., Zhang, H., Wu, C.-J., Yingchun Liu, S., Lawrence, M.S., Zou, L., Sivachenko, A., Lin, P., Stojanov, P., Jing, R., Cho, J., Sinha, Raktim, Park, R.W., Nazaire, M.-D., Robinson, J., Thorvaldsdottir, H., Mesirov, J., Park, P.J., Chin, L., Reynolds, S., Kreisberg, R.B., Bernard, B., Bressler, R., Erkkila, T., Lin, J., Thorsson, V., Zhang, W., Shmulevich, I., Ciriello, G., Weinhold, N., Schultz, N., Gao, J., Cerami, E., Gross, B., Jacobsen, A., Sinha, Rileen, Arman Aksoy, B., Antipin, Y., Reva, B., Shen, R., Taylor, B.S., Ladanyi, M., Sander, C., Anur, P., Spellman, P.T., Lu, Y., Liu, W., Verhaak, R.R.G., Mills, G.B., Akbani, R., Zhang, N., Broom, B.M., Casasent, T.D., Wakefield, C., Unruh, A.K., Baggerly, K., Coombes, K., Weinstein, J.N., Haussler, D., Benz, C.C., Stuart, J.M., Benz, S.C., Zhu, J., Szeto, C.C., Scott, G.K., Yau, C., Paull, E.O., Carlin, D., Wong, C., Sokolov, A., Thusberg, J., Mooney, S., Ng, S., Goldstein, T.C., Ellrott, K., Grifford, M., Wilks, C., Ma, S., Craft, B., Yan, C., Hu, Y., Meerzaman, D., Gastier-Foster, J.M., Bowen, J., Ramirez, N.C., Black, A.D., Pyatt, R.E., White, P., Zmuda, E.J., Frick, J., Lichtenberg, T.M., Brookens, R., George, M.M., Gerken, M.A., Harper, H.A., Leraas, K.M., Wise, L.J., Tabler, T.R., McAllister, C., Barr, T., Hart-Kothari, M., Tarvin, K., Saller, C., Sandusky, G., Mitchell, C., Iacocca, M. V, Brown, J., Rabeno, B., Czerwinski, C., Petrelli, N., Dolzhansky, O., Abramov, M., Voronina, O., Potapova, O., Marks, J.R., Suchorska, W.M., Murawa, D., Kycler, W., Ibbs, M., Korski, K., Spychała, A., Murawa, P., Brzeziński, J.J., Perz, H., Łaźniak, R., Teresiak, M., Tatka, H., Leporowska, E., Bogusz-Czerniewicz, M., Malicki, J., Mackiewicz, A., Wiznerowicz, M., Van Le, X., Kohl, B., Viet Tien, N., Thorp, R., Van Bang, N., Sussman, H., Duc Phu, B., Hajek, R., Phi Hung, N., Viet The Phuong, T., Quyet Thang, H., Zaki Khan, K., Penny, R., Mallery, D., Curley, E., Shelton, C., Yena, P., Ingle, J.N., Couch, F.J., Lingle, W.L., King, T.A., Maria

Gonzalez-Angulo, A., Mills, G.B., Dyer, M.D., Liu, S., Meng, X., Patangan, M., Network, T.C.G.A., Louis, G. sequencing centres: W.U. in S., Agency, G. characterization centres: B.C.C., Institute, B., School, B.& W.H.& H.M., University of North Carolina, C.H., Hopkins, U. of S.C., Medicine, G. data analysis: B.C. of, Biology, I. for S., Center, M.S.-K.C., University, O.H.& S., Center, T.U. of T.M.D.A.C., University of California, S.C.I., NCI, Resource, B. core resource: N.C.H.B.C., ABS-IUPUI, T. source sites:, Christiana, Cureline, Center, D.U.M., Centre, T.G.P.C., ILSBio, Consortium, I.G., Clinic, M., MSKCC, Center, M.D.A.C., 2012. Nature 490, 61–70.

Kowalczyk, A., Krajczewski, J., Kowalik, A., Weyher, J.L., Dzięcielewski, I., Chłopek, M., Góźdź, S., Nowicka, A.M., Kudelski, A., 2019. Biosens. Bioelectron. 132, 326–332.

Kuijjer, M.L., Paulson, J.N., Salzman, P., Ding, W., Quackenbush, J., 2018. Br. J. Cancer 118, 1492–1501.

Langer, J., Jimenez de Aberasturi, D., Aizpurua, J., Alvarez-Puebla, R.A., Auguié, B., Baumberg, J.J., Bazan, G.C., Bell, S.E.J., Boisen, A., Brolo, A.G., Choo, J., Cialla-May, D., Deckert, V., Fabris, L., Faulds, K., García de Abajo, F.J., Goodacre, R., Graham, D., Haes, A.J., Haynes, C.L., Huck, C., Itoh, T., Käll, M., Kneipp, J., Kotov, N.A., Kuang, H., Le Ru, E.C., Lee, H.K., Li, J.-F., Ling, X.Y., Maier, S.A., Mayerhöfer, T., Moskovits, M., Murakoshi, K., Nam, J.-M., Nie, S., Ozaki, Y., Pastoriza-Santos, I., Perez-Juste, J., Popp, J., Pucci, A., Reich, S., Ren, B., Schatz, G.C., Shegai, T., Schlücker, S., Tay, L.-L., Thomas, K.G., Tian, Z.-Q., Van Duyne, R.P., Vo-Dinh, T., Wang, Y., Willets, K.A., Xu, C., Xu, H., Xu, Y., Yamamoto, Y.S., Zhao, B., Liz-Marzán, L.M., 2019. ACS Nano 14, 28–117.

Li, J., Wang, L., Mamon, H., Kulke, M.H., Berbeco, R., Makrigiorgos, G.M., 2008. Nat. Med. 14, 579–584.

Li, X., Yang, T., Li, C.S., Song, Y., Lou, H., Guan, D., Jin, L., 2018. Theranostics 8, 1678–1689.

Luo, X., Xing, Y., Galvan, D.D., Zheng, E., Wu, P., Cai, C., Yu, Q., 2019. ACS Sensors 4, 1534–1542.

Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J., Turner, D.J., 2010. Nat. Methods 7, 111.

Martincorena, I., Raine, K.M., Gerstung, M., Dawson, K.J., Haase, K., Van Loo, P., Davies, H., Stratton, M.R., Campbell, P.J., 2017. Cell 171, 1029.

Misale, S., Yaeger, R., Hobor, S., Scala, E., Janakiraman, M., Liska, D., Valtorta, E., Schiavo, R., Buscarino, M., Siravegna, G., Bencardino, K., Cercek, A., Chen, C.-T., Veronese, S., Zanon, C., Sartore-Bianchi, A., Gambacorta, M., Gallicchio, M., Vakiani, E., Boscaro, V., Medico, E., Weiser, M., Siena, S., Di Nicolantonio, F., Solit, D., Bardelli, A., 2012. Nature 486, 532–536.

Mitsudomi, T., Morita, S., Yatabe, Y., Negoro, S., Okamoto, I., Tsurutani, J., Seto, T., Satouchi, M., Tada, H., Hirashima, T., Asami, K., Katakami, N., Takada, M., Yoshioka, H., Shibata, K., Kudoh, S.,

Shimizu, E., Saito, H., Toyooka, S., Nakagawa, K., Fukuoka, M., 2010. Lancet Oncol. 11, 121–128.

Muzny, D.M., Bainbridge, M.N., Chang, K., Dinh, H.H., Drummond, J.A., Fowler, G., Kovar, C.L., Lewis, L.R., Morgan, M.B., Newsham, I.F., Reid, J.G., Santibanez, J., Shinbrot, E., Trevino, L.R., Wu, Y.-Q., Wang, M., Gunaratne, P., Donehower, L.A., Creighton, C.J., Wheeler, D.A., Gibbs, R.A., Lawrence, M.S., Voet, Douglas, Jing, R., Cibulskis, K., Sivachenko, A., Stojanov, P., McKenna, A., Lander, E.S., Gabriel, S., Getz, G., Ding, L., Fulton, R.S., Koboldt, D.C., Wylie, T., Walker, J., Dooling, D.J., Fulton, L., Delehaunty, K.D., Fronick, C.C., Demeter, R., Mardis, E.R., Wilson, R.K., Chu, A., Chun, H.-J.E., Mungall, A.J., Pleasance, E., Robertson, A.G., Stoll, D., Balasundaram, M., Birol, I., Butterfield, Y.S.N., Chuah, E., Coope, R.J.N., Dhalla, N., Guin, R., Hirst, C., Hirst, M., Holt, R.A., Lee, D., Li, H.I., Mayo, M., Moore, R.A., Schein, J.E., Slobodan, J.R., Tam, A., Thiessen, N., Varhol, R., Zeng, T., Zhao, Y., Jones, S.J.M., Marra, M.A., Bass, A.J., Ramos, A.H., Saksena, G., Cherniack, A.D., Schumacher, S.E., Tabak, B., Carter, S.L., Pho, N.H., Nguyen, H., Onofrio, R.C., Crenshaw, A., Ardlie, K., Beroukhim, R., Winckler, W., Getz, G., Meyerson, M., Protopopov, A., Zhang, J., Hadjipanayis, A., Lee, E., Xi, R., Yang, Lixing, Ren, X., Zhang, H., Sathiamoorthy, N., Shukla, S., Chen, P.-C., Haseley, P., Xiao, Y., Lee, S., Seidman, J., Chin, L., Park, P.J., Kucherlapati, R., Auman, J.T., Hoadley, K.A., Du, Y., Wilkerson, M.D., Shi, Y., Liquori, C., Meng, S., Li, L., Turman, Y.J., Topal, M.D., Tan, D., Waring, S., Buda, E., Walsh, J., Jones, C.D., Mieczkowski, P.A., Singh, D., Wu, J., Gulabani, A., Dolina, P., Bodenheimer, T., Hoyle, A.P., Simons, J. V, Soloway, M., Mose, L.E., Jefferys, S.R., Balu, S., O'Connor, B.D., Prins, J.F., Chiang, D.Y., Hayes, D.N., Perou, C.M., Hinoue, T., Weisenberger, D.J., Maglinte, D.T., Pan, F., Berman, B.P., den Berg, D.J., Shen, H., Jr, T.T., Baylin, S.B., Laird, P.W., Getz, G., Noble, M., Voet, Doug, Saksena, G., Gehlenborg, N., DiCara, D., Zhang, J., Zhang, H., Wu, C.-J., Liu, S.Y., Shukla, S., Lawrence, M.S., Zhou, L., Sivachenko, A., Lin, P., Stojanov, P., Jing, R., Park, R.W., Nazaire, M.-D., Robinson, J., Thorvaldsdottir, H., Mesirov, J., Park, P.J., Chin, L., Thorsson, V., Reynolds, S.M., Bernard, B., Kreisberg, R., Lin, J., Iype, L., Bressler, R., Erkkilae, T., Gundapuneni, M., Liu, Y., Norberg, A., Robinson, T., Yang, D., Zhang, W., Shmulevich, I., De Ronde, J.J., Schultz, N., Cerami, E., Ciriello, G., Goldberg, A.P., Gross, B., Jacobsen, A., Gao, J., Kaczkowski, B., Sinha, R., Aksoy, B.A., Antipin, Y., Reva, B., Shen, R., Taylor, B.S., Chan, T.A., Ladanyi, M., Sander, C., Akbani, R., Zhang, N., Broom, B.M., Casasent, T., Unruh, A., Wakefield, C., Hamilton, S.R., Cason, R.C., Baggerly, K.A., Weinstein, J.N., Haussler, D., Benz, C.C., Stuart, J.M., Benz, S.C., Sanborn, J.Z., Vaske, C.J., Zhu, J., Szeto, C., Scott, G.K., Yau, C., Ng, S., Goldstein, T., Ellrott, K., Collisson, E., Cozen, A.E., Zerbino, D., Wilks, C., Craft, B., Spellman, P., Penny, R., Shelton, T., Hatfield, M., Morris, S., Yena, P., Shelton, C., Sherman, M., Paulauskis, J., Gastier-Foster, J.M., Bowen, J., Ramirez, N.C., Black, A., Pyatt, R., Wise, L., White, P., Bertagnolli, M., Brown, J., Chan, T.A., Chu, G.C., Czerwinski, C., Denstman, F., Dhir, R., Doerner, A., Fuchs, C.S., Guillem, J.G., Iacocca, M., Juhl, H., Kaufman, A., Kohl III, B., Van Le, X., Mariano, M.C., Medina, E.N., Meyers, M.,

Nash, G.M., Paty, P.B., Petrelli, N., Rabeno, B., Richards, W.G., Solit, D., Swanson, P., Temple, L., Tepper, J.E., Thorp, R., Vakiani, E., Weiser, M.R., Willis, J.E., Witkin, G., Zeng, Z., Zinner, M.J., Zornig, C., Jensen, M.A., Sfeir, R., Kahn, A.B., Chu, A.L., Kothiyal, P., Wang, Z., Snyder, E.E., Pontius, J., Pihl, T.D., Ayala, B., Backus, M., Walton, J., Whitmore, J., Baboud, J., Berton, D.L., Nicholls, M.C., Srinivasan, D., Raman, R., Girshik, S., Kigonya, P.A., Alonso, S., Sanbhadti, R.N., Barletta, S.P., Greene, J.M., Pot, D.A., Shaw, K.R.M., Dillon, L.A.L., Buetow, K., Davidsen, T., Demchok, J.A., Eley, G., Ferguson, M., Fielding, P., Schaefer, C., Sheth, M., Yang, Liming, Guyer, M.S., Ozenberger, B.A., Palchik, J.D., Peterson, J., Sofia, H.J., Thomson, E., 2012. Nature 487, 330–337.

Nagai, Y., Miyazawa, H., Huqun, Tanaka, T., Udagawa, K., Kato, M., Fukuyama, S., Yokote, A., Kobayashi, K., Kanazawa, N., Hagiwara, K., 2005. Cancer Res. 65, 7276–7282.

Network, C.G.A. (Cancer Genome Atlas Network), 2012. Nature 487, 330–337.

Network, C.G.A.R. (Cancer Genome Atlas Research Network), 2011. Nature 474, 609–615.

Panneerselvam, R., Liu, G.-K., Wang, Y.-H., Liu, J.-Y., Ding, S.-Y., Li, J.-F., Wu, D.-Y., Tian, Z.-Q., 2018. Chem. Commun. 54, 10–25.

Quick, J., Grubaugh, N.D., Pullan, S.T., Claro, I.M., Smith, A.D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T.F., Beutler, N.A., Burton, D.R., Lewis-Ximenez, L.L., de Jesus, J.G., Giovanetti, M., Hill, S.C., Black, A., Bedford, T., Carroll, M.W., Nunes, M., Alcantara, L.C., Sabino, E.C., Baylis, S.A., Faria, N.R., Loose, M., Simpson, J.T., Pybus, O.G., Andersen, K.G., Loman, N.J., 2017. Nat. Protoc. 12, 1261–1276.

Rudin, C.M., Poirier, J.T., Byers, L.A., Dive, C., Dowlati, A., George, J., Heymach, J. V, Johnson, J.E., Lehman, J.M., MacPherson, D., Massion, P.P., Minna, J.D., Oliver, T.G., Quaranta, V., Sage, J., Thomas, R.K., Vakoc, C.R., Gazdar, A.F., 2019. Nat. Rev. Cancer 19, 289–297.

Salk, J.J., Fox, E.J., Loeb, L.A., 2010. Annu. Rev. Pathol. Dis., Annual Review of Pathology-Mechanisms of Disease 5, 51–75.

Shi, C.J., Eshleman, S.H., Jones, D., Fukushima, N., Hua, L., Parker, A.R., Yeo, C.J., Hruban, R.H., Goggins, M.G., Eshleman, J.R., 2004. Nat. Methods 1, 141–147.

Sidaway, P., 2017. Nat. Rev. Clin. Oncol. 14, 587.

Thierry, A.R., Mouliere, F., El Messaoudi, S., Mollevi, C., Lopez-Crapez, E., Rolet, F., Gillet, B., Gongora, C., Dechelotte, P., Robert, B., Del Rio, M., Lamy, P.-J., Bibeau, F., Nouaille, M., Loriot, V., Jarrousse, A.-S., Molina, F., Mathonnet, M., Pezet, D., Ychou, M., 2014. Nat. Med. 20, 430.

Tian, S., Neumann, O., McClain, M.J., Yang, X., Zhou, L., Zhang, C., Nordlander, P., Halas, N.J., 2017. Nano Lett. 17, 5071–5077.

van't Veer, L.J., Dai, H.Y., van de Vijver, M.J., He, Y.D.D., Hart, A.A.M., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteveen, A.T., Schreiber, G.J., Kerkhoven, R.M., Roberts, C., Linsley, P.S., Bernards, R., Friend, S.H., 2002. Nature 415, 530–536.

Wang, J., Koo, K.M., Wang, Y., Trau, M., 2019. Adv. Sci. 6, 1900730.

Wang, X., liu, B., Xiao, M., Zou, Y., Lai, W., Pei, H., Alam, M.F., Zhang, W., Wan, Y., Li, L., 2020. Biosens. Bioelectron. 156, 112130.

Wang, X., Park, S.-G., Ko, J., Xiao, X., Giannini, V., Maier, S.A., Kim, D.-H., Choo, J., 2018. Small 14, 1801623.

Wang, Y., Yan, B., Chen, L., 2013. Chem. Rev. 113, 1391–1428.

Wang, Z., Zong, S., Wu, L., Zhu, D., Cui, Y., 2017. Chem. Rev. 117, 7910–7963.

Wee, E.J.H., Wang, Y., Tsao, S.C.-H., Trau, M., 2016. Theranostics 6, 1506–1513.

Wu, L., Garrido-Maestu, A., Guerreiro, J.R.L., Carvalho, S., Abalde-Cela, S., Prado, M., Diéguez, L., 2019. Nanoscale 11, 7781–7789.

Wu, L., Wang, Z., Fan, K., Zong, S., Cui, Y., 2015. Small 11, 2798–2806.

Wu, Y., Choi, N., Chen, H., Dang, H., Chen, L., Choo, J., 2020. Anal. Chem. 92, 2628–2634.

Xu, L.-J., Lei, Z.-C., Li, J., Zong, C., Yang, C.J., Ren, B., 2015. J. Am. Chem. Soc. 137, 5149–5154.

Zhang, J., Zheng, J., Yang, Y., Lu, J., Gao, J., Lu, T., Sun, J., Jiang, H., Zhu, Y., Zheng, Y., Liang, Z., Liu, T., 2015. Sci. Rep. 5, 18678.

Zhang, Xingang, Zhang, Xiaolei, Luo, C., Liu, Z., Chen, Y., Dong, S., Jiang, C., Yang, S., Wang, F., Xiao, X., 2019. Small 15, 1805516.

Zhang, Y., Wang, Z., Wu, L., Zong, S., Yun, B., Cui, Y., 2018. Small 14, 1704433.

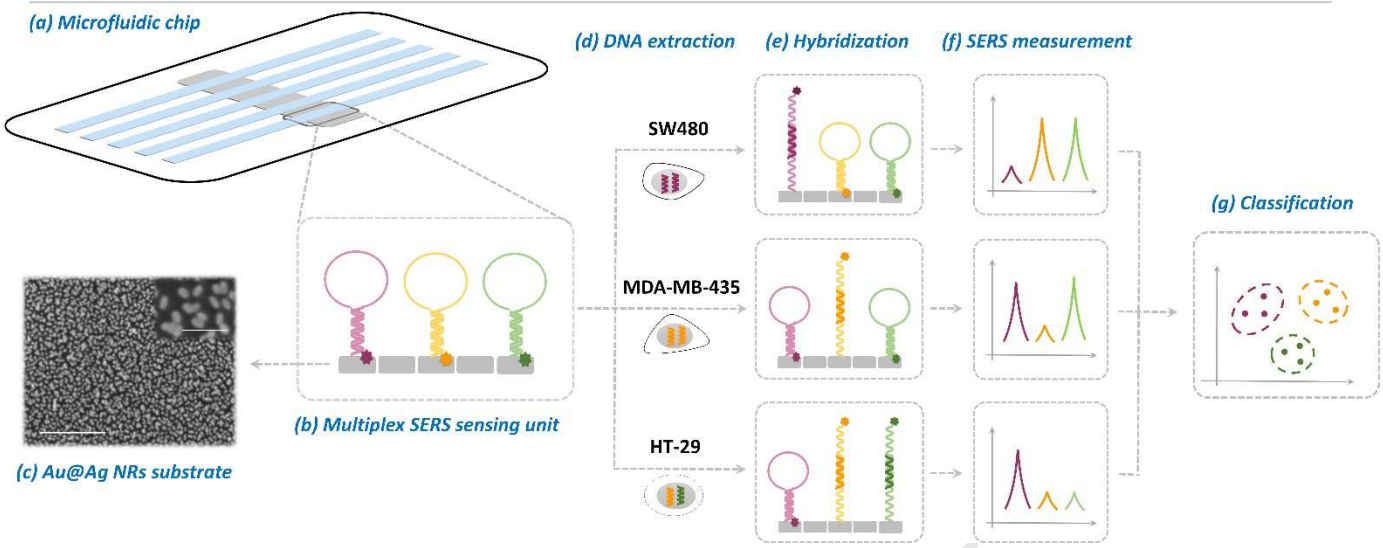


Figure 1. Schematic design of the multiplex SERS mutation assay for cancer classification. (a) Microfluidic chip fabricated by bonding PDMS microchannels (represented by the blue bars) with SERS nanosubstrate (represented by the grey bars). **(b)** Enlarged view of a single sensing unit. Three hairpin-structured molecular beacon probes targeting at *KRAS G12V* (violet), *BRAF V600E* (yellow) and *PIK3CA P449T* (green) were assembled onto the SERS substrate fabricated with Au@Ag NRs (represented by the grey rods). The Raman reporters (represented by the stars)) on the molecular beacon probes were Cy5 (dark violet), FAM (orange) and ROX (dark green) respectively. **(c)** SEM image of the Au@Ag NRs substrate. The scale bar indicates a length of 1 μm. (Inset: enlarged SEM image. The scale bar indicate a length of 250 nm) **(d)** DNAs extracted from three cancer cell lines were used as the analytes, including SW480 cells carrying *KRAS G12V* (represented by the dark violet helix), MDA-MB435 cells carrying *BRAF V600E* (represented by the orange helix) and HT-29 cells carrying *BRAF V600E* and *PIK3CA P449T* (represented by the dark green helix). **(e)** Hybridization of target sequences with the molecular beacon probes. The specific interaction between target DNA with the corresponding molecular beacon probes makes the Raman reporter moving away from the surface of the nanorods, during which the SERS signal switch from 'on' to 'off'. **(f)** Acquisition of the SERS spectra. The characteristic Raman bands of different reporters were represented in different colours. **(g)** Classification of cell types using supervised spectrum analysis.

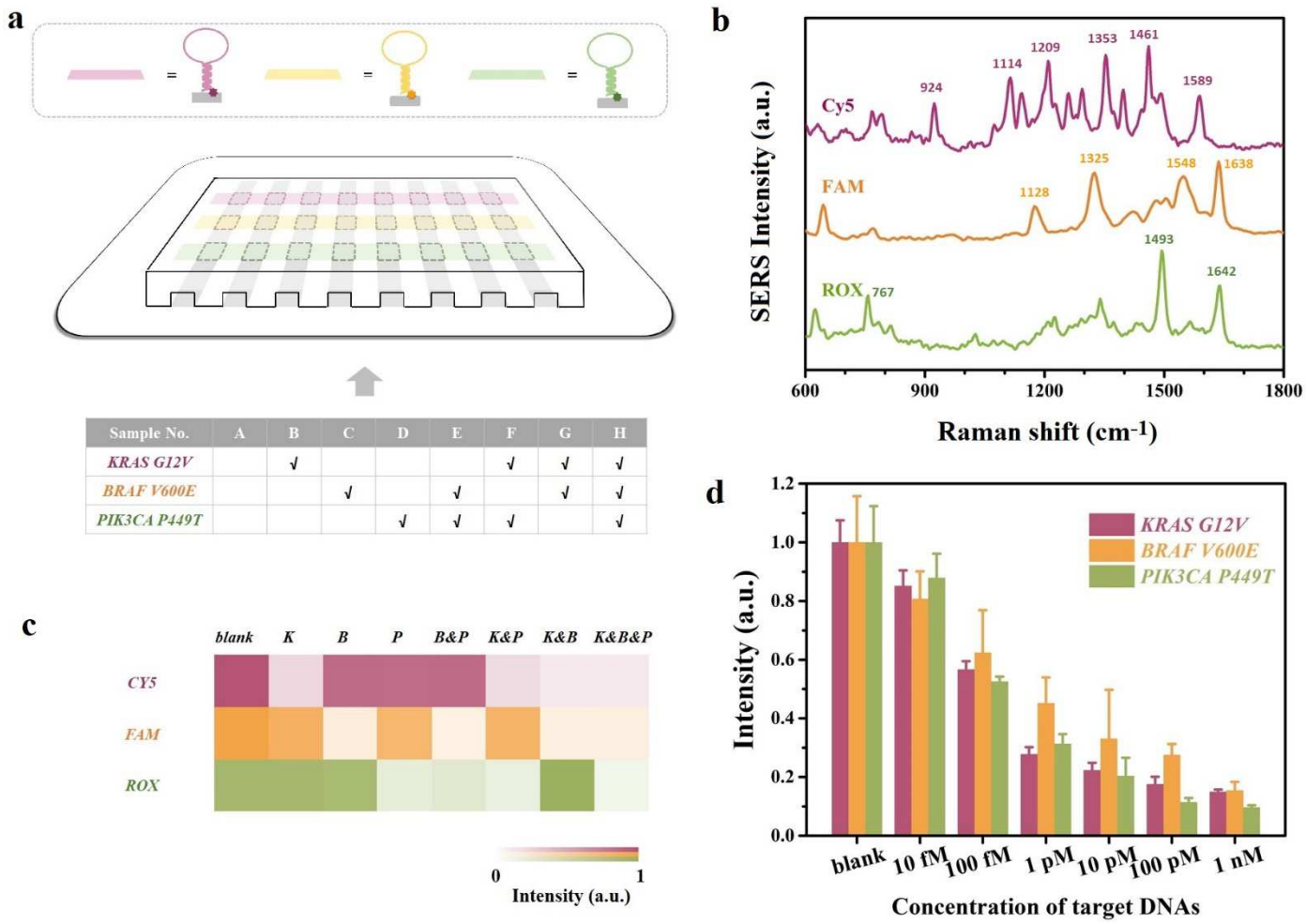
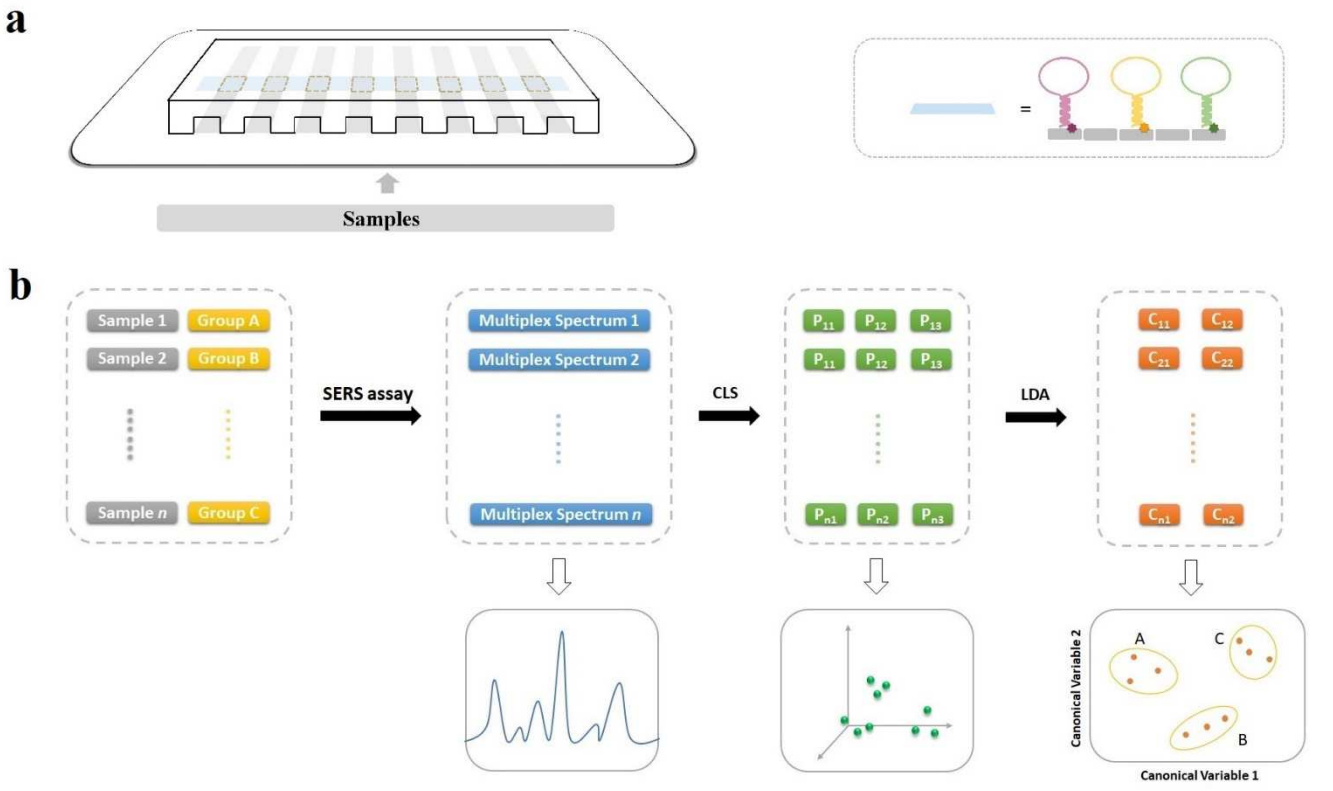
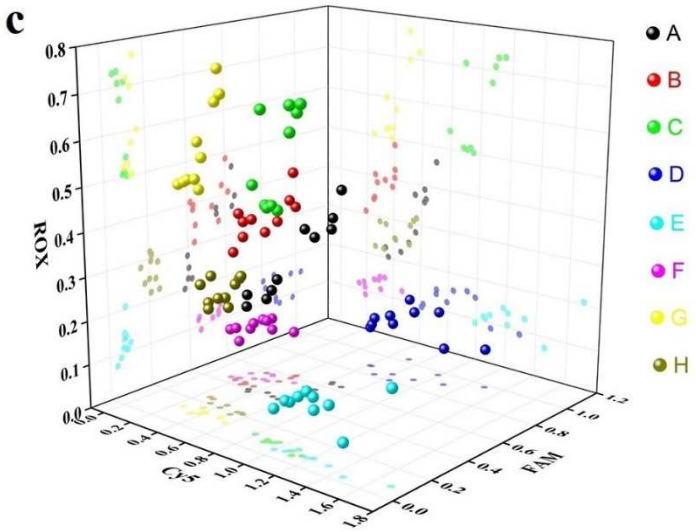
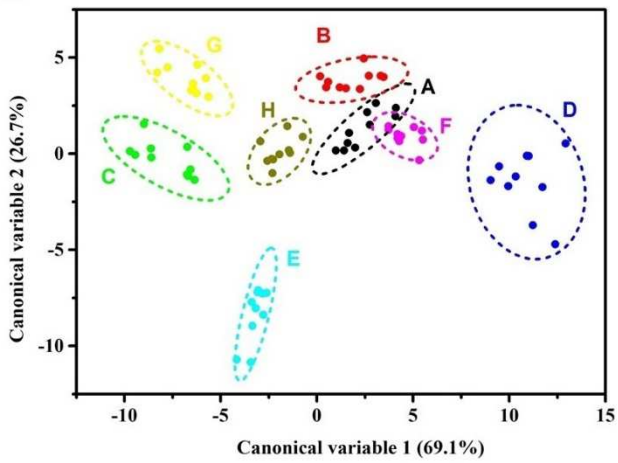
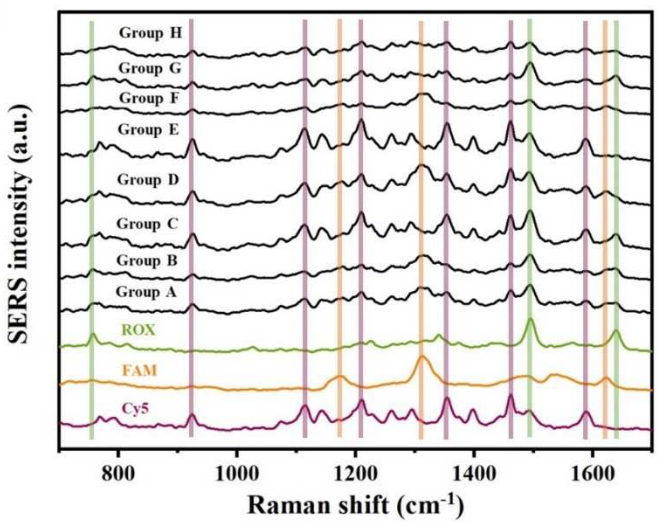


Figure 2. Specificity test of the molecular beacon probes. (a) Schematic view of the array sensor. PDMS chip with 8 microchannels was bonded against glass slides patterned with 3 stripes of SERS substrate, which were modified with *KRAS* (violet), *BRAF* (yellow) and *PIK3CA* (green) molecular beacons respectively. Sample A-H (table) were flowed through the channels following the direction of the arrow. **(b)** SERS spectra of the Cy5, FAM and ROX molecules labelled on the *KRAS*, *BRAF* and *PIK3CA* mutation probes. **(c)** Results of the specificity test plotted in colour map. 8 samples containing different combinations of 3 mutations were tested on the array sensor. The peak area at 1589 (violet), 1325 (orange) and 1493 (green) cm^{-1} were used to characterize the SERS intensity of Cy5, FAM and ROX, which corresponds to the *KRAS*, *BRAF* and *PIK3CA* mutation probes. **(d)** Quantitative detection of synthetic *KRAS G12V*, *BRAF V600E* and *PIK3CA P449T* sequences. Plot of peak area as a function of the concentration for three targets. Error bars represent the standard deviation of 3 measurements.







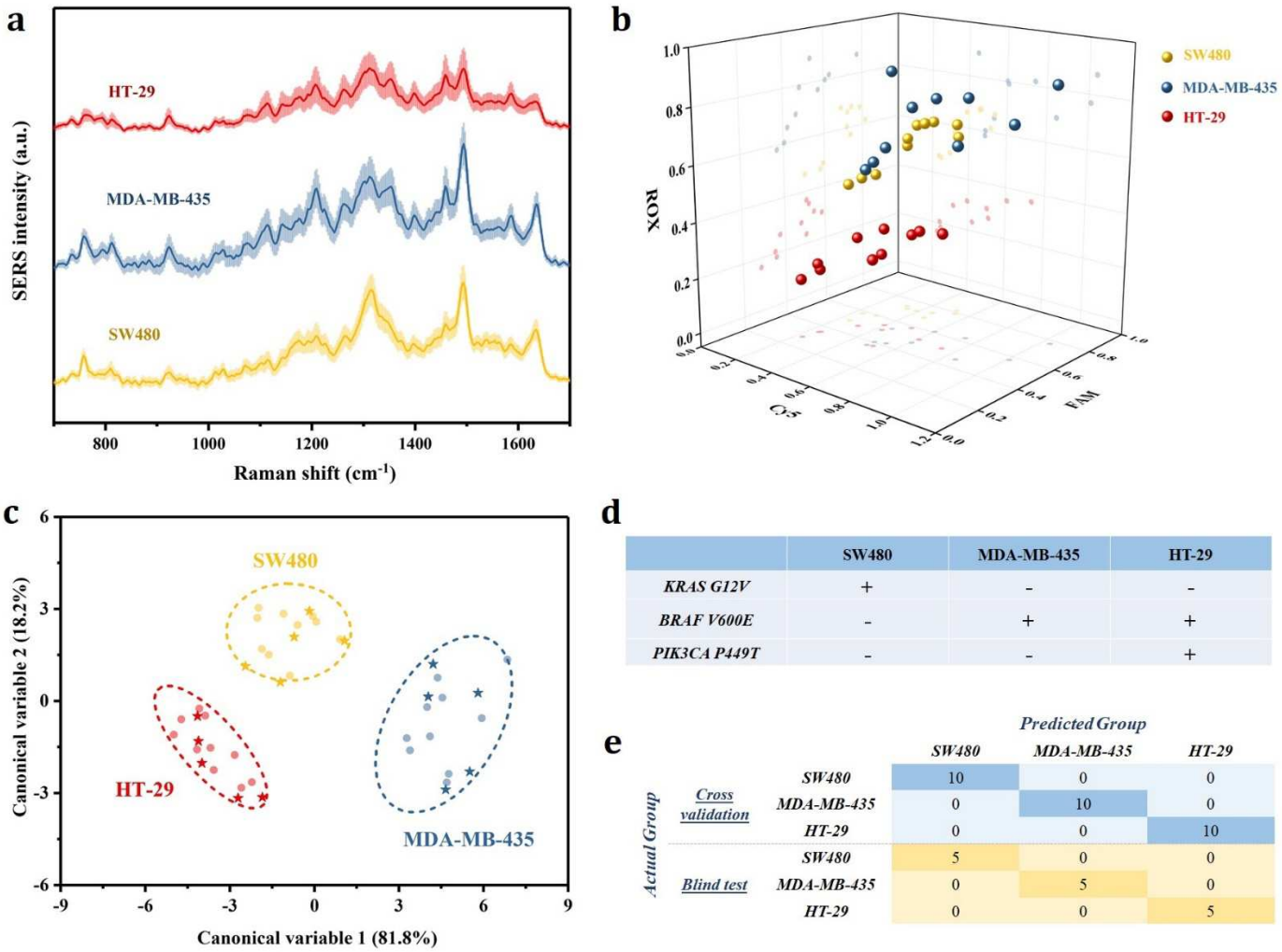


f

Predicted Group

Actual Group	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
A	9	1	0	0	0	0	0	0
B	0	10	0	0	0	0	0	0
C	0	0	10	0	0	0	0	0
D	0	0	0	10	0	0	0	0
E	0	0	0	0	10	0	0	0
F	0	0	0	0	0	10	0	0
G	0	0	0	0	0	0	10	0
H	0	0	0	0	0	0	0	10

Figure 3. Multiplex mutation analysis and sample classification. (a) Schematic view of the multiplex detection based on SERS encoding. The SERS substrate was modified with a mixture of 3 molecular beacon probes for simultaneous detection. **(b)** Workflow for the spectrum analysis. The acquired SERS spectra were successively processed with CLS and LDA algorithm. CLS was employed to calculate the contribution of each Raman reporter, the results of which were plotted as the scattering diagram. LDA algorithm was then applied to the results of CLS analysis to establish a new algorithm which classifies samples into groups. **(c)** 3D scattering plot of the contribution of each Raman reporter calculated with CLS algorithm. 8 groups of samples (A-H, in different colours), which contain a combination of *KRAS G12V*, *BRAF V600E* and *PIK3CA P449T*, were tested (10 tests for each sample). The component of each group is the same as shown in Figure 2a. The plot shows the results of 10 tests for each of the 8 groups. **(d)** Scattering plot of the canonical variables calculated with the LDA algorithm, in which 8 groups of samples were classified. **(e)** Average SERS spectra for 8 groups of samples. **(f)** Cross-validation results for the 8 groups of samples.



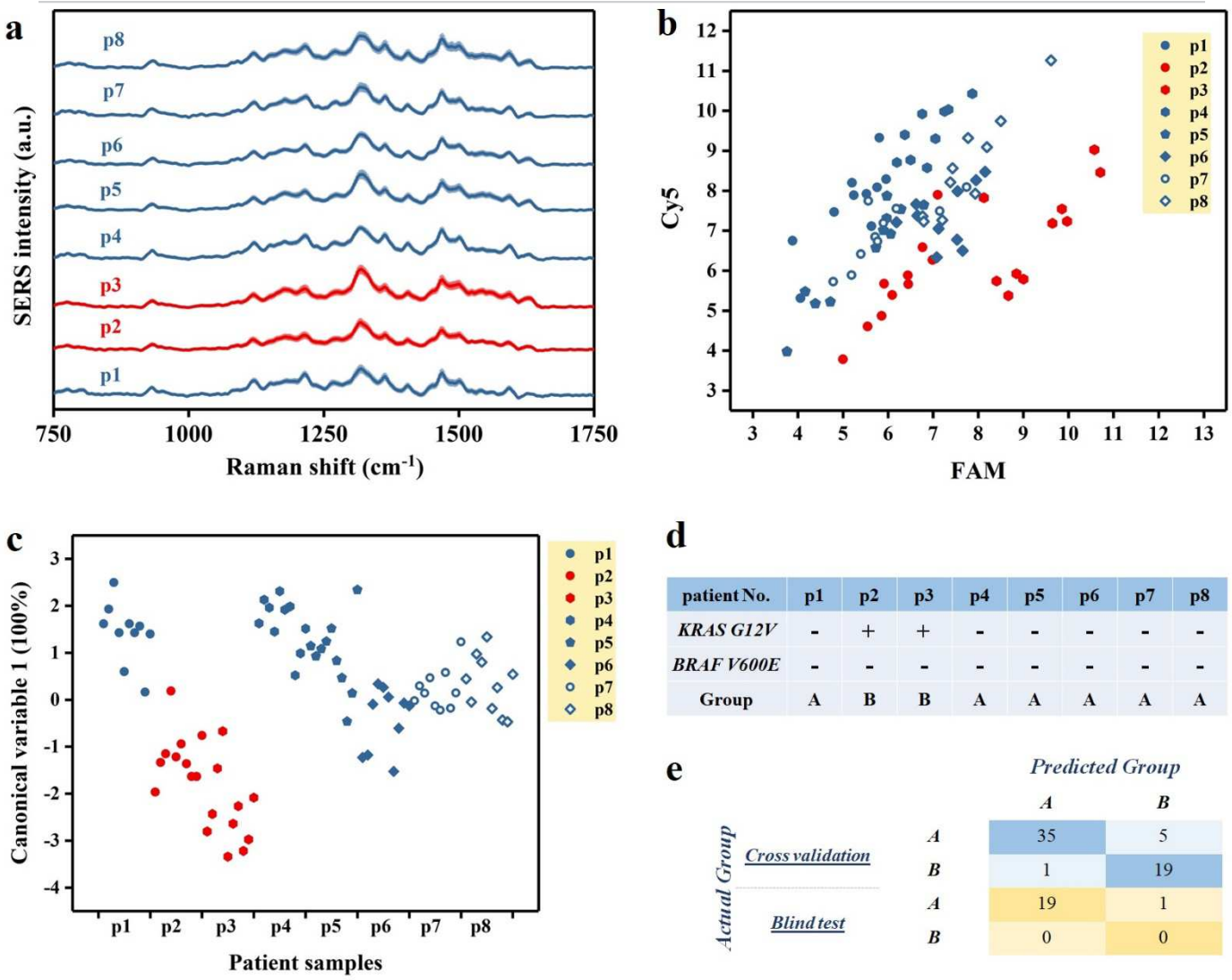
d

	SW480	MDA-MB-435	HT-29
KRAS G12V	+	-	-
BRAF V600E	-	+	+
PIK3CA P449T	-	-	+

e

Actual Group		*Predicted Group* SW480	MDA-MB-435	HT-29
Cross validation	*SW480*	10	0	0
	MDA-MB-435	0	10	0
	HT-29	0	0	10
Blind test	*SW480*	5	0	0
	MDA-MB-435	0	5	0
	HT-29	0	0	5

Figure 4. Multiplex detection of cellular mutations and cancer classification. (a) SERS spectra for DNA lysates from SW480, MDA-MB-435 and HT-29 cells. The shadows indicate the standard deviation among spectra. **(b)** 3D scattering plot of each Raman reporter's contribution calculated with CLS algorithm. DNAs extracted from SW480, MDA-MB-435 and HT-29 cells were tested (10 tests for each cell line). **(c)** Scattering plot of the canonical variables calculated with the LDA algorithm, in which 3 cancer cell lines were classified (10 training tests, plotted as dots, and 5 blind tests, plotted as stars, were tested for each cell line). **(d)** Mutation status of cell lines. **(e)** Cross-validation and blind test results for 3 cancer cell lines.



patient No.	p1	p2	p3	p4	p5	p6	p7	p8
KRAS G12V	-	+	+	-	-	-	-	-
BRAF V600E	-	-	-	-	-	-	-	-
Group	A	B	B	A	A	A	A	A

		Predicted Group A	Predicted Group B
Cross validation	A	35	5
	B	1	19
Blind test	A	19	1
	B	0	0

Figure 5. Patient sample test and classification. (a) SERS spectra for tissue DNA lysates extracted from 8 colorectal cancer patients (p1-p8). The shadows indicate the standard deviation among spectra. **(b)** Scattering plot of each Raman reporter's contribution calculated with CLS algorithm. The *KRAS G12V* and *BRAF V600E* mutations were tested (10 tests for each patient sample). **(c)** Scattering plot of the canonical variables calculated with the LDA algorithm. 10 training tests (plotted as solid labels, p1-p6) and 10 blind tests (plotted as hollow labels, p7-p8) were performed with samples from each patient. **(d)** Mutation status of patient samples. **(e)** Cross-validation and blind test results for patient samples.

Highlights:

- A SERS biochip for analyzing cellular DNA mutation is developed.
- Amplification-free detection of multiple cellular point mutations is achieved.
- Supervised learning algorithm is employed for cancer classification.
- The chip is capable to analyze real clinical samples from colorectal cancer patients.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: