ACSNANO

www.acsnano.org

ARTICLE

Ratiometric 3D DNA Machine Combined with Machine Learning Algorithm for Ultrasensitive and High-Precision Screening of Early Urinary Diseases

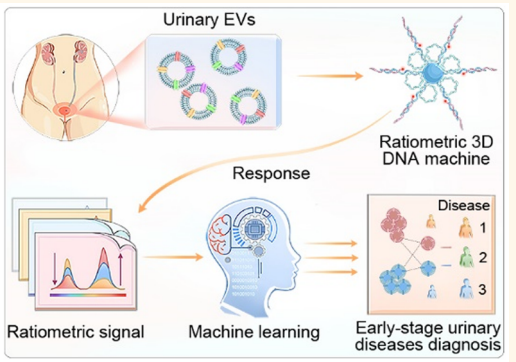
Na Wu, Xin-Yu Zhang, Jie Xia, Xin Li, Ting Yang,* and Jian-Hua Wang

Cite This: *ACS Nano* 2021, 15, 19522−19534

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Urinary extracellular vesicles (uEVs) have received considerable attention as a potential biomarker source for the diagnosis of urinary diseases. Present studies mainly focus on the discovery of biomarkers based on high-throughput proteomics, while limited efforts have been paid to applying the uEVs' biomarkers for the diagnosis of early urinary disease. Herein, we demonstrate a diagnosis protocol to realize ultrasensitive detection of uEVs and accurate classification of early urinary diseases, by combing a ratiometric three-dimensional (3D) DNA machine with machine learning (ML). The ratiometric 3D DNA machine platform is constructed by conjugating a padlock probe (PLP) containing cytosine-rich (C-rich) sequences, anchor strands, and nucleic-acid-stabilized silver nanoclusters (DNA_{AgNCs}) onto the magnetic nanoparticles (MNPs). The competitive binding of uEVs with the aptamer releases the walker strand, thus the ratiometric 3D DNA machine was activated to undergo an accurate amplification reaction and produce a ratiometric fluorescence signal. The present biosensor offers a detection limit of 9.9×10^3 particles mL^{-1} with a linear range of 10^4−10^8 particles mL^{-1} for uEVs. Two ML algorithms, K-nearest neighbor (KNN) and support vector machine (SVM), were subsequently applied for analyzing the correlation between the sensing signals of uEV multibiomarkers and the clinical state. The disease diagnostic accuracy of optimal biomarker combination reaches 95% and 100% by analyzing with KNN and SVM, and the disease type classification exhibits an accuracy of 94.7% and 89.5%, respectively. Moreover, the protocol results in 100% accurate visual identification of clinical uEV samples from individuals with disease or as healthy control by a t-distributed stochastic neighbor embedding (tSNE) algorithm.



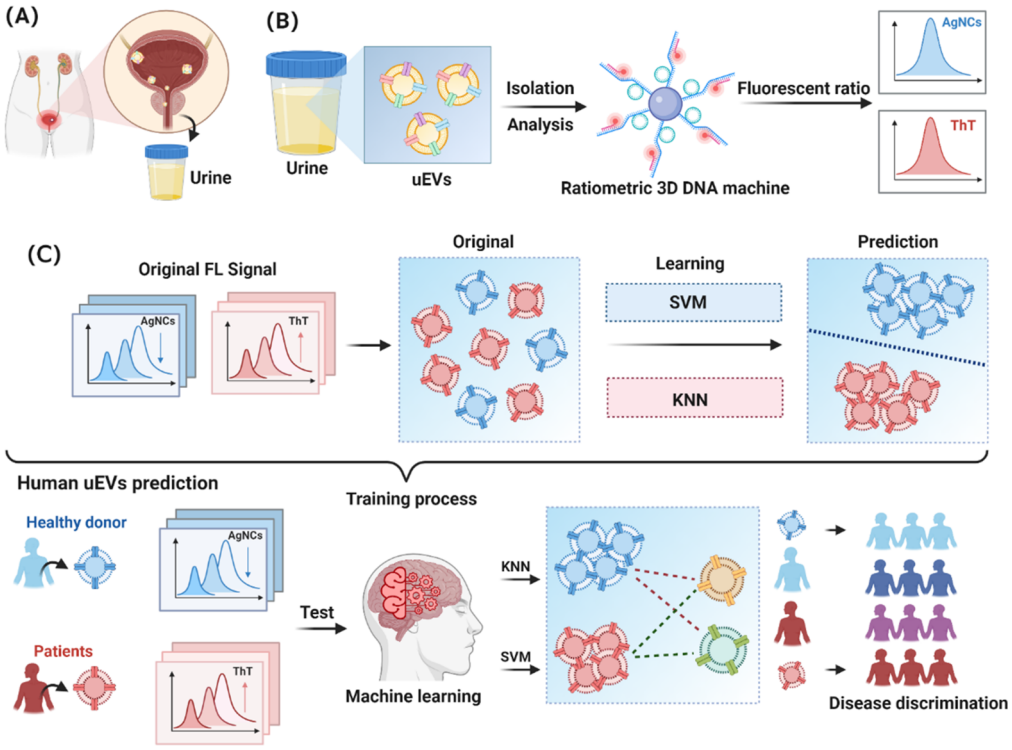
KEYWORDS: *uEVs, DNA machine, fluorescence detection, machine learning, urinary diseases*

Urinary system tumors, ranking 11th among all malignancies, have high incidences of 9.0/100 000 for males and 2.2/100 000 for females according to the World Health Organization.[1] Moreover, the first treatment recurrence rate for urinary system tumors is reported to be as high as 50% in 10 years.[1] It is recognized that effective prevention and screening of cancers can improve the cure rate and reduce the mortality rate.[1,2] Traditional medical checkups, e.g., urine cytology and cystoscopy, still remain the predominant clinical tools for cancer diagnosing and monitoring,[3] which are time-consuming and invasive, with low sensitivity and a high false positive rate. These drawbacks largely limit the popularization of early screening for urinary tract malignancy.[4] Therefore, it is of great importance to develop noninvasive diagnosis systems with high accuracy and minimal false results.

With respect to a circulating blood sample, urine is an information-rich fluid that may be more suitable for POCT diagnosis.[4,5] Moreover, many urine proteins are produced and/or shed in the kidneys and urogenital tract, making urine a promising proximal source of biomarkers.[6] However, the variability, complexity, and the wide dynamic concentration range of urinary proteins bring a significant challenge for biomarker analysis.[7−9]

Received: July 28, 2021
Accepted: November 16, 2021
Published: November 23, 2021

Scheme 1. Schematic Illustration of ML-Based uEV Analysis for Urinary Disease Diagnosis and Classification[a]



[a](A) EVs diffuse passively from bladder tissue to the urethra and were collected from naturally voided urine. (B) Collection of fluorescence spectroscopic data of uEVs by ratiometric 3D DNA machine. (C) Overview of ML algorithm-based uEV classification and multidisease diagnosis using uEV fluorescence signal patterns. This illustration is created by BioRender apps with and authorized license.

Extracellular vesicles (uEVs) are membrane-bound nanovesicles secreted by diverse cell types and are present in various biological fluids.[10−12] They carry a large amount of molecular information originating from their parental cells and participate in numerous physiological and pathophysiological processes, e.g., immune regulation and cancer development.[13−15] Moreover, increasing evidence shows that a number of disease-associated protein biomarkers have been identified in EVs from urine, and their types and expression levels are correlated with the presence and progression of certain diseases, e.g., CD63 for bladder cancer, PSA for prostate cancer, and EpCAM for kidney stones.[16−18] Therefore, urinary EVs (uEVs) are believed to be potential biomarkers for early disease diagnosis.

Nowadays, most studies regarding uEV surface proteins mainly focus on the discovery of biomarkers, and little attention is paid to their applications in the diagnosis of early urinary system diseases.[6,13] Most recently, Gu et al. successfully designed immunoaffinity detection platform of urinary exosomes for prostate cancer diagnosis.[6] Lee et al. presented integrated kidney exosome analysis for the detection of kidney transplant rejection.[19] Despite the impressive progress, most of the studies only employ a single biomarker and ignore the heterogeneity of uEV surface proteins.[6,19,20] In addition, these diagnosis strategies are still limited by the sensitivity and accuracy of current detection technology, especially with the consideration that the concentration of uEV biomarkers is ultralow at the early stage of tumors.[21] It is noteworthy that human uEVs have a more complex and heterogeneous composition than the cell line derived EVs, and it is a key problem to decipher the huge amount of garbled information within EVs.[22−26] Machine learning (ML) algorithms are capable of modeling complicated networks and leveraging valuable information within observed data for the accurate estimation and prediction of practical samples.[27−29] They have shown superior performance to analyze spectroscopic signals in complex samples,[30,31] thus becoming trustworthy solutions to this issue.

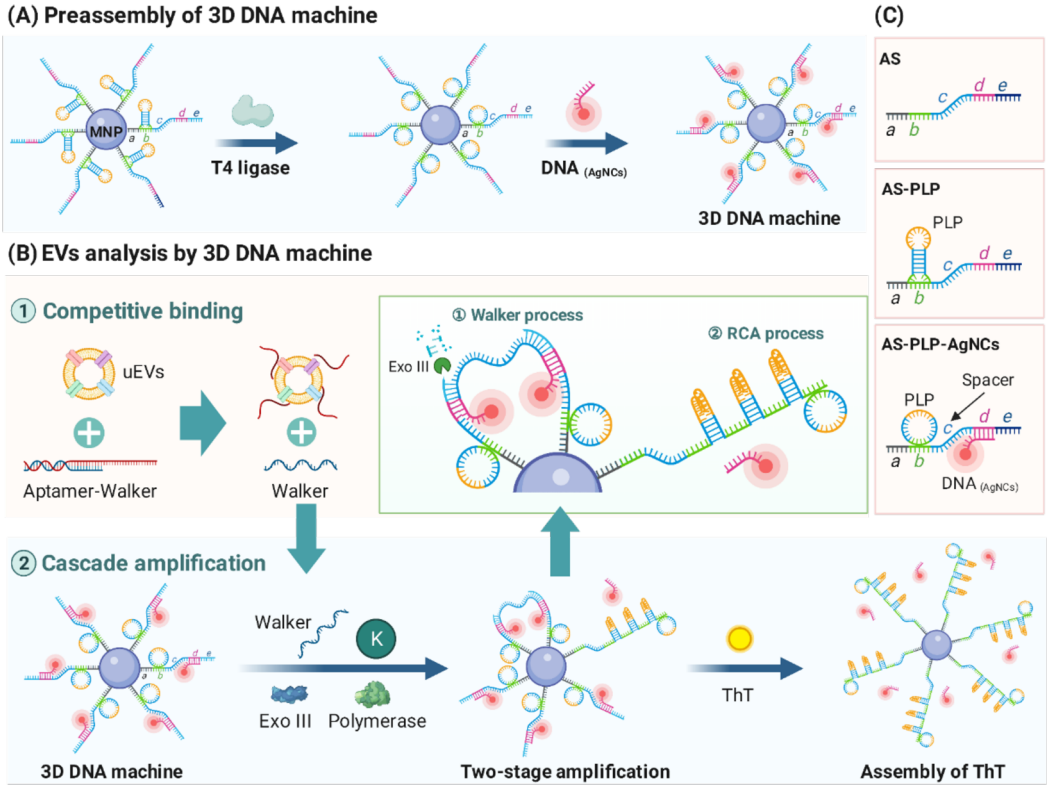
Here, we demonstrate an early stage urinary disease diagnosis method based on a ratiometric 3D DNA machine with the assistance of ML algorithm analysis. First, we hypothesized that the fluorescence signal of disease-derived EVs can be distinguished from that of normal uEVs. Then, we isolated uEVs from the collected human urine samples and obtained the fluorescence signals by the ratiometric 3D DNA machine (Scheme 1, part B). The spectral data set of uEVs is used to train the KNN and SVM models for binary and multiple classification of disease types (Scheme 1, part C). The supervised model classifies the uEV data into two clusters. The uEVs from healthy controls and three types of urinary diseases are also predicted by different combination of biomarkers. Further, tSNE was used to map the high-dimensional data collected from each individual uEV onto a two-dimensional plane and then to proceed with automated classification across three diseases types.[32,33] On one hand, this diagnosis model generates high sensitivity and anti-interference capability by a cascading integration amplification technique and ratiometric strategy; on the other hand, it achieves an accurate diagnosis and classification of multitype diseases through the introduction of ML algorithms.

RESULTS AND DISCUSSION

The Principle for uEV Detection. Scheme 2 illustrates the mechanism of the ratiometric 3D DNA machine for the

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534

Scheme 2. Schematic Illustration of the Mechanism for Profiling of uEV Surface Proteins and Multidisease Classification Based on a Ratiometric 3D DNA Machine[a]



[a](A) The ratiometric 3D DNA machine constructed by functionalizing the anchor strand (AS) (consisting of five parts: a, spacer to reduce the steric hindrance effect; b, partly complementary pairing with PLP; c, spacer to prevent excessive digestion; d, partly complementary pairing with $DNA_{(AgNCs)}$; e, partly complementary pairing with WS), padlock strand (PLP), and $DNA_{(AgNCs)}$ strand (signal probe 1) on a single MNP. (B) The ratio fluorescence for uEV detection via competitive reaction and cascade amplification (DNA walkers and RCA) of the 3D DNA machine. (C) DNA strands (AS, PLP, and $DNA_{(AgNCs)}$) required in the preassembly process. This illustration is created by BioRender apps with an authorized license.

detection of urinary disease-derived uEVs. CD63 protein, a biomarker enriched on the surface of uEVs, was chosen as a proof-of-concept target for this purpose. The prepared protocol mainly consists of two steps: (1) uEVs competitively bind to the CD63 aptamer and release a walker strand (WS) from the aptamer−walker hybridization complex, converting the information on uEV content to that of WS; (2) the released WS triggers the 3D DNA machine to generate ratiometric fluorescent intensity as signal output. The WS sequence consists of two parts: one is partially complementary with the aptamer sequence and another is complementary to the portion of the anchor strand (AS) of the 3D DNA machine (Scheme 2C). This design facilitates a higher affinity between the aptamer and uEVs with respect to that between the aptamer and WS. For the accurate detection of low level uEVs, a cascade amplification strategy combining DNA walking with RCA was introduced for the construction of the ratiometric 3D DNA machine based on an AS-decorated MNP (Scheme 2A). A C-rich padlock probe (PLP) together with a template DNA for the *in situ* growth of AgNCs was prehybridized with AS on MNPs. The PLP was then circularized upon the addition of T4 DNA ligase. AgNCs were then grown on the strand of template DNA via the reduction of $AgNO_3$ by $NaBH_4$.

When a free WS strand was introduced, it partially hybridized with AS on the MNPs and formed a blunt end at the 3′-end, triggering the stepwise digestion of Exo III from the blunt 3′-end. During Exo III digestion, a large amount of AgNCs were released from the surface of MNPs, resulting in the decrement of fluorescence for AgNCs (signal probe 1) on the surface of MNPs. As WS was also released, a set of reactions described above were cycled. Exo III digestion stops at a spacer (A 6 nt-oligo segment, indicated as portion c in Scheme 2C), thus protecting PLP from digestion. When adding phi29 DNA polymerase, the 6-nt oligo tail was progressively hydrolyzed, and the RCA reaction on the surface of MNPs was activated with AS as primers to produce long, single-stranded DNA. As PLP is subtly designed to compose a C-rich region, the RCA process produces large numbers of G-rich units, which forms stable ThT-G-quadruplex structures upon the incubation with K^+ and ThT and results in a significant fluorescence enhancement of ThT (signal probe 2) via aggregation-induced effects. Overall, these changes result in an obvious rise in the fluorescence intensity of ThT (F_{ThT}), along with a dramatic decline of the fluorescence intensity of AgNCs (F_{AgNCs}), which produce significantly amplified ratiometric fluorescence signals of F_{ThT}/F_{AgNCs} to facilitate sensitive and selective detection of uEVs. It is worth noting that the design of the DNA machine is universal; that is, when a different target protein is involved, the operation is feasible by simply changing the aptamer strands and the corresponding WS without altering the sequence of AS on MNPs. Therefore, phenotype analysis of uEVs can be facially achieved by the introduction of multiple aptamers. With the

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534

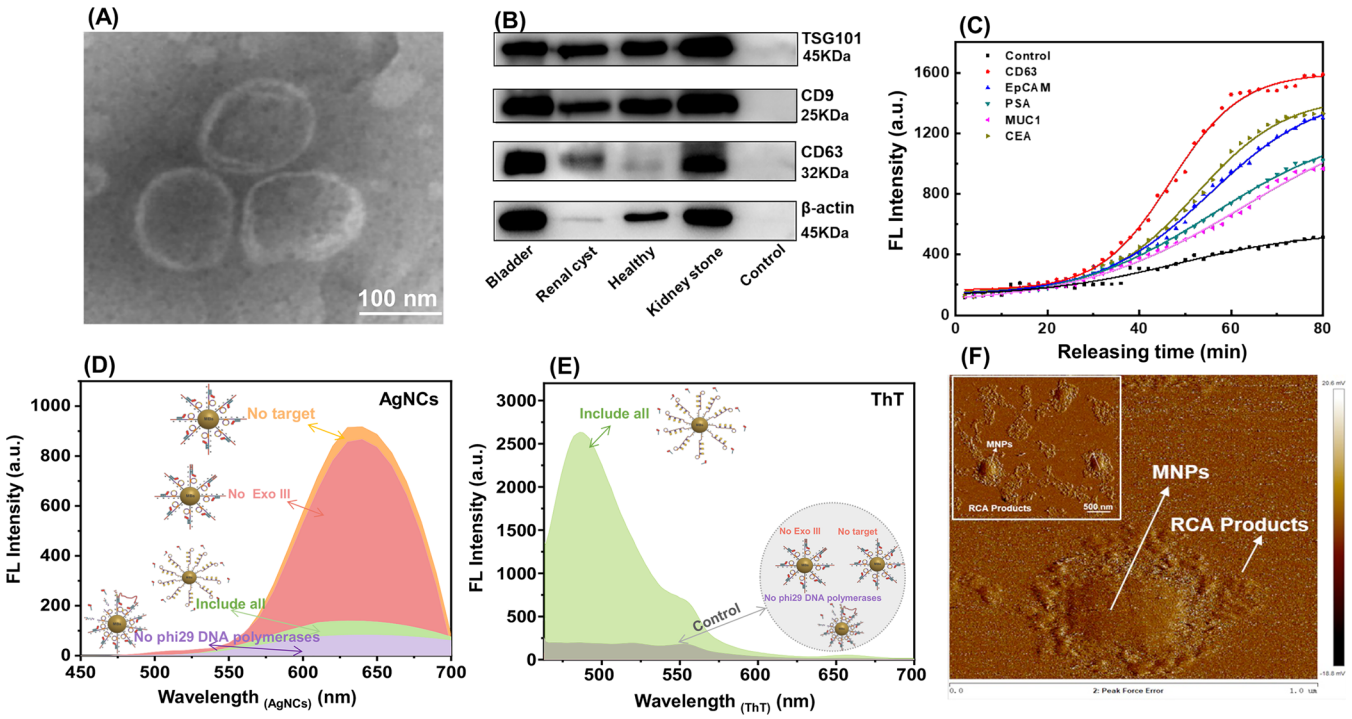


Figure 1. Characterization of uEVs: (A) TEM image of uEVs; (B) Western blotting analysis of the uEVs purified from different urinary samples (control: untreated urinary sample). The expression of CD63, CD9, β-Actin, and TSG 101 in the healthy donors and three types of patients (bladder cancer, kidney stones, and renal cysts). (C) The kinetics analysis of uEV−aptamer interaction. Fluorescence intensity of all five aptamer-based molecular beacons mixed with the extracted uEVs with different incubation times (experimental conditions: aptamer 10 μM; uEVs 1 × 10^7 particles mL^{-1}; incubation temperature, 37 °C). (D, E) Fluorescence responses of AgNCs and ThT in the presence of all component combinations (green) and in the absence of Exo III (red), target (yellow) or phi29 DNA ploymerase (purple), respectively (The gray part in Figure 1E represents either without Exo III, phi29 DNA ploymerase, or target). uEV concentration: 5 × 10^6 particles mL^{-1}. (F) AFM image of ratiometric 3D DNA machine after cascade amplification. Scale bar: 1 μm (experimental conditions: uEVs 5 × 10^6 particles mL^{-1}; RCA reaction time, 100 min).

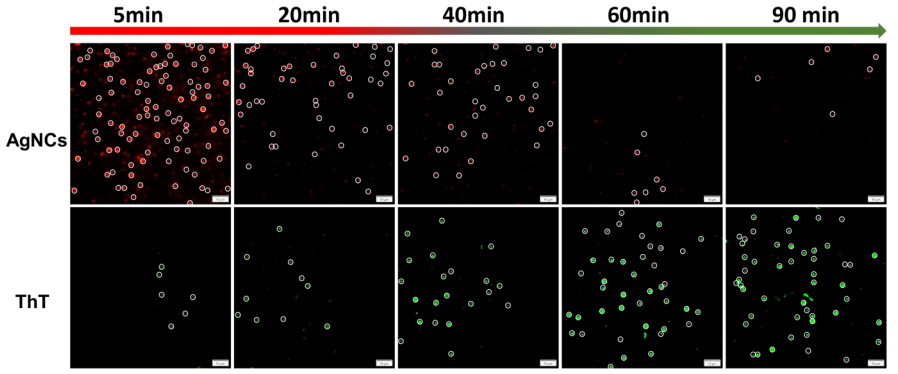


Figure 2. TIRF imaging of the MNP surface of AgNCs (red) and ThT (green) at various reaction times with uEVs. The concentration of uEVs was 5 × 10^6 particles mL^{-1}. Scale bar 10 μm. λ_{em} (AgCNs), 633 nm; λ_{em} (ThT), 488 nm.

assistance of machine leaning, binary or multiple classification of urinary disease can be realized through the training of spectral data sets of uEVs from different sources.

uEV Extraction and Feasibility Verification. By using an ultracentrifugation-based protocol (Figure S2A), uEVs with sizes ranging from 40 to 350 nm were isolated from human urine samples of healthy donors and patients with renal cyst disease, kidney stones, and bladder cancer. Nanoparticle tracking analysis (NTA) results reveal an average size of ∼100 nm of uEVs from all types of urine samples (Figure S2B−F). A flatted round shape was observed in the transmission electron microscopy (TEM) images of uEVs (Figure 1A). The presence of EV-specific markers within the uEVs was further verified by Western blotting (WB). We observed a higher expression level of CD63 in patients than in healthy donors. In comparison with healthy donors, β-actin protein shows lower expression in renal cyst patients but higher expression in both bladder cancer and kidney stone patients. In addition, CD9 and TSG101 showed similar expression trends in the four samples. From the observations herein, it can be inferred that the heterogeneous express level of uEV protein biomarkers associates with the clinical urinary disease status, illustrating the possibility to discriminate patients from healthy donors by analyzing the expression difference of uEV surface proteins.

The competitive binding of aptamer with its corresponding target protein was verified by monitoring the fluorescence

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534

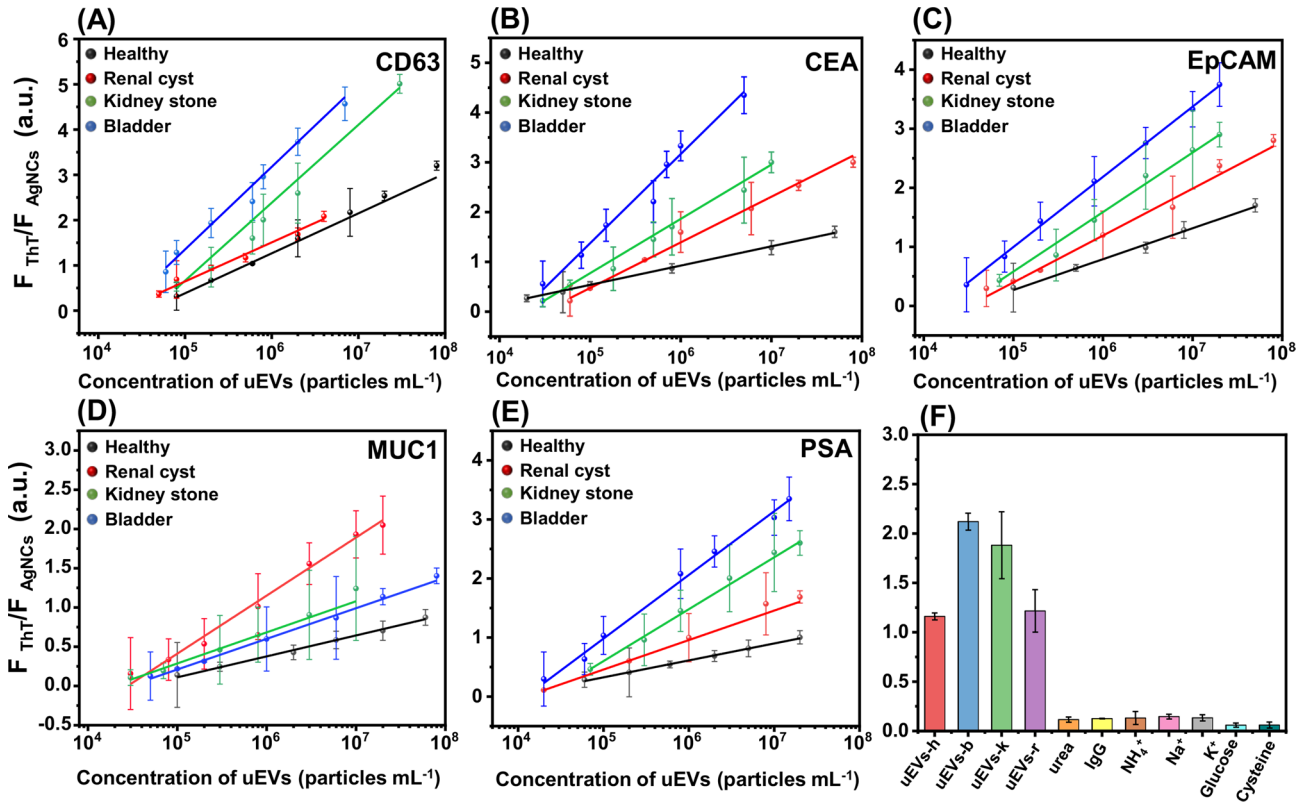


Figure 3. Linear regression between the fluorescence ratio (F_{ThT}/F_{AgNCs}) and the logarithm of uEV concentration for protein biomarkers of (A) CD63, (B) CEA, (C) EpCAM, (D) MUC1, and (E) PSA. (F) Specificity of ratiometric 3D DNA machine toward uEVs. F_{ThT}/F_{AgNCs} of uEVs (5×10^6 particles mL^{-1}), urea (1000 μg mL^{-1}), IgG (100 μg mL^{-1}), Na^+ (0.1 mg mL^{-1}), K^+ (0.05 mg mL^{-1}), NH_4^+ (0.05 mg mL^{-1}), glucose (6 mM), and cysteine (6 mM). The error bars show the standard deviation from three replicates. uEVs-h, uEVs-b, uEVs-k, and uEVs-r represent uEVs derived from healthy donors, bladder cancer, kidney stones, and renal cysts, respectively.

change during the incubation of uEVs with aptamer-based molecular beacons. As shown in Figure 1C, the increase of incubation time results in the gradual recovery of fluorescence, indicating that uEVs can effectively bind to aptamers. As the ratiometric 3D DNA machine is based on the cascade amplification of DNA walkers and RCA, their feasibility to proceed in solution was first demonstrated by gel electrophoresis and AFM (Figure S3 with detailed discussion). Further evidence for the cascade amplification to proceed on MNPs was shown by fluorescence analysis and total internal reflection fluorescence microscopy (TIRF). As shown in Figure 1D,E, the presence of all elements needed for thr DNA walker process (uEVs, DNA machine, and Exo III) leads to the release of AgNCs and the consequent fluorescence decrement (green and purple line in Figure 1D), while the increment on fluorescence of ThT is observed only in the presence of all elements required for both DNA walkers and RCA (green line in Figure 1E). The AFM image of the DNA machine after cascade amplification (Figure 1F) illustrated abundant RCA amplification products wrapped around the MNPs. TIRF results clearly showed that cascade amplification leads to a gradual decrease in the fluorescence intensity of AgNCs, whereas that of ThT exhibits an obvious increase at the same time (Figure 2). These results well demonstrated that the ratiometric 3D DNA machine assay may be used for the detection of uEVs.

Single-Biomarker Analysis for Ultrasensitive Detection of uEVs. For illustrating the potential of the ratiometric 3D DNA machine approach in uEV identification, the extracted uEVs were quantified by using NTA (Figure S2). Thereafter, the subtle changes of the expression levels of CD63 in different urinary disease-derived uEVs were determined (Figure 3) under optimal detection conditions (Figures S4 and S5). It showed a concentration-dependent fluorescence response, in which the fluorescence intensity ratio F_{ThT}/F_{AgNCs} gradually increases with the rising of uEVs concentrations due to the specific interaction between CD63 and its corresponding aptamers (Figure 3A). The linear relationship between F_{ThT}/F_{AgNCs} and the logarithm of uEV concentration are shown in Table S2. It is clear that the response slopes are significantly different for three types of urinary disease-derived uEVs and the healthy control. The bladder cancer-derived uEVs present the most abundant CD63, followed by kidney-stone-derived uEVs, renal-cyst-derived uEVs, and healthy-control-derived uEVs. This observation is well consistent with that derived in the WB results. By using CD63 as a biomarker, a LOD of 9.9×10^3 particles mL^{-1} was achieved, which stands out from the enzyme-linked immunosorbent assay (ELISA; Figure S6) and other methods with similar systems (Table S3). The average signal-to-noise ratio (SNR) of the ratiometric 3D DNA machine was ~3. Subsequently, by independently introducing other aptamer−walker complex probes, the expression of EpCAM, PSA, MUC1, and CEA on uEVs from four sources was also profiled (Figure 3B−E), with the related correlation equations summarized in Table S2. Similarly, the difference in the response slopes indicates that the expression level of EpCAM, PSA, MUC 1, and CEA vary obviously on different source-derived uEVs. It indicated that the ratiometric 3D DNA machine has the potential for distinguishing the subtle changes of different

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534

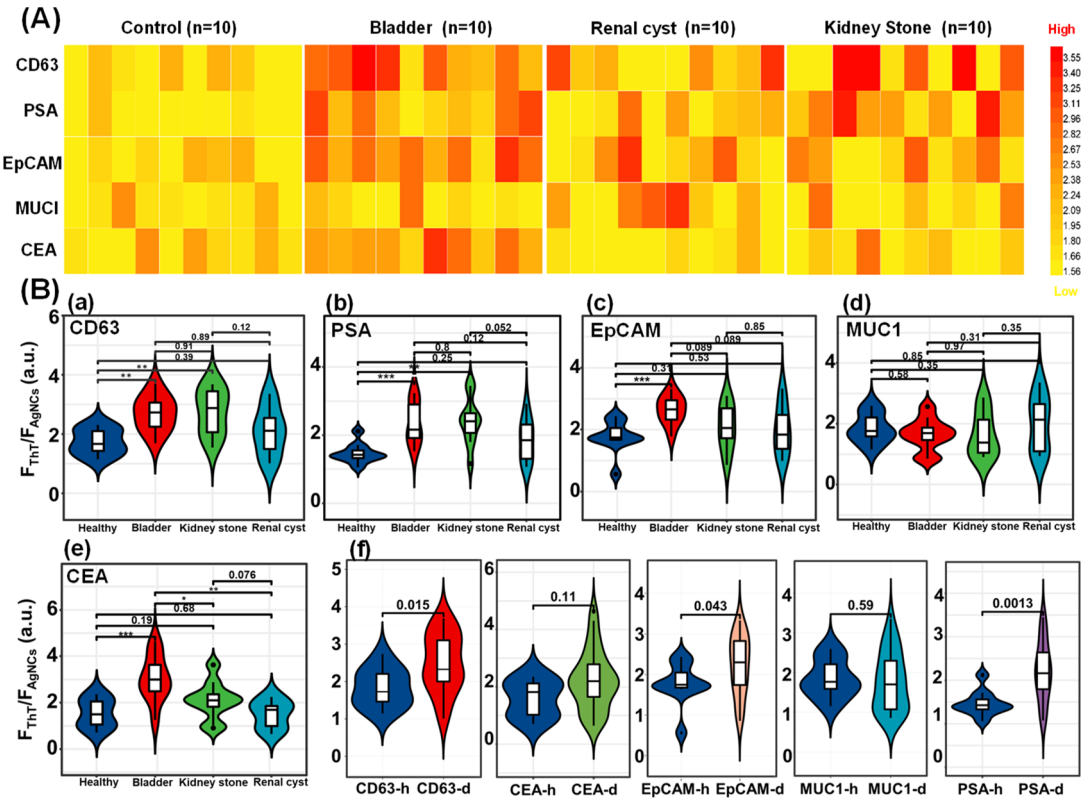


Figure 4. (A) Heat map of uEV surface protein profiles in a cohort (n = 40; 10 cases each for bladder cancer, kidney stone, renal cyst, and healthy controls). The color intensity is based on a single measurement of uEV surface protein biomarkers. (B) The estimated average expression of (a) CD63, (b) PSA, (c) EpCAM, (d) MUC1s and (f) CEA in uEVs derived from a healthy control (blue) and three types of urinary disease patients (bladder cancer (red), kidney stone (green,) and renal cyst (indigo naturalis). (f) Summary of the total expression of five proteins on the surface of uEVs of three urinary diseases (d: disease) and healthy controls (h: healthy). **p < 0.01, ***p < 0.001, ****p < 0.0001 (*Student's t test*).

biomarker levels in different uEVs. In addition, the stage study of bladder cancer showed that the ratiometric 3D DNA machine is able to classify urinary diseases at different stages, confirming a potential early diagnostic ability (Figure S7).

The favorable anti-interference capability and specificity are crucial for the sensing strategy in complex biological samples. To evaluate the disturbance resisting capability, some potential interfering substances which might be involved in human urine samples, e.g., protein metabolism (urea), protein (IgG), inorganic cations (Na^+, K^+ and NH_4^+), urine sugar (glucose), and amino acid (cysteine), were assessed. The concentrations of these substances were selected by referring to the clinically diagnosed positive level. As shown in Figure 3F, the F_{ThT}/F_{AgNCs} responses caused by these coexisting substances are negligible compared to the obvious F_{ThT}/F_{AgNCs} responses for uEVs (5×10^6 particles mL^{-1}). This observation well demonstrates the favorable selectivity of the ratiometric 3D DNA machine for the detection of uEVs in a biological sample matrix.

Assay of Clinical Sample-Derived uEVs Protein for Disease Identification. The utility of the ratiometric 3D DNA machine was extended to profile uEV protein in clinical urine samples. First, comparison tests were performed using conventional NTA and our established approach to directly quantify the extracted uEVs. uEVs derived from healthy urine samples, bladder cancer, kidney stones, and renal cysts were extracted according to the standard ultracentrifugation procedures described in section S1.4 and measured by NTA. The uEV detection results by using the 3D DNA machine approach exhibit a high consistency with those of NTA and ELISA (Figure S8). This demonstrates that the 3D DNA machine analytical strategy has a favorable potential in the assay of clinical specimens. The previous studies indicated that the expression of EV surface proteins was highly heterogeneous between different types of disease and healthy control, which could provide conclusive information in disease discrimination against a healthy specimen. In the present study, we attempted to find the association between the protein biomarker on uEVs and urinary disease by profiling the surface protein expressions on clinical sample-derived uEVs. We collected uEVs from 40 naturally voided urine samples of volunteers enrolled in a cohort involving 10 cases each for bladder cancer, kidney stones, renal cysts, and healthy controls. In all 30 patients, tissues were available for clinical pathology interpretation. We obtained 200 mL of urine for each patient on the day of the preceding surgery for uEV extraction and detection. Figure 4A summarizes the expression of uEV surface protein (CD63, PSA, EpCAM, MUC1, and CEA) from patients and healthy groups. As shown in the heat map of uEV surface protein profiles (Figure 4A), each kind of urinary disease showed a heterogeneous level of EV surface protein, whereas low expression of the surface proteins was observed on uEVs from healthy donors. The violin chart details the different expression level of uEV surface proteins. As shown in Figure 4B, CD63, PSA, EpCAM, and CEA were obviously overexpressed on the surface of uEVs from both bladder cancer and kidney stone disease, while those biomarkers were slightly overexpressed in the case of renal cyst disease. Compared with four other proteins, the MUC1 showed relatively high expression in the case of renal cyst disease,

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534

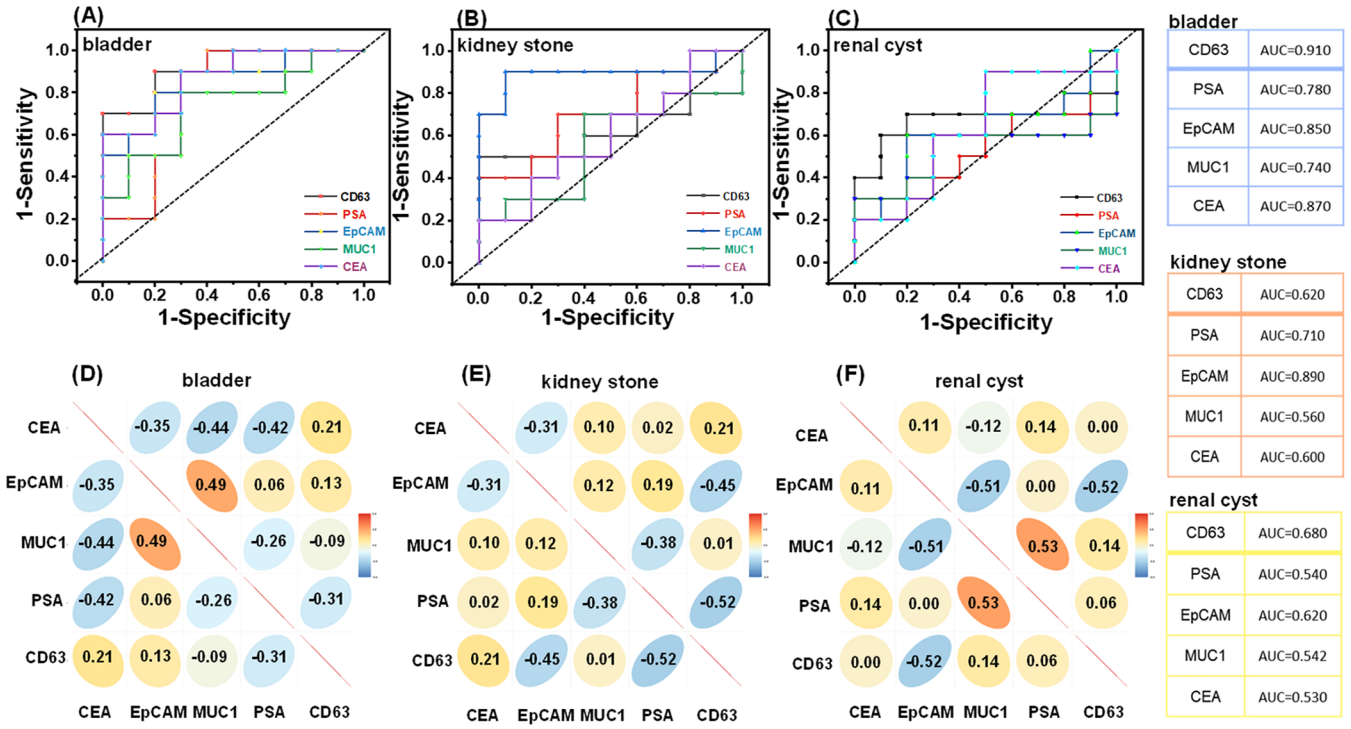


Figure 5. (A−C) ROC curve analysis for evaluating the diagnostic power of ratiometric the 3D DNA machine in measuring three urinary-disease-derived uEVs. (A) Bladder cancer, (B) kidney stone, (C) renal cyst. AUC, area under the curve. (D−F) Correlation matrix of marker pairs in three urinary-disease-derived uEVs. (D) Bladder cancer, (E) kidney stone, (F) renal cyst (*r*: correlation coefficient).

while it provided the lowest diagnostic performance in the present model. Overall, the five uEV surface protein biomarkers in the urinary disease groups showed higher expression than healthy groups (Figure 4Bf).

The receiver operating characteristic (ROC) analysis was used to further investigate the diagnostic accuracy of the ratiometric 3D DNA machine using a single marker. The area under the ROC curve (AUC) represents the model's sensing performance, i.e., the larger the AUC value is, the higher the diagnostic capability to correctly identify patients with the corresponding disease. As shown in Figure 5A−C, the AUC values for the diagnostic of three types of urinary diseases using five markers fall into a range of 0.53 to 0.91, indicating that the ratiometric 3D DNA machine is able to a certain extent to discriminate different urinary diseases. The result also indicated that single-biomarker analysis lacks sufficient sensitivity and specificity. This may be achieved by formulating a strategy to improve both specificity and accuracy by introducing multiple markers. Previous studies illustrated that a lower correlation between different markers may help to enhance the content of the message and improve the diagnostic performance.[23] In the present study, the Pearson correlation coefficient was used to evaluate the correlation between these biomarkers. In Figure 5D−F, overall low correlations were given; i.e., the correlation of CD63+CEA pairing for the renal cyst was 0.00, while that of CD63+MUC1 pairing for kidney stones was 0.01. In the highest correlation case derived from PSA+MUC1 pairing of renal cysts, the value was found to be only 0.53. It can be concluded that the information of disease diagnostics, which are decided by each biomarker, do not overlap significantly. Therefore, it may be an opportunity to improve the ratiometric 3D DNA machine diagnostic power if the sensing signals from each biomarker are blended properly.

ML Algorithms Assisted Multibiomarker Joint Diagnosis. The heterogeneity information of uEVs by single-biomarker-based diagnosis is easy to mask with bulk measurement.[32] In the opposite, multibiomarker-based diagnosis can offer sufficient heterogeneity information; however, the conventional data handling method can hardly extract the effective multibiomarker information and achieve accurate/precise diagnosis. Therefore, the ML algorithm diagnosis model was used for this issue due to its excellent performance in dealing with complex clinical data. Logically, the diagnosis decision accuracy is based on the diagnostic model accuracy, i.e., the closer the model accuracy value is to 100%, the more excellent the certainty of the decision is.[23,24] In order to ensure the accuracy of the diagnostic model with limited sets of data, an advanced machine learning method was implemented to analyze heterogeneous signal patterns of urinary disease-derived uEVs for highly precise diagnosis.[34] Two popular classification ML algorithms, K-nearest neighbor (KNN) and support vector machine (SVM), were chosen to optimize the multiplexed marker diagnostic model (Figure 6). The KNN algorithm mainly depends on the surrounding limited adjacent samples, rather than the algorithm of discriminating the class domain to determine the category,[35] and thus KNN is more suitable than other algorithms to divide the sample set with a lot of crossovers or overlap of the class domain.[36] As for the SVM algorithm, it is a kind of binary classification algorithm to classify the sample by building a hyperplane function, which has the advantage in handling a small set of samples and overfitting issues. Moreover, KNN and SVM can provide useful information in terms of understanding variables, e.g., the key factor of each biomarker in its decision-making process.[35]

The training data were collected from interaction between the ratiometric 3D DNA machine and five kinds of uEV surface proteins of an independent validation cohort (total sample

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534

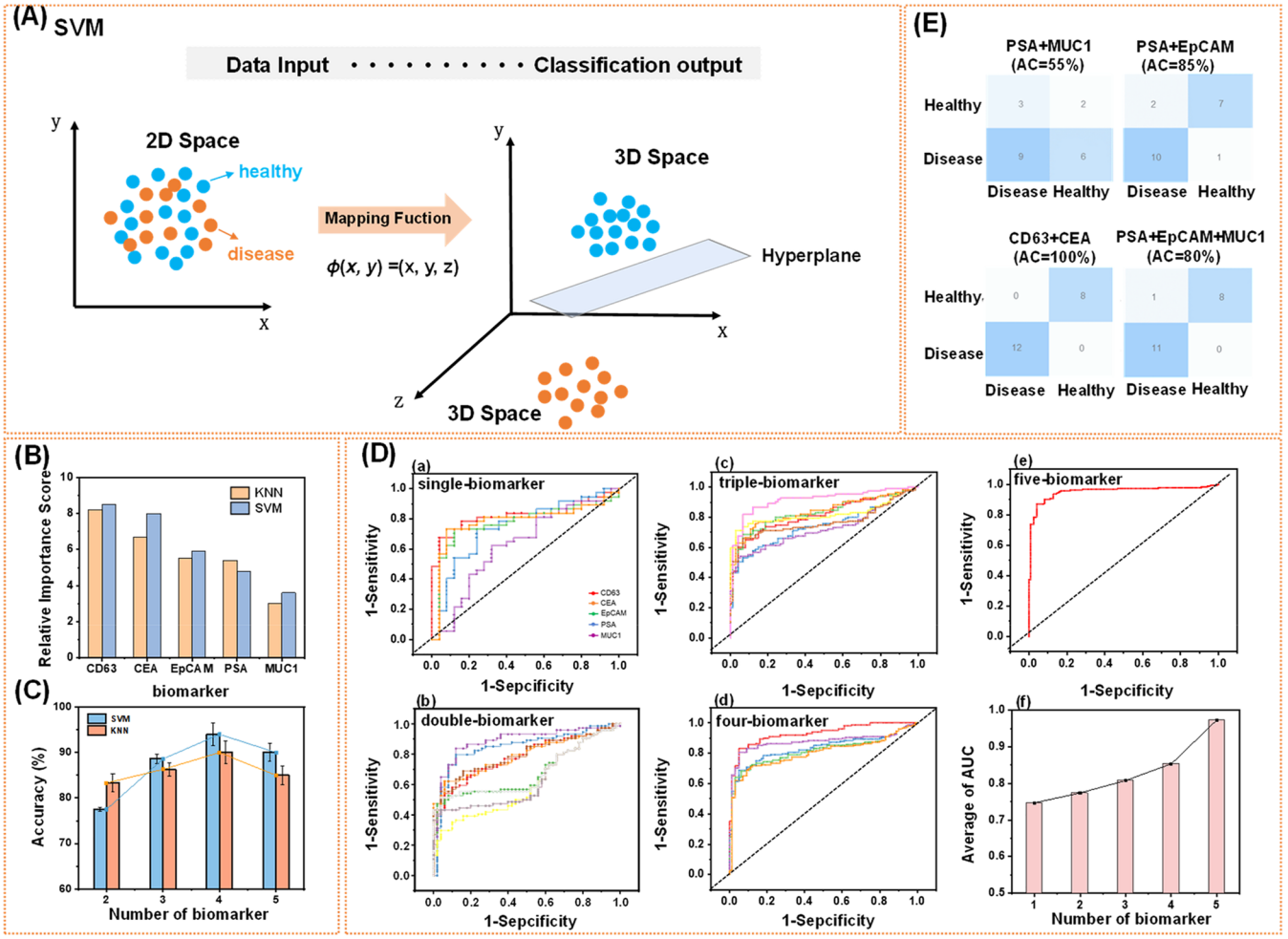


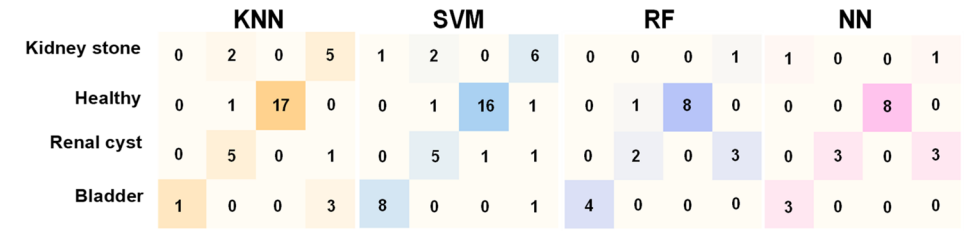
Figure 6. ML-based diagnosis model of urinary disease patients by using KNN and SVM algorithms. A total of 63 clinical samples were separated randomly into a training set (70%) and a test set (30%). (A) By taking the algorithm model architecture of SVM, the input space is transformed into a high-dimensional space by the nonlinear transformation defined by the inner product function, and then the optimal classification is obtained in this space. (B) Relative importance score of the single biomarkers (CD63, CEA, EpCAM, PSA, and MUC1) in KNN and SVM algorithms. (C) Average accuracy predicted by KNN and SVM algorithms with different numbers of biomarkers. (D) ROC curves of the multibiomarker diagnostic model with different biomarker combinations. The ROC curves with one (a), two (b), three (c), four (d), and five (e) biomarkers, illustrating a better diagnostic performance with an increased number of biomarker combinations. The AUC, which represents the diagnostic performance, is the single most major parameter. (E) Confusion matrix of SVM algorithm of binary classification (discrimination between urinary disease patients and healthy control) with different biomarker combinations.

capacity (n) = 63, including bladder cancer n = 13, kidney stone patients n = 13, renal cyst patients n = 12, and healthy donors n = 25). Then, the data were shuffled randomly to ensure that each batch of the data set was composed of both healthy donor-derived data and urinary disease-derived data. To avoid the overfitting problem of KNN/SVM-trained diagnosis models, 70% of the total data set was used as the training set for supervised learning, and the remaining 30% of the data set was used as a blind test set to validate the diagnostic model performance on urinary disease screening. Subsequently, the test data corresponding to each biomarker was input into the diagnostic models to generate an output of binary classification (healthy state and disease) and multiple classification (bladder cancer, kidney stones, and renal cysts).

Figure 6B revealed the relative importance of each biomarker in the five-biomarker combination system investigated by KNN and SVM. The CD63, CEA, PSA, and EpCAM biomarkers showed a similar sensing ability yet had a notably different importance in the KNN and SVM algorithms. As illustrated in the algorithm of KNN, PSA and EpCAM showed similar importance (Figure 6B, 5.4 vs 5.5), while in SVM the importance obtained from EpCAM was higher than PSA (Figure 6B, 5.9 vs 4.8). In addition, each biomarkers' importance was not necessarily related to its accuracy. Although PSA and EpCAM have a similar importance for KNN, their combinations with other biomarkers showed different diagnosis accuracy, as shown in Figure S9, the accuracy is 78.9% for CD63+EpCAM, whereas that for CD63+PSA is 90%. As expected from the calculations in Figure 6B, MUC1 showed the least importance (only ~3), which suggests that some biomarkers barely contribute to the diagnostic performance.

The diagnostic performance of binary classification by KNN and SVM using different numbers of biomarkers at the marker panel level was then investigated. In general, the accuracy of the confusion matrix output can represent the performance of ML models. In both KNN and SVM, the average accuracy increased with the increase of the number of biomarkers, indicating that the multibiomarker approach provides a far more accurate

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534



			Predicted Class		Sensitivity	Specificity	Accuracy
			Positive	Negative			
KNN	Actual Class	Positive	11	1	91	100	84.2
		Negative	0	17			
SVM	Actual Class	Positive	19	2	90	94	89.5
		Negative	1	16			
RF	Actual Class	Positive	7	1	87	100	78.4
		Negative	0	8			
NN	Actual Class	Positive	7	0	100	100	78.4
		Negative	0	8			

All predicted value were determined at the optimal cut-off value. (Sensitivity and specificity only consider disease and healthy control; accuracy: across three different multi-type diseases and healthy control.)

Figure 7. Comparison of accuracy with different algorithms by using five biomarkers. Those models were verified using split data of the training set of 70% and test set of 30%. Test conditions: 5-fold cross validation (CV). Table: the predicted score represents a score value of discriminant result.

clinical output (Figure 6C). Figure 6C showed that the average accuracy of KNN and SVM algorithms with a double marker was 83% and 77%, respectively, while the average accuracy improved dramatically to 90% and 95% for KNN and SVM, respectively, by increasing the number of biomarkers to four. The same trend was observed with the ROC curve (Figure 6D). Those results support the idea that pathologically independent biomarker combination broadens the disease-specific information.

Figure 6D,a showed the diagnostic capability of a single biomarker obtained through the ROC curve. CD63 gives rise to the largest AUC value, followed by CEA, EpCAM, PSA, and MUC1. We hypothesized the fact that when these biomarkers were combined, the diagnostic performance of the model may vary significantly depending on the biomarker's combination. To validate this assumption, SVM and KNN algorithms were used to predict the accuracy of the diagnostic model under different combinations of biomarkers. As shown in Figure 6D, the diagnostic performance of models was indeed significantly different among various biomarker combinations. It was observed in Figure 6E that the pair of CD63 + CEA (AC = 100%) led to a better diagnostic performance than PSA + MUC1 (AC = 55%) in the SVM algorithm. In addition, the accuracy of PSA + EpCAM (AC = 85%) was higher than that of PSA + EpCAM + MUC1 (AC = 80%) in the SVM algorithm, indicating that the variation factor (the variation in the ways that biomarkers combine) was often greater than the effect of the number of biomarkers. In fact, KNN and SVM algorithms contain complex nonlinear function calculation. This means that increasing the number of biomarkers may not necessarily ensure a monotonous increase of the diagnostic accuracy. Therefore, when the five biomarkers were combined, the average accuracy of the model's predictions was even reduced compared to the four biomarkers combination, indicating that an optimal biomarker combination exists (Figure 6C). As expected, the best sensing output in terms of accuracy was achieved from combinations of CD63 + PSA + MUC1 + CEA (AC = 100%) and CD63 + EpCAM + MUC1 + CEA (AC = 95%) for SVM and KNN in terms of accuracy on average (Figure S10), yet this is not statistically significant. In addition, for both models, the diagnostic performance showed similar trends.

Clinical Diagnosis of Different Types of Urinary Diseases. The accurate diagnosis of different types of urinary disease in the early stages is of great importance. However, the correlation and the information overlap of these diseases bring challenges to their accurate discriminations. By incorporating SVM and KNN algorithms into the diagnosis models, an average accuracy of 69.1% and 66.3% was achieved for the SVM and KNN algorithms, respectively (Figure S10), revealing the ability of the ML algorithms for predicting multitype urinary disease. However, the average predicting accuracy in multiple classification was significantly lower than that in binary classifications (88.5% and 85.4% for the SVM and KNN algorithms, respectively). It may be due to the fact that the overlapping distribution of multimarker information has a coupling effect, making the multiple classifications more difficult. As shown in Figure S9, the double-biomarker combination of PSA + CEA exhibited a much better diagnostic performance than that of the triple-biomarker combination of CDE + PSA + MUC1 in a SVM algorithm (the AC values were 84.2% and 57.9%, respectively). This observation indicated that the biomarker combination itself has a higher impact on the model than the number of biomarkers. However, Figure S10 illustrated that the two best diagnosis outputs from the perspective of accuracy were obtained from the four-biomarker combinations, i.e., CD63 + EpCAM + PSA + CEA (AC = 94.7%) in SVM and CD63 + EpCAM + MUC1 + CEA (89.5%) in KNN, confirming to a certain extent that the number of biomarkers still makes sense in the modeling process. Overall, the diagnosis results of multitype urinary diseases showed a similar trend to that of the binary classification, that is, with the increase of the number of

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534

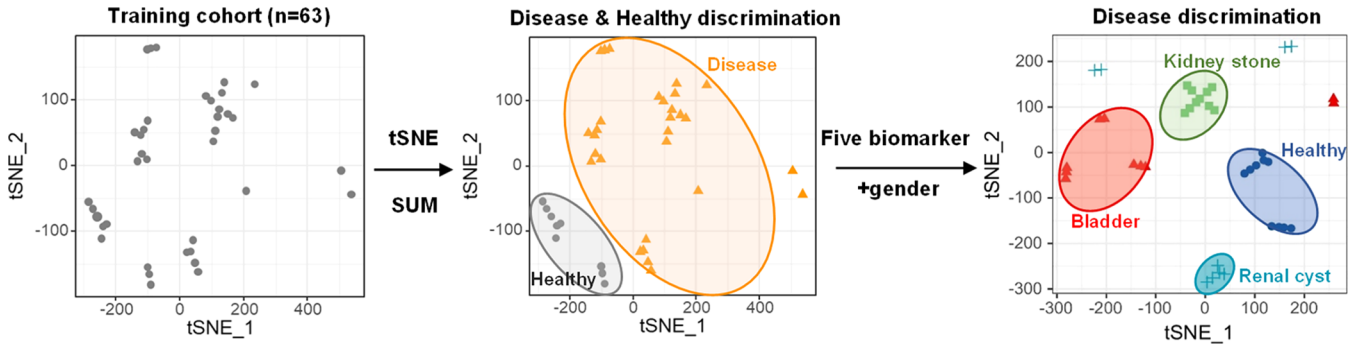


Figure 8. uEVs analysis data derived from multitype diseases and healthy controls are visually classified through tSNE mapping. A two-step tSNE proceeded in the training cohort (n = 63). The first tSNE discriminated between urinary disease patients (n = 38) and healthy controls (n = 25) using the SUM signature as input, while the second tSNE classified disease individuals into three disease types (bladder cancer n = 13; kidney stone n = 13; renal cyst n = 12) and healthy control (n = 25), using the five biomarkers (CD63, EpCAM, MUC1, CEA, and PSA) and gender as inputs.

biomarkers, the accuracy of the diagnostic model is nonlinearly improved (Figure S11). Therefore, it is especially important to take into consideration whether the clinical resources are sufficient during actual modeling, followed by a comprehensive design and the determination of research priorities.

In addition, considering the diversity of information handling techniques and the corresponding algorithms often leads to differences among predicted results, four common ML algorithms, i.e., KNN, SVM, random forest (RF), and neural network (NN), were used to classify multitype urinary diseases. As shown in Figure 7, the average classification accuracy of KNN, SVM, RF, and NN algorithms for multitype urinary diseases were 84.2%, 89.5%, 78.4%, and 78.4%, respectively. It is worth noting that the diagnostic ability of RF and NN algorithms is slightly inferior compared to SVM and KNN. The main reason may be that the optimizing parameters cause the overfitting of the NN algorithm and the imbalance of actual sample data weaken the diagnostic ability of the RF.[37] In the four algorithms, SVM showed the best diagnostic performance in term of accuracy, although the result is not statistically significant.

Visualization of the model output is conducive to simplifying the complex confusion matrix and to intuitively understanding the distribution of multitype diseases. To visualize the subpopulation signature of uEV with varied expressions of CD63, EpCAM, PSA, MUC1, and CEA from different types of urinary diseases, the tSNE algorithm, the currently available best approach of dimension reduction visualization, was applied. Subsequently, all of the analytical information of the five biomarkers from each clinical samples was used as inputs to generate the predicted results with disease groups or the healthy control. Notably, the disease group and healthy group formed two distinct clusters (Figure 8), suggesting the classification results can be well visualized on a 2D plane map by using tSNE. Moreover, we tried to use the information of five biomarkers and the gender as inputs and changed the prediction results as the type of urinary disease or healthy control. As visually shown in Figure 8, all clusters of the three types of urinary disease and healthy control were clearly separated without any overlap under a uEV concentration of 5×10^6 particles mL^{-1}. The results demonstrated that the tSNE algorithm was successfully applied in visualization of the model output of early urinary disease diagnosis, providing a promising perspective on the development of sensitive and powerful diagnostic/analytical strategy.

CONCLUSIONS

In summary, a feasible and affordable diagnosis protocol composed of a ratiometric 3D DNA machine and machine learning (ML) algorithms was developed for an ultrasensitive assay of urinary EVs and the accurate classification of early urinary diseases. The ratiometric 3D DNA machine with high sensitivity and favorable anti-interference was fabricated by combining the RCA amplification-based 3D DNA walker with two kinds of label-free fluorescent probes serving as the ratiometric signal sources. Based on the universality of the ratiometric 3D DNA machine, the expression of dysregulated proteins can be profiled with the corresponding signal characteristics by targeting multiple biomarkers on the uEVs. Subsequently, the correlation between multibiomarker signals and clinical status was analyzed by applying two typical ML algorithms, K-nearest neighbor (KNN) and support vector machine (SVM). Under the supervision of both algorithms, the accuracies for the discrimination of individuals of urinary disease or the healthy state and the classification of urinary disease types reached 100% and 86.1% in a clinical validation cohort (n = 63), respectively. In general, the average accuracies of both algorithms exhibited a monotonic increase with the number of biomarkers used. Yet introducing more biomarkers did not guarantee an improvement in diagnosis performance, confirming the necessity of synergistic biomarker choosing. All results confirmed the great potential of combining a ratiometric 3D DNA machine and ML algorithms for screening early urinary diseases, which provided a promising perspective on the development of sensitive and powerful diagnostic/analytical strategy in life science assays.

METHODS

Preparation of Anchor Strands (AS)-Padlock Probe (PLP) Co-Conjugated MNPs. The NH_2-modified AS strands were annealed at 95 °C for 6 min and then slowly cooled down to 25 °C to form stable structures. In the following, COOH-MNPs (approximately 200 nm; Figure S1) were decorated with AS: 100 μL of COOH-MNPs (10 mg·mL^{-1}) were washed and resuspended in 500 μL of MES (100 mM, pH 5.0) for activating by reacting with 200 μL of EDC (200 mg mL^{-1}) and 200 μL of NHS (200 mg mL^{-1}) for 60 min at room temperature. After washing with MES buffer three times, the activated MNPs were resuspended in 500 μL of hybridization buffer (10 mM PBS, 0.05% tween-20, pH 7.4) and mixed with 100 μL of NH_2-teminated AS (8 μM) for an overnight incubation at 25 °C. The coupling efficiency of AS with −COOH on the MNP surface is calculated to be 41.25% (Figure S12). Afterward, the AS decorated MNPs (AS-MNPs) were repeatedly

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534

washed and incubated in 500 μL of blocking buffer (10 mM PBS, 0.05% tween-20, 1% BSA pH 7.4) for 30 min at 37 °C to block the nonspecific binding sites. Finally, AS-MNPs were repeatedly washed and resuspended in 500 μL of hybridization buffer. Subsequently, 5 μL of the padlock probes (8 μM), 5 μL of 10 × T4 buffer, 35 U/μL of T4 DNA ligase, and 10 μL of BSA (2 mg mL^{-1}) were mixed with AS-MNPs suspension and incubated at 16 °C for 3 h. Thereafter, the resulting MNPs (labeled as AS-PLP-MNPs) were washed three times with a hybridization buffer and resuspended in 500 μL of the same buffer.

Construction of 3D DNA Machine. The DNA-templated AgNCs ($DNA_{(AgNCs)}$) were prepared according to the literature.[38] Briefly, 100 μL of $DNA_{(AgNCs)}$ (10 μM) and 20 μL of $AgNO_3$ (600 μM) were mixed and incubated in an ice−water bath for 30 min. Then, 20 μL of the freshly prepared $NaBH_4$ (600 μM) was added with vigorous shaking for 40 s. The mixture was allowed to incubate at 4 °C in the dark for 5 h, followed by an overnight incubation at 37 °C for hybridizing with 500 μL of AS-PLP-MNPs. Finally, the excessive $DNA_{(AgNCs)}$ was removed by magnetic separation. The resulting AgNCs-MNPs were repeatedly washed and resuspended in 500 μL of TE buffer (pH 7.4), followed by characterization by energy dispersive spectrum (EDS) analysis (Figure S1B).

Detection of uEVs Based on the Ratiometric 3D DNA Machine. The aptamer strands (1 μM) and walking strands (1 μM) were annealing at 95 °C for 6 min, and then slowly cooled down to 25 °C to form a stable dsDNA structure. A total of 10 μL of different concentrations of uEVs were separately mixed with 50 μL of the above formed dsDNA, 10 μL of Exo III (8 U/μL), 50 μL of AgNCs-MNPs, and 10 μL of the reaction buffer (10 mM PBS, 0.05% tween-20, pH 7.4) and incubated at 37 °C with gentle shaking at 156 rpm for 40 min. Subsequently, an RCA reaction was conducted in a mixture containing 10 μL of phi29 DNA polymerase (40 U/μL), 10 μL of 10 × phi29 DNA polymerase buffer, 3 μL of dNTP (10 mM), 10 μL of BSA (10 mg/mL), and the above-mentioned mixture. The reaction mixture was incubated at 37 °C for 100 min and denatured at 65 °C for 15 min. The final product (100 μL) was mixed with 6 μM of ThT (50 μL), 10 mM of K^+ (50 μL), and 300 μL of TE buffer (40 mM Tris-HCl, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$) at 37 °C for 30 min. The ultimate products were subject to fluorescence measurement. All fluorescence spectra were recorded using on a Hitachi F-7000 fluorescence spectrophotometer (Tokyo, Japan). The excitation wavelength was 440 nm, and the emission spectra were recorded over the wavelength from 450 to 500 nm with a slit width of 5 nm for both excitation and emission.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.1c06429.

Materials and instrumentations used in this work; supplemental experimental details; sequences of the oligonucleotides used in this study; the linear relationship between the fluorescence intensity ratio (F_{ThT}/F_{AgNCs}) and the logarithm of concentration of uEVs for five protein biomarkers from the healthy group and three types of urinary diseases; comparison of the capability of the present method with some common DNA machine approaches for a uEV assay; summary of the heatmap cohort; summary of the algorithms training cohort; performance of urinary uEVs biomarkers to differentiate diseases and the healthy control in the ROC cohort; SEM image of MNPs; EDS spectrum of AgNCs-MNPs; evaluation of the isolated uEVs; native PAGE characterization of the DNA walker process; agarose gel (1%) electrophoresis for the characterization of the RCA process; atomic force microscopy (AFM) images of the RCA product; the effect of various parameters on the preassembling of the ratiometric 3D DNA machine. The effect of various parameters on the performance of the ratiometric 3D DNA machine; the sensitivity comparison between the ratiometric 3D DNA machine and commercial ELISA kit with CD63 as a biomarker; linear relationship between the absorbance and the logarithm of the number of uEVs of the commercial ELISA kit for quantification of the surface CD63 of uEVs; quantification of uEVs from the healthy control and three types of urinary diseases (bladder cancer, kidney stones, and renal cysts) by NTA, DNA machine, and ELISA; optimization of biomarker combination for the training cohort. The performances of SVMs and KNNs using five markers as the input for binary classification (healthy and disease); optimization of biomarker combination for the training cohort; the performances of SVMs and KNNs using five markers as the input for multitype urinary diseases (the three kinds of urinary diseases and healthy donors); confusion matrix of KNN (A−D) and SVM (E−H) algorithms for diagnosis output as the best different biomarker combinations; and calibration curve of the FAM-conjugated anchor strand (PDF)

AUTHOR INFORMATION

Corresponding Author

Ting Yang − *Research Center for Analytical Sciences, Department of Chemistry, College of Sciences, Northeastern University, Shenyang 110819, China;* orcid.org/0000-0001-5517-5658; Email: yangting@mail.neu.edu.cn

Authors

Na Wu − *Research Center for Analytical Sciences, Department of Chemistry, College of Sciences, Northeastern University, Shenyang 110819, China*

Xin-Yu Zhang − *General Hospital of Northern Theater Command, Shenyang 110015, China; Dalian Medical University, Dalian 116044, China*

Jie Xia − *Product Research Institute, Research and Development Center, Huayou Nonferrous Industrial Group, Zhejiang Huayou Cobalt Co., Ltd., Quzhou 324000, China*

Xin Li − *General Hospital of Northern Theater Command, Shenyang 110015, China*

Jian-Hua Wang − *Research Center for Analytical Sciences, Department of Chemistry, College of Sciences, Northeastern University, Shenyang 110819, China;* orcid.org/0000-0003-2175-3610

Complete contact information is available at: https://pubs.acs.org/10.1021/acsnano.1c06429

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support from the Natural Science Foundation of China (21874014, 21727811, 22074011), the Fundamental Research Funds for the Central Universities (N180505021, N2005017, N2005027), and the Liaoning Revitalization Talents Program (XLYC2007102, XLYC1802016) are highly appreciated. Special thanks are due to the instrumental analysis from the Analytical and Testing Center, Northeastern University.

REFERENCES

(1) Welton, J. L.; Khanna, S.; Giles, P. J.; Brennan, P.; Brewis, I. A.; Staffurth, J.; Mason, M. D.; Clayton, A. Proteomics Analysis of Bladder CancerExosomes. *Mol. Cell. Proteomics.* **2010**, *9*, 1324−1338.

(2) Pisitkun, T.; Shen, R. F.; Knepper, M. A. Identification and Proteomic Profiling of Exosomesin Human Urine. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 13368−13373.

(3) Wang, Q.; Sun, Y.; Yang, Y.; Li, C.; Zhang, J.; Wang, S. Quantitative Proteomic Analysis of Urinary Exosomes in Kidney Stone Patients. *Transl. Androl. Urol.* **2020**, *9*, 1572.

(4) Fujita, K.; Nonomura, N. Urinary Biomarkers of Prostate Cancer. *Int. J. Urol.* **2018**, *25*, 770−779.

(5) Raj, D. A. A.; Fiume, I.; Capasso, G.; Pocsfalvi, G. A Multiplex QuantitativeProteomics Strategy for Protein Biomarker Studies in Urinary Exosomes. *Kidney Int.* **2012**, *81*, 1263−1272.

(6) Li, P.; Yu, X. Y.; Han, W. J.; Kong, Y.; Bao, W. Y.; Zhang, J. Q.; Zhang, W. C.; Gu, Y. Q. Ultrasensitive and Reversible Nanoplatform of Urinary Exosomes for Prostate Cancer Diagnosis. *ACS Sensors.* **2019**, *4*, 1433−1441.

(7) Miranda, K. C.; Bond, D. T.; McKee, M.; Skog, J.; Paunescu, T. G.; Da Silva, N.; Brown, D.; Russo, L. M. Nucleic Acids within Urinary Exosomes/MicrovesiclesAre Potential Biomarkers for Renal Disease. *Kidney Int.* **2010**, *78*, 191−199.

(8) Caplin, B.; Nitsch, D. Urinary Biomarkers of Tubular Injury in Chronic Kidney Disease. *Kidney Int.* **2017**, *91*, 21−23.

(9) Shurtleff, M. J.; Temoche-Diaz, M. M.; Schekman, R. Extracellular Vesiclesand Cancer: Caveat Lector. *Annu. Rev. Cancer Biol.* **2018**, *2*, 395−411.

(10) Tarasov, V. V.; Svistunov, A. A.; Chubarev, V. N.; Dostdar, S. A.; Sokolov, A. V.; Brzecka, A.; Sukocheva, O.; Neganova, M. E.; Klochkov, S. G.; Somasundaram, S. G.; Kirkland, C. E.; Aliev, G. Extracellular Vesicles in Cancer Nanomedicine. *Semin. Cancer Biol.* **2021**, *69*, 212−225.

(11) Yang, Q. S.; Cheng, L. M.; Hu, L.; Lou, D. D.; Zhang, T.; Li, J. Y.; Zhu, Q. F.; Liu, F. An Integrative Microfluidic Device for Isolation and Ultrasensitive Detection of Lung Cancer-Specific Exosomes from Patient Urine. *Biosens. Bioelectron.* **2020**, *163*, 112290.

(12) LeBleu, V. S.; Kalluri, R. Exosomes as a Multicomponent Biomarker Platform in Cancer. *Trends. Cancer.* **2020**, *6*, 767−774.

(13) Hoshino, A.; Kim, H. S.; Bojmar, L.; Gyan, K. E.; Cioffi, M.; Hernandez, J.; Zambirinis, C. P.; Rodrigues, G.; Molina, H.; Heissel, S.; et al. Extracellular Vesicle and Particle Biomarkers Define Multiple Human Cancers. *Cell* **2020**, *182*, 1044.

(14) Liu, C.; Zhao, J. X.; Tian, F.; Chang, J. Q.; Zhang, W.; Sun, J. S. λ-DNA-and Aptamer-Mediated Sorting and Analysis of ExtracellularVesicles. *J. Am. Chem. Soc.* **2019**, *141*, 3817−382.

(15) Huang, M. J.; Yang, J. J.; Wang, T.; Song, J.; Xia, J. L.; Wu, L. L.; Wang, W.; Wu, Q. Y.; Zhu, Z.; Song, Y. L.; Yang, C. Y. Homogeneous, Low-Volume, Efficient, and Sensitive Quantitation of Circulating Exosomal PD-L1 for Cancer Diagnosis and Immunotherapy Response Prediction. *Angew. Chem., Int. Ed.* **2020**, *59*, 4800−4805.

(16) Wang, Y. T.; Shi, T. J.; Srivastava, S.; Kagan, J.; Liu, T.; Rodland, K. D. Proteomic Analysis of Exosomes for Discovery of Protein Biomarkers for Prostate and Bladder Cancer. *Cancers* **2020**, *12*, 2335.

(17) Bryan, R. T.; Shimwell, N. J.; Wei, W.; Devall, A. J.; Pirrie, S. J.; James, N. D.; Zeegers, M. P.; Cheng, K. K.; Martin, A.; Ward, D. G. Urinary EpCAM in Urothelial Bladder Cancer Patients: Characterization and Evaluation of Biomarker Potential. *Br. J. Cancer* **2014**, *110*, 679−685.

(18) Wu, G.; Datar, R. H.; Hansen, K. M.; Thundat, T.; Cote, R. J.; Majumdar, A. Bioassay of Prostate-Specific Antigen (PSA) Using Microcantilevers. *Nat. Biotechnol.* **2001**, *19*, 856−860.

(19) Park, J.; Lin, H.-Y.; Assaker, J. P.; Jeong, S.; Huang, C.-H.; Kurdi, A.; Lee, K.; Fraser, K.; Min, C.; Eskandari, S.; Routray, S.; Tannous, B.; Abdi, R.; Riella, L.; Chandraker, A.; Castro, C. M.; Weissleder, R.; Lee, H.; Azzi, J. R. Integrated Kidney Exosome Analysis for theDetection of Kidney Transplant Rejection. *ACS Nano* **2017**, *11*, 11041−11046.

(20) Karpman, D.; Tontanahal, A. Extracellular Vesicles in Renal Inflammatory and Infectious Diseases. *Free Radical Biol. Med.* **2021**, *171*, 42−54.

(21) Qin, Z. Y.; Xu, Q. W.; Hu, H. H.; Yu, L. S.; Zeng, S. Extracellular Vesicles in Renal Cell Carcinoma: Multifaceted Roles and Potential Applications Identified by Experimental and Computational Methods. *Front. Oncol.* **2020**, *10*, 00724.

(22) Cao, Y.; Wang, Y.; Yu, X. M.; Jiang, X. J.; Li, G.; Zhao, J. Identification of Programmed Death Ligand-1 Positive Exosomes in Breast Cancer Based on DNA Amplification-Responsive Metal-Organic Frameworks. *Biosens. Bioelectron.* **2020**, *166*, 112452.

(23) Kim, H. J.; Park, S. W.; Jeong, I. G.; Song, S. H.; Jeong, Y.; Kim, C. S.; Lee, K. H. Noninvasive Precision Screening of Prostate Cancer by Urinary Multimarker Sensor and Artificial Intelligence Analysis. *ACS Nano* **2021**, *15*, 4054−4065.

(24) Poudineh, M.; Sargent, E. H.; Pantel, K.; Kelley, S. O. Profiling CirculatingTumour Cells and Other Biomarkers of Invasive Cancers. *Nat. Biomed. Eng.* **2018**, *2*, 72−84.

(25) Karpman, D.; Stahl, A. L.; Arvidsson, L. Extracellular Vesicles in Renal Disease. *Nat. Rev. Nephrol.* **2017**, *13*, 545−562.

(26) Lee, K.; Fraser, K.; Ghaddar, B.; Yang, K.; Kim, E.; Balaj, L.; Chiocca, E. A.; Breakefield, X. O.; Lee, H.; Weissleder, R. Multiplexed Profiling of Single ExtracellularVesicles. *ACS Nano* **2018**, *12*, 494−503.

(27) Eraslan, G.; Avsec, Ž.; Gagneur, J.; Theis, F. J. Deep Learning: New Computational Modelling Techniques for Genomics. *Nat. Rev. Genet.* **2019**, *20*, 389−403.

(28) Liu, C.; Zhao, J. X.; Tian, F.; Cai, L. L.; Zhang, W.; Feng, Q.; Chang, J. Q.; Wan, F. N.; Yang, Y. J.; Dai, B.; Cong, Y. J.; Ding, B. Q.; Sun, J. S.; Tan, W. H. Low-Cost Thermophoretic Profiling of Extracellular-Vesicle Surface Proteins for the Early Detection and Classification of Cancers. *Nat. Bio. Eng.* **2019**, *3*, 183−193.

(29) Ko, J.; Bhagwat, N.; Yee, S. S.; Ortiz, N.; Sahmoud, A.; Black, T.; Aiello, N. M.; McKenzie, L.; O'Hara, M.; Redlinger, C.; Romeo, J.; Carpenter, E. L.; Stanger, B. Z.; Issadore, D. Combining Machine Learning and Nanofluidic Technology to Diagnose Pancreatic Cancer Using Exosomes. *ACS Nano* **2017**, *11*, 11182−11193.

(30) Nicoliche, C. Y. N.; de Oliveira, R. A. G.; da Silva, G. S.; Ferreira, L. F.; Rodrigues, I. L.; Faria, R. C.; Fazzio, A.; Carrilho, E.; de Pontes, L. G.; Schleder, G. R.; Lima, R. S. Converging Multidimensional Sensor and Machine Learning toward High-Throughput and Biorecognition Element-FreeMultidetermination of Extracellular Vesicle Biomarkers. *ACS Sensor.* **2020**, *5*, 1864−1871.

(31) Shin, H.; Oh, S.; Hong, S.; Kang, M.; Kang, D.; Ji, Y.-g.; Choi, B. H.; Kang, K.-W.; Jeong, H.; Park, Y.; Hong, S.; Kim, H. K.; Choi, Y. Early-Stage Lung Cancer Diagnosis by Deep Learning-Based Spectroscopic Analysis of Circulating Exosomes. *ACS Nano* **2020**, *14*, 5435−5444.

(32) Zhang, J. L.; Shi, J. J.; Zhang, H. L.; Zhu, Y. F.; Liu, W.; Zhang, K. X.; Zhang, Z. Z. Localized Fluorescent Imaging of Multiple Proteins on IndividualExtracellular Vesicles Using Rolling Circle Amplification for CancerDiagnosis. *J. Extracell. Vesicles* **2020**, *10*, No. e12025.

(33) Yang, K. S.; Im, H.; Hong, S.; Pergolini, I.; del Castillo, A. F.; Wang, R.; Clardy, S.; Huang, C.-H.; Pille, C.; Ferrone, S.; Yang, R.; Castro, C. M.; Lee, H.; del Castillo, C. F.; Weissleder, R. Multiparametric Plasma EV Profiling Facilitates Diagnosis of Pancreatic Malignancy. *Sci. Transl. Med.* **2017**, *9*, 391.

(34) Tzelepis, C.; Mezaris, V.; Patras, I. Linear Maximum Margin Classifier for Learning from Uncertain Data. *IEEE. T. Patternanal.* **2018**, *40*, 2948−2962.

(35) Mary-Huard, T.; Robin, S. Tailored Aggregation for Classification. *IEEE. T. Patternanal.* **2009**, *31*, 2098−2105.

(36) Jayadeva; Khemchandani, R.; Chandra, S. Twin Support Vector Machines for Pattern Classification. *IEEE. T. Patternanal.* **2007**, *29*, 905−910.

(37) Liu, C.; Sun, J. S. AI in Measurement Science. *Annu. Rev. Anal. Chem.* **2021**, *14*, 1−19.

(38) Wu, N.; Wang, K.; Wang, Y. T.; Chen, M. L.; Chen, X. W.; Yang, T.; Wang, J. H. Three-Dimensional DNA Nanomachine Biosensor by IntegratingDNA Walker and Rolling Machine Cascade Amplification

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534

for Ultrasensitive Detection of Cancer-Related Gene. *Anal. Chem.* **2020**, *92*, 11111−11118.

https://doi.org/10.1021/acsnano.1c06429
ACS Nano 2021, 15, 19522−19534