The genetic source tracking of human urinary exosomes

Qingfu Zhu[a,1], Liming Cheng[b,1], Chunyu Deng[c,1], Liu Huang[d], Jiaoyuan Li[b], Yong Wang[e], Meng Li[e], Qinsi Yang[e], Xianjun Dong[f], Jianzhong Su[a,e], Luke P. Lee[g,h,i,j,2], and Fei Liu[a,e,2]

[a]Eye Hospital, School of Ophthalmology and Optometry, School of Biomedical Engineering, Wenzhou Medical University, Wenzhou 325035, China; [b]Department of Laboratory Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China; [c]School of Life Science and Technology, Harbin Institute of Technology, Harbin 150080, China; [d]Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China; [e]Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou 325001, China; [f]Genomics and Bioinformatics Hub, Precision Neurology Program, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115; [g]Renal Division and Division of Engineering in Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115; [h]Department of Bioengineering, University of California, Berkeley, CA 94720; [i]Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720; and [j]Institute of Quantum Biophysics, Department of Biophysics, Sungkyunkwan University, Suwon, Gyeonggi-do 16419, Korea

Edited by James A. Wells, University of California, San Francisco, CA, and approved August 30, 2021 (received for review May 20, 2021)

The genetic origins of nanoscale extracellular vesicles in our body fluids remains unclear. Here, we perform a tracking analysis of urinary exosomes via RNA sequencing, revealing that urine exosomes mostly express tissue-specific genes for the bladder and have close cell-genetic relationships to the endothelial cell, basal cell, monocyte, and dendritic cell. Tracking the differentially expressed genes of cancers and corresponding enrichment analysis show urine exosomes are intensively involved in immune activities, indicating that they may be harnessed as reliable biomarkers of noninvasive liquid biopsy in cancer genomic diagnostics and precision medicine.

exosome origin | RNA sequencing | cancer diagnosis | biomarker discovery

Exosomes (30 to 150 nm in size) are the smallest type of extracellular vesicles (~30 to 1,000 nm) secreted by living cells, acting as intercellular messengers via the circulation system of all body fluids such as blood, urine, and tears (1). Urinary exosomes have been proved to reflect pathophysiology and have provided a biological understanding of renal-related dysfunction, which represent a promising source of biomarkers for urological cancers (2, 3). However, the genetic networks between the components of urine exosomes and urinary/nonurinary organs or cells are still unclear. Thus, the source analysis of exosomal RNAs is essential for investigating exosome origins and applying exosomes to disease diagnostics and treatments.

Bladder and kidney cancers are the top two prevalent urological malignancies. Although noninvasive biomarkers such as circulating tumor cells and circulating tumor DNA have been reported in recent decades, their clinical applications for the early detection and monitoring of bladder and kidney cancers remain limited by their inability to identify tumors precisely (4). Therefore, novel biomarker discovery based on exosomes is of paramount importance for the noninvasive liquid biopsy–based diagnosis of bladder and kidney cancers.

Here, we performed a tracking analysis for investigating the origins of urinary exosomes at both tissue and cell levels through exosomal RNA sequencing (RNA-seq), intending to find genetic connections between urine exosomes and cancers. As a demonstration, we further explored exosomal RNA markers for the diagnosis of bladder and kidney cancers.

Results and Discussion

The high-purity exosomes were obtained by EXODUS that we have recently demonstrated (5). Fig. 1*A* shows the deconvoluted tissue-proportion analysis of exosome sources based on the expression of tissue-specific genes (TSGs) in the 116 individual urine samples. We found that urinary exosomes were mainly contributed by the bladder tissue, which might be attributed to the bladder being the storage of urine.

Taking cancer types into consideration, based on analysis of the single-sample gene set enrichment analysis (ssGSEA) scores (6), we observed that TSGs of the kidney and stomach were differentially expressed in bladder cancer and control samples (Fig. 1*B*). In contrast, TSGs of bladder, lung, brain, and liver were differentially expressed in kidney cancer samples and their controls (Fig. 1*C*). We then constructed a cell-type matrix and investigated the cell-level sources of exosomes (Fig. 1*D*). The endothelial cell and basal cell were the two major sources, possibly because they are closely related to bladder and urethra system. Notably, we noticed that the immune cells such as monocytes, dendritic cells, and B cells showed relatively high gene expression levels (Fig. 1*D*). This indicates the urinary exosomes were intensively involved in immune activities, which might offer a potential way to evaluate immune functionality through exosomes existing in human urine.

We next investigated the source of differentially expressed genes (DEGs) of cancers and the functionality of exosomal RNAs for cancer diagnosis. Initially, we used a single-cell mapper (scMappR) (7) to identify the cell type traced for differentiating cancer and control based on both exosomal RNAs and scRNA-seq datasets. We then assigned cell types contributed by the DEGs and determined the cell types with the highest cell-weighted fold (cwFold) change. Consequently, the clear gene signature sets were observed from the endothelial cell (endothelial-to-mesenchymal transition), neutrophil, proliferating T cell, and monocyte both for kidney cancer (Fig. 1*E*) and bladder cancer (Fig. 1*G*); and the corresponding most enriched pathways of the reranked DEGs are shown in Fig. 1 *F* and *H*. The data reveals that a certain number of DEGs of urinary exosomes are from immune cells and shows a common immune function at the pathway level, reflecting that the urine exosomes intensively participated in immune activities during cancer initiation and progression.

We then investigated whether the DEGs can be used as cancer biomarkers using a random forest model. We first searched for an optimal subset from the training cohort, which was then applied to the testing set, as described in a previous study (8).

Author contributions: L.P.L. and F.L. designed the research; Q.Z., L.H., Y.W., M.L., Q.Y., and F.L. performed the experiments; L.C. provided clinical samples; L.H. and J.L. collected clinical samples; Q.Z., L.C., C.D., X.D., J.S., L.P.L., and F.L. analyzed data; Q.Z., L.P.L., and F.L. wrote the paper; and all experiments were conducted under the supervision of L.P.L. and F.L.

The authors declare no competing interest.

[1]Q.Z., L.C., and C.D. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: lplee@bwh.harvard.edu or feiliu@wmu.edu.cn.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2108876118/-/DCSupplemental.

Published October 18, 2021.

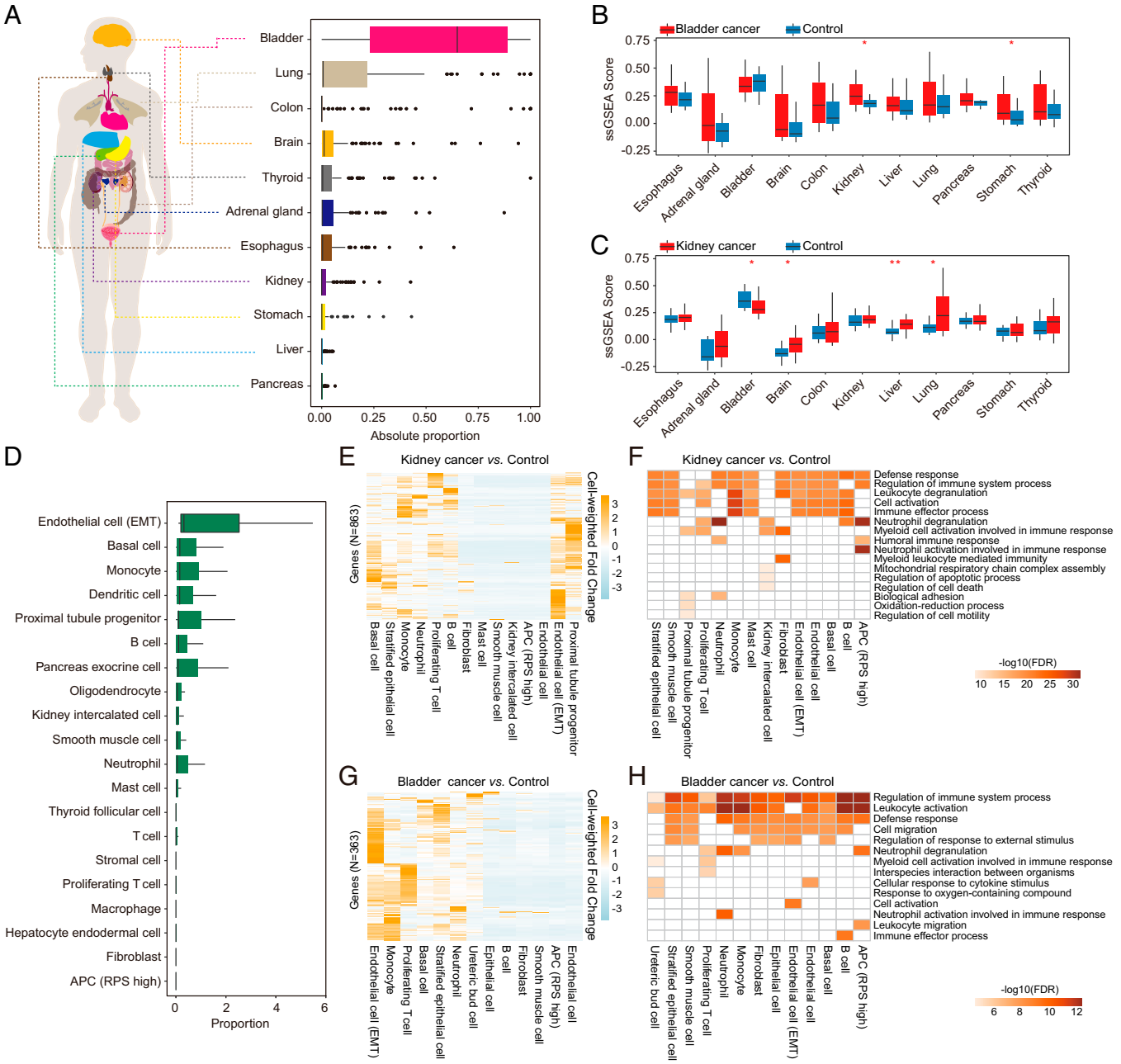

Fig. 1. The genetic source analysis of exosomes in human urine. (*A*) The deconvoluted tissue-proportion analysis of urinary exosome sources. (*B*) The differential expression of TSGs between bladder cancer and controls and (*C*) kidney cancer and controls. *P value < 0.05, **P value < 0.01 by unpaired two-sided Wilcoxon test. (*D*) Tracking the sources of urinary exosomes at the cell-type level. (*E*) Heatmap analysis of the traced cell types sorted by the normalized cwFold change of the DEGs of kidney cancer and bladder cancer (*G*). (*F*) The cell-type–specific pathway analysis of DEGs and the top five most enriched pathways were listed for kidney cancer and bladder cancer (*H*).

We finally selected 26 genes as a marker panel to predict kidney cancer and 19 genes to predict bladder cancer. The top 10 significant DEGs from the marker panels for kidney cancer and bladder cancer in the random forest classifier are shown in Fig. 2 *A* and *C*, respectively. The predicted probability of cancers based on the selected marker gene panels could clearly distinguish cancer samples from noncancer control samples derived from patients with other urinary diseases. The diagnostic receiver operating characteristic (ROC) curves showed excellent area under curve (AUC) values of 90.0% for kidney cancer (Fig. 2*B*) and 89.8% for bladder cancer (Fig. 2*D*), indicating the model's good prediction capability. It is worth mentioning that healthy individuals were not included in these experiments, making cancer diagnosis prediction even more challenging, considering the similar symptoms between cancers and noncancerous urinary diseases.

In the marker gene panels, *S100A10* (9) and *CCAR1* (10) in the kidney cancer marker panel and *CD248* (11) and *MT-ATP* (12) in the bladder cancer marker panel have been previously shown to be up-regulated and to possess a tumor-promoting function in urinary carcinomas. Notably, the expression level of *DDX17* showed a trend (i.e., *DDX17* expression in muscle-invasive bladder cancer > *DDX17* expression in nonmuscle-invasive bladder cancer > *DDX17* expression in control samples) that might be attributed to the fact that the human RNA helicase DDX17 contributes to tumor cell invasiveness by regulating the

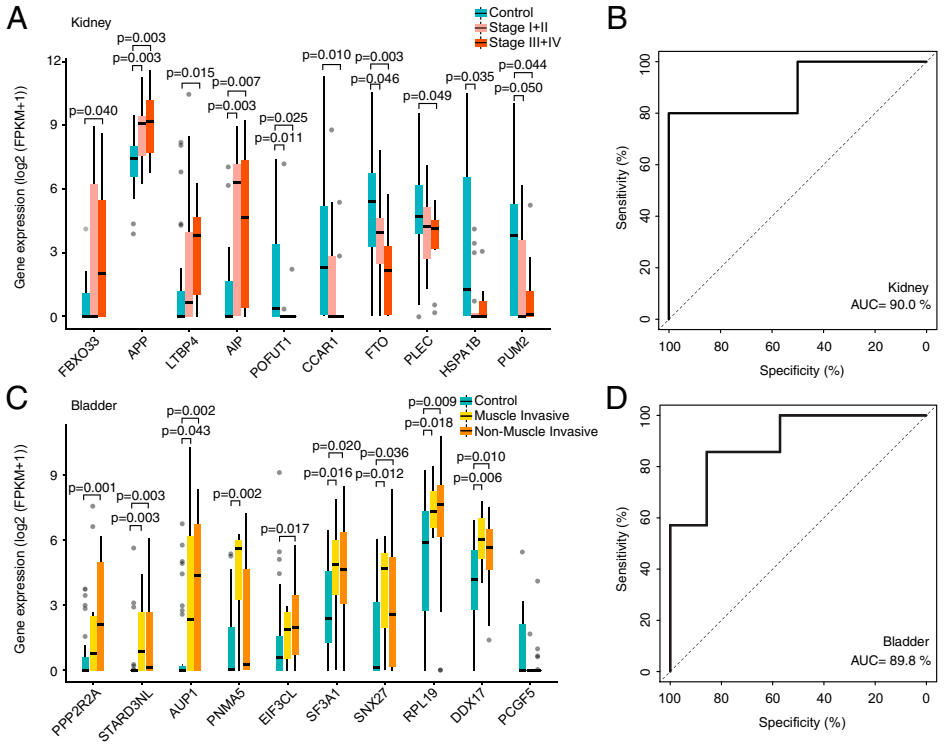

Fig. 2. Model performance based on the selected DEGs as cancer prediction markers. Top 10 significant DEGs in the marker panel for diagnosing kidney cancer (*A*) and bladder cancer (*C*), sorted by *P* value. The two-tailed Wilcoxon signed-rank test was used to determine significance. Construction of ROC curves indicates the predictive ability of the panels using the selected gene markers for kidney cancer (*B*) and bladder cancer (*D*) (kidney cancer, $n = 5$ and controls, $n = 6$; bladder cancer, $n = 7$ and controls, $n = 7$).

alternative splicing of several DNA and chromatin-binding factors (13). The combination of the genes in each panel can distinguish tumor samples from samples of related benign diseases with high accuracy, which may have guiding significance for the early diagnosis of these two urinary carcinomas.

In summary, we investigated the genetic sources of urinary exosomes both at the levels of organs and cells, showing that the bladder, endothelial cell, basal cell, monocyte, and dendritic cell may closely participate in the formation of urine exosomes. By tracking DEGs of urological cancers at cell levels and analyzing their enriched pathways, we reported that the urinary exosomes are intensively involved in immune activities in cancer development. Further biomarker investigation from exosomal RNAs resulted in two marker sets that could distinguish cancer from noncancer urinary diseases with AUC values > 89.8%. The exosome tracking analysis could provide a practical, noninvasive method for diagnosis and prognosis at the molecular level using human urine.

Materials and Methods

Clinical Samples. Participants were recruited with informed consent. The study was approved by the Institutional Review Board of Tongji Hospital in the Tongji Medical College at Huazhong University of Science and Technology (TJ-IRB20190914). For detailed clinical information, refer to Dataset S1.

Exosome Source Analysis. The R package GSVA (14) with default settings was applied for the ssGSEA. We used the absolute mode of CIBERSORT (https://cibersortx.stanford.edu/) to calculate the abundances of cell types and tissues, which generates a score (absolute proportion) in arbitrary units for comparison.

scMappR. scMappR combines DEGs, cell-type expression matrix, and bulk exosomal RNA expression to generate cwFold change. The cwFold change for each DEG is ordered across cell types to determine which cell types were most likely responsible for the differential expression.

Prediction Modeling. The kidney or bladder datasets were randomly divided into a training set and validation set at a ratio of 4:1 (training set *vs.* validation set). The marker gene panels refer to Datasets S5 and S6.

Statistical Analyses. For graphical representation and statistical analyses, the software Prism 5, Origin 2018, and Adobe Illustrator 2021 were used. We used the R package "cluster profile" for enrichment analysis and the R package (version 3.0.2) for ROC curve analysis.

Data Availability. All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. The work was primarily supported by research funds provided by the Zhenan Technology City Research Fund, the Zhejiang Provincial and Ministry of Health Research Fund for Medical Sciences (WKJ-ZJ-1910), the Wenzhou Medical University (89218012), and the Wenzhou Institute, University of Chinese Academy of Sciences (WIBEZD2017006-05).

1. D. Karpman, A. L. Ståhl, I. Arvidsson, Extracellular vesicles in renal disease. *Nat. Rev. Nephrol.* **13**, 545–562 (2017).
2. M. L. Merchant, I. M. Rood, J. K. J. Deegens, J. B. Klein, Isolation and characterization of urinary extracellular vesicles: Implications for biomarker discovery. *Nat. Rev. Nephrol.* **13**, 731–749 (2017).
3. J. J. Morrissey, A. N. London, J. Luo, E. D. Kharasch, Urinary biomarkers for the early diagnosis of kidney cancer. *Mayo Clin. Proc.* **85**, 413–421 (2010).
4. A. Di Meo, J. Bartlett, Y. Cheng, M. D. Pasic, G. M. Yousef, Liquid biopsy: A step forward towards precision medicine in urologic malignancies. *Mol. Cancer* **16**, 80 (2017).
5. Y. Chen *et al.*, Exosome detection via the ultrafast-isolation system: EXODUS. *Nat. Methods* **18**, 212–218 (2021).
6. D. A. Barbie *et al.*, Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108–112 (2009).
7. D. J. Sokolowski *et al.*, Single-cell mapper (scMappR): Using scRNA-seq to infer the cell-type specificities of differentially expressed genes. *NAR Genom. Bioinform.* **3**, lqab011 (2021).
8. Q. Feng *et al.*, Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat. Commun.* **6**, 6528 (2015).
9. T. Domoto *et al.*, Evaluation of S100A10, annexin II and B-FABP expression as markers for renal cell carcinoma. *Cancer Sci.* **98**, 77–82 (2007).
10. S. Y. Ha *et al.*, The overexpression of ccar1 in hepatocellular carcinoma associates with poor prognosis. *Cancer Res. Treat.* **48**, 1065–1073 (2016).
11. J. P. Volkmer *et al.*, Three differentiation states risk-stratify bladder cancer into distinct subtypes. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 2078–2083 (2012).
12. E. Reznik, Q. Wang, K. La, N. Schultz, C. Sander, Mitochondrial respiratory gene expression is suppressed in many cancers. *eLife* **6**, e21592 (2017).
13. E. Dardenne *et al.*, Splicing switch of an epigenetic regulator by RNA helicases promotes tumor-cell invasiveness. *Nat. Struct. Mol. Biol.* **19**, 1139–1146 (2012).
14. S. Hänzelmann, R. Castelo, J. Guinney, GSVA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).