Electronic Supplementary Material (ESI) for Lab on a Chip.

Supporting Information for

Integrated Nanoscale Deterministic Lateral Displacement Arrays for Separation of Extracellular Vesicles from Clinically-Relevant Volumes of Biological Samples

Joshua T. Smith,[*,†] *Benjamin H. Wunsch,*[†] *Navneet Dogra,*[†] *Mehmet E. Ahsen,*[‡] *Kayla Lee,*[†] *Kamlesh K. Yadav,*[ƨ] *Rachel Weil,*[ƨ] *Michael A. Pereria,*[†] *Jyotica V. Patel,*[†] *Elizabeth A. Duch,*[†] *John M. Papalia,*[†] *Michael F. Lofaro,*[†] *Mantu Gupta,*[ƨ] *Ashutosh K. Tewari,*[ƨ] *Carlos Cordon-Cardo,*[‡,§] *Gustavo Stolovitzky,*[†,‡] *and Stacey M. Gifford* [*,†]

[†] IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, United States.

[‡] Department of Genetics and Genomics Sciences, Icahn School of Medicine at Mount Sinai, New York 10029, United States.

[ƨ] Department of Urology, Icahn School of Medicine at Mount Sinai, New York 10029, United States.

[§] Department of Oncology Sciences and Pathology, Icahn School of Medicine at Mount Sinai, New York 10029, United States.

* Corresponding authors: joshsmit@us.ibm.com, smgifford@us.ibm.com

Table of contents

1. Experimental section

1.1 Integrated nanoDLD chip fabrication (Scheme S1)

Two types of a chips were fabricated for the purpose of this study, referred to as Design I (22 mm x 30 mm) and Design II (26 mm x 33 mm). Design I chips contain 1,024 arrays (1.44 billion pillars per chip), each having W = 150 μm and L = 1.5 mm dimensions with collection walls separating them, and Design II chips have 3,840 arrays (2.4 billion pillars per chip), each with W = 100 μm, L = 1 mm, and being conjoined to its neighboring arrays (no collection walls separating them). Despite these design differences, the fabrication process flow for both Design I and II chips is identical.

Pillar arrays and other shallower microfluidic features (filters, pre-array loading features, outlet microchannels, etc.) were defined using 193-nm lithography and reactive-ion etching (RIE) (Scheme S1, step 1). Prior to resist coating, 200-mm silicon wafers were RCA cleaned and a 300 nm-thick SiO_2 layer was grown on the surface to provide a hard mask (HM) for pillar array definition during RIE processing. Next, a tri-layer stack was spin coated onto the wafers to prepare them for lithographic exposure, which included a 500 nm HM8006 organic planarization layer (OPL) (JSR Microelectronics), a 70 nm SHBA-940-L35 Si-containing anti-reflective coating (Si ARC) (Shin-Etsu MicroSi, Inc., AZ), and finally a 126 nm-thick negative tone resist AIM7946-126 (JSR Microelectronics). A 0.75 NA 193-nm wavelength step-and-scan PAS 5500/1100 B-4X scanner system (ASML) was used to expose the resist layer at dose and focus ranges of 16 – 30 mJ and -0.1 to 0 μm, respectively, to vary G as desired. Exposure within this process window led to G = 102 – 202 nm features as defined in the AIM7946-126 resist (Figure S3). After developing the exposed wafers in NBA FN-DP001 developer (Fuji), downstream RIE was used to selectively transfer the exposed pattern into silicon. Pillar definition was carried out in a DPSII ICP etch chamber (Applied Materials, CA) using a 5-step process to etch pillars to a depth of ~1 μm (Figure 1f). In step (1), Si ARC breakthrough was achieved using an CF_4/CHF_3 chemistry at 500 W source power, 100 W bias power and 30 mTorr pressure at 65 °C. For step (2), an $N_2/O_2/Ar/C_2H_4$ chemistry at 400 W source power, 100W bias power, and 4 mTorr pressure at 65 °C was then applied to break through the OPL. Next, in step (3), the pattern was defined in the SiO_2 HM using the same conditions and chemistry applied during the step (1) Si ARC etch. Steps (1) – (3) utilized end-point (EP) detection as a quality control (QC) affirmation of material breakthrough with etch times adjusted around EP for each wafer to ensure reproducibility. After HM patterning, the OPL

carbon resist was stripped at step (4) in an Axiom downstream asher (Applied Materials, CA) using an O_2/N_2 chemistry. Finally, step (5) involved transferring the HM pattern into silicon. This was accomplished using the DPS II ICP etch system (Applied Materials, CA) by applying a 5 sec CF_4/C_2H_4 native oxide open RIE followed by a $Cl_2/HBr/CF_4/He/O_2/C_2H_4$ main etch using 650 W source power, 85W bias power, and 4 mTorr pressure at 65 °C to etch all features to a depth of ~ 1 µm. After RIE etching to define the integrated nanoDLD pillar arrays, residual organics were removed in a bath containing a 10:1 mixture of sulfuric/nitric acid at 150 °C, after which the HM was removed completely in a 10:1 dilute hydrofluoric (DHF) acid bath.

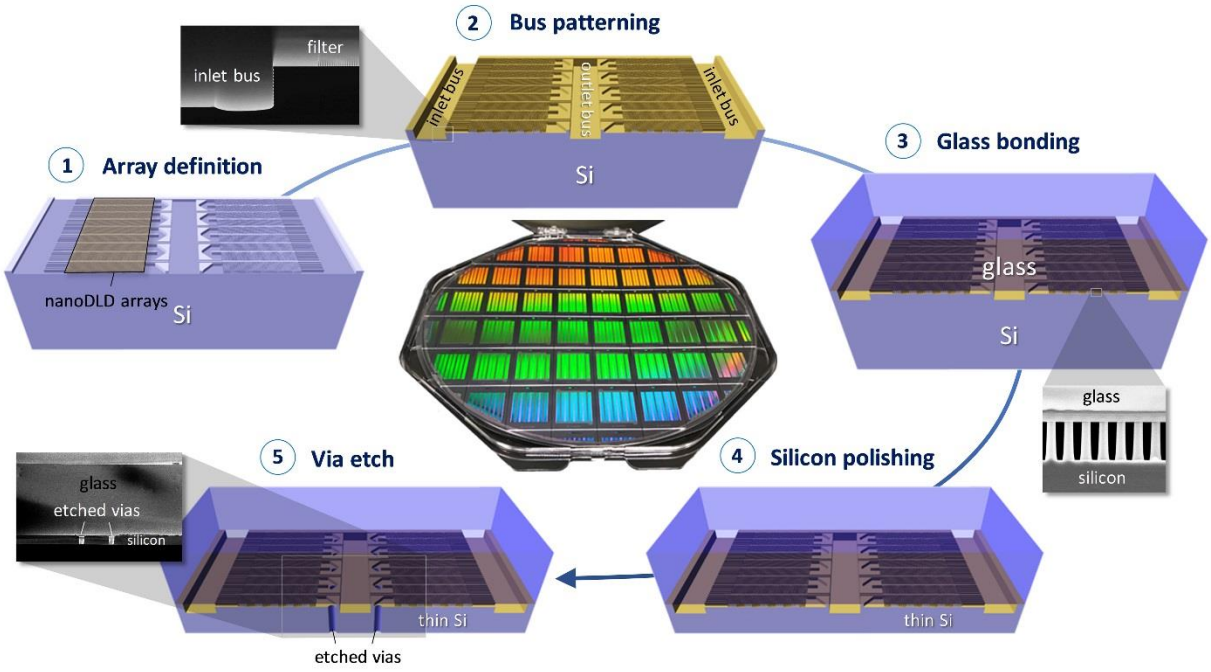
As shown in step 2 of Scheme S1, a second lithography layer was used to etch the deeper bus network. For this task, an optical contact MA8 mask aligner (Karl Suss, Germany) was used to lithographically expose the inlet / outlet bus network features with an energy of 24 mJ in a 10 µm-thick AZ4620 (AZ Electronic Materials) resist that was spin coated on top of a hexamethyldisilazane (HMDS) adhesion promoter and subsequently developed with an AZ4620 developer (AZ Electronic Materials). Once defined in resist, bus features were transferred into Si by an Alcatel 601E inductively coupled plasma etcher (Alcatel Micro Machining Systems, France) using a Bosch process with alternating pulses of SF_6 300 sccm and C_4F_8 150 sccm at a temperature of 20 °C with source and bias powers of 1800 W and 80 W, respectively. Design I and II chip bus networks were etched to depth targets of ~10 µm and ~20 µm, respectively, to ensure sufficient fluidic conductivity during operation. The post-etched resist was removed in AZ300T Photoresist Stripper (AZ Electronics Materials) at 60 °C, and the wafers were cleaned in a 3-step process, including a 10 min piranha clean (5:1 volume ratio of 98% H_2SO_4 and 30% H_2O_2) to remove organic residue, 60 sec 100:1 DHF dip to remove any native oxide formed, and an SC1 clean (1:1:5 H_2O_2:NH_4OH:deionized water (DI)) at 65 °C with intermittent rinse steps. A subsequent RCA clean prepared wafers for a thin thermal oxide growth of 5 - 50 nm of SiO_2, which simultaneously prepared the substrate for subsequent anodic bonding and fine-tuned *G* to the final target. While a continuum of gap sizes ranging from G = 80 - 225 nm is on offer through this tuning process, devices with G = 80, 150, and 225 nm were prepared for this study (see Figure 2e).

Step 3 of Scheme S1 illustrates the next major process of anodically bonding 200-mm diameter glass wafers to the thin SiO_2 layer grown on the silicon wafers. Borosilicate glass wafers (Swift Glass, NY) with a measured thickness of ~700 µm were SC1 cleaned (1:1:5 H_2O_2:NH_4OH:DI) for 10 min at 65 °C and spin-rinse-dried (SRD) to prepare them for bonding to the silicon wafers. An

EVG 301 Wafer Cleaning System (EV Group, Austria) dispensing DI water through a 1 MHz megasonic nozzle further cleaned the bonding surfaces of both the silicon and glass wafers just prior to bonding them, a process followed by a spin dry to remove any water on the surface. Anodic bonding was carried out in an EVG 520IS 200 mm Semi-Automated Wafer Bonding System (EV Group, Austria) in a 900 mbar N_2 environment at 260 °C with a voltage of 600 V and down force = 3200 N applied to the wafer stack.

After monolithically bonding the glass and silicon wafers, the silicon was thinned and polished from the unbonded side to a thickness of ~100 μm in a 2-stage chemical mechanical polish (CMP) process (see Scheme S1, step 4). A course silicon grind was accomplished using a DAG810 Automatic In-Feed Surface Grinder (DISCO Corporation, Japan) to thin bulk silicon, removing ~575 μm of the nominally 725 μm-thick silicon wafer followed by a post grind cleaning in DCS1440 Disco Cleaning System (DISCO Corporation, Japan), and a final 50 μm polish in an IPEC-Westech 372M (Axus Technology, AZ) was utilized for post-grind CMP to meet final thickness and surface quality targets by applying Nalco 2358, a colloidal silica abrasive slurry, to polish the silicon to a ~100 μm final thickness and a mirror-like surface finish. A brush clean with oxalic acid and rinse was used to clean the wafers after polish as well as a downstream SRD to ensure the removal of slurry particles.

Step 5 of Scheme S1 shows the final process of creating open through-silicon vias (TSVs) that constitute the sample inlet, bump outlet, and zigzag outlet bank vias, or fluidic access points for routing fluid to and from the integrated nanoDLD chips. A 3 μm-thick layer of TOKip 3250-27 cp resist (Tokyo) was applied to the polished silicon and an optical contact MA8 mask aligner (Karl Suss, Germany) with front-to-back alignment was used to pattern the via positions. The open pattern features were used to etch the vias with a deep silicon RIE process. Deep silicon etching was achieved with an Alcatel 601E inductively coupled plasma etcher (Alcatel Micro Machining Systems, France) using a Bosch process with alternating pulses of SF_6 300 sccm and C_4F_8 150 sccm at a temperature of 20 °C with source and bias powers of 1800 W and 80 W, respectively. Via breakthrough was verified visually with backlighting and with an optical microscope. An O_2 ash system (Plasma-Therm, FL) was used to remove resist from the silicon wafer after RIE. Wafers were subsequently diced while attached to a high-tack dicing film to prevent liquid from wetting the chips before use. Design I resulted in 32 usable chips per wafer whereas the larger Design II yielded a maximum of 20 usable chips per wafer.



Scheme 1. Process flow for integrated nanoDLD chips on 200 mm wafers. (1) Wafers are step printed using 193nm lithography and reactive-ion etched to a depth of ~1 μm to fabricate parallel nanoDLD pillar array features. (2) Bus network is patterned with photoresist and deep etched to a depth of ~10 μm, producing a low-resistance fluidic distribution network. The etched features are RCA cleaned and a thin oxide is grown over the pillar arrays and bus features, preparing the surface for glass bonding. (3) 700 μm-thick glass is anodically bonded to the silicon wafers that fluidically isolates and protects the patterned features from downstream processing. (4) Chemical mechanical polishing (CMP) is applied to thin and polish the nominally 725 μm-thick silicon wafer to a thickness of ~100 μm. (5) Back side wafer patterning using front-to-back alignment is used with subsequent deep silicon etching to form vias that act as fluidic access points to the bus network and zigzag outlet features on the front side of the wafer. The center photo shows a fully fabricated and diced 200 mm wafer, which produces 32, ready-to-use, integrated nanoDLD chips.

1.2 Chip and flow cell protocol

Fabricated and diced integrated nanoDLD chips with gap sizes G = 80, 150 and 225 nm needed to be wetted and primed before running any sample type. Chips were prepared in batches of 6-8 by standing them upright within a slotted, custom quartz glass holder and submerged in 300 mL of DI water (Millipore) in a 500 mL beaker. The beaker was sealed with aluminum foil and autoclaved (Tuttnauer, 2540 E-B/L, Heidolph Brinkmann) at 130°C for 90 min. By following this process, it was found that chips are fully wet without the presence of bubbles in the fluidic bus network, nanoDLD arrays, or TSVs. Chips were stored in the DI water until use, typically within a week.

Individual chips were run in custom-built flow cells, constructed from clear acrylic (Figure S1). The cell consists of a top plate with a viewing window for *in situ* microscopy and a base plate with a chip pocket, an inlet port, and outlet ports for the sorted (bump) and unsorted (zigzag) fluids. The base pocket contains an o-ring seat that conforms to the perimeter of the backside of the chip to create a fluidic seal. A segment of 1 mm-wide o-ring cord is fitted within the seat with the ends joined together with cyano acrylic glue. The volume defined by the chip / o-ring cord / base pocket is termed the *drain space* and is where the unsorted (zigzag) fluid collects. Separate o-rings with a 3 mm outer diameter (OD) and 1 mm inner diameter (ID) isolate the sample inlet and bump outlet via fluid streams from the drain space. Flow cells are cleaned in 15% v/v hydrogen peroxide (H_2O_2) solution for 10 min, then soaked in DI water for 10 min. A syringe is used to pump fluid through the inlet port to clear any accumulated material. To load, cleaned flow cells are submerged in fresh DI water and the nanoDLD chip is quickly transferred from the beaker to the submerged flow cell to prevent it from drying out. Once the chip is securely sealed within the flow cell using hex-nut screws, the flow cell is removed from the water and dried off with compressed air. This process leaves the drain space filled completely with DI water.

1.3 Characterization of flow rates in integrated nanoDLD chips

For testing the fluidic conductance, a F-100 microtight connector (Upchurch Scientific) with a 1/16” OD FEP, ~30 cm long tube (Upchurch Scientific) was fitted to the inlet port on the flow cell. The tubing was filled with DI water and attached to a Qmix syringe pump (Cetoni GmBH) with an in-line pressure sensor module. Care was taken not to introduce bubbles within the fluid line. The Qmix pump was then set at a fixed pressure with the injection rate controlled by feedback,

and the dispensed volume was recorded as a function of time. The dispense rate was equated to the injected flow rate and recorded for each pressure setting. Stepping the pressure values discretely from low to high pressure, the system was run for 10 min at steady state for each pressure value tested. The reported applied pressure P_{app} (Figure 2h) is modified to account for offset error in the pressure sensor.

1.4 Running fluorescent beads to calibrate deflection efficiency

For running fluorescent beads, it is necessary to exchange the DI water in the integrated nanoDLD chips with a buffer solution to prevent precipitation. A buffer of 2% TWEEN-20 (Sigma Aldrich) in 1X Tris-EDTA buffer (Sigma Aldrich), which has been 0.02 μm filtered (Whatman), was used in this capacity. A 1 mL syringe was prepared with the buffer solution and attached to the tubing via a Luer Lok adapter (Upchurch Scientific). The inlet has a second port to allow the purge and removal of any bubbles that may exist in the line when loading buffer. Buffer solution is pushed through the tubing and inlet port until it exits from this second, 'burp' port, after which the secondary port is sealed with an o-ring and screw. The seal on the inlet port was tested manually by pushing the syringe to make sure no liquid was leaking from the flow cell. The excess DI water was removed from the bump and zigzag outlet ports (note this does not remove the fluid from the drain space). The buffer syringe was set on a Pump 11 Pico Plus Elite syringe pump (Harvard Apparatus) and run at a fixed injection rate corresponding to ~ 7.8 bar inlet pressure: 1 μm·min^{-1} for G = 80 nm, 4.9 μm·min^{-1} for G = 150 nm, 12.4 μm·min^{-1} for G = 225 nm. For the first 5 min, the bump outlet is sealed with a micro-plug (Upchurch Scientific) to force all fluid through the drain TSVs, then after 5 min the bump outlet is left unsealed so fluid flows through both the bump and zigzag outlets. For G = 80 nm chips, the run times were doubled to collect sufficient fluid for downstream analysis.

The following fluorescent beads were obtained from Thermo Fischer: FluoSpheres carboxylate-modified, 580 nm excitation / 605 nm emission, 0.2 μm (0.22 μm actual), FluoSpheres, carboxylate-modified, 350 nm excitation / 440 nm emission, 0.1 μm (0.097 μm actual), Fluoro-Max, 468 nm excitation / 508 nm emission, 0.075 μm (0.075 μm actual), Fluoro-Max, 468 nm excitation / 508 nm emission, 0.050 μm (0.047 μm actual), FluoSpheres carboxylate, 505 nm excitation / 515 nm emission, 0.020 μm (0.025 μm actual). Bead solutions were prepared by diluting as-purchased stocks in 2% TWEEN-20 in 1X TE buffer which has been 0.02 μm filtered,

in 1:100 dilutions, 1-2 mL total volume. A 1 mL syringe with the bead sample was prepared and attached to the inlet tubing. The tubing was purged until sample fluid exited from the secondary port, or 'burp' port, and then the port is sealed. The flow cell mounts on a Scope.A1 epi-fluorescent microscope (Zeiss) with a 10x N-Achroplan objective (Zeiss). Fluorescence was measured using light emitting diode sources and dichroic mirror sets (Zeiss) at 365/445 nm, 470/510 nm, and 560/610 nm excitation/emission. Microscope images and videos were captured with an Andor iXon Ultra 897 EMCCD camera (Oxford Instruments). See the microscope setup in Figure S4. Bead fluid was injected at the rates given above. Steady state flow was typically reached within 60 s. Captured images of the array inlet and outlet, including the junction between the zigzag outlet vias, or drain TSVs, and bump outlet channel, were further analyzed to determine the deflection angle, θ, as a function of bead diameter, D_p, and gap size, G. Image analysis was carried out using an in-house developed python script. For each bead experiment, the intensity profile at the array outlet was acquired. The deflection angle was taken as the inflection point where the intensity profile drops to the background. The inflection point was determined from the first derivative of the intensity profile, using a Savitzky-Golay filter (SciPy Community) to smooth the curve (window = 11, polynomial = 10). For all devices, the geometric angle (maximum angle) $\theta_{max} = 5.7°$. For each chip, 12 arrays were measured and the average reported as $P = \theta / \theta_{max}$ (see Figure 3g). Display images were generated using ImageJ (National Institute of Health) with false color merging of the bright field and fluorescent channels. The contrast was adjusted to enhance the fluorescent channel clarity.

1.5 Biological sample preparation for nanoDLD

Human serum (AB, male, USA origin, sterile-filtered, Sigma Aldrich, H4255) was thawed, and aliquoted into 1 mL samples for long-term storage at -80 °C. Just prior to use, 1 mL samples were rapidly thawed and labeled with 10 μL 100X SYBRgold (ThermoFisher, S11494). Samples were diluted to 5 mL total volume with 1X PBS and 0.22 μm-filtered with a cellulose acetate syringe-top filter (Whatman). Samples were stored at 4 °C prior to nanoDLD isolation for no longer than 1 week.

Human urine was collected from donor patients and pooled prior to being aliquoted into 1 mL samples for long-term storage at -80 °C. Just prior to use, 1 mL samples were rapidly thawed and labeled with 10 μL 100X SYBRgold (ThermoFisher, S11494). Samples were 0.22 μm-filtered

with a cellulose acetate syringe-top filter (Whatman). Samples were stored at 4 °C prior to nanoDLD isolation for no longer than 1 week.

1.6 Biological Sample Run Protocol

Before running biological samples, the integrated nanoDLD chips were first primed with bovine serum albumin, BSA. A F-100 microtight connector with a 1/16" OD, ~2-4 cm long tube was fitted to the inlet port on the flow cell for this priming step. Additionally, a 1 mL syringe was prepared with a solution of 5% w/v BSA, (Sigma Aldrich) in a 1X phosphate buffer saline (Sigma Aldrich), the solution having been previously 0.02 μm filtered (Whatman). The syringe was attached to the tubing via a Luer Lock adapter. The excess DI water was then removed from the bump and zigzag outlet ports. The BSA syringe was set on the syringe pump and run at a fixed injection rate of 8 μm·min^{-1} for G = 225 nm (P_{app} ~ 5 bar). For the first 5 min, the bump outlet is sealed with a microplug to ensure all drain TSVs are cleared of residual bubbles, then after 5 min the bump is left unsealed so fluid flows through both ports.

After priming with BSA, a 1 mL syringe with the sample fluid (human serum or urine) was attached to the inlet tubing. For the G = 225 nm devices used, samples were 0.2 μm filtered. The tubing was purged, as in the bead case, until sample fluid exited the secondary port. The outlet ports were then cleared of any collected fluid, and the chip was allowed to run at 8 μL·min^{-1}. At this point, the sample run time clock was initiated, allowing sample to accumulate for 60 min. For the zigzag fluid, a F-100 microtight was fitted to the zigzag outlet port with a short piece of tubing (~5 cm), the end of which was inserted into a 1.5 mL, tared centrifuge tube sealed with parafilm to collect zigzag fluid as it exited the device. Zigzag fluid was allowed to continuously collect for the entire run time. For the bump outlet, fluid was removed every 30 min manually, using a pipet, and collected in a separate, tared tube. Samples were weighed to determine the volume collected from the bump and zigzag reservoirs and then stored in a 4 °C fridge for analysis, typically within 24 hr.

Ultracentrifugation: Human serum (AB, male, USA origin, sterile-filtered, Sigma Aldrich, H4255) was thawed, and aliquoted into 5 mL samples for long-term storage at -80 °C. Just prior to use, serum was rapidly thawed, diluted to 12 mL with 1X PBS, and 0.22 μm-filtered with a cellulose acetate syringe-top filter (Whatman). Filtered serum was centrifuged in a benchtop ST-

16R Sorvall centrifuge in 15 mL conicals at 2000 x g at 4 °C for 30 min. Supernatant was centrifuged in a Thermo Fisher tabletop ultracentrifuge in polycarbonate tubes and S150AT fixed angle rotor at 12,000 x g at 4 °C for 45 minutes. Supernatant was ultracentrifuged in S150AT fixed angle rotor at 110,000 x g at 4 °C for 70 min. Pellets were resuspended in 1 mL cold PBS, pooled, and diluted to a total volume of 8 mL with PBS. Samples were repelleted in S150AT fixed angle rotor at 110,000 x g at 4 °C for 70 min. Pellet was resuspended in 600 µL cold PBS. 100 µL was reserved for analysis of UC isolation alone and the remaining 500 µL was processed using density gradient centrifugation as described below.

Human urine was collected from donor patients and pooled prior to being aliquoted into 20 mL samples for long-term storage at -80 °C. Just prior to use, urine was rapidly thawed, and centrifuged in a benchtop ST-16R Sorvall centrifuge in 15 mL conicals at 300 x g at 4 °C for 10 min. Supernatants were 0.22 µm-filtered with a cellulose acetate syringe-top filter (Whatman) and centrifuged in ST-16R Sorvall centrifuge in 50 mL conicals at 2000 x g at 4 °C for 20 min. Supernatant was centrifuged in a Beckman Coutler ultracentrifuge Type 45-Ti fixed angle rotor in polycarbonate tubes at 10,000 x g at 4 °C for 30 min. Supernatant was ultracentrifuged in Type 45-Ti fixed angle rotor at 100,000 x g at 4 °C for 70 min. Pellets were resuspended in 1 mL cold PBS, pooled, and diluted to a total volume of 50 mL with PBS. Samples were repelleted in Type 45-Ti fixed angle rotor at 100,000 x g at 4 °C for 60 min. Pellet was resuspended in 600 µL cold PBS. 100 µL was reserved for analysis of UC isolation alone and the remaining 500 µL was processed using density gradient centrifugation as described below.

Density gradient centrifugation: Stocks of 40, 20, 10 and 5% Optiprep (v/v) in 0.25 M sucrose, 10 mM Tris pH 7.5 were chilled at 4 °C and used to create a 11.5 mL discontinuous gradient in SW 40 Ti swinging bucket rotor tubes. 0.5 mL of a resuspended UC sample from urine or serum was applied to the top. Sample was centrifuged in a Beckman Coulter ultracentrifuge in a SW 40 Ti rotor for 16 hours at 100,000 x *g* at 4 °C. 1 mL fractions were collected starting at the top of the tube and analyzed by NTA for particle concentration. The three most concentrated fractions were pooled and diluted to 8 mL with PBS. Pool was centrifuged in a Thermo Fisher tabletop ultracentrifuge in polycarbonate tubes and S150AT fixed angle rotor at 100,000 x g at 4 °C for 2 hours. Pellets were resuspended in 200 µL cold PBS.

Size exclusion chromatography: Human serum (AB, male, USA origin, sterile-filtered, Sigma Aldrich, H4255) or urine (collected from donor patients and pooled) was thawed, and aliquoted into 1 mL samples for long-term storage at -80 °C. Just prior to use, serum or urine was rapidly thawed and 500 µL centrifuged at 14,000 rpm in a microcentrifuge for 20 min at room temperature. A qEV original (Izon Science) column was equilibrated with 10 mL 0.02 µm-filtered (Anotop) 1X PBS. With the bottom capped, the top layer of buffer was removed and 500 µL sample applied. The first 3 mL was collected as void volume and 12 x 500 µL fractions were collected. These fractions were analyzed with NTA for particle concentration and highest fractions pooled (generally fractions 7-9). The pools were concentrated to 150 µL in a 50K MWCO spin concentrator (Corning, Spin-X UF).

QIAGEN ExoEasy Maxi Kit: Human serum (AB, male, USA origin, sterile-filtered, Sigma Aldrich, H4255) was thawed, and aliquoted into 4 mL samples for long-term storage at -80 °C. Human urine was collected from donor patients and pooled prior to being aliquoted into 16 mL samples for long-term storage at -80 °C. Just prior to use, serum or urine was rapidly thawed and 0.8 µm-filtered with a cellulose acetate syringe-top filter (Whatman). All subsequent steps were performed at room temperature according to the manufacturer's instructions and eluted in 400 µL elution buffer into a clean collection tube.

Nanoparticle tracking analysis: Samples were diluted to approximately 10^6-10^7 particles/mL in Millipore DI water. Particle concentration, size, and zeta potential were measured using ZetaView (Particle Metrix). Particle size and concentration were measured using the built-in EMV protocol and zeta potential was measured using the built-in EMV Zeta protocol.

Enzyme-linked immunosorbent assay (ELISA): 50 µL of each EV sample per well was adsorbed to a high-adsorption 96-well microtiter plate (Greiner) overnight at 37 °C in triplicate. PBS was used as a control. Wells were washed 3 times with 200 µL PBST (PBS + 0.05% Tween-20) and blocked with 200 µL blocking buffer (PBST + 6% nonfat dried milk) for 1 hour at room temperature. Wells were washed three times with 200 µL PBST. Wells were incubated with 50 µL of primary antibody (TSG101 mouse monoclonal, ThermoFisher, MA1-23296, 1.5 ug/mL and calnexin goat polyclonal, Thermo Fisher PA5-191-69, 2 ug/mL) in blocking buffer for 1 hour at room temperature. Wells were washed three times with 200 µL PBST. Incubate with secondary

horseradish peroxidase (HRP)-conjugated exosomes (goat anti-mouse-HRP, Abcam, ab6789 and donkey anti-goat-HRP, Abcam, ab97110) at a dilution of 1:50,000 in blocking buffer for 1 hour at room temperature. Wells were washed three times with 200 µL PBST. Wells were then incubated with 100 µL SuperSignal ELISA femto solution (Thermo Fisher, 37075) and read using a Spectramax i3x Multi-Mode Microplate Reader (Molecular Devices) with 1 min pre-mixing and luminescence reading at 425 nm.

Cryo-electron microscopy: Samples were shipped overnight on ice for analysis by cryo-electron microscopy at the John M. Crowley Center for High Resolution Electron Microscopy at Arizona State University. Exosomes were applied to holey carbon films, vitrified, and imaged by TEM. Specifically, C-Flat 1.2/1.3 TEM grids (Protochips, Research Triangle, NC) were glow discharged for 20 seconds at 25 mA to create a hydrophilic surface to receive the 3 µL of exosome solution. A thin film was formed by wicking away the excess liquid with Whatman #1 filter paper. This thin vicinal film containing exosomes was rapidly immersed into liquid ethane cooled to -180° C by liquid nitrogen. These thin films were examined cryogenically at -178° C in a FEI (Hillsboro, OR) Titan Krios microscope operating at 300 KeV at magnifications indicated in the figure legends. Images were recorded on a Gatan (Warrendale, PA) K2 direct electron detector or hybrid CMOS FEI CETA camera.

1.7 RNA extraction, library preparation, and next-generation sequencing

EVs were isolated from 0.5 mL serum collected from high-grade prostate cancer patients by nanoDLD. Total RNA was extracted from the bump fraction using the Total Exosome and Protein Isolation Kit (Invitrogen 4478545). 50 µL serum from bump fraction was resuspended in 50 µL ice cold exosome resuspension buffer. 2X denaturing solution was added to the final exosomes solution on ice. Equal volume of acid-phenol:chloroform solution was added to each sample. The final solution was vortexed for 60 seconds and centrifuged at 10,000 x g. The top aqueous phase was carefully isolated without disturbing the lower organic phase. The top aqueous phase was transferred to the provided filter cartridge in collection tubes. Bound RNA was washed 3 times using included wash solution. Finally, preheated elution solution was used to elute the RNA in 100 µL. RNA was stored at -20 °C until further use and RNA quality was assessed by bioanalyzer (Agilent 2100 Bioanalyzer, RNA 6000 Pico Kit, Agilent Technologies).

cDNA Libraries were prepared for small RNAs using the SMARTer smRNA-seq Kit for Illumina (Takara Bio 635030). A total of 18 cycles of PCR we carried out to obtain a good yield of libraries. Final library quality was verified with Qbit and bioanalyzer. Negative (no RNA) and positive controls provided expected results.

Next-generation RNA sequencing was performed using a HiSeq 4000 (Illumina), 100 base pair, single end reads at the New York Genome Center. For quantification of gene expression, raw reads were aligned to GENCODE v26 (GRCh38) using a STAR aligner (version 2.5.4b).[1] STAR was also used to map the aligned reads to the GENCODE v26 primary gene annotation, which includes transcripts corresponding to lncRNA, miRNA as well as protein-coding RNA. To reduce the noise, only reads mapping to a single location in quantification of gene expression were used. miRge 2.0 was utilized for quantification of the fraction of different RNA species.[2] Comparison of correlation coefficients was performed using the R package cocor.[3]

2. Supplementary figures and tables

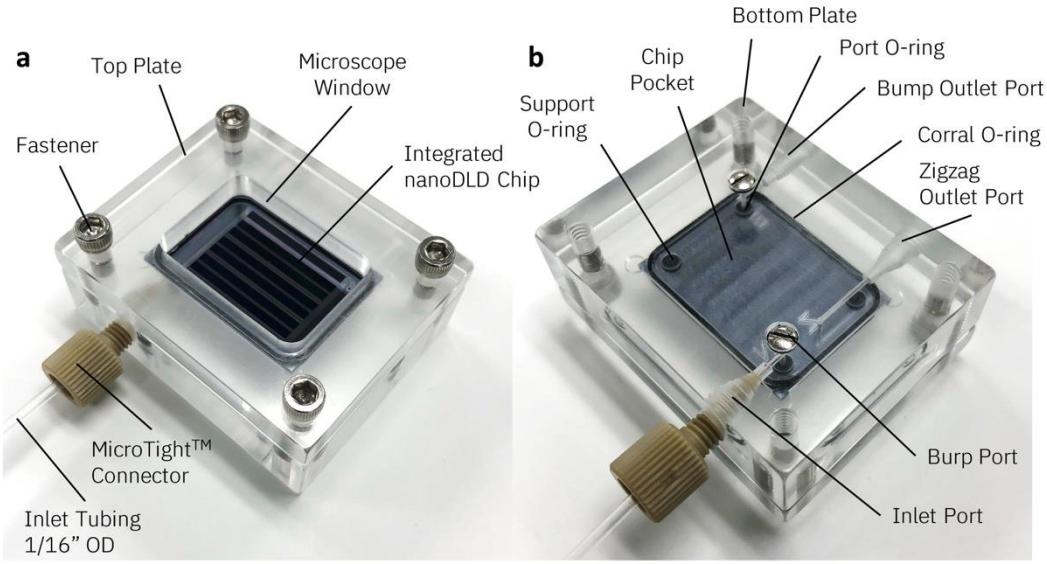


Figure S1. Overview of flowcell used to operate integrated nanoDLD chips. The flowcell consists of (a) top and (b) bottom plates, machined from clear acrylic plastic. A chip pocket in the bottom plate is machined with the necessary inlet and outlet ports, as well as o-ring seats. Four, symmetric o-ring seats (accepting o-rings size 001 ½) are positioned to form seals with the inlet and bump outlet TSVs. A thin trench is machined around the chip pocket's perimeter to fit a 1 mm OD cord stock. This forms a corral o-ring that seals the backside of the chip. The small gap formed between the pocket floor and chip compressing the cord stock defines the drain space volume into which the zigzag fluid drains. A small burp port is provided at the inlet to remove excess bubbles from the sample line prior to fluid injection.

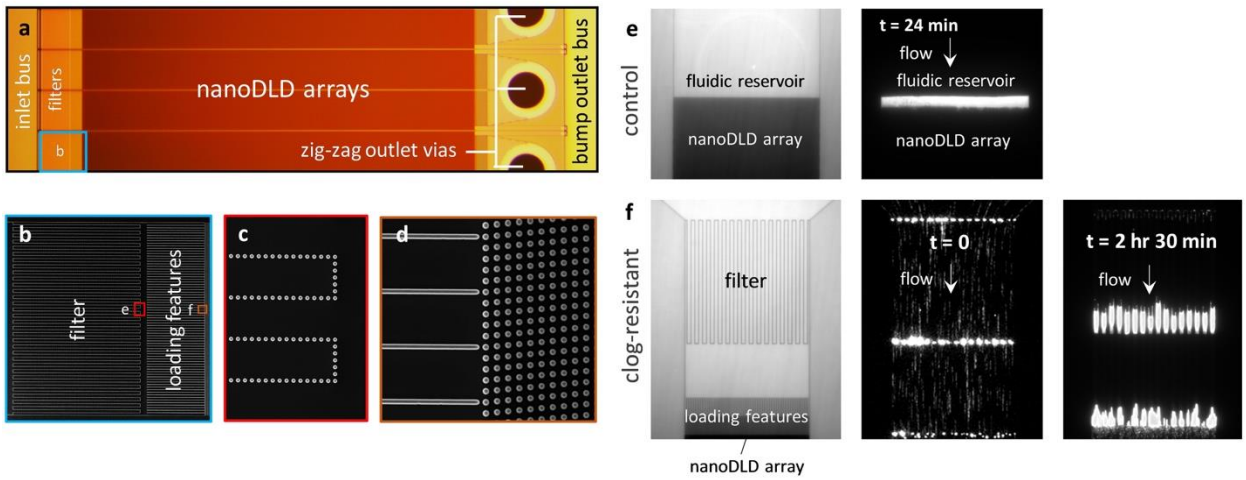


Figure S2. Structure and operation of clog-resistant features upstream from nanoDLD arrays. (a) Optical microscope image showing four parallel nanoDLD arrays relative to the surrounding inlet bus, filters, zigzag outlet vias, and bump outlet bus features with the filter section highlighted. (b) Scanning electron microscope (SEM) image of the filter region in (a), including serpentine filters and downstream pre-loading bladed features. (c) Zoom-in SEM image of the serpentine filter exit from (b). Filters are made up of a single row of pillars with G equal to gaps within in the downstream nanoDLD pillar array. The snake-like boundary of the serpentine filter catches particles with diameter $D_P > G$ without impeding flow, allowing larger particles to pile-up in the catch while particles with $D_P < G$ are allowed to flow through the filter in a rerouted side flow. (d) Zoom-in SEM image from (b) of the pre-loading blade features at the interface of the nanoDLD array. Pre-loading features compartmentalize, or localize, clogs along the array interface to prevent proliferation. (e) SEM (left) and fluorescence microscope (right) images of a single array, G = 240 nm device control in operation having no filters or loading features. D_P = 110 nm beads completely clog the array within 24 min of operation. (f) SEM (left) and fluorescence microscope (center, right) images of a single array, G = 240 nm device in operation containing both serpentine filters and loading features along the nanoDLD array interface. D_P = 110 nm beads flowing through the array at t = 0 min (middle) and t = 2 hr and 30 min (right). Device ran for over 5 hours (not pictured) before catastrophically clogging, a more than 12X improvement in longevity. Note: Carboxylated polystyrene beads provide a conservative estimate of run time as biological samples were typically capable of running for longer periods of time.

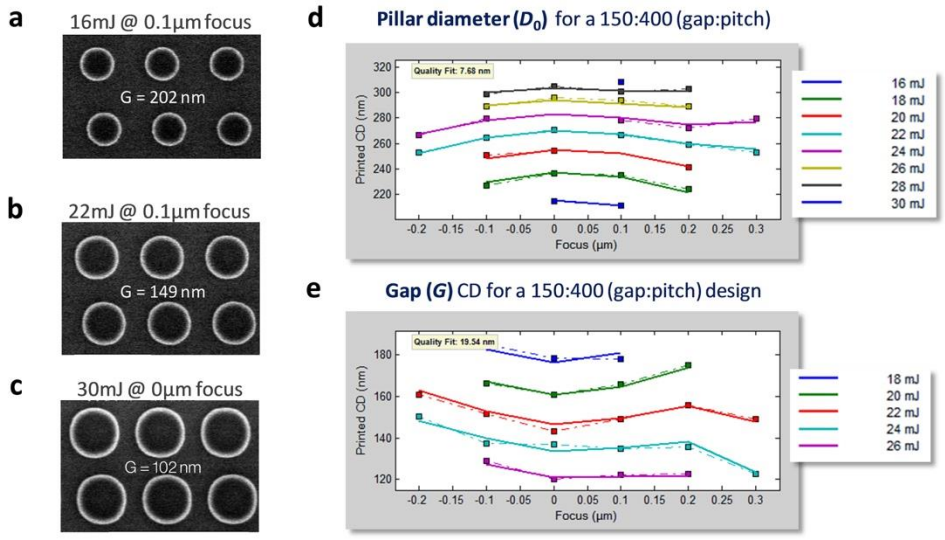


Figure S3. Focus-exposure matrix (FEM) and metrology characterization of pillars in developed AIM7946-126 resist. Scanning electron microscope (SEM) images after FEM exposure/develop of the resist in a 193-nm wavelength step-and-scan PAS 5500/1100 B-4X scanner system (ASML) at (a) dose = 16 mJ / focus = 0.1 µm (G = 202 nm), (b) dose = 22 mJ / focus = 0.1 µm (G = 149 nm), and (c) dose = 30 mJ / focus = 0 µm (G = 103 nm). Dose testing ranged from 16 – 30 mJ in 2 mJ increments and focus ranged from -0.2 to 0.3 µm in 0.1 µm increments as shown in (d, e). Critical dimensions (CD) of the (d) pillar diameter, D_0, and (e) gap, G as measured by autoscan for a focus exposure matrix in a NanoSEM 3D Automated CD Metrology System (AMAT) with the measured CD plotted for each set of exposure conditions printed on a test wafer. Mask features were 150 nm square pillars with pitch λ = 400 nm (not shown). This characterization provide a reference for achieving G of a target size in combination with reactive-ion etch (RIE) pillar formation data and thermal oxidation touch-up. Using this approach, a continuum of G ranging from 80 to 225 nm was possible using a single lithography mask.

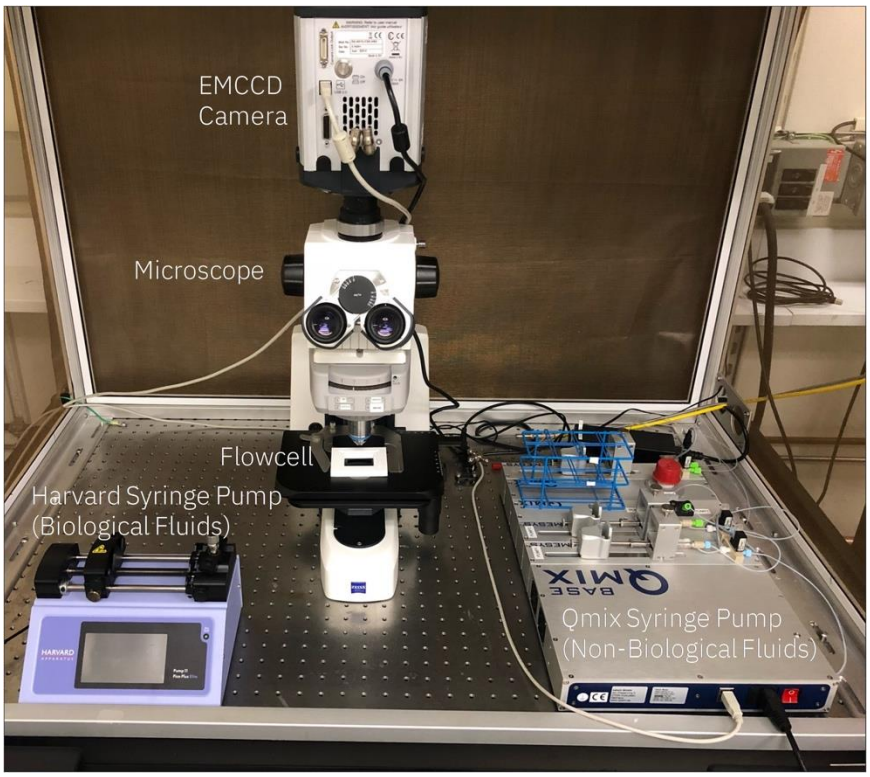


Figure S4. Instrumentation for imaging and running nanoDLD chips. All bright and fluorescence field images were taken using the central microscope (Scope.A1, Zeiss) equipped with a EMCCD camera (Andor iXon Ultra 897, Oxford Instruments). The flowcell, loaded with the requisite nanoDLD chip, is stationed on a holding plate fixed to the x,y stage of the microscope. Pressure drive is provided by either the Qmix syringe pump (for non-biological fluids) or the Harvard syringe pump (for biological fluids). The Qmix pump is used for testing fluidic conductivity of each chip type, while the Harvard pump is used for processing biological fluids for exosome enrichment.

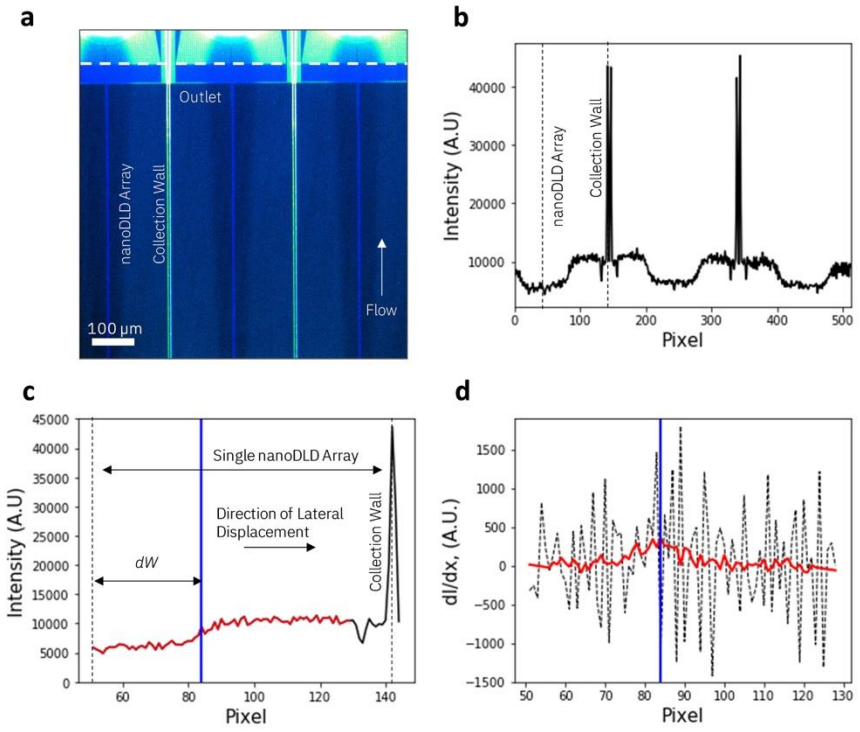


Figure S5. Diagram of image processing steps for extracting the displacement angle in an integrated nanoDLD device for fluorescent beads. The sequence of steps is labeled in order of execution. (a) False color overlay of bright field and fluorescence images near the outlet of a section of integrated nanoDLD arrays. The full widths of 4 arrays are visible with 2 additional partial arrays at the edges. The green shaded fluorescence corresponds to D_P = 75 nm beads flowing in a G = 225 nm device, showing partial bumping. The dashed line indicates where the fluorescence intensity profile is taken to determine the bead deflection. (b) Fluorescence intensity profile taken at the outlet of a set of nanoDLD arrays for the 75 nm beads in (a). The collection walls are denoted by the abrupt spikes in intensity. The two arrays between each set of collection walls is determined by the known width of each array, converted into pixels. The intensity profile is cropped for each array, as shown in (c). The average lateral displacement of the bead flux is taken as the inflection point in the sigmoidal profile, as indicated by the vertical line. Within a given profile, the lateral displacement moves the bead flux towards the collection wall. To avoid mis-assignment by the analysis program due to large deviations, the region around the collection wall is excluded, leaving only the profile shown in red. (d) First derivative of the intensity profile shown in (c) (dashed line) and Savitzky-Golay filtered result (solid line). The inflection point is taken as the local maximum in the filtered (smoothed) line, as detected and indicated by the vertical line (blue). The width deflected, dW, is determined by converting the pixel range into distance, and from this the deflection angle, θ, and displacement efficiency, P, can be calculated from the geometry of the array (see text).

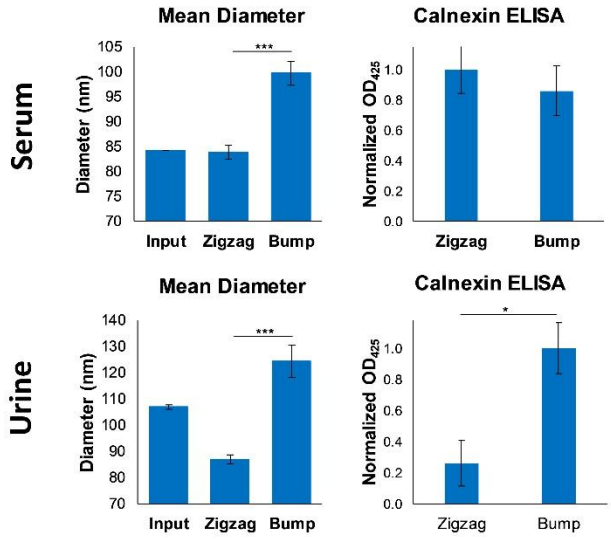


Figure S6. Mean diameter and ELISA results for urine and serum EVs separated by nanoDLD. Mean diameters show an increase in particle size for EVs measured in the bump fraction by NTA (left) for both serum (top) and urine (bottom). Semi-quantitative calnexin ELISAs show approximately the same level of calnexin in serum bump and zigzag fractions (top, right) while urine shows some enrichment in the bump fraction (bottom, right).

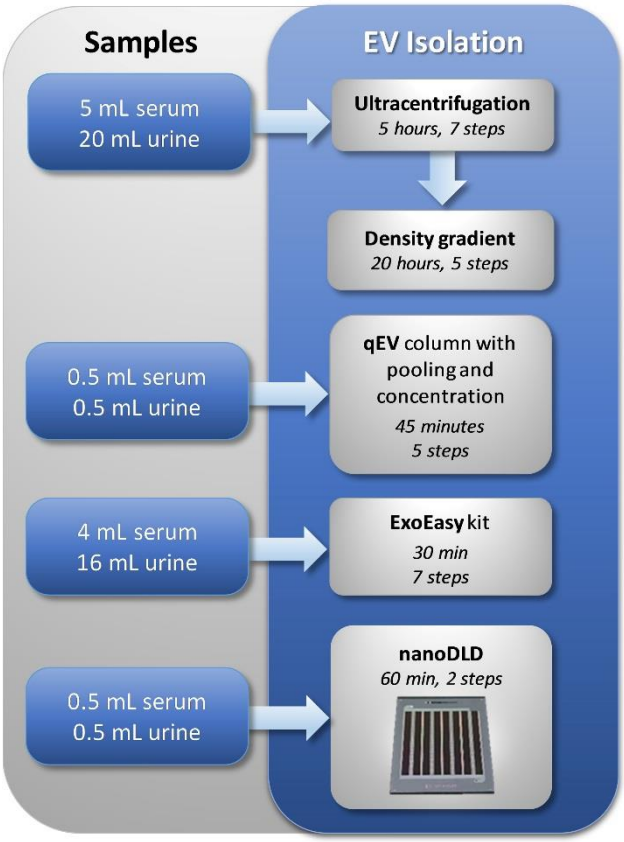


Figure S7. Resource requirements for EV isolation methods used in this study. Required input volumes of urine and serum are indicated on the left panel for each method. Run times and number of steps required for each method are indicated on the right panel.

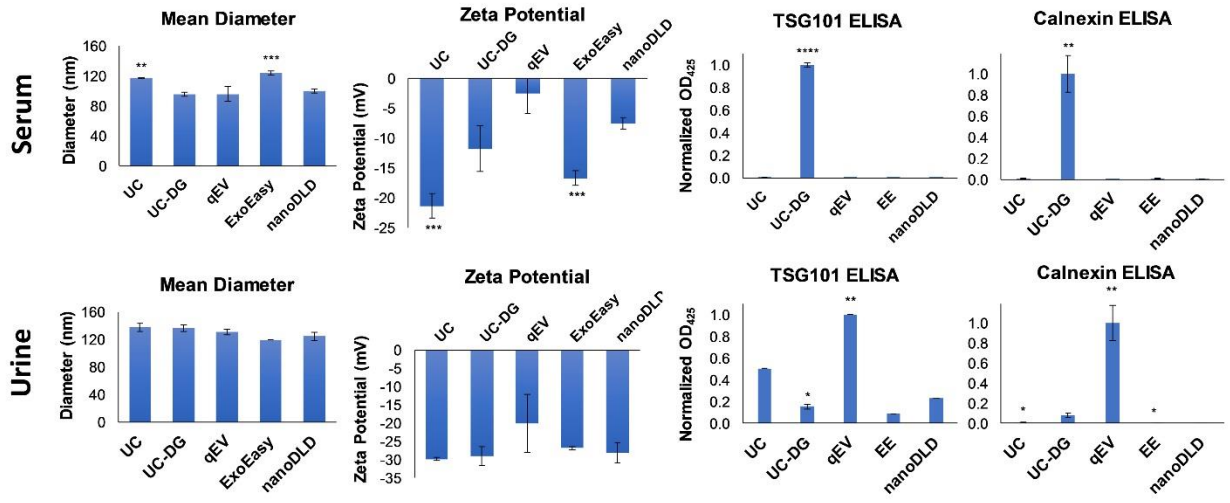


Figure S8. Additional benchmark results comparing nanoDLD against existing EV-isolation methods. Mean diameters of isolated by EVs measured by NTA show little variation between methods. Zeta potentials are negative for all EVs measured. Zeta potential analysis shows negatively charged serum EVs for all isolation methods, but considerable variability in the absolute value. TSG101 and calnexin ELISAs show that while UC-DG produces the highest TSG101 signal for serum and qEV for urine, both also have comparatively high calnexin levels indicating higher levels of contamination. For urine, nanoDLD and UC both have comparatively high TSG101 compared to their relative calnexin levels.

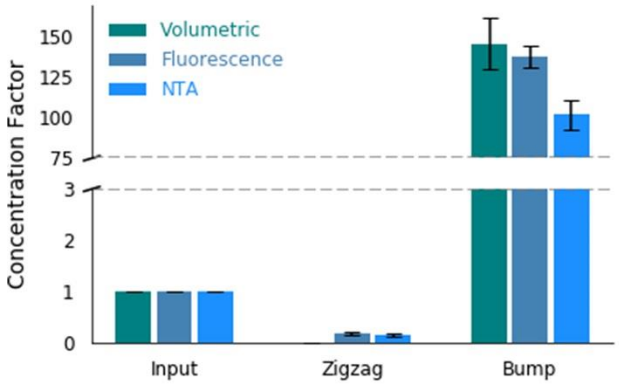


Figure S9. Experimentally measured concentration factors for 95 nm polystyrene beads in a *G* = 225 nm Design II nanoDLD device. 95 nm beads, dissolved in 1X TE buffer with 2% v/v TWEEN-20 surfactant, were run for 60 min at 26 µL/min injection rate. Collected fluid was analyzed using a 96-well plate reader with end-point fluorescence, 350 nm excitation / 450nm emission, and nanoparticle tracking analysis. The fluorescence concentration factor is the ratio of bump to input fluorescence density. Nanoparticle tracking concentration factor is the ratio of bump and input particle concentrations. The volumetric concentration ratio is taken as the total collected volume divided by the bump fraction, assuming 100% displacement given that 95 nm beads show total bump mode in a *G* = 225 nm array. The concentrations factors show reasonable correspondence across the three measurements, demonstrating the concentration effect of the modified Design II. Trials were run in triplicate; error bar represents standard error of measurement (SEM).

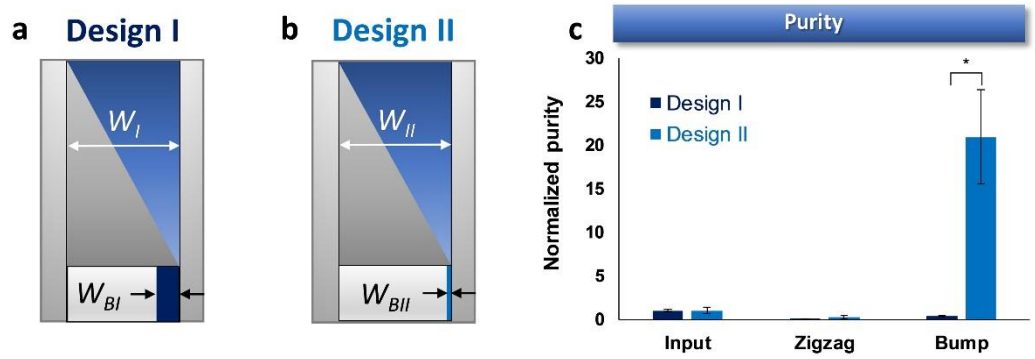


Figure S10. Design layout and purity comparison of different nanoDLD chip designs. Schematic representations of a chip Design I (a) and Design II (b), showing bump outlet channels widths W_{BI} and W_{BII}, respectively, at the exit of a single array where the width of the bump outlet channel narrows such that $W_{BII} << W_{BI}$ as a fraction of the overall array width W. The wedge shape represents the region occupied by $D_P \geq D_C$ particles bumping at θ_{max}. (c) Purity is measured as the number of particles (NTA) per mg of total protein (Bradford assay, ThermoFisher). A greater than 50-fold increase in purity is observed in the bump fraction of Design II vs. Design I.

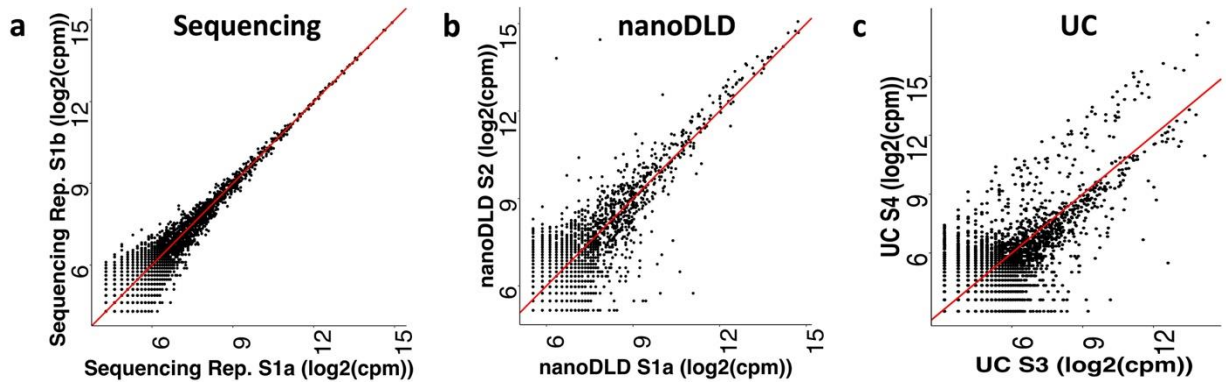


Figure S11. RNA analysis from the serum of a single patient comparing sequencing only replicates with nanoDLD and UC-isolated EV-RNA replicates. Plots of $\log_2$ counts per million (CPM) for RNA sequencing data of (a) sequencing replicates (S1a and S1b – see Fig. 6a), establishing the highest gene expression correlation baseline possible of 0.93 as only the final sequencing step differs between replicates, (b) two nanoDLD isolates (S1a and S2) with a correlation of 0.85, and (c) UC isolates (S3 and S4) with a correlation of 0.78. Plots (b) and (c) are duplicated from Fig. 6 of the text to provide a visual comparison with nanoDLD showing better agreement with the sequencing baseline in (a).

coding RNA			non-coding RNA	
AGR2	FASN	MYL9	MALAT1	MIRLET7BHG
ATP1A2	FLNC	MYLK	MIR106B	MIRLET7C
CAMK2G	FMOD	MYOF	MIR141	MIRLET7D
CDC42EP3	GLI3	NKX3-1	MIR145	MIRLET7F1
CFL2	IRS1	PAGE4	MIR221	MIRLET7F2
CLIP4	KLK3	PI15	MIR222	MIRLET7G
CNN1	MCL1	SEMA3C	MIR22HG	MIRLET7I
E2F1	MSRB3	STAT6	MIR25	NEAT1
EN2	MXRA7	TNS1	MIR32	PCA3
ERG	MYH11	WFDC1	MIR375	PCAT6

Table S1. List of 50 of the most abundant PCa markers found in RNA sequencing data for nanoDLD and UC isolation methods. 20 ncRNA, including miR-25,[4] LET-7 family of miRNA precursor,[4,5] MALAT1 and PCA3,[6,7] and 30 mRNA, including KLK3,[8] ERG,[7] GATA2,[9] and STAT6,[8,10] all of which show differential expression in PCa.

LIVE SUBJECT STATEMENT:

Serum from Human Subjects. Whole blood samples (2 to 5 ml) were collected by the team of Dr. Ashutosh Tewari at the Icahn School of Medicine at Mount Sinai, New York, Department of Urology by venipuncutre from 9 consenting adult male Prostate Cancer patients under Institute Review Board approved protocols (GCO # 06-0996, 14-0318, and surgical consent) in purple capped tubes. After blood collection, serum was isolated using BD Vacutainer blood collection tubes, stored in serum separation tubes (Fisher Scientific, Cat # 368016) and kept at -80°C until further steps were taken for exosome isolation. Serum was rapidly thawed prior to EV isolation with both nano-DLD and UC.

Urine samples from Human Subjects. Urine samples (~20 ml per subject) from 10 consenting adult male kidney stone patients were collected by the team of Dr. Mantu Gupta at the Icahn School of Medicine at Mount Sinai, New York, Department of Urology, under Institutional Review Board approved protocols (GCO #15-1135). After collection the fresh urine was centrifuged at 300 g and the supernatant was collected and kept at -80°C until further steps were taken for EV isolation. Prior to EV isolation with nano-DLD, urine was rapidly thawed and the urine of different patients pooled.

REFERENCES

1. Dobin, A. *et al.* STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **29,** 15–21 (2013).
2. Lu, Y., Baras, A. S. & Halushka, M. K. miRge 2.0 for Comprehensive Analysis of microRNA Sequencing Data. *BMC Bioinformatics* **19,** 275 (2018).
3. Diedenhofen, B. & Musch, J. cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLoS One* **10,** 1–12 (2015).
4. Cochetti, G. *et al.* Different Levels of Serum microRNAs in Prostate Cancer and Benign Prostatic Hyperplasia: Evaluation of Potential Diagnostic and Prognostic Role. *Onco. Targets. Ther.* **9,** 7545–7553 (2016).
5. Ozen, M., Creighton, C. J., Ozdemir, M. & Ittmann, M. Widespread Deregulation of microRNA Expression in Human Prostate Cancer. *Oncogene* **27,** 1788–1793 (2008).
6. Ren, S. *et al.* Long Noncoding RNA MALAT-1 is a New Potential Therapeutic Target for Castration Resistant Prostate Cancer. *J. Urol.* **190,** 2278–2287 (2013).
7. Alford, A. V. *et al.* The Use of Biomarkers in Prostate Cancer Screening and Treatment. *Rev. Urol.* **19,** 221–234 (2017).
8. Prensner, J. R., Rubin, M. A., Wei, J. T. & Chinnaiyan, A. M. Beyond PSA: The Next Generation of Prostate Cancer Biomarkers. *Sci. Transl. Med.* **4,** 127rv3 (2012).
9. Böhm, M., Locke, W. J., Sutherland, R. L., Kench, J. G. & Henshall, S. M. A Role for GATA-2 in Transition to an Aggressive Phenotype in Prostate Cancer through Modulation of Key Androgen-Regulated Genes. *Oncogene* **28,** 3847–3856 (2009).
10. Das, S., Roth, C. P., Wasson, L. M. & Vishwanatha, J. K. Signal Transducer and Activator of Transcription-6 (STAT6) Is a Constitutively Expressed Survival Factor in Human Prostate Cancer. *Prostate* **67,** 1550–1564 (2007).