Lab on a Chip

Versatile Exclusion-based Sample Preparation Platform for Integrated Rare Cell Isolation and Analyte Extraction

Journal:	*Lab on a Chip*
Manuscript ID	LC-ART-06-2018-000620.R1
Article Type:	Paper
Date Submitted by the Author:	18-Sep-2018

Complete List of Authors:	Pezzi, Hannah; University of Wisconsin, Biomedicla Engineering Guckenberger, David ; University of Wisconsin, Biomedicla Engineering Schehr, Jennifer; University of Wisconsin - Madison, Medicine Rothbauer, Jacob; University of Wisconsin, Biomedical Engineering Stahlfeld, Charlotte; University of Wisconsin - Madison, Medicine Singh, Anupama; University of Wisconsin - Madison, Medicine Horn, Sacha; University of Wisconsin - Madison, Medicine Schultz , Zachery; University of Wisconsin - Madison, Medicine Bade, Rory; University of Wisconsin - Madison, Medicine Sperger, Jamie; University of Wisconsin - Madison, Medicine Berry, Scott; University of Wisconsin - Madison, Biomedical Engineering Lang, Joshua; University of Wisconsin - Madison, Medicine Beebe, David; University of Wisconsin, Biomedical Engineering

Journal Name

ARTICLE

Versatile Exclusion-based Sample Preparation Platform for Integrated Rare Cell Isolation and Analyte Extraction

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Hannah M. Pezzi [a], David J. Guckenberger [a], Jennifer L. Schehr [b c], Jacob Rothbauer [a], Charlotte Stahlfeld [b c], Anupama Singh [b c], Sacha Horn [b c], Zachery D. Schultz [b c], Rory M. Bade [b c], Jamie M. Sperger [b c], Scott M. Berry [a], Joshua M. Lang [a b c], David J. Beebe [a c]

Rare cell populations provide a patient-centric tool to monitor disease treatment, response, and resistance. However, understanding rare cells is a complex problem, which requires cell isolation/purification and downstream molecular interrogation – processes challenged by non-target populations, which vary patient-to-patient and change with disease. As such, cell isolation platforms must be amenable to a range of sample types while maintaining high efficiency and purity. The Multiplexed Technology for Automated Extraction (mTAE) is a versatile magnetic bead-based isolation platform that facilitates positive, negative, and combinatorial selection with integrated protein staining and nucleic acid isolation. mTAE is validated by isolating circulating tumor cells (CTCs) – a model rare cell population – from breast and prostate cancer patient samples. Negative selection yielded high efficiency capture of CTCs while positive selection yielded higher purity with an average of only 95 contaminant cells captured per milliliter of processed whole blood. With combinatorial selection, an overall increase in capture efficiency was observed, highlighting the potential significance of integrating multiple capture approaches on a single platform. Following capture (and staining), on platform nucleic acid extraction enabled the detection of androgen receptor-related transcripts from CTCs isolated from prostate cancer patients. The flexibility (e.g. negative, positive, combinatorial selection) and capabilities (e.g. isolation, protein staining, and nucleic acid extraction) of mTAE enable users to freely interrogate specific cell populations; a capability required to understand the potential of emerging rare cell populations and readily adapt to the heterogeneity presented across clinical samples.

Introduction

Emerging discoveries have begun to highlight the biological and clinical significance of rare, discrete cell populations (e.g., minority 'stem' populations[1], circulating fetal cells[2,3], and circulating tumor cells[4]). Yet, rare cells are often masked within larger, more diverse backgrounds of cells (e.g., the bloodstream), complicating isolation[5,6] and analysis of rare cell populations. Each of these rare populations may serve as valuable biomarkers and provide actionable clinical information to improve patient care[7,8,9]. However, patient-to-patient variation introduces diversity in both the rare populations and the background population(s) in which these rare cells reside, coincidently complicating interrogation. In order to evaluate the informative potential of these rare populations and improve patient care, rare populations must first be isolated and analyzed – requiring technologies to separate rare target cells from background.[10]

There are two primary approaches in the growing field of antibody-based cell isolation: positive and negative selection[11,12]. The dominant method, positive selection, typically utilizes antibodies to capture cells in an antigen-dependent manner, yielding a captured population specific to a chosen cellular marker (through antibodies[13,14], carbohydrate receptors[15], etc.). While precise, positive selection requires the marker to be specific to the target population and known *a priori*. As such, positive selection becomes limiting if distinguishing markers are unknown or non-differential (i.e., shared by neighboring cell populations), even if expressed at differing levels. Negative selection leverages known non-target markers to deplete background populations. In this approach, the target cells remain uncaptured, enabling a true "discovery" approach to isolation.

[a] Department of Biomedical Engineering, Wisconsin Institutes for Medical Research, University of Wisconsin-Madison, 1111 Highland Avenue, Madison, Wisconsin 53705, United States
[b] Department of Medicine, University of Wisconsin-Madison, Madison, WI 53705
[c] Carbone Cancer Center, University of Wisconsin-Madison, Madison, WI 53705

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

Despite these advantages, negative selection typically results in incomplete background removal, yielding relatively low purity[16]. Largely, platforms have been forced into a trade-off between "richness" of data (e.g., number of endpoints), specificity (higher with positive selection), and sensitivity (higher with negative selection); these tradeoffs may limit the information collected from rare cells, impairing understanding at a research level and limiting utility in a clinical setting.

Limitations in existing positive and negative selection technologies have risen to the forefront with recent interest in patient-based rare cell isolation applications. One such application pushing the limits of cell isolation technologies is circulating tumor cells (CTCs). CTCs are cancer cells, which separate from a primary tumor or metastatic site and enter the bloodstream. The tumor origin of CTCs paired with their easy, minimally invasive accessibility (e.g., blood draw), make CTCs a uniquely poised asset with which to monitor response to anti-cancer therapies. Following treatment, CTC enumeration in prostate cancer and breast cancer patients has demonstrated (based on EpCAM-captured CTCs) prognostic potential in informing treatment outcome[12,17]. However, it is now clear that enumeration alone is unlikely to revolutionize cancer monitoring and additional endpoints will likely be necessary to expand the clinical utility of CTCs[10]. Furthermore, populations of CTCs deviating from the 'classic' EpCAM-positive CTCs have highlighted the need for flexibility in isolation approaches, even for enumeration.

Similar to other patient-based cell isolations, CTC isolation is fundamentally challenged by the heterogeneity that exists between and within cancer types, including: variability in expression of capture markers, differing marker profiles (e.g., EpCAM-positive CTCs[18,19], CTCs undergoing epithelial-mesenchymal transition (EMT)[20-26], cancer stem cells[26,27]), emerging markers[28], and varying background populations[29,30]. The heterogeneity of patient-based cell isolations and the pursuit of rare cell populations requires technologies that are adaptable - not limited to a single, specific marker – and maintain the capability of discovery-based negative selection. While existing platforms have facilitated development in the CTC field, growing understanding of the complexity of CTCs largely enabled by these platforms has highlighted the need for adaptability.[31] Ultimately, platforms limiting users to rigid isolation protocols inherently screen and bias the information obtained, leaving researchers and clinicians unable to fully assess the clinical value of rare cell populations. Recently, the CTC-iChip introduced the capability to switch between positive selection and negative selection, allowing users to benefit from both techniques independently.[32-33] However, the iChip process sequesters the sample restricting the user to one selection method. While enabling, to be truly versatile platforms should allow users to integrate selection methodologies on single samples to maximize information gain, at the highest quality. In the CTC field, interrogation beyond enumeration will be required to deliver on the full clinical potential of CTCs. In moving beyond enumeration, captured populations will need to meet a new level of purity in order to facilitate integration with downstream molecular analyses (RT-PCR, whole genome amplification (WGA), sequencing, etc.) to ensure the target signal is not masked by background populations. This transition from enumeration to purity-driven endpoints is challenging as the rarity of CTCs (1 in 1-10 million PBMCs), paired with the diverse cell populations found in peripheral circulation, makes CTCs a difficult target to capture and isolate with high purity; yet purity remains a prerequisite for accurate downstream interrogation and analysis. Existing platforms alone will likely be inadequate to meet the purity demands of the next generation of rare cell analysis endpoints, ultimately limiting the ability of researchers and clinicians to discern the population's full potential to inform patient care.[17]

Building on a suite of exclusion-based sample preparation (ESP™) technologies[34-36], we have developed an automated multi-sample cell isolation platform termed the **M**ultiplexed **T**echnology for **A**utomated **E**xtraction (mTAE) to enable users to perform serial positive and negative selections on multiple samples in parallel while reducing user-to-user variation through automation. To achieve both positive and negative cell isolation, antibodies are bound to small, magnetically responsive particles termed paramagnetic particles (PMPs). PMP-bound cells are removed from the high-background sample population using the Sliding Lid for Immobilized Droplet Extraction (SLIDE™) technology – a low shear method for achieving high purity extraction of PMP-bound analytes[36]. PMPs and bound cells are pulled to the top of sample wells and collected on a hydrophobic surface for removal. Due to the limited interaction of the surface and sample, SLIDE leaves the sample readily available for re-interrogation. In other words, SLIDE does not dilute, wash away, or otherwise manipulate a sample during cell selection, leaving it available for subsequent positive or negative selection steps. In this manner, mTAE can achieve both high specificity (positive) selection and high sensitivity (negative) selection on a single rare sample. Once extracted, PMP-bound cells can then be deposited into 1) wash wells to improve purity, 2) protein staining wells for cell identification, or 3) wells for PMP-based nucleic acid (NA) extraction. Integrated staining and NA extraction capabilities will facilitate downstream analytical pipelines (cell identification/enumeration, qRT-PCR, sequencing, WGA). mTAE allows users to tailor their cell isolation protocols to best facilitate their endpoints, as they are no longer limited to a single selection methodology or downstream analysis method. mTAE is scalable to perform isolations and downstream processing on four samples in parallel, with the capacity to easily expand to eight, thus increasing throughput and decreasing sample-to-sample variability. Here, mTAE's capabilities are demonstrated by performing cell selection with on-chip immunofluorescent protein staining and/or NA extraction. mTAE's ability to perform both positive and negative selection of rare cells is evaluated using CTCs as a model rare cell system. Using patient samples, we are able to evaluate the platform's performance on complex "real world" samples with high inter-patient variability. Additionally, we demonstrate the capacity to perform serial selections, specifically sequential negative and

positive selection that significantly improved sample purity in a subset of samples to achieve a high signal-to-noise ratio needed for high content molecular analyses.

Experimental

Automated ESP Platform Overview

PMP-based manipulation of cells and NA was achieved by integrating two technologies on a Gilson PIPETMAX automated liquid handler (Gilson Inc.): SLIDE liding Lid for Immobilized Droplet Extraction)[36] and a custom magnetic box component[37]. SLIDE is a technique for isolating and purifying PMP-bound analytes. As the SLIDE technology has been integrated into Gilson's EXTRACTMAN, EXTRACTMAN extraction plates (#22100008, Gilson) and collection strips (#22100007, Gilson) were used to achieve these isolations. The developed, automated SLIDE leverages convex droplets and an automated pipette head modified to house magnets to capture and transfer PMPs between wells. To isolate PMPs and PMP-bound targets (e.g., NA, cells), samples are placed in a custom well plate (#22100008, Gilson), filled so that each well contains a convex droplet. Then the magnetic pipette head, covered with a plastic collection strip (#22100007, Gilson), can be brought into contact with the convex meniscus and the magnets lowered towards the plate. The lowered magnets attract the PMPs, driving \collection on the strip. The strip is then removed from the fluid and introduced to a new well (e.g., containing wash buffer). By retracting the magnet from the plastic consumable the beads are released into the well with the assistance of a custom magnetic box located below the plate [37]. The magnetic box is a magnetic technology that operates in unison with the magnetic pipette head leveraging magnet proximity to manipulate the PMPs; the box allows PMPs to move up to the magnetic pipette head (capture) or into the well (release) depending on relative distances between the magnets (**S1**). By controlling the distance of the magnets in the pipet head, paired with the magnetic box contained below the plate, the PMPs can be readily released off the collection strip and into a new well. PMPs and bound target can readily be manipulated in and out of wells enabling cell isolation, washing, staining, fixation, and lysis depending on the buffer of the well.

For the developed cell isolation protocol, PMP-bound cells are captured from a sample well and carried to, released in, mixed, and recaptured in a series of sample wells, then released in an output well for image analysis. All cell fixation, permeabilization, and staining (both intra- and extracellular) is performed in the wash wells. For protocols including RNA or DNA extraction, the PMP-bound analytes are then carried on the collection strip to a separate plate, lysed, washed, and eluted. Additional information is available in **SI**.

Cell Culture

All cell lines used for characterization of the automated platform (LNCaPs (gift from Dr. Douglas McNeel, University of Wisconsin-Madison), HCC2218 (ATCC), PC3-MM2 (gift from Dr. C. Pettaway, MD Anderson Cancer Centre, TX, USA), 22RV1 (gift from Dr. Douglas McNeel)) were cultured in RPMI1640 media (#11875-093, Thermo Fisher Scientific) supplemented with 10% Fetal Bovine Serum (Gibco) and 1% Penicillin Streptomycin (Gibco). Cells were maintained under sterile conditions at 37° C in 5% CO_2.

Human Subject Blood Processing and PBMC Isolation

Whole blood was collected via venipuncture and processed via Ficoll Paque PLUS (#17-1440-02, GE Healthcare) to enrich mononucleated cells. For characterization, whole blood from healthy donors (Biological Specialty Corporation) was received and processed within 24 hours of collection. Clinical peripheral blood specimens were collected at the University of Wisconsin with informed written consent under a University of Wisconsin Health Sciences Institutional Review Board (HS-IRB) approved protocol **(S9-11)**. The HS-IRB complies with the applicable requirements of the Department of Health and Human Services (DHHS) regulations, 45 CFR Part 46; the Food and Drug Administration (FDA) regulations, 21 CFR Parts 50, 56, 312, and 812; Veteran's Administration (VA) Regulations pertaining to the protection of human subjects, 38 CFR Part 16; and the privacy requirements of the Health Insurance Portability and Accountability Act of 1996 implemented by 45 CFR Parts 160 and 164 (Privacy Rule). The blood was collected in EDTA tubes, and processed within 5 hours of collection. Briefly, whole blood was mixed 1:1 with 2mM EDTA 1x PBS. 35 mL of diluted whole blood was overlaid on 15 mL of Ficoll. The tubes were centrifuged according to the manufacturer's instructions and the buffy coat diluted in wash buffer (1x PBS, 2mM EDTA, 0.1% BSA, 2.5% FBS). The cells were washed twice at 200 x g for 10 minutes.

PMP Conjugation and Binding

For positive selection via EpCAM, anti-EpCAM antibody (clone VU-1D9) (#ab98003, Abcam) was conjugated to Dynabeads M-270 Epoxy using the Dynabeads Antibody Coupling Kit (#14311D, Thermo Fisher Scientific) at a concentration of 10 µg Ab / mg PMP (250 µg of PMPs per sample). Prior to use, the PMPs washed by collecting the PMPs to the side of a tube, removing the supernatant, and resuspending in twice the volume of PBS supplemented with 0.1% Tween20 (PBST). After recollecting the PMPs and removing the PBST, PMPs were resuspended in wash buffer. For experiments involving negative selection or depletion of PBMCs, M-270 PMPs were coupled with antibodies against CD45 (clone HI30) (#304002, Biolegend), CD14 (clone M5E2) (#301802, Biolegend), CD34 (clone 581) (#343502, Biolegend), and CD11b (clone M1/70) (#101202, Biolegend) using the manufacturer's recommended protocol at a concentration of 10 µg Ab / mg PMP. The samples were bound to PMPs on ice for 30 minutes, with mixing at minute 5, 15, and 25 minutes.

Characterization of the Automated Platform

To validate the platform and select ideal operating conditions for positive selection from liquid biopsies, we assessed the impact mixing rate, cell populations, and cell phenotypes have on loss, purity, and capture efficiency. As CTCs are known to have variable EpCAM expression, the platform was characterized with three, variable EpCAM-positive cell lines

(**S2**): LNCaPs, HCC-2218, and PC3-MM2 (sub-clone of PC3 cell line). Cell line viability was assessed via a Live/Dead assay (#L3324, Thermo Fisher Scientific). Target cells were placed into a PBMC background, creating a pseudo-sample. To differentiate between target and non-target cells the two were pre-stained with Cell Tracker Red and Calcein AM (Life Technologies), following the manufacturer's protocol. The samples were mixed with PMPs – pre-conjugated with anti-EpCAM antibody, binding for 30 minutes on ice, and transferred to mTAE,

Samples consisting of ~500 target cells in a background of 1 million non-target PBMCs (from a healthy individual) were used to evaluate the cell loss as a result pipette mixing. The PMPs were first collected from a sample well and transferred to a small wash well, leaving behind unbound target and non-target cells. Without mixing, the PMPs were recollected, released, and transferred to a second wash well, excluding any non-bound cells that were non-specifically carried with the PMPs (i.e., in the interstitial space). In the second wash well, the PMPs were mixed for four (aspirate/dispense) cycles at flow rates ranging 1 – 20 mL min^{-1}, with a no mixing control (0 mL min^{-1}). After mixing, the PMPs were transferred from the second wash well to the output well, leaving behind cells that detached from the PMPs as a result of the mixing. Loss of target cells was quantified by collecting the contents of each well and counting the number cells. Similarly, the loss of non-target cells, as a result of shear, was quantified by counting the number of PMBCs present in each well. A mixing flow rate of 5 mL min^{-1} was utilized for all subsequent experiments.

To assess how the quantity of target cells impacts purity and capture efficiency, pseudo-samples of 10, 100, and 1000 target cells in a constant background of 10 million PBMCs were utilized. PMPs and bound cells were collected from the sample well, washed in three wash wells, the released in the output well; all wells were then collected for imaging. Conversely, to assess how the quantity of non-target/background cells effects purity and capture efficiency, the amount of non-target cells was varied from 0 to 20 million cells while the number of target cells was held constant at 1000 cells. In both cases, percent capture represents the number of target cells in the output well divided by the total number of target cells spiked into the sample. The purity is the number of target cells captured divided by the total number of cells in the output well (target and non-target cells).

Patient Sample Cell Staining

In mTAE, PMP-bound cells were washed in one well then transferred into 100 μL of extracellular staining buffer and the plate transferred to on ice for 30 minutes. The extracellular staining buffer contains: anti-EpCAM-PE (#ab112068, Abcam) and anti-CD45 (#304002, Biolegend), anti-CD14 (#301802, Biolegend), anti-CD11b (#101202, Biolegend), and/or anti-CD34 (#343502, Biolegend), all diluted at 1:100 in wash buffer. Antibodies against CD45 (PBMC marker), CD14 (monocyte marker), CD11b (NK, monocyte, neutrophil marker), and CD34 (endothelial marker) were conjugated to Alexa Fluor 647 (#A-20186, Thermo Fisher Scientific) and comprised what is referred to as the 'exclusion channel'. The markers beyond CD45 were included to ensure the most accurate identification of CTCs and reduce potential false positive identifications. [38] In the next well, cells were fixed at room temperature in 4% PFA (diluted in PBS) for 15 minutes (100 μL). The PMP-bound cells were then moved into permeabilization buffer (PBS supplemented with 1% Tween20 and 0.05% Saponin) for 30 minutes at room temperature. PMP-bound cells were then transferred into 100 μL of an intracellular staining buffer (Hoechst (diluted 1:250) and a pan-cytokeratin antibody (FITC) (clone C-11) (#ab78478, Abcam) (diluted 1:100)) for two hours. The cells were then transferred to a new well for imaging. With patient samples, imaging was often performed in a "sieve device" as described by Zasadil et al.[39] and Casavant et al.[35] Following processing and staining in mTAE, cells were transferred into the sieve device, and a magnet applied to the back, to deplete unbound PMPs to enhance image clarity.

Imaging and Image Analysis

All imaging was done on a Nikon TI Eclipse inverted microscope. For cell line characterization of the automated platform, the cells were transferred to a 96-well plate, allowed to settle, and then imaged with a 10x APO objective. Enumeration of the cells for platform characterization was accomplished using the "Find Maxima" function in ImageJ. All patient sample target populations (either positively or negatively selected) were imaged using a 20x or 40x APO objective. For patient samples, an ImageJ macro was developed to first identify the location of cells based on positive nuclear staining then measure the mean fluorescence intensity of each marker (exclusion channel, nuclear, CK, and EpCAM). Each identified cell was plotted based on exclusion channel intensity and CK intensity; then, using the entire population, thresholds were created to differentiate between negative and positive fluorescence intensity. CTC events were defined as cells containing a nucleus, positive CK stain, and negative for any contaminant population markers (CD45, CD14, CD34, CD11b). Due to patient-to-patient variability, thresholds were determined on a patient-to-patient basis using the entire population of cells. Each identified CTC event was then visually re-inspected to ensure that: (a) the event was indeed a cell (verified by bright field images) and (b) the event did not have any nearby staining artifacts artificially biasing quantified signal (e.g., enhance the CK intensity). To test the specificity and sensitivity of both the platform and image analysis program, multiple healthy patients were analyzed, all of which did not identify any CTC events.

Nucleic Acid Extraction and Quantification

Both RNA and DNA extraction processes were integrated into mTAE and validated. To validate the DNA protocol, samples of 10, 100, and 1000 LNCaPs were lysed on mTAE in RLT (Qiagen) along with 5 μL stock Magnesil KF PMPs (Promega). The bound DNA was then subjected to three washes in wash buffer (10 mM Tris-HCl pH 7.5, 0.15 M LiCl, 1 mM EDTA) with mixing and eluted in nuclease-free water. For comparison, DNA was extracted from identical samples using the same reagents in a manual tube-based approach as well as the commercial

QIAamp DNA Mini spin columns (#51304, Qiagen). For elution, 15 µL was used across all platforms. Extracted DNA was quantified for a housekeeping gene, GAPDH (#402869, Thermo Fisher Scientific) on a real-time thermal cycler (LightCycler 480 II, Roche). In brief, 2 µL of eluted DNA was mixed with 5 µL of Roche LightCycler480 Master Mix, 0.5 µL of mRNA specific primers to GAPDH (Taqman, FAM), and 2.5 µL of nuclease-free water. The mix was amplified for 45 cycles (95 °C for 15 seconds, 60 °C for 30 seconds). Cycle threshold values for each result were then calculated based on 2nd derivative maximum function (LightCycler 480 Software).

For validation of the RNA extraction, 10, 100, and 1000 LNCaPs were used. On mTAE, LNCaPs were added to Lysis Binding Buffer (Dynabeads® mRNA DIRECT™, ThermoFisher) along with suspension Dynabeads® Oligo $(dT)_{25}$ PMPs. The PMPs and bound RNA were then subjected to two wash wells (wash buffer-see above) with mixing and eluted in nuclease-free water. For comparison, both a manual tube-based method was performed using identical reagents as well as a commercial spin column (RNeasy Plus Micro Kit, #74034, Qiagen). All samples were eluted in 15 µL for consistency. Quantification of the RNA was performed using primers specific to mRNA; no reverse transcription controls demonstrated RNA specificity and minimal DNA amplification (>8 cycle delay). RT-PCR was performed on a real time thermocycler (LightCycler480, Roche) using a one-step RT-PCR mix (Taqman Mastermix 1-step Master Mix, Life Technologies). Detection of RNA was done following reverse transcription (50 °C for 5 minutes) followed by inactivation of the RT enzyme (95 °C for 20 seconds) and then 45 amplification cycles (95 °C for 3 seconds followed by 30 seconds at 60 °C).

For evaluation of AR-specific transcripts, cells were captured with anti-EpCAM M-270s, released into Lysis/Binding Buffer (Dynabeads® mRNA DIRECT™, ThermoFisher) along with 20 µL Dynabeads® Oligo $(dT)_{25}$ PMPs. The PMPs were then collected, transferred through two wash wells and dropped in nuclease-free water. The eluted mRNA sample (both cell capture and RNA-specific PMPs) was reverse transcribed using a High Capacity cDNA Reverse Transcriptase Kit (Life Tech, USA) according to the manufacturer's directions (20 µL total reaction). 12.5 µL of cDNA then underwent 10 amplification cycles using TaqMan® PreAmp (Life Tech, USA) according to the manufacturer's directions. Following PreAmp, the amplified product was diluted 1:5 in 1x TE buffer (10 mM Tris-HCL pH8, 1 mM EDTA). For detection, 5µL of diluted cDNA template was mixed with 10 µL iTaq® master mix (Bio-Rad, USA), 1 µL TaqMan® Gene Expression Assay (**S8**) (Life Technologies, USA) and 4 µL nuclease free (NF) water. Each reaction was amplified for 45 cycles (denatured at 95 °C for 15 seconds, annealing at 60°C for 1 minute) using a CFX Connect® Real-Time PCR System (Biorad, USA). A table of primers used is available in **S6**.

mTAE was designed as a multiplexed, versatile rare-cell isolation platform enabling positive and negative selection (either alone or in combination), on-chip cell fixation, protein staining, and extraction of RNA or DNA from a single sample. mTAE builds upon an automated liquid handler (PIPETMAX™, Gilson, Inc.), leveraging the SLIDE technology described by Casavant et al.[36], the magnetic system described by Guckenberger et al.[37], and EXTRACTMAN™ (Gilson) consumables (i.e., strips and plates).

To evaluate mTAE's performance as a rare cell isolation platform, CTCs provide a relevant and challenging example of a rare population masked within a large, diverse background. Isolated from the bloodstream, CTC isolation requires separation of the target population from a vast, diverse background, which varies with disease and treatment. Furthermore, CTCs are a population complicated by heterogeneity (cancer heterogeneity, patient-to-patient heterogeneity, and intra-patient CTC variation), making CTCs a difficult population to isolate and purify.

To facilitate multiple approaches to cell isolation with mTAE, cells are isolated from a sample via: (i) a *positive selection* mode, whereby CTCs are selected based on an extracellular marker (i.e., EpCAM), and (ii) a *negative selection* mode, whereby cells (i.e., PBMCs) are selected for removal based on expression of contaminant markers (i.e., CD45 (lymphocytes), CD14 (monocytes), CD11b (myeloid), and CD34 (endothelial cells)).

Following positive selection, PMP-bound cells are carried through a series of washes, wherein PMPs can be gently mixed via magnetic mixing or vigorously mixed via pipette mixing. We assessed the impacts of pipette mixing by quantifying the loss of both target cells **(Fig. 2A)** and non-target cells **(S1)**. After mixing PMP-bound target cells at various flow rates and assessing loss, a mixing rate of 5 mL/min was selected for all subsequent experiments. Wash wells contained washing buffers, stains, permeabilization buffers, and fixation buffers as described in the methods section. For enumeration, cells were either imaged directly in the plate via a modified plate containing a glass-bottom well, or transferred, via pipette, to a secondary imaging platform. Non-fixed cells, intended for nucleic acid extraction, were magnetically recaptured after imaging and brought to a subsequent well for lysis.

mTAE was tested and characterized with commercially available PMPs and antibodies, enabling straightforward adaptation to new targets by future users. After an ESP capture step, the sample remains available for re-interrogation, despite repeated PMP additions and cell capture steps. As such, specific cellular populations can be serially isolated using any desired combination of positive and negative selection at the discretion of the user making mTAE a truly open, adaptable platform.

Results & Discussion

Automated cell isolation platform design and optimization

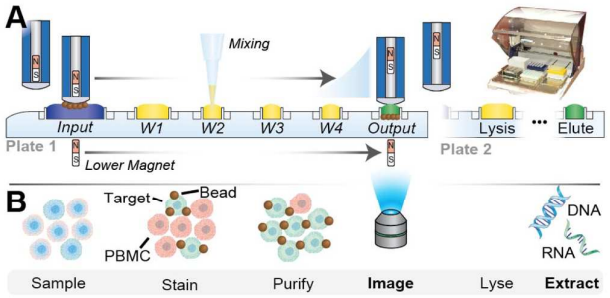

Figure 1: mTAE system overview. (A) Cell extractions are performed on a modified Gilson PIPETMAX. The extraction plate contains six wells per sample with four samples per plate. The input is filled with sample (PBMC sample) and anti-target PMPs while the remaining wash wells (W1-W4) contain a combination of wash buffers, fixatives, permeabilization buffers, or fluorescent stains depending on the application. A second plate is added to the system if the user requires additional processing steps (i.e., multiple wash, fixation, permeabilization, and staining steps), RNA extraction, or DNA extraction. A magnetic head moves the PMP-bound analyte between wells and adjacent plates through a balancing, and opposing magnetic force located below the plate. (B) Schematic overview of the process. Cells are stained and purified in the wash wells transferred to an output well for imaging, and, if applicable, transferred to a second plate for NA extraction.

Cell Capture Characterization and Validation

mTAE was first evaluated for capture efficacy and purity using EpCAM-positive cells lines in a background of healthy donor PBMCs. Three EpCAM-positive cells lines served as model target cells: LNCaPs (human prostate adenocarcinoma cell line), PC3-MM2 (MM2) (highly metastatic PC3 derivative), and HCC2218 (HCC) (derived from a primary ductal adenocarcinoma). Capture efficacy (i.e., captured target cells/starting total number of target cells) was first individually assessed for each of the three EpCAM positive cell lines (LNCaPs, HCC, PC3-MM2) **(S3)** from a background of 10 million PBMCs **(Figure 2B)**. Each cell line demonstrated differing capture efficacy, ranging from ~40% (HCCs) to >95% (LNCaPs). Capture was reflective of EpCAM expression (via immunohistochemistry) with capture correlating to observed EpCAM expression **(S3)**. The lower capture of HCCs may also have been artificially influenced by viability; as a suspended cell line, dead cells were maintained in the culture at higher frequencies than the adherent target cell lines, potentially lowering capture efficiency (LNCaP and PC3-MM2 viability >95%; HCC viability 65%). Temporal variation in capture efficacy was also evaluated. Cells were collected across five different days and isolated on mTAE. LNCaPs and PC3-MM2s, the two adherent cell lines, consistently captured (on average 98% and 84% respectively) with less than 4% standard deviation across the five different days assayed (2% for LNCaPs, 4% for PC3-MM2). The suspended cell line captured with greater variation at 11% standard deviation across the five days. However, this variation in part may be explained by an increased variability in the viability of the HCC line compared to the adherent cells (viability standard deviation of 11% for HCC, 3% for LNCaP, 2% for PC3-MM2). Next, we varied the quantities of target cells in the sample to assess how order of magnitude differences in target cells impacts capture efficacy and purity **(Figure 2C,E)**. The capture efficacy remained consistent (LNCaPs: <3% variation, HCCs: <5% variation), while purity (i.e., the number of target cells captured/total number of cells captured) increased with increasing target cells. Notably contaminant cells also increased with increasing target; however, the increase in contaminant cells was much lower than the fold increase in target cells leading to an overall increase in purity.

Next, to assess how background populations impact capture efficacy and purity, 1,000 target cells were isolated from 0-20 million background cells (PBMCs) **(Figure 2D,F)**. For both LNCaPs and HCCs, capture from increasing background had limited impact on capture efficiency; only LNCaPs showed a statistically significant decrease in capture efficiency in a background of 20 million compared to no background ($p<0.05$). Purity, however, decreased as the background population increased, suggesting that an integrated negative selection may prove advantageous for applications requiring high purity. To validate cells remain viable post-isolation, PMP-captured cells were cultured; viability remained within 10% of the cell only control following a 5-day culture **(S4)**. These results begin to characterize performance parameters and validate mTAE as a gentle, effective method of isolating rare cells from diverse background populations.

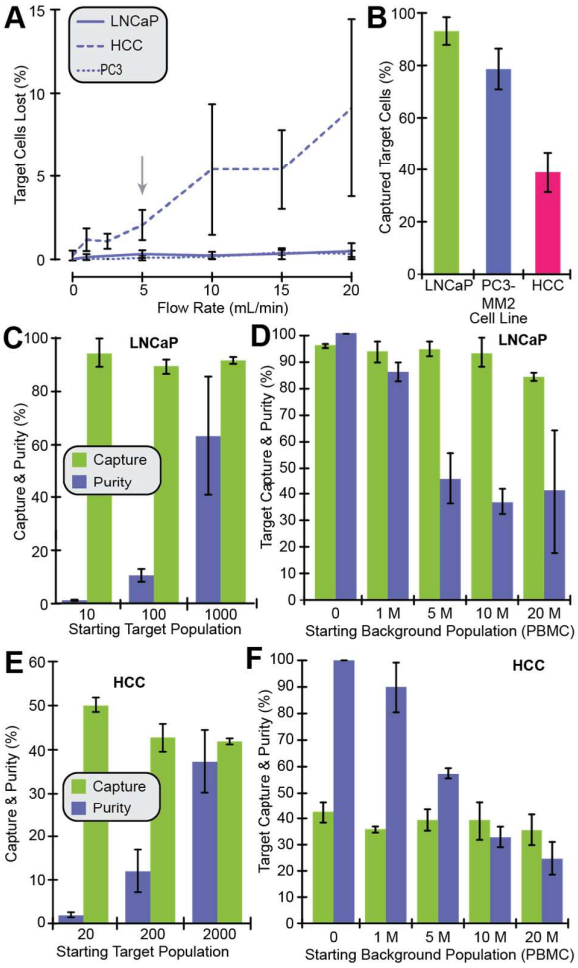

Figure 2: Platform characterization and cell line validation. (A) Loss of target-cells as a result of flow introduced by pipette with flow rates ranging from 0 to 20 mL min-1. A flow rate of 5 mL min-1 was used for all subsequent experiments as noted by the arrow. (B) Capture efficacy of 500 target cells spiked into a background of 10 million PBMCs. (C, E) Impact of target-cell quantity on purity and capture efficacy of HCCs (C) and LNCaPs (E). In both cell lines, specific quantities of target-cells were spiked into a constant background population of 10 million PBMCs. (D, F) Impact of background population on purity and capture efficacy of HCCs (D) and LNCaPs (F). For both cell lines, 500 target cells were spiked into a population of 0 to 20 million PBMCs.

Positive Selection: CTC purification from patient samples via EpCAM

Positive selection enables target cell isolation via known target markers. In the CTC field, the extracellular marker EpCAM has been extensively utilized as a positive selection CTC marker for both prostate[40,41] and breast cancer[19,42,43]; within a subset of cancers, EpCAM-based CTC enumeration has provided predictive insight into prognosis and overall survival.[44,45] To evaluate positive selection on mTAE, we assessed EpCAM-based PMP isolation of CTCs from samples (i.e., PBMCs isolated from 5-10 mL of whole blood) obtained from patients with prostate cancer (n=16), breast cancer (n=8), and healthy donors (n=4). Once captured, the cell-bound PMPs were washed in a series of wells, containing: i) a nuclear stain and extracellular stains for EpCAM, CD45, CD14, CD34, and CD11b, ii) a fixative, iii) a permeabilization buffer, and iv) an anti-pan cytokeratin (pCK) intracellular stain. CTCs were identified as cells staining positive for pCK and a nucleus, but negative for all exclusionary markers (i.e., CD45, CD14, CD34, and CD11b). Exemplary CTCs and PBMC images from both breast and prostate cancer patients are shown in **Figures 3A,B**. While most CTCs were found as individual cells, clumps of CTCS **(Figure 3C)** were identified in a subset of patients, a phenomenon previously observed in other CTC isolation platforms.[46,32] Capture was reported by quantifying the positively identified CTC events per milliliter of whole blood assayed **(Figure 4D)**. Variable CTC counts were identified for both prostate and breast cancer patients, ranging from 0 to ~13 CTCs per milliliter of whole blood. Patient-to-patient variability in CTC count is known based on a patient's treatment, disease progression, and treatment response.[47,48] Importantly, no CTCs were identified in healthy patients (n = 4 patients), validating the specificity of our CTC identification parameters and exclusionary markers. Along with captured CTCs, purity (**Figure 3F**) and purity's relationship with CTC count was also evaluated **(Figure 3E)**. Within the evaluated samples, purity was highly variable ranging from <1% up to ~37%. Purity and CTC number did not appear to be strongly correlated. Rather, the inconsistency across patients highlights the patient-to-patient heterogeneity and limited capacity for users to predict sample purity both in advance of isolation and in the absence of any identification staining.

While capture for enumeration provides clinically relevant information, downstream processes often require a threshold level of purity for successful analysis. Thus, in addition to maximizing target capture, purity also requires minimizing contaminant cells. With positive selection, we observed an average of ~4.3 $\log_{10}$ fold reduction of non-target cells across healthy, prostate cancer, and breast cancer patients **(Fig. 3G)**. Overall, samples demonstrated a carryover, contaminant population (i.e., cells not identified as a CTC), ranging from ~30 to ~280 cells per milliliter of whole blood (average: ~95; standard deviation: 64) (**S6**). Notably, this carryover population did not correlate to the number of CTCs captured and likely is a reflection of patient-to-patient variability. Furthermore, this non-CTC captured population was typically excluded due to positive staining of exclusion markers (CD45, CD11b, CD34). To evaluate if this non-target capture population was specific to the anti-EpCAM antibody, healthy samples were spiked with both capture (anti-EpCAM M-270) and blank (unconjugated M-270) PMPs; both yielded similar numbers of contaminant cells (Data not shown). Thus, we hypothesize a subset of the PBMC fraction nonspecifically adheres to the PMP surface (rather than through the capture antibody) and is resultantly captured. The exact mechanism by-which the contaminant cells non-specifically adhere to the anti-EpCAM PMPs and unconjugated PMPs is currently unknown; yet, the existence of this non-specifically captured population impacts the purity of the extracted population, a consideration in any PMP-based platform. While powerful, positive selection is incapable of capturing analytes for which an identifiable marker is lacking or not yet known, a frequent problem in emerging rare populations. Thus negative selection may enable discovery-based approaches and separation of cells with unknown markers.

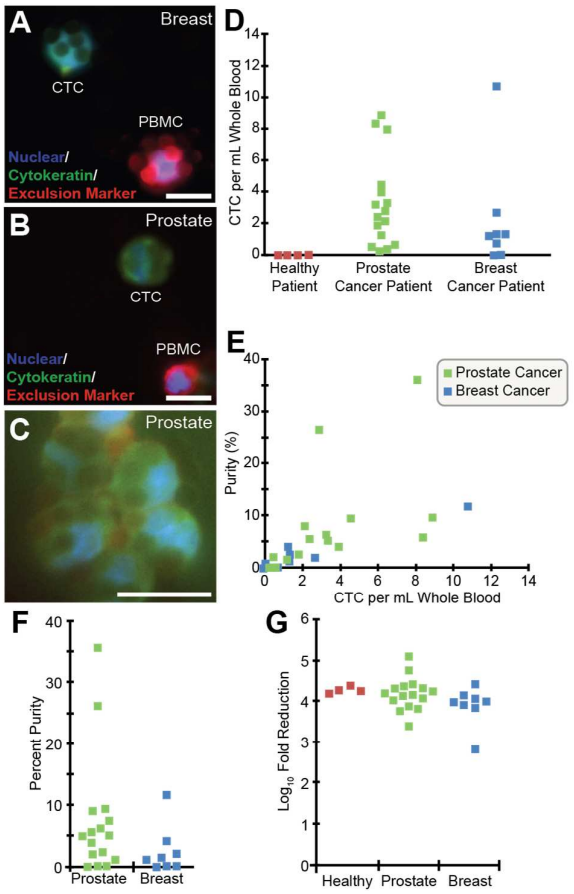

Figure 3: Positive Selection. (A, B) Image of a CTC and PBMC captured from (A) a breast cancer sample and (B) a prostate cancer sample. (C) Image of a CTC clump captured from a prostate cancer sample. Small circular distortions in the stains are artifacts of the PMPs partially attenuating the fluorescent signal. (D) Quantity of CTCs captured per milliliter of whole blood. (E) The relationship between purity and CTCs captured per mL of whole blood, using the same patients from the previous plot. (F) Percent purity observed in prostate and breast cancer patients. (G) Fold reduction of PBMCs. (A-C) Scale bars represent 10 μm. (D-G) Each dot represents a single patient and the same patients are represented across plots.

Negative Selection: CTC purification from patient samples via CD45, CD14, CD34, and CD11b depletion

Negative selection enables a discovery-based approach to rare cell isolation. To isolate the target population, negative selection removes contaminant cells via non-target markers, making the approach especially useful when target markers are unknown or shared with contaminant populations. To transition the platform to negative selection, M-270s were conjugated to antibodies targeting known non-CTC markers including CD45, CD14, CD34, and CD11b. When the negative selection mode was applied to a selection of breast cancer patient samples, negative selection yielded an average ~2 fold $\log_{10}$ reduction in PBMCs **(Fig. 4A)**. When duplicate samples underwent positive selection in parallel, an average of ~4.2 fold $\log_{10}$ reduction in PBMCs occurred (n=5). Due to the number of cells that must be successfully targeted in negative compared to positive selection, negative selection mode yielded a less pure population, similar to other platforms.[33] Using this approach, non-target cells, expressing low quantities of the selected depletion markers, may be been missed, contributing to the higher contaminant population.

Negative selection is not dependent on expression of specific markers in the target population (i.e., EpCAM may not capture CTCs undergoing EMT); thus negative selection mitigates the risk of incomplete capture of target cells, including low expressers. In patient samples, negative selection resulted in a greater number of CTCs identified (defined as nuclear events positive for cytokeratin (CK), negative for exclusion markers) (**Figure 4B**). This increase could be due to a number of factors. With positive selection, an additional selection criteria of a capture marker (i.e., capture by anti-EpCAM antibodies) is placed on CTCs; thus the additional CTCs identified with negative selection may not express (or very lowly express) the capture marker, EpCAM, preventing their capture via positive selection. Due to the suspected heterogeneity of CTCs both within a patient and patient-to-patient, the potential for EpCAM-low or EpCAM-negative CTCs could explain the increase in CTCs identified with negative selection. Additionally, the depleted population is more likely to house populations with low expression of contaminant markers. Thus, some of these cells may stain very low for contaminant markers; if the cells also stain positive for CK, they would be incorrectly included in the CTC population. Although negative selection is less effective at removing contaminants (i.e., lower purity than positive selection), negative selection minimizes the risk of missed, un-captured target cells providing a potentially larger, more complete target population to analyze.

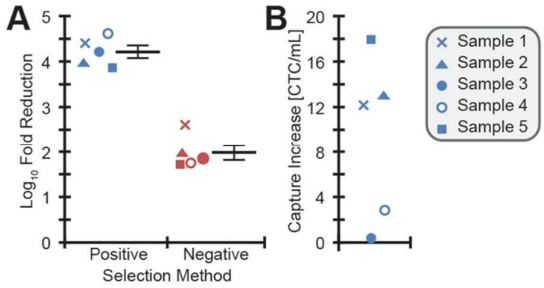

Figure 4: Negative Selection. (A) A comparison between positive and negative selection demonstrating the efficacy in reducing the background PBMC population. Each point represents a single (breast cancer) patient sample (n=5). The large line represents the average and error bars represent standard error. (B) A comparison between positive and negative selection demonstrating the increase in captured CTCs with negative selection. Capture increase was defined as the number of CTCs identified per milliliter of whole blood using positive selection subtracted from those identified using negative selection. All samples saw an increase in identified CTCs using negative selection.

Combinatorial Selection: Negative and Positive Selection

While positive selection enables the extraction of targeted, specific populations and negative selection allows for the discovery of cells with unknown identifiers, combining negative and positive selection allows 1) increased removal of contaminant populations prior to target capture, 2) the removal of contaminant cells which may also express the intended target selection marker (i.e., EpCAM), and 3) the

evaluation of potential CTC events negative for various capture markers. mTAE's non-dilutive and non-destructive approach to cell isolation and open, easily accessed format (i.e., open platform), enables selection methodologies to be readily combined. To evaluate the impact of combining positive and negative selection on the same sample, PBMCs from breast (n=6) and prostate cancer patients (n=5) were split into two samples, one for positive selection and one for sequential (negative, positive) selection. For sequential selection, depletion PMPs specific to CD45, CD14, CD34 and CD11b were incubated with the sample and depleted on mTAE followed by EpCAM selection. While, negative selection of undesired populations followed by positive selection is highlighted, readily combining positive selection markers could similarly be approached.

When positive selection was compared to sequential selection (i.e., negative selection followed by positive selection), the majority of samples yielded an increase in CTC capture with sequential selection (4/6 breast samples, 3/5 prostate samples) (**Figure 5A**). This result contrasted with the cell line characterization; characterization with cell lines identified little impact of background (0-20 million PBMCs) on specific capture (**Figure 2F, 2G**) (i.e., capture efficiency was unchanged by increasing background populations). Thus, spiking cell lines into a PBMC background may not fully mimic the complexity of patient samples. We hypothesize the reason behind the increased CTC yields following negative selection of background is due to improved interactions between the CTCs, blood cells, and PMPs. Interactions between CTCs and background blood cells may block or impair contact between PMPs and CTCs. For example, there is evidence that platelets may "cloak" or hide CTCs from the immune system that may in turn mask the CTCs from anti-EpCAM PMPs, reducing capture.[49,50] Thus, a depletion step may help to reduce these potential interactions and improve interactions between the PMPs and CTCs.

The impact of sequential selection on overall purity was more variable. Sequential selection resulted in an increase in purity in only ~45% of samples tested. The decrease in purity in a subset of samples was largely due to more contaminant cells per milliliter of whole blood being captured during the EpCAM portion of the sequential selection (compared to no sequential selection) method. The increased in nonspecific contaminant capture may be due to activation of cells during the introduction of additional PMPs leading to increased cell-PMP interactions and response (e.g., phagocytosis[51,52]).

While combinatorial selection often resulted in increased capture of contaminant cells, notably, the increase in CTCs captured was significant enough to still yield an improved purity population with dual selection in a subset of samples (5/11 samples) (**Figure 5B**). The samples that did not see benefit from depletion typically had a low initial purity with EpCAM capture alone (<1%). In contrast, a subset of patients, which generally had positive selection purity of >1%, benefited greatly from depletions, with sequential EpCAM selection improving the end purity of the sample as well as CTC yields (**Figure 5C**). While it is unlikely that a sample with <1% purity would have much value beyond enusmeration, the enhanced purity and CTC capture of dual selection to a subset of the samples may be sufficient to integrate with additional downstream processes making dual selection a key asset in moving beyond enumeration for these samples. Similarly interesting, would be the use of the platform to pursue multiple positive selection markers to fully assess the potential heterogeneity of CTC markers available.

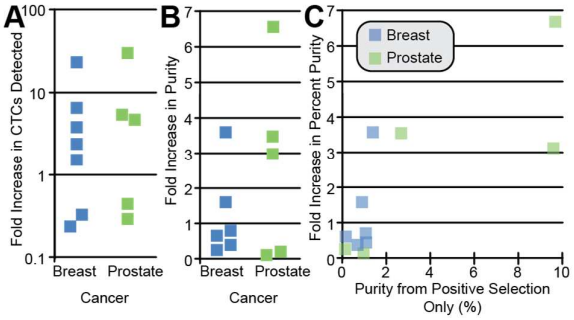

Figure 5: Combinatorial selection. (A) Fold increase in CTCs detected when comparing CTCs captured with combinatorial to positive selection. The majority of samples from both breast and prostate saw an increase in CTCs detected. (B) Fold increase in the captured population's percent purity from combinatorial selection to positive selection. Variable improvements in percent purity were observed across samples. (C) Relationship between positive selection and combinatorial selection purity demonstrating samples with higher positive selection purities (>1%) were more likely to see improved purity with combinatorial selection.

Automated Nucleic Acid Extraction Protocols: Development and Validation

To obtain the most information from rare cell populations, isolation is only the first step, often followed by molecular extraction and analysis (e.g., RNsA, DNA). These downstream processes and analyses often come with added requirements in order to provide accurate cellular information unbiased by contaminant populations; requirements often include high yields (i.e., low loss) and high purity. Pipette-based transfer of samples between systems is inherently prone to loss (residual volume in the pipette tips, lost volume in wells resulting from transfer), reducing material with each transfer. Thus to minimize sample loss and fully utilize the platform's capabilities, RNA and DNA extraction processes were both integrated into mTAE.

To validate NA extraction, mTAE was benchmarked against commercial, non-PMP NA isolation products (e.g., spin columns) and a manual tube-based alternative, using samples of ~10 to ~1,000 LNCaPs **(Fig. 6A & 6B)**. Extracted RNA and DNA were quantified by qRT-PCR or qPCR respectively. Upon comparison, RNA yields via mTAE extraction were within one cycle of alternative methods. Each method of DNA extraction resulted in statistically indistinguishable yields (p-value > 0.05). Next, we tested whether the presence of M-270s - used for cell isolation - impacted NA acid yields. NA was extracted from cells (LNCaPs) that were first bound to and captured by M-270s; the yields were then compared to NA extraction of the cells alone (i.e., without M-270s). On average, the presence of M-270s increased RNA yield **(S7)**; however, the presence of M-270s had an adverse impact on DNA, significantly impacting

DNA yield **(S7)**. We hypothesized the loss in DNA was due to DNA irreversibly binding to the M-270s, preventing detection and quantification. We circumvent this issue by first lysing the M-270-bound cells in LiDS buffer, magnetically removing the M-270s, and adding the DNA lysis buffer and PMPs. This method of pre-lysis depletion resulted in equivalent yields to DNA extraction from cells alone. As the platform continues through development for potential clinical studies, internal assay controls could be integrated. To monitor reproducibility of DNA or RNA extraction and detection, a non-human DNA/RNA sequence could be added to the lysate, validating NA capture and extraction for every sample.

In CTCs from patients with prostate cancer, detection and quantification of AR gene expression, AR variants, and downstream targets can provide insight into the pathways being used by the cancer and therefore potential therapeutic targets (i.e., prediction of patient response to AR-targeted therapy).[53] Combining cell isolation with RNA extraction, ~10 LNCaPs were captured via EpCAM and RNA extracted on mTAE for detection of AR transcripts and, similarly 22Rv1 cells, a prostate epithelial cell line known to express AR splice variants V1 and V7. Results highlighted the ability of mTAE to combine capture of low numbers of target cells (e.g. down to 10 cells) with RNA extraction and variant detection. Results helped to support the specificity of the detected transcripts to specific cell types (**S12**), known to express specific variants, specifically demonstrating increased detection of the AR splice variants V1 and V7 in 22Rv1s in comparison to LNCaPs (Figure 6C).

Following cell line characterization, mRNA was extracted from cells captured on the platform from prostate cancer patients **(S11)** following positive selection via EpCAM. Detected transcripts (**S12**, **S8**) include AR variants and downstream AR transcripts. The detection of these transcripts supports mTAE's ability to not only capture CTCs (through EpCAM enrichment), but also detect therapy relevant transcripts from the captured cell population if present (Figure 6D). While the detection of AR transcripts from patients serves as a first step towards validation of the platform's ability to integrate cell isolation with rare transcript detection, future evaluations against baseline AR variant expression and existing assays, which detect AR and AR transcripts (e.g., AR V7) would further benchmark the presented assay against existing clinical assays (e.g., Epic Sciences' AR-V7 CTC Liquid Biopsy Test).

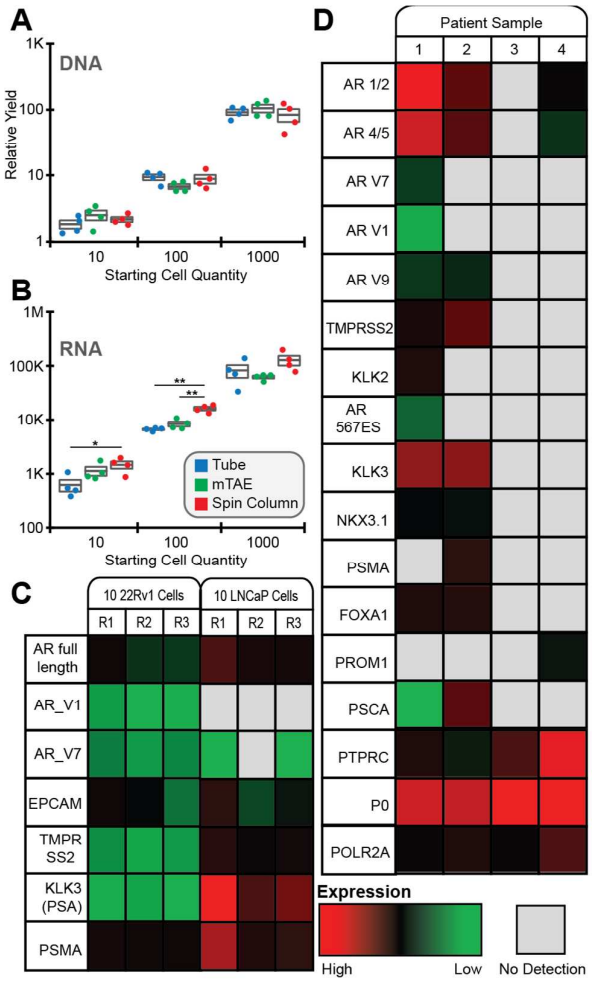

Figure 6: Automated NA acid extraction on mTAE. (A) DNA extraction from 10, 100, and 1,000 cells compared across three methods: manual tube-based method (i.e., PMP-based capture followed by washing of NA bound PMPs performed in a micro-centrifuge tube using a magnetic tube rack), mTAE, and Qiagen QIAmp DNA Mini spin columns. (B) RNA extracted from 10, 100, 1000 cells compared across three methods: manual tube-based method, mTAE, and Qiagen RNeasy Plus Micro. (A,B) Each point represents a replicate (n = 4 per condition). Boxes demonstrate average and standard error; statistical significance based on T-Test is represented by * $P < 0.05$ and ** $P < 0.005$ (C) Normalized relative detection of prostate-related transcripts from approximately 10 22Rv1s and 10 LNCaPs demonstrating greater AR V1 and AR V7 detection in 22Rv1s (normalized to RPII housekeeping). (D) Quantitative RT-PCR detection of prostate-related transcripts from prostate cancer patient CTC samples including AR V1 and AR V9 (Ct values represented as a heat map). Cells were captured on mTAE followed by on-chip RNA extraction.

Conclusions

An automated PMP-based sample preparation platform, mTAE facilitates flexibility in cell isolation while reducing user-to-user variation through process automation. Able to process four samples in parallel, mTAE combines positive and negative selection through exclusion-based sample preparation (ESP), a purification method where PMPs are drawn through a phase interface to cleanly isolate bound material and leave behind a minimally perturbed sample. CTCs – identified as cellular events that contained a nuclei, stained positive for CK, and

negative for contaminant markers (CD45, CD14, CD11b, CD34) – were successfully isolated from prostate and breast cancer patients samples via EpCAM-based positive selection. Positive selection resulted in improved depletion of contaminant cells and thus higher purity when compared to negative selection on mTAE. Negative selection via removal of CD45, CD14, CD34, and CD11b populations, yielded a lower purity than positive selection, but a greater number of identified CTCs. Arguably most importantly, mTAE allows users to integrate positive and negative selection on a single sample, enabled by the non-dilutive, non-destructive method of sample preservation during processing.

The flexibility of the platform expands beyond cell selection methodologies, allowing integration of additional washing for higher stringency, staining for downstream analysis or cellular identification, or additional processing to isolate specific biomarkers, including RNA and DNA. By integrating these capabilities, samples can be stained within the platform, minimizing sample loss. Using mTAE, CTC isolation and identification was evaluated from prostate cancer and breast cancer patients; specificity was confirmed via healthy donor blood. The platform's capacity to perform RNA and DNA extraction was then validated against commercial alternatives demonstrating comparable results. mTAE was then used to extract RNA from CTCs isolated from prostate cancer patients, leading to detection of AR (including AR variants V1, V7 and AR-driven transcripts. The flexibility provided to the user from cell selection methodologies to downstream analysis paired with the compatibility of accessible, commercial components (e.g., commercial magnetic beads, antibodies, plastic consumables) makes mTAE a readily accessible platform in the cell isolation field.

Acknowledgements

We would like to thank all patients who participated in this study. We are also grateful for the help of the UWCCC clinical research group, especially Jamie Wiepz, Kelly Bush, Amy Forsyth, Dorothea Horvath, Jane Straus, Mary Jane Staab, Dr. Glenn Liu, Dr. Douglas McNeel, Dr. Christos Kyriakopolous, Dr. C. Pettaway, and Dr. George Wilding.

This work was supported by a Movember-Prostate Cancer Foundation Challenge Award (to J.M. Lang and D.J. Beebe), the Bill & Melinda Gates Foundation through the Grand Challenges in Global Health Initiative (to S.M. Berry and D.J. Beebe), NIH grant #1R01CA181648 (to S.M. Berry and J.M. Lang), Department of Defense PCRP grant #W81XWH-12-1-0052 (to J.M. Lang), Department of Defense Synergistic Idea Development Award BC 150425 (to J.M. Lang), NIH grant #5R33CA137673 (to D.J. Beebe), University of Wisconsin Carbone Cancer Center Support Grant NIH P30 (to D.J. Beebe, S.M. Berry, D.J. Guckenberger) and NSF GRFP DGE-0718123 (to D.J. Beebe).

Author Contributions:

Contributions by Hannah M. Pezzi included leading investigation and writing – original draft. David J. Guckenberger contributed by supporting investigation and leading methodology. Jennifer L. Schehr assisted in validation. Jacob Rothbauer supported methodology. Charlotte Stahlfeld, Anupama Singh, and Sacha Horn all assisted in resources (e.g., patient sample preparation). Jamie M. Sperger provided validation. Scott M. Berry was involved in supporting conceptualization and funding acquisition. Joshua M. Lang contributed through funding acquisition, project administration, and supervision. David J. Beebe contributed in funding acquisition and project administration. David J. Guckenberger, Jennifer L. Schehr, Scott M. Berry, Joshua M. Lang, and David J. Beebe all contributed to writing – review & editing.

Notes and references

1 Massberg, S. *et al.* Immunosurveillance by Hematopoietic Progenitor Cells Trafficking through Blood, Lymph, and Peripheral Tissues. *Cell* **131**, 994-1008, doi:http://dx.doi.org/10.1016/j.cell.2007.09.047 (2007).

2 Bianchi, D. W., Zickwolf, G. K., Weil, G. J., Sylvester, S. & DeMaria, M. A. Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum. *Proceedings of the National Academy of Sciences* **93**, 705-708 (1996).

3 Bianchi, D. W., Flint, A. F., Pizzimenti, M. F., Knoll, J. H. & Latt, S. A. Isolation of fetal DNA from nucleated erythrocytes in maternal blood. *Proceedings of the National Academy of Sciences* **87**, 3279-3283 (1990).

4 Plaks, V., Koopman, C. D. & Werb, Z. Circulating Tumor Cells. *Science* **341**, 1186-1188, doi:10.1126/science.1235226 (2013).

5 Chun, T.-W. *et al.* Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 13193-13197 (1997).

6 Simpson, J. L. & Elias, S. Isolating fetal cells from maternal blood. Advances in prenatal diagnosis through molecular technology. *Jama* **270**, 2357-2361 (1993).

7 Saker, A. *et al.* Genetic characterisation of circulating fetal cells allows non-invasive prenatal diagnosis of cystic fibrosis. *Prenatal diagnosis* **26**, 906-916, doi:10.1002/pd.1524 (2006).

8 Seppo, A. *et al.* Detection of circulating fetal cells utilizing automated microscopy: potential for noninvasive prenatal diagnosis of chromosomal aneuploidies. *Prenatal diagnosis* **28**, 815-821, doi:10.1002/pd.1987 (2008).

9 Chun, T. W. *et al.* Decay of the HIV reservoir in patients receiving antiretroviral therapy for extended periods: implications for eradication of virus. *J Infect Dis* **195**, 1762-1764, doi:10.1086/518250 (2007).

10 Lang, J. M., Casavant, B. P. & Beebe, D. J. Circulating tumor cells: getting more from less. *Science translational*

medicine **4**, 141ps113, doi:10.1126/scitranslmed.3004261 (2012).

11 Casavant, B. P. *et al.* The VerIFAST: an integrated method for cell isolation and extracellular/intracellular staining. *Lab on a Chip* **13**, 391-396 (2013).

12 Danila, D. C. *et al.* TMPRSS2-ERG status in circulating tumor cells as a predictive biomarker of sensitivity in castration-resistant prostate cancer patients treated with abiraterone acetate. *European urology* **60**, 897-904, doi:10.1016/j.eururo.2011.07.011 (2011).

13 Pope, N. M., Alsop, R. C., Chang, Y.-A. & Smith, A. K. Evaluation of magnetic alginate beads as a solid support for positive selection of CD34+ cells. *Journal of Biomedical Materials Research* **28**, 449-457, doi:10.1002/jbm.820280407 (1994).

14 Schuler, P. J., Harasymczuk, M., Schilling, B., Lang, S. & Whiteside, T. L. Separation of human CD4+CD39+ T cells by magnetic beads reveals two phenotypically and functionally different subsets. *Journal of Immunological Methods* **369**, 59-68, doi:http://dx.doi.org/10.1016/j.jim.2011.04.004 (2011).

15 Rye, P. D. Sweet and Sticky: Carbohydrate-Coated Magnetic Beads. *Nat Biotech* **14**, 155-157 (1996).

16 Lara, O., Tong, X., Zborowski, M. & Chalmers, J. J. Enrichment of rare cancer cells through depletion of normal cells using density and flow-through, immunomagnetic cell separation. *Experimental hematology* **32**, 891-904, doi:10.1016/j.exphem.2004.07.007 (2004).

17 King, J. D., Casavant, B. P. & Lang, J. M. Rapid translation of circulating tumor cell biomarkers into clinical practice: technology development, clinical needs and regulatory requirements. *Lab Chip* **14**, 24-31, doi:10.1039/c3lc50741f (2014).

18 Konigsberg, R. *et al.* Detection of EpCAM positive and negative circulating tumor cells in metastatic breast cancer patients. *Acta oncologica (Stockholm, Sweden)* **50**, 700-710, doi:10.3109/0284186x.2010.549151 (2011).

19 Armstrong, A. J. *et al.* Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers. *Molecular cancer research : MCR* **9**, 997-1007, doi:10.1158/1541-7786.mcr-10-0490 (2011).

20 Kallergi, G. *et al.* Epithelial to mesenchymal transition markers expressed in circulating tumour cells of early and metastatic breast cancer patients. *Breast Cancer Res* **13**, doi:10.1186/bcr2896 (2011).

21 Mego, M. *et al.* Expression of epithelial-mesenchymal transition-inducing transcription factors in primary breast cancer: the effect of neoadjuvant therapy. *Int J Cancer* **130**, doi:10.1002/ijc.26037 (2012).

22 Yu, M. *et al.* Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. *Science* **345**, 216-220, doi:10.1126/science.1253533 (2014).

23 Poudineh, M. *et al.* Tracking the dynamics of circulating tumour cell phenotypes using nanoparticle-mediated magnetic ranking. **12**, 274-281, doi:10.1038/nnano.2016.239 (2017).

24 Gorges, T. M. *et al.* Circulating tumour cells escape from EpCAM-based detection due to epithelial-to-mesenchymal transition. *BMC Cancer* **12**, 178, doi:10.1186/1471-2407-12-178 (2012).

25 Gradilone, A. *et al.* Circulating tumour cells lacking cytokeratin in breast cancer: the importance of being mesenchymal. *Journal of cellular and molecular medicine* **15**, 1066-1070, doi:10.1111/j.1582-4934.2011.01285.x (2011).

26 Kasimir-Bauer, S., Hoffmann, O., Wallwiener, D., Kimmig, R. & Fehm, T. Expression of stem cell and epithelial-mesenchymal transition markers in primary breast cancer patients with circulating tumor cells. *Breast Cancer Research* **14**, R15, doi:10.1186/bcr3099 (2012).

27 Kasimir-Bauer, S. *et al.* Survival of tumor cells in stem cell preparations and bone marrow of patients with high-risk or metastatic breast cancer after receiving dose-intensive or high-dose chemotherapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* **7** (2001).

28 Arafat, W. *et al.* Intra-patient heterogeneity in urothelial cancer (UC) circulating tumor cells (CTC) and PDL1 expression to identify biomarkers of response and new therapeutic targets: A pilot study. *Journal of Clinical Oncology* **35**, 4537-4537, doi:10.1200/JCO.2017.35.15_suppl.4537 (2017).

29 Mackall, C. L. *et al.* Lymphocyte depletion during treatment with intensive chemotherapy for cancer. *Blood* **84**, 2221-2228 (1994).

30 Savary, C. A. *et al.* Multidimensional flow-cytometric analysis of dendritic cells in peripheral blood of normal donors and cancer patients. *Cancer Immunology, Immunotherapy* **45**, 234-240 (1997).

31 Sperger, J. M. *et al.* Integrated analysis of multiple biomarkers from circulating tumor cells enabled by exclusion-based analyte isolation. *Clinical Cancer Research*, doi:10.1158/1078-0432.ccr-16-1021 (2016).

32 Stott, S. L. *et al.* Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. *Proceedings of the National Academy of Sciences* **107**, 18392-18397, doi:10.1073/pnas.1012539107 (2010).

33 Ozkumur, E. *et al.* Inertial Focusing for Tumor Antigen–Dependent and –Independent Sorting of Rare Circulating Tumor Cells. *Science translational medicine* **5**, 179ra147-179ra147, doi:10.1126/scitranslmed.3005616 (2013).

34 Berry, S. M. *et al.* AirJump: Using Interfaces to Instantly Perform Simultaneous Extractions. *ACS applied materials & interfaces* **8**, 15040-15045, doi:10.1021/acsami.6b02555 (2016).

35 Casavant, B. P. *et al.* The VerIFAST: an integrated method for cell isolation and extracellular/intracellular staining. *Lab Chip* **13**, 391-396, doi:10.1039/c2lc41136a (2013).

36 Casavant, B. P., Guckenberger, D. J., Beebe, D. J. & Berry, S. M. Efficient Sample Preparation from Complex Biological Samples Using a Sliding Lid for Immobilized Droplet Extractions. *Analytical Chemistry* **86**, 6355-6362, doi:10.1021/ac500574t (2014).

37 Guckenberger, D. J. *et al.* Magnetic System for Automated Manipulation of Paramagnetic Particles. *Analytical Chemistry* **88**, 9902-9907, doi:10.1021/acs.analchem.6b02257 (2016).

38 Schehr, J. L. *et al.* High Specificity in Circulating Tumor Cell Identification Is Required for Accurate Evaluation of

Programmed Death-Ligand 1. *PLOS ONE* **11**, e0159397, doi:10.1371/journal.pone.0159397 (2016).
39 Zasadil, L. M. *et al.* High rates of chromosome missegregation suppress tumor progression but do not inhibit tumor initiation. *Molecular biology of the cell* **27**, 1981-1989, doi:10.1091/mbc.E15-10-0747 (2016).
40 Stott, S. L. *et al.* Isolation and characterization of circulating tumor cells from patients with localized and metastatic prostate cancer. *Science translational medicine* **2**, 25ra23-25ra23 (2010).
41 Miller, M. C., Doyle, G. V. & Terstappen, L. W. Significance of circulating tumor cells detected by the CellSearch system in patients with metastatic breast colorectal and prostate cancer. *Journal of oncology* **2010** (2009).
42 Königsberg, R. *et al.* Detection of EpCAM positive and negative circulating tumor cells in metastatic breast cancer patients. *Acta oncologica* **50**, 700-710 (2011).
43 Cohen, S. J. *et al.* Isolation and characterization of circulating tumor cells in patients with metastatic colorectal cancer. *Clinical colorectal cancer* **6**, 125-132 (2006).
44 Cohen, S. J. *et al.* Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. *J Clin Oncol* **26**, 3213-3221, doi:10.1200/jco.2007.15.8923 (2008).
45 Schulze, K. *et al.* Presence of EpCAM-positive circulating tumor cells as biomarker for systemic disease strongly correlates to survival in patients with hepatocellular carcinoma. *Int J Cancer* **133**, 2165-2171, doi:10.1002/ijc.28230 (2013).
46 Aceto, N. *et al.* Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110-1122, doi:10.1016/j.cell.2014.07.013 (2014).
47 Cristofanilli , M. *et al.* Circulating Tumor Cells, Disease Progression, and Survival in Metastatic Breast Cancer. *New England Journal of Medicine* **351**, 781-791, doi:10.1056/NEJMoa040766 (2004).
48 de Bono, J. S. *et al.* Circulating Tumor Cells Predict Survival Benefit from Treatment in Metastatic Castration-Resistant Prostate Cancer. *Clinical Cancer Research* **14**, 6302-6309, doi:10.1158/1078-0432.ccr-08-0872 (2008).
49 Leblanc, R. & Peyruchaud, O. Metastasis: new functional implications of platelets and megakaryocytes. *Blood* **128**, 24-31, doi:10.1182/blood-2016-01-636399 (2016).
50 Lou, X. L. *et al.* Interaction between circulating cancer cells and platelets: clinical implication. *Chinese journal of cancer research = Chung-kuo yen cheng yen chiu* **27**, 450-460, doi:10.3978/j.issn.1000-9604.2015.04.10 (2015).
51 Cannon, G. J. & Swanson, J. A. The macrophage capacity for phagocytosis. *Journal of Cell Science* **101**, 907-913 (1992).
52 Dunn, P. A. & Tyrer, H. W. Quantitation of neutrophil phagocytosis, using fluorescent latex beads. *The Journal of Laboratory and Clinical Medicine* **98**, 374-381, doi:10.5555/uri:pii:0022214381900433.
53 Li, Y. *et al.* Androgen receptor splice variants mediate enzalutamide resistance in castration-resistant prostate cancer cell lines. *Cancer Res* **73**, 483-489, doi:10.1158/0008-5472.can-12-3630 (2013).