Developing Tools for Interrogation of Clinically Significant Rare Cell Populations

By

Hannah M. Pezzi

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Biomedical Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of final oral examination: 12/15/2017

The dissertation is approved by the following members of the Final Oral Committee:
    David J. Beebe, Professor, Biomedical Engineering
    Joshua Lang, Assistant Professor, Hematology/Oncology
    Robert Striker, Associate Professor, Infectious Disease
    Nathan Sherer, Associate Professor, Molecular Virology and Oncology
    William Rehrauer, Associate Professor, Pathology and Laboratory Medicine

Acknowledgements

Throughout this process I have had the support of incredible individuals to help me through, both professionally and personally. I am extremely grateful for the opportunity I have had in Dr. David Beebe's laboratory. Under Dave's guidance I have developed into an independent scientist able to adapt to new ideas and integrate solutions into scientific problems. I am grateful for the provided freedom, which allowed me to pursue my own ideas and develop as a scientist. Thank you to Dr. Joshua Lang - both for your participation on my committee and for your ability to motivate, reminding me of the patients behind the science. To Dr. Robert Striker, thank you for your mentorship both on my committee and your continual support and enthusiasm in pursuing new ideas.

To my committee members specifically – Dr. David Beebe, Dr. Joshua Lang, Dr. Robert Striker, Dr. Nathan Sherer, and Dr. William Rehrauer - thank you for your enthusiasm and critical evaluation of my research.

To all the members of the Microtechnology, Medicine and Biology Lab, past and present, thank you for creating a thought-provoking and positive work environment in which to problem solve. The work environment that was built by the members of this lab has made my graduate school both a productive, enjoyable, and innovative experience that I am extremely grateful for. I would like to especially thank those that I had the privilege of working closely with in various settings: Dr. Scott Berry, Molly Morgan, Pete Geiger, Dr. Farsh Moussavi-Harami, Dr. Won Hong, Dr. Jennifer Schehr, Dr. Jamie Sperger, Jennifer Loeb, and Dr. David Niles. I would also like to personally thank Alice Puchalski; her hard work behind the scenes ensured the continuation of my position as well as made my research possible.

Thank you also to those who supported me from the Department of Biomedical Engineering at the University of Wisconsin-Madison for my undergraduate career. Without a strong foundation in both science and engineering I would have been unable to progress to this level. Additionally, thank you to Dr. Pamela Kreeger, for providing me with my first laboratory position and instilling a me a strong foundation in biological techniques and basic science, which has helped me every step of the way.

To my friends and family, thank you for your constant support and dedication to my own goals and as always, your loving support. You have installed in me an ability to work hard and believe that anything I set my

mind to is possible. Above all thank you to my parents, Al and Karen Pezzi, and my brother, Paul Pezzi for all of your support and love throughout the years. You truly made what I have accomplished possible and have been there every step of the way. I owe much of this incredible journey to you and your guidance making me into the person I have become.

To John Guckenberger, thank you for being both my colleague within the lab as well as my partner outside. Without you this would have been an incredibly daunting task – your support throughout this process has been incredible.

Abstract

Rare cell populations are a potential new tool for assessing disease progression, treatment, and monitoring while also providing the opportunity for new prognostic and diagnostic biomarkers. Hidden within large, diverse contaminating populations, rare cell populations present challenges in obtaining and assessing the information they contain. Physically or analytically, rare cell populations must be identified and distinguished from contaminant populations to obtain and apply the information they hold.

Chapter 1 serves as an introduction to rare cell populations highlighting the challenges associated with rare cell populations. The chapter is presented as a review, highlighting the obstacles in specifically accessing and identifying rare cells and adapting to the heterogeneity presented in and across rare cell populations, exemplified through case studies. Despite their clinical potential, utilization of rare cell populations requires assays able to overcome the challenges presented by each rare cell population, challenges that change and emerge with each population. The information highlighted in the review serves as a foundation for designing and adapting the next generation of rare cell technologies.

The remaining articles focus on engineering, technologies, and processes to better target and assess both rare cell populations and the analytes they contain. Magnetic beads have proven a critical advance in the specific targeting and manipulation of rare cell populations; through combination with antibodies, magnetic beads enable the targeting of nearly any population. Chapter 2 highlights the variable efficacy in integrating different commercial magnetic beads into rare cell isolation protocols, including analysis of their impact on downstream applications through assessment of fluorescent immunostaining for protein expression and localization, post-capture culture, and impact on nucleic acid isolations. Utilizing the bead type that demonstrated the highest specific capture (and lowest nonspecific binding of non-target cells), Chapter 3 introduces a rare cell isolation platform, which is evaluated with circulating tumor cells. This platform leverages Exclusion-based Sample Preparation (ESP) to facilitate positive, negative, and combinatorial selection of cell populations, and combines cell isolation with downstream immunostaining and nucleic acid extraction capabilities. Expanding on the targeting of rare cell populations, Chapter 4 and

5 investigate an alternative rare cell population: the HIV viral reservoir. Lacking differential external protein markers, the HIV viral reservoir requires alternative approaches (typically specific targets following activation) to identify and quantify this rare cell population. In Chapter 4, ESP is used to facilitate multiplexed RNA extractions, which increases sensitivity in detecting rare RNA transcripts from large cellular inputs. In Chapter 5, an alternative approach to reservoir quantification is pursued, specifically detection of reservoir cells capable of producing infectious virus. Here, a dual fluorescent reporter cell line is used to provide a sensitive readout enabling the live visualization of a single infection event. Building on the reporter readout, an alternative culture approach was shown to enhance infection rates, further enhancing sensitivity of the readout.

Table of Contents:

Table of Figures:

Chapter 1   Introduction[1]

The discovery of a new rare cell population continually excites both the scientific community and general public as a new facet of clinical information. A potential tool for diagnosis, treatment monitoring, or prognosis, rare cell populations provide an edge with which to augment clinical care. Despite the early-envisioned potential, the transition of rare cell populations into the clinic is challenged by *access, identification, and heterogeneity*; strongly interdependent, these concepts challenge our understanding of rare cell populations. Access, both in obtaining the sample from the patient and in accessing the RcP within, meter the population's clinical potential. Even once obtained, the rarity of these populations and diversity of background populations complicates the ability to successfully identify and target the rare population of interest. Once identified, heterogeneity (patient-to-patient to cell-to-cell) adds inherent variation across the population - variation, which could complicate identification or provide clinically significant information. Here we review the core concepts – access, identification, and heterogeneity – that impact the use and understanding of rare cell populations and provide case studies to evaluate the lessons hidden within a specific rare cell population.

---

1 This chapter has been adapted from the following manuscript in preparation: "Embracing the Rare Cell Chalange" Hannah M. Pezzi, Molly M. Morgan, and David J. Beebe

2 This chapter has been adapted from the following manuscript under revision at ACS Omega: "Integration

1.1   Introduction

The human body is a consortium of cell populations (e.g., immune, epithelial, endothelial…), existing in continuous flux. Driven by patient-specific conditions - such as cancer[1], pregnancy[2,3], and viral infections[4,5] - cell populations undergo modifications, shifts in size, and population divisions. The complex, ever-changing assortment of cell populations within the body has facilitated the continual discovery of new rare cell populations (RcPs). Often connected to a clinical condition (i.e., cancer[1], pregnancy[2,3], viral infections[4,5]), RcPs provide a new front with which to inform, influence, and modulate condition understanding and clinical care. Despite their potential, RcPs are small populations of cells vastly outnumbered by surrounding populations, often on the order of one rare cell per million (or even billion) background cells. Hidden by the overwhelming surrounding contaminant cell populations, the rarity of RcPs challenges both basic understanding and interrogation of these populations.

Driven by the potential power of RcPs in augmenting clinical care, a suite of RcP-focused technologies has emerged. Frequently due to the unique challenges presented by each population, technologies are specifically designed and tailored to a single, highly specific rare cell population (e.g., Herringbone chip [6], CTC-ichip [7], VERSA [8], CellSearch target circulating tumor cells (CTC) [9]; list assays/tech target the HIV viral reservoir; TRIC[10] targets fetal cells -TRIC). These technologies have facilitated an initial understanding and the utility of RcPs, with some success in transitioning RcPs from the bench to the clinic (e.g., FDA-approved CellSearch platform for CTC enumeration). Though much information and understanding has been made from the initial wave of rare cell platforms and technologies, these technologies highlighted the true challenges associated with understanding, targeting, and utilization of RcPs. As a result, RcPs still face an arduous path to clinical implementation.

The next generation of RcP technologies will have to overcome the challenges made evident in early RcP technology approaches: *access*, *identification*, and *heterogeneity* (**Fig. 1.1**). Dependent on both sample location and type, *access* to not only the sample, but also the RcP is required; to utilize the RcP, *identification* of the RcP is required to differentiate the RcP from a vast background of similar cells; yet,

heterogeneity makes RcPs an ever evolving target. Here, we evaluate and guide readers through the obstacles associated with RcP *access, identification, and heterogeneity*. Case studies are integrated into the discussions to improve understanding of the role these challenges have in shaping a clinically relevant RcP. With an improved understanding of the challenges facing RcPs, the next generation of RcP technologies may facilitate the transition of RcPs from research to clinics for improved patient care.
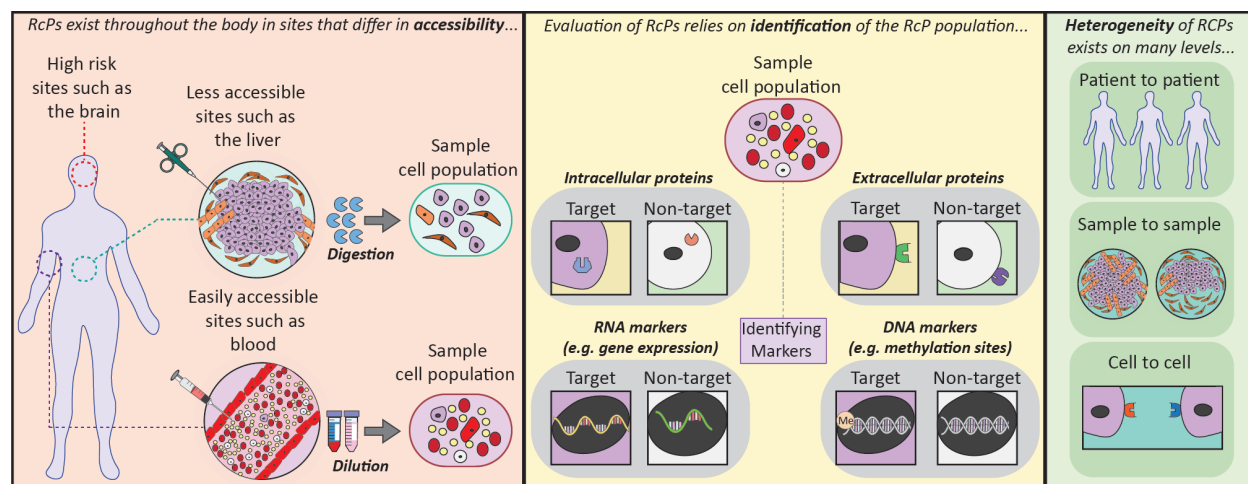
Fig. 1.1 Schematic overview of the challenges facing rare cell populations. Challenges in rare cell population access exist both in obtaining a sample from the patient and from accessing the specific cell population within the sample. To identify rare cell populations, differential features (e.g., in extracellular or intracellular protein, RNA, DNA) are necessary to determine target from non-target populations. Heterogeneity presents challenges in adapting to the variation introduced into the population, patient-to-patient, sample-to-sample, and cell-to-cell.

1.2 Accessing Rare Cell Populations

RcP accessibility dictates both the research and future clinical potential of the RcPs. Residing in a range of tissues and organs - including the liver [11], lung [12], brain [13,14], lymph nodes [15,16], cervix[10], and blood [17]– RcPs must first be accessed to utilize the carried information. To fully access the population, researchers and clinicians must consider both primary (extraction of the tissue from the body) and secondary (extraction of RcPs from the tissue) modes of acquisition to determine the potential value (e.g., risk factor, information quality, information value) in pursuing each RcP for clinical applications.

While RcPs exist throughout the body, the ultimate clinical potential of RcP is a balance of the sample acquisition risk to the patient and the potential value and quality of information obtained. Accessing the RcPs within essential organs (e.g., myocardial stem cells in the heart [18], cancer stem cells in

the brain [13,14]) has little clinical appeal as their location limits their ability to be safely sampled. Even if not located within a critical tissue, accessibility can be limited by the invasiveness of the sampling technique (i.e., risk of surgical complication, blood clot, infection). Due to the invasiveness associated with sample many conditions directly (e.g., risk to the fetus in direct sampling), the most sought-after RcPs are the ones that provide insight into conditions that are difficult or dangerous to directly or repeatedly sample (e.g., fetal health, cancer). Clinically, the most exciting RcPs emerge out of an easy-to-sample tissue, requiring minimally invasive acquisition techniques with the potential to provide information on an otherwise challenging condition to monitor.

After a tissue sample is acquired, direct access to the RcP is not guaranteed. Human samples are a complex assortment of matrix (e.g., collagen, fibronectin), biological factors (e.g. mucus), and varied populations of cells (e.g., immune, endothelial, epithelial). Consequently extracting information, let alone a specific cell population, from this background can be extremely difficult. If existing within a solid matrix (i.e., solid tissue sample), digestion may be required to extract the RcP. Digestion requires an extensive understanding of the target population, the tissue in which it resides, and tissue-specific digestion protocols to physically access the population. Tissue composition can also be impacted by disease state [19,20], which impacts digestion efficiencies across patient cohorts. Fundamentally, tissue digestion protocols are often labor intensive, long, and require extensive manipulation of the sample, leading to diminished cell yields and viability. From the time the sample is obtained until the population is analyzed, alterations can occur in the RcP, impacting protein expression and RNA transcripts within the cells (SOURCE). To assess the most representative population, handling of the sample and RcP manipulation must be streamlined to reduce the potential risk of modifying the population prior to analysis. Due to their complexity, solid tissue samples are likely to contain yet undiscovered RcPs; despite their potential for RcPs, discovery and progress are likely, at least initially, to be hindered by processing limitations.

One tissue, the blood, has emerged as the predominant frontier for RcPs with high potential for clinical utility. Peripheral blood can be accessed in large volumes (tens to hundreds of milliliters), with little time-to-recovery for the patient (on the order of days). Acquired through low-risk minimally invasive venipuncture, blood provides a snapshot into the body that can be repeatedly captured. As a cell suspension, the populations are easily accessible with minimal manipulation of the sample (e.g., no enzymatic digestion). Interrogation of the bloodstream has provided numerous clinically intriguing "circulating" rare cell populations with known and anticipated clinical readouts (e.g., fetal cells [21,22], stem cells [23], circulating tumor cells [17], viral reservoirs [24,25]). While a physically accessible sample, the complexity and diversity of blood populations also presents challenges in accessing the RcPs – a challenge exemplified in the case of prenatal testing.

1.3   Accessing Rare Cell Populations Case Study: Prenatal Diagnosis

Prenatal testing for fetal abnormalities - including chromosomal abnormalities (e.g., Down's Syndrome, Fragile X Syndrome) as well as inherited genetic disorders (e.g., cystic fibrosis, Tay-Sachs disease)[26,27] – requires assessment of fetal material. Due to the risk to the fetus in traditional methods of accessing fetal material[28,29], the field of prenatal testing has driven the discovery and evaluation of alternative sources of fetal material (e.g., bloodstream, cervical canal). The tradeoffs between 1) the risk to the patient (e.g., the mother as well as the fetus), 2) the availability of material (e.g., tissue digestions, contaminant maternal information), and 3) the quality of the information received (e.g. confidence in the results) underscores the clinical challenges with RcP accessibility. Overall, identifying the best method - the method that poses the least risk and offers the highest quality information - is key to maximizing the resultant benefit for the patient. Here we compare different sources of fetal material for prenatal diagnostics and evaluate the tradeoffs in sample acquisition risk, accessibility of the target analyte for assessment, and the quality of information obtained.

Amniocentesis and chorionic villus sampling - Traditionally, prenatal diagnosis is performed via invasive sampling of fetal material through amniocentesis or chorionic villus sampling (CVS).

Amniocentesis, performed during the second trimester, requires sampling of the amniotic fluid and is used to detect neural tube defects, genetic disorders, as well as aneuploidies. CVS evaluates the chorionic villi of the placenta and can be conducted earlier than amniocentesis (between 9.5 and 12.5 weeks of gestation), but is less informative than amniocentesis. Both sampling approaches provide extremely pure fetal material and highly reliable results. Unfortunately, both approaches are associated with high risk factors incurred by the invasiveness of the procedure; fetal loss is observed in approximately 1% of cases that use amniocentesis or CVS [28,29]. Due to the risks associated with accessing fetal material through CVS and amniocentesis, there is interest in developing alternative, lower risk tests to diagnose fetal abnormalities [30]. Much of this effort has led to the discovery and assessment of rare fetal analytes within the maternal bloodstream and cervical canal.

Cell free fetal DNA in maternal blood - In 1997, fetal DNA was found to circulate in the maternal bloodstream [31]. Collected via venipuncture, assessment of cell free fetal DNA (fetal cfDNA) presents an attractive alternative to traditional, invasive techniques. Fetal cfDNA is detectable as early as week 4 of gestation [32], has a higher positive predictive value and a lower false positive rate compared to other noninvasive screening approaches such as ultrasonography or maternal serum analyte screening, and is commonly used in clinical practice [33-35]. However, fetal cfDNA is less predictive than amniocentesis or CVS due to the large percentage of contaminating maternal DNA in the sample; maternal DNA makes up the bulk of cfDNA, where fetal cfDNA constitutes on average 11%-13.4%, although with a large variance across patients [36]. The concentration of fetal cfDNA influences the accuracy of the test, requiring a fetal fraction of at least 4% after losses associated with extraction [37]. Even in samples with an acceptable fetal cfDNA fraction, the majority of cell free DNA is still maternal. Consequently, fetal cfDNA testing is limited to evaluating characteristics that the mother does not possess, such as a Y chromosome, rhesus D status, or aneuploidies [38]. Altogether, fetal cfDNA testing is a screening test rather than a diagnostic test; thus, after a positive screening result, fetal cfDNA testing must be followed up with more invasive procedures such as amniocentesis or CVS [39], which present a risk to the patient/fetus. Thus, though highly

accessible, fetal cfDNA is limited by the quality of the material extracted as the assay is unable to easily differentiate fetal from maternal sequences. To increase the quality of information obtained from fetal cfDNA testing, some groups have attempted to enrich fetal cfDNA within the bulk maternal DNA population. One approach to isolating fetal cfDNA is through size separation, as fetal cfDNA is often longer than maternal cfDNA [40]. Alternatively, other groups have attempted to use epigenetic markers to differentiate fetal cfDNA from maternal cfDNA [41]. While there has been some success at enriching fetal cfDNA, assays integrating fetal cfDNA have yet to transition into clinics.

Fetal cells in maternal blood - Long before the discovery of fetal cfDNA, researchers speculated about the circulation of fetal cells [42-44]. Indeed, rare fetal cells such as trophoblasts [45], lymphocytes [46], granulocytes [46], stem cells [47], and red blood cells [48] circulate the mother's bloodstream, able to be more safely accessed than amniocentesis or CVS [49]. However, in blood, fetal cells are a RcP, circulating the maternal bloodstream at a frequency of 1 in 10,000-1,000,000 [50-52]. In order to ensure that the genetic signatures evaluated are representative of the fetus, techniques have been developed to identify and isolate fetal cells from the surrounding maternal material. Most commonly, researchers have turned to flow activated sorting (FACS) [43], antibody-based magnetic beads [53], or size based exclusion [21,54] to separate fetal cells from the bulk population. Following enrichment, fetal cells may be analyzed via FISH to evaluate the presence of aneuploidies or a Y chromosome. However, a 5 year study on fetal cells isolated from maternal blood via FACS or magnetic bead separation and analyzed via FISH reported a false positive rate of 11% for gender determination, and a false positive rate of 0.6-4.1% for detecting fetal aneuploidies [55]. Overall, the scarcity of fetal cells in maternal blood has limited the technique's sensitivity, and in turn clinical adoption [56].

Trophoblasts in cervical canal - Emerging between the infrequent bloodstream fetal analytes (cffDNA and fetal cells) and the invasive, high-risk methods associated with amniocentesis or CVS sampling, is the collection, enrichment, and analysis of fetal cells via sampling of the cervix. Shed from the chorionic villi, a population of fetal trophoblasts reside in the cervical canal [57]. Cervical specimens are

easily accessible as they can be collected by endocervical brushing with a cytobrush (same protocol as a pap smear), which poses minimal threat to mother or fetus [58]. In contrast to fetal cells in maternal circulation, for women with normal intrauterine pregnancies fetal trophoplasts constitute approximately 1 in 2000 cells in the cervical canal [59]. Recently, Armant et. al found that fetal trophoblasts can be retrieved from cervical swabs via magnetic beads targeting human leukocyte antigen G (HLA-G) [59], a process referred to as trophoblast retrieval and isolation from the cervix (TRIC). Following TRIC, the isolated population has enabled fetal sex determination (via PCR and FISH) [60]. Furthermore, given the high purity of the TRIC-isolated fetal trophoblasts, extracted DNA (purity 85-99%) has proven compatible with next-generation sequencing [61], which has the potential to improve the predictive value of non-invasive prenatal screening.
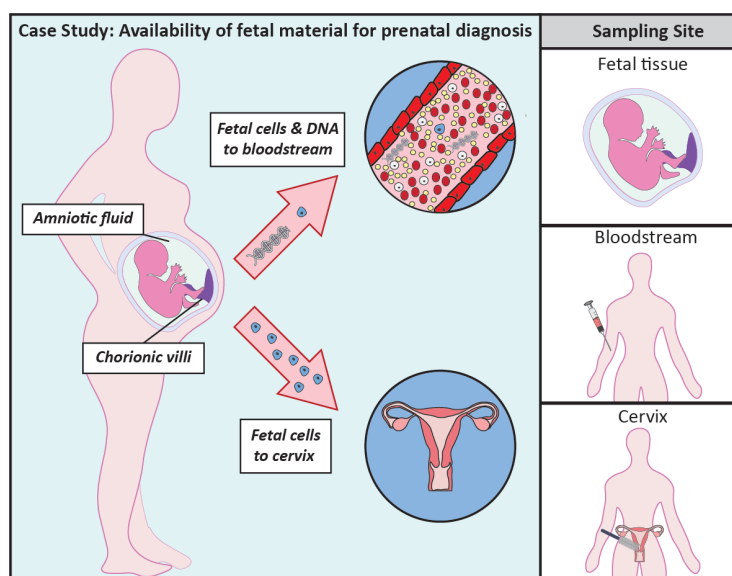
Fig. 1.2 Accessing rare fetal material. Fetal material can be found across the mother's body during pregnancy including direct sampling of amniotic fluid, the maternal bloodstream, and the mother's cervix. Accessing this material for use in prenatal testing requires balancing the risk associated with the procedure acquire the material with the quality of information able to be gathered.

Prenatal testing provides insight into the challenges associated with sample accessibility, particularly the trade-off between patient risk and information quality. Traditional methods to evaluate fetal material, such as amniocentesis or CVS, offer high quality information but have lower accessibility due to the significant risk to the fetus. The most accessible sample, blood, presents challenges in

obtaining the quality of material required (e.g., limited purity, high levels of contaminant maternal material), and despite detection of rare, relevant materials (e.g., fetal cfDNA, circulating fetal populations) remains a screening rather (fetal cfDNA) than a diagnostic test at best. Collection of fetal cells from the cervix appears promising, balancing low risk access with high quality and quantity of material obtained; material able to meet the clinical benchmark required for downstream genetic endpoints.

1.4   Identification of Rare Cell Populations

Whether for enumeration, enrichment, or molecular interrogation, a population needs to be defined (e.g., RNA, DNA, or protein signatures) to differentiate the target RcP from surrounding background populations (e.g., circulating tumor cells are defined as epithelial cells (express EpCAM and cytokeratin protein) in the bloodstream, which are negative for contaminant blood cell markers (express CD45)). Early identification criteria, proposed long before a population is well understood, runs the risk of narrowing researchers' perspectives and complicating the researchers' abilities to openly observe the RcP. Ultimately, due to the complexity of assessing RcPs, RcP identities evolve under continued study with the addition, removal, and modification of identification features. Despite the continually morphing identity, in order to study the RcP, an identification criteria, developed around differential protein expression (extracellular and intracellular) and alternative markers (e.g., specific RNA, DNA), has to be defined.

Identification is often and potentially most easily achieved through population-specific patterns of extracellular protein expression. Antibodies enable the specific targeting of cell features (e.g., proteins, glycolipids) as well as feature visualization through fluorescence (either secondary fluorescent antibodies or direct fluorescent labeling of the primary antibody) via microscopy or flow activated sorting (FACS). While antibodies are highly enabling, facilitating assessment of protein expression on live cells if targeting only extracellular markers, pinpointing the optimal antibody to identify and target RcPs presents challenges due to a lack of positive controls with which to assess the antibody. The rarity of target events require large sample quantities in order to compare antibodies; further, heterogeneity across samples can

add complexity to comparing antibodies (e.g., varying expression, glycosylation of proteins). Despite these challenges, antibodies can provide a useful tool in RcP identification as well as enhance the specificity and robustness of the cell identification process; antibodies can enable the simultaneous identification of contaminant populations. By multiplexing antibodies across several fluorescent channels, multiple protein targets can be visualized in combination.

The use of antibodies to differentiate extracellular protein expression across populations also enables physical sorting of RcPs (or background cells). In clinical practice, RcP identification is often paired with a need to enrich the RcP as, without enrichment, contaminating information from background cells would mask any information obtained in the target RcP. The challenge, however, is that each new RcP presents different complexities (e.g., differing target markers, contaminant populations, purity requirements), which require researchers to constantly adapt. For example, magnetic bead-based antibody capture represents a powerful tool for targeting cell populations by utilizing proteins which are differentially expressed by either the background population (negative selection) or on the population of interest (positive selection). While readily utilized in bulk isolation processes, applying a bead-based capture approach to rare cell populations has resulted in the emergence of numerous application-specific magnetic bead technologies. These refined cell isolation tools often utilize multiple approaches to concentrate a target of interest, while reducing contaminant populations (i.e., Ficoll Paque for peripheral blood mono-nuclear cell isolation, MACS columns for contaminant depletion [62], red blood cell lysis). Even with magnetic bead-based enrichment, complicated interactions of background cells and magnetic beads yields highly variable purity, requiring layering with fluorescent protein staining to further identify the captured populations.

In cases where higher specificity is necessary or where extracellular markers are insufficient for identification (e.g., cytokeratin expression in circulating tumor cells [63]), researchers often turn to intracellular markers. Intracellular proteins for RcP identification require fixation and permeabilization (e.g., detergents) of the sample to allow antibodies to enter the cell and target the protein(s) of interest.

However, fixation alters the cell, a problem in antibody-based recognition of target proteins.[64] The resultant modifications to cellular proteins can impact recognition of antibody binding to target proteins,[64] requiring re-validation of antibodies between live and fixed populations. Many fixative buffers including the most common fixatives, formalin-based buffers[65], impact nucleic acid, which limits compatibility of the fixed population with downstream nucleic acid-based endpoints (e.g., sequencing, PCR).[66,67]

While differential protein expression (or lack of) is frequently used to include or exclude cells from a population (e.g., EpCAM expression in prostate circulating tumor cells), protein expression, both intracellular and extracellular, is often not binary (i.e., on/off, present/absent); thus researchers must rely on analytics and image processing to define whether a cell is negative or positive [68]. This need for user-identified expression thresholds (i.e., defining a cutoff between low expression and no expression) leaves room for artificial inclusion or exclusion of events. While not traditionally a problem in bulk cell identification (i.e., T-cell fraction of PBMCs), even low frequency inclusions of false positive events can impact RcP identification and readouts (through enumeration and purity). The inclusion or exclusion of events in image processing thresholds further challenges benchmarking of identified populations across platforms, complicating RcP understanding.

Oftentimes the most challenging RcP identification situation is when populations lack differential protein markers (e.g., HIV viral reservoir). RcPs, due to immune evasion or other mechanisms, may not exhibit differential extracellular protein expression from surrounding contaminant cells (or sub-population). This lack of differential protein expression profiles limits physical separation of the cells, complicating RcP targeting and purification. Alternative differentiating characteristics are then required to identify the target cell (e.g., size, expression of RNA variants, genomic modifications, functional response, RNA expression patterns). On occasion, RcPs must be manipulated (e.g., chemical activators) to bring about a detectable change in the population, which can then be measured. While alternative methods exist to detect intracellular RcP identifying traits, clever assays are required to understand the frequency and behavior of the masked RcPs.

While identification is inherently required to define the population for continued study, the identification parameters set (i.e., proteomic, transcriptomic, genomic) shape downstream research. In pursuing a RcP, researchers must decide between building on 'accepted' knowledge or challenging the existing/published identification criteria. In this way identification becomes fluid, as one criterion is exchanged for another, complicating the benchmark of previous results against one another as subtle identification differences yield significant population differences, especially in RcPs. The HIV viral reservoir presents one case study where subtle variations in identification criteria yield strikingly different RcPs complicating clinical adoption of HIV viral reservoir quantitation assays.

1.5   Identification of Rare Cell Populations Case Study: The HIV Viral Reservoir

Due to successful treatment on Antiretroviral Therapy (ART), patients are able to control their HIV infection, suppress plasma HIV RNA levels[69,70], and greatly reduce their risk of death from AIDS [71]. Despite this accomplishment, HIV remains uncured due to the early establishment of a viral reservoir. The HIV viral reservoir consists of a small population of HIV infected cells, which persist even after decades of HIV treatment[72-75]. The latent, infected population is capable of being activated, leading to production of infectious, replication-competent virus [76,77]. While uncertainty remains over the reservoir location(s) (e.g., lymph nodes[78,79], brain[80,81], gut[82]), a portion of the HIV viral reservoir is known to circulate the bloodstream[76,77])[83] providing a readily-accessed fraction for clinical monitoring. In pursuit of a cure and improved treatment management, targeted approaches are emerging to deplete the reservoir. One approach, often termed 'kick and kill', is the activation of latent cells in patients on ART[84], enabling targeting of the activated cells[85], resulting in reservoir depletion. In order to apply such targeted methods to patients, a clinical assay to monitor the reservoir is required, driving the need for clinically amenable assays to identify and quantifying the reservoir.

As a RcP, the HIV viral reservoir presents a fundamental identification challenge due to the absence of extracellular differential proteins by which to identify the RcP from background cells. The lack of external markers, which helps infected cells avoid detection by the immune system, complicates

any physical, antibody-based sorting or visualization of a reservoir cell from a normal cell. While not able to specifically identify reservoir cells using extracellular markers, antibody-based cell sorting enables targeted depletion of suspected non-reservoir populations, enriching for potential reservoir cells (resting CD4-positive T-cells). From the enriched population, identification of the HIV viral reservoir requires an internal differentiator be found. Reservoir assays targeting different analytes in the HIV life cycle have been developed aiming to connect clinically compatible assays with clinically relevant quantitation.

HIV DNA - Upon entry into a host cell, the HIV RNA genome is reverse transcribed into DNA, capable of integrating into the host cell's genome (provirus). Infected cells, including reservoir cells, are thus likely to carry HIV DNA sequences. Through the high sensitivity of quantitative polymerase chain reaction (PCR) paired with the specificity of sequence-specific HIV probes, HIV DNA events can be detected from amongst large numbers of background, HIV-negative cells. By performing limiting dilutions of cells, or single cell PCR, the frequency of HIV DNA-positive events in a population can be quantified. The stability of DNA paired with the simplicity of DNA extraction results in a relatively rapid assay transferable to a clinical laboratory. Despite the assay simplicity, the specificity and clinical relevance of HIV DNA in identifying the reservoir is debated. Reverse transcription and integration of the resultant DNA into the host genome is an error-prone process oftentimes resulting in provirus, which will never lead to replication-competent HIV production (e.g., hypermutated[86], deletions[86,87], or silencing[87,88]). In an attempt to integrate more specificity, Alu-PCR was developed, limiting detection of HIV DNA to events occurring within the host cell's genome[89,90]. However, Alu-PCR still falls short in ensuring the reservoir identified is representative (or correlative) of events capable of reversing latency and leading to the active production of replication-competent HIV[89]. As a result of the limitations of HIV DNA endpoints, assays have continued to emerge, redefining the viral reservoir identification criteria with each assay.

Multiply-spliced tat/rev RNA - Upon activation, a reservoir cell exits latency and begins transcribing the HIV DNA. The resultant mRNA then undergoes splicing for downstream protein

production. The splicing includes the production of multiply-spliced tat/rev RNA (ms-HIV RNA). Unlike unspliced HIV RNA, which is frequently detected amongst the PBMCs of virally suppressed subjects on ART[91-94], ms-HIV RNA correlates with the ability of a cell to produce virus[91,93-98]. A step beyond DNA, ms-HIV RNA is closer to ensuring production of infectious virus, overcoming some of the shortcomings of HIV DNA (i.e., latency reversal, faulty integration). To measure ms-HIV RNA, the Tat/rev induced limiting dilution assay (TILDA) relies on a serial dilution of cells in a nested RT-PCR reaction. While slightly more labor intensive than the HIV DNA assays, TILDA maintains a level of simplicity appropriate for a clinical laboratory. Yet, ms-HIV RNA still falls short of ensuring production of infectious virus, and due to TILDA's early stage there is limited clinical data.

Production of Infectious Virus- Following production of HIV RNA transcripts, viral proteins are manufactured, assembled into virions, and released from the cell. If correctly produced and assembled, the released virus infects another cell, driving continued HIV production, and ultimately viral rebound within the patient. The Quantitative Viral Outgrowth Assay (QVOA) aims to quantify only reservoir cells capable of producing infectious virus by detecting virus propagation[72,77,99,100]. Largely considered the gold standard in HIV viral reservoir quantitation, the assay is limited by its lack of sensitivity and complexity hindering adoption outside of research settings. In spite of the QVOA's relevance in specifically identifying reservoir events most likely to impact the patient, the required assay is not compatible with clinical care (i.e., expense, labor, sample volume) resulting in a lack of adoption and widespread use. In contrast with QVOA, a more recent assay, called the TZA assay, uses a reporter rather than virus propagation to detect infectious virus production, detecting initial infection[101]. Using a luciferase reporter, the TZA assay combines a simplistic endpoint (i.e., luciferase endpoint on a plate reader) with a shorter duration (i.e., 48-hour culture).
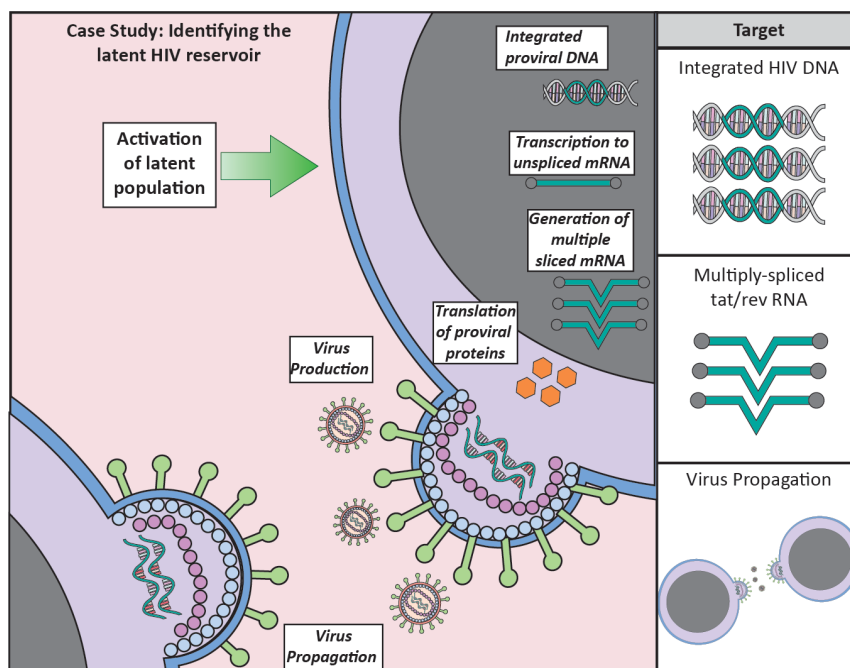
Fig. 1.3 Identification of the HIV viral reservoir. Identification of the HIV viral reservoir can be achieved by assessing a variety of HIV-specific targets harbored by cells contributing to the viral reservoir. Starting from cell-associated HIV DNA within the latent cells, cell activation and latency reversal results in the downstream production of multiple potential HIV-specific analytes to use in identifying and subsequently quantifying the reservoir; these analytes include HIV DNA, multiply-spliced tat/rev RNA, and virus propagation.

Identification of the RcP composing the circulating HIV viral reservoir is a balance of identification criteria, clinical relevance, and clinical feasibility of the assay. Challenged by the hidden nature of the reservoir, alternative approaches to assay the target population have resulted in the emergence of multiple reservoir identification criteria and targetable analytes (e.g., HIV DNA, ms-HIV RNA, production of infectious virus). While ideally the most "relevant" analyte would prevail, assay complexity, sensitivity, and clinical compatibility all complicate clinical potential. As reservoir quantitation will be required to facilitate the next generation of HIV treatment, the identification parameters must not only provide clinically significant information, but also translate into a clinically compatible assay.

1.6 Heterogeneity in Rare Cell Populations

Heterogeneity is one of the most challenging aspects of RcPs. Occurring at all levels – patient-to-patient, sample-to-sample, and cell-to-cell – RcPs may exhibit heterogeneity through subtle differences in

protein expression, RNA, and DNA. These differences complicate identification, assessment, and understanding of RcPs, and limit the ubiquity of technologies in targeting RcPs; however, the heterogeneity may also provide valuable information in understanding the underlying clinical condition. To overcome the implicit challenges of heterogeneity and capitalize on the information hidden within heterogeneity, the potential sources of heterogeneity must be understood.

Heterogeneity is often first encountered at a patient level, resulting in patient-to-patient variation in populations. The resultant patient-to-patient heterogeneity, however, can frequently be used to differentiate and categorize patients. Within patients, differential protein expression on red blood cells is used to stratify blood types in healthy patients (e.g., A-positive, A-negative, AB, etc.). In the event of a blood transfusion, ensuring compatible blood types is critical to ensuring a successful transfusion. Similarly, understanding heterogeneity in disease patients can impact treatment. For example, breast cancer can be stratified into multiple subtypes (e.g., estrogen receptor positive or triple negative populations). Despite a common diagnosis, these subtypes contain populations with specific cellular compositions and identifying markers, which impacts disease progression, targetable biomarkers, and clinical approaches. Thus, when approaching RcPs across multiple patients, one must consider the potential for RcPs to stratify patients into subtypes; such stratification presents two main challenges: 1) ensuring the RcP is identified across patients and 2) ensuring the stratifying markers are identified. To be broadly utilized, RcP identification must ensure detection of the relevant population across patients; once achieved identification of patient-to-patient heterogeneity could provide information applicable to patient care.

Sample-to-sample heterogeneity presents a challenge when collecting from a single patient. Whether cellular factors (e.g. various populations, differing densities) or sample composition (density, matrix, components), samples are known to change both locationally (e.g., sampling site) and temporally. For instance, recent studies have revealed differing gene expression signatures dependent on the sampling site of the tumor.[102] Even in blood – a seemingly homogeneous sample type – location of the blood draw

can impact the frequency of different blood cell populations[103] or cell expression profiles[104]. Together, these locational variations can impact or bias RcP results and must be accounted for by means of consistent sampling or multiple sampling locations. A more complex form of heterogeneity is temporal heterogeneity. A patient's body is transient; disease states and external influences (e.g., treatments, stress, diet, exercise) that occur on a daily, hourly, or minute-to-minute basis can have major effects on both the RcP and the surrounding populations. For instance, immune cell populations in the blood stream change as a result of chemotherapy[105,106]; chemokine activation of lymphocyte integrins can occur in milliseconds and modulate behavior (e.g., adhesion[107]). Further, samples can be affected by the amount of time and conditions in which they are handled after removal from the body. Understanding the impact of these variables may be significant in understanding RcPs; thus, while challenging, constraining factors such as diets, treatments, and sample handing may be crucial in accurately (and consistently) obtaining information from RcPs. Overall, understanding sample-to-sample heterogeneity reduces potential sampling bias, which could otherwise muddle the clinical value elucidated from the targeted RcP.

Within a single sample, cell-to-cell heterogeneity is arguably the most challenging. Prevailing in RNA variants, DNA modification, or differences in protein expression or localization, cell-to-cell heterogeneity can impact capture, identification, and analysis of RcPs. Many methods rely on antibody-protein interaction to capture or identify cells (as discussed in the identification section and highlighted in the heterogeneity case study). Thus, heterogeneity between cells can complicate both capture and identification, leading to non-specific inclusion or exclusion of the RcP (dependent on presence/absence of markers). For example (as elaborated in the following case study), EpCAM-negative CTCs would not be captured or identified in an EpCAM-based selection approach. Similarly, population-based endpoints can mask cell-to-cell heterogeneity by yielding a single readout rather than multiple cell-specific ones. However, just as cell-to-cell heterogeneity presents challenges in the capture and identification of RcPs, clinically identifying heterogeneity in targetable biomarkers could be advantageous in identifying

treatment routes. Thus, while heterogeneity is a challenge that must be considered, heterogeneity itself may be a source of information, giving insight into the RcP and potential clinical outcomes.

Improving understanding of the mechanisms with which to evaluate populations has continued to reveal the heterogeneity that exists within seemingly homogeneous populations. When applied to cell populations, heterogeneity can occur at numerous biological levels (i.e., phenotypic, transcriptomic, and genotypic) as well as system levels (i.e., patient-to-patient, sample-to-sample, cell-to-cell). Solutions to heterogeneity in targeting a RcP will have to be integrated into future RcP technologies and platforms to facilitate ubiquity of the RcP across patients. The field of circulating tumor cells presents clear insight into the impact of heterogeneity on evaluating RcPs, challenging both identification and clinical insight both within and across cancers.

1.7   Heterogeneity in Rare Cell Populations Case Study: Circulating Tumor Cells

Circulating tumor cells (CTCs) are broadly defined as cells, which break off of a metastatic site and enter the bloodstream. An accessible link to a cancer occurring deep within the body, CTCs have the potential to provide insight into cancer progression, treatment, and drug response. While universally identified under the 'umbrella' term of CTCs, this RcP faces challenges due to the cellular heterogeneity uncovered cancer-to-cancer, patient-to-patient, and cell-to-cell. Specifically, the CTC field has faced challenges in defining markers for identifying and targeting CTCs alongside challenges in understanding the role of disease in heterogeneity. Thus, while heterogeneity can provide insight into disease progression and response, CTC heterogeneity presents obstacles in utilizing CTCs.

EpCAM-based Identification – Despite originating from a single tumor, research has demonstrated significant CTC variability on a cancer-to-cancer, patient-to-patient, and, as technologies advance, cell-to-cell basis. One key embodiment of CTC heterogeneity is extracellular protein expression markers, which are used to identify and capture target CTCs. The only Food and Drug Administration approved CTC platform, CellSearch, maintains a rigid definition of CTCs. Using anti-EpCAM magnetic bead-based

capture, CellSearch is approved to enumerate circulating cells of epithelial origin (CD45-, EpCAM+, and cytokeratins 8+, 18+, and/or 19+) in metastatic breast[108], colorectal[109], and prostate cancer (limited to patients with androgen-independent, hormone-resistant, or castration-resistant prostate cancer).[110] While EpCAM has enabled CTC capture with moderate success, tumor cells from other major cancers - including melanoma and pancreatic cancer – and EpCAM-negative CTCs will escape EpCAM-based detection. Heterogeneity in CTCs through EpCAM-negative cells likely limits (and underestimates) CTCs, demonstrating CTC detection in only 57% of prostate cancers (188 total), 37% breast cancers (1,316 total), and only 20% (168 total) of lung cancers [63]. Thus, from identification to downstream analyte extraction and analysis, CTC heterogeneity must be considered to access the entire population, rather than just a fraction.

Heterogeneity in Extracellular Protein Expression - As the CTC field is largely developed from the targeting of external cell markers for CTC capture and enrichment, heterogeneity in extracellular protein expression can influence downstream CTC understanding by impacting the 'types' of CTCs captured. Heterogeneity can occur protein-to-protein as a single protein can take many forms due to subtle modifications (e.g., levels of glycosylation [111-113], phosphorylation, dimerization). These modifications can impact antibody capture efficacy and efficiency by modifying how the protein epitope is recognized and bound by an antibody. As CTC capture platforms largely rely on antibody-based capture approaches,[6,7,63] to ensure robustness in antibody-based targeting, the antibody must target an epitope conserved across populations. Similarly, levels of expression (e.g., low to high EpCAM expression) can influence the ability to identify and target a CTC as low expression may be inadequate for antibody-targeted approaches, resulting in incomplete capture of the entire population. Fundamentally, CTCs may also decrease EpCAM expression as tumor cells are thought to undergo epithelial-mesenchymal transition (EMT) prior to entering the bloodstream.[114,115] Concurrently, cells undergoing EMT begin to express mesenchymal markers (e.g., Muc-1), resulting in another challenge in CTC heterogeneity - variable expression markers (e.g., EpCAM [108-110], Muc-1 [114], PDL-1, Trop-2). Based on the diversity of known,

suspected, and yet to be discovered CTC markers cell-to-cell CTCs likely express multiple independent targets including: epithelial markers (EpCAM), mesenchymal markers (Muc-1), and alternative markers as in the case of cancer stem cells (e.g., ALDH1 [116]). Due to the high heterogeneity occurring in cell surface markers, how CTCs are identified and subsequently targeted (e.g., EpCAM only) impacts the population observed. Incomplete CTC targeting can produce an incomplete view of the CTC population, which ultimately may be misrepresentative of disease the CTC population is reflective of. Clinically, incomplete understanding and targeting of CTCs could translate to inaccurate enumeration of CTCs, and if paired with a molecular endpoint (i.e., DNA sequencing, RNA variants), offer an incomplete view of the cancer and the pathways being exploited.

Heterogeneity in Intracellular Protein Expression and Localization – Heterogeneity within cells, such as through intracellular localization (e.g., nuclear vs. cytoplasmic) and relative expression of proteins, can provide insight into the pathways being exploited by the cancer. Through fluorescence imaging, protein localization and expression can be directly visualized and assessed independently for each cell. Within a cell, localization of proteins can provide information on the pathways being modulated by the cancer allowing assessment of effectiveness of protein- and pathway-targeted therapies. In prostate cancer, the androgen receptor (AR), a nuclear hormone receptor, plays a role in disease development, response to hormone therapy, and eventual development of resistance. To personalize treatment in metastatic castration-resistant prostate cancer (mCRPC), which often targets AR, it is important to understand the AR activity occurring within the patient. In EpCAM captured CTCs from mCRPC patients, staining of AR protein in CTCs demonstrated heterogeneity in AR staining intensity (expression) both cell-to-cell within a patient and patient-to-patient.[117] Due to the role of AR in downstream gene transcription, nuclear localization of AR may be suggestive of AR activity when compared to cytoplasmic AR. Similar inter- and intra-patient variability was observed in AR nuclear localization.[117] Understanding the heterogeneity in AR expression and localization within CTCs could prove a valuable biomarker in evaluating AR-based treatment approaches for each patient.

Heterogeneity in Intracellular Transcripts – Just as heterogeneity can be observed at the protein level, heterogeneity can also be detected through mRNA transcripts. Heterogeneity indicative of treatment response can be assessed through the presence of androgen receptor splice variants (expressed in mRNA), which could serve as a biomarker for treatment decisions. In prostate cancer, resistance to AR-targeted therapy is characterized by continuous AR activity driving cancer progression. One mechanism behind the persistent AR activity is AR gene rearrangements leading to production of modified versions of the AR receptor known as splice variants; splice variants lack the ligand-binding domain enabling continuous activity independent of ligand binding. The AR splice variant 7 (AR-V7) was detected (via mRNA expression) in EpCAM-positive CTCs of men with progressive mCRPC. The presence of AR-V7 was associated with resistance to abiraterone[118] and enzalutamide[119],[120] two AR-directed therapies for castration-resistant prostate cancer. Yet, AR-V7 did not predict response to taxane chemotherapies (docetaxel[121], cabazitaxel[122]).[123,124] Thus, the heterogeneity across patients and cells in the presence or lack of AR-V7 could provide a biomarker to assist in the treatment selection for men with castrate resistant prostate cancer.
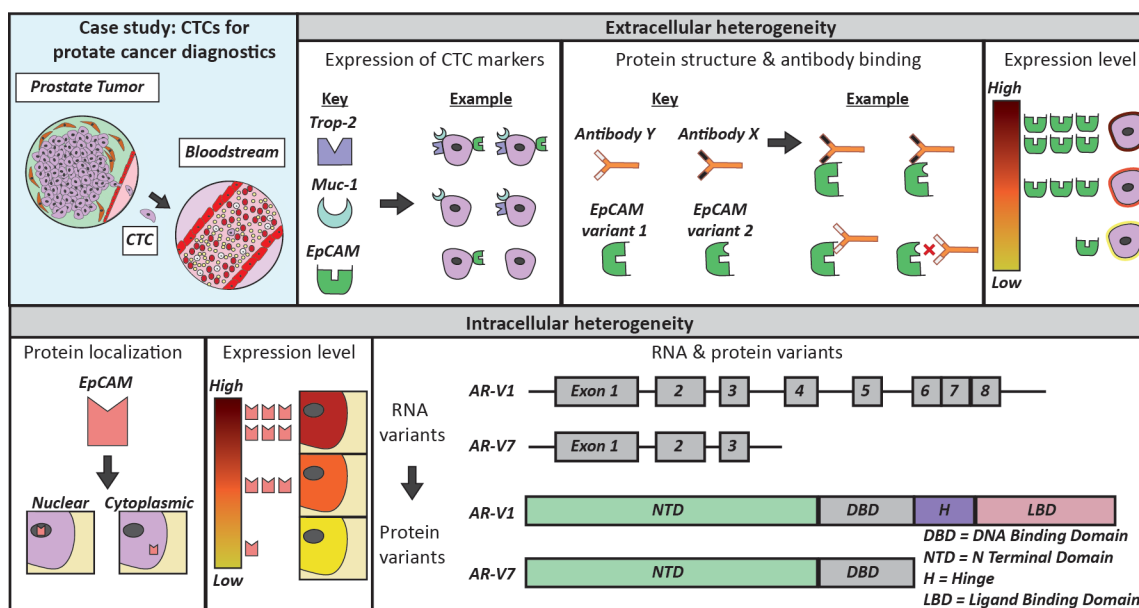
Fig. 1.4 Heterogeneity in circulating tumor cells. Circulating tumor cells provide insight into the implications and potential value in assessing heterogeneity. As a population of cells, which is typically targeted due to extracellular marker expression, cell-to-cell, heterogeneity in extracellular protein expression can impact successful capture of the

population (e.g., variable CTC markers, protein marker variation, expression level). While a challenge in capturing the population, intracellular heterogeneity in prostate cancer CTCs has also provided clinically actionable information on treatment efficacy through the localization and expression of AR and AR variants.

CTCs provide a valuable case study of heterogeneity, highlighting the potential differences across cancers, patients, and cells. Due to the high clinical potential of CTCs, capturing these populations largely through extracellular protein-based capture (or sorting) remains a priority. However, heterogeneity in CTC extracellular expression has challenged the efficacy 'one marker' approaches to CTC capture (e.g., EpCAM only capture) as new, alternative CTC markers are being uncovered (e.g., Muc-1, PDL1, Trop-2). Despite the challenges heterogeneity presents in identifying and targeting CTCs, heterogeneity within CTCs has the potential to provide insight into the pathways being targeted and monopolized by the cancer. However, the targets must be known to utilize them for analysis, presenting challenges due to the breadth of cancers, pathways, and targets encompassed by the term CTCs. In treating CRPC specifically, the androgen receptor demonstrates the potential value, using heterogeneity through protein expression, protein localization, and variant detection (e.g., mRNA) to inform clinical treatments. Thus, while heterogeneity can challenge CTC identification and targeting, heterogeneity can also provide clinical context to a cancer.

1.8   Conclusions

Frequently associated with concerning clinical conditions (e.g., virus infections, cancer, pregnancy), RcPs provide a promising new tool with which to obtain information and improve clinical care. Despite clear potential clinical value (e.g., diagnostic, prognostic biomarkers), the path of a RcP to the clinic is challenged by seemingly simple concepts: *access*, *identification*, and *heterogeneity*. Overcoming the obstacles to RcP access requires excising not only the sample, but also the RcP within requiring the potential patient risked to be weighed against the obtainable information. To then truly interrogate the RcP, the RcP needs to be identified. The rarity of such a population (i.e., 1 in a million to 1 in a billion cells) combined with the often-diverse masking contaminant populations fundamentally challenges identification (and targeting) of the RcP. Heterogeneity complicates RcP understanding by introducing variation to the population. Heterogeneity can span from patient-to-patient down to cell-to-cell, embodied through subtle differences in protein, RNA, or DNA. If not accounted for, heterogeneity can cause RcP events to fall outside of the identification parameters or be lost based on the endpoints and targets

evaluated. Existing RcP technologies have highlighted the role of these challenges in complicating implementation of RcP in the clinic. The next generation of RcP-targeted technologies will need to consider and adapt to these challenges to transition the RcP from research into clinical care.

Chapter 2   Integration of Magnetic Bead-Based Cell Selection into Complex Isolations[2]

Magnetic bead-based analyte capture has emerged as a ubiquitous method in cell isolation, enabling highly specific capture of target populations through simple magnetic manipulation. To date, no "one-size fits all" magnetic bead has been widely adopted leading to an overwhelming number of commercial beads. Ultimately, the ideal bead is one that not only facilitates cell isolation but also proves compatible with the widest range of downstream applications and analytic endpoints. Despite the diverse offering of sizes, coatings, and conjugation chemistries, little literature exists to benchmark the performance characteristics of different commercially available beads; importantly, these bead characteristics ultimately determine the ability of a bead to integrate into the user's assay. In this report, we evaluate bead-based cell isolation considerations, approaches, and results across a subset of commercially available magnetic beads (Dynabeads FlowComps, Dynabeads CELLection, GE Healthcare Sera-Mag SpeedBeads Streptavidin-Blocked Magnetic Particles, Dynabeads M-270s, Dynabeads M-280s), to compare and contrast both capture specific traits (i.e., purity, capture efficacy, contaminant isolations) and endpoint compatibility (i.e., protein localization, fluorescence imaging, nucleic acid extraction). We identify specific advantages and contexts of use in which distinct bead products may facilitate experimental goals and integrate into downstream applications.

2.1 Introduction

Cell isolation provides a foundation for both clinical assays and basic biological research. The isolation of a subset of cells from a large, diverse population enables the enrichment of a specific population, unmasking the isolated population for continued analyses. In clinical assays – whether a tissue biopsy or blood draw - cell isolation is a critical step as patient-derived samples yield a complex mixture consisting of a broad spectrum of cell types, matrices, and biological factors. Cell isolation is required to 1) access a target population hidden within the sample and 2) assess specific (and often rare) analytes contained within target cells (i.e., RNA, DNA, protein).[125,126] Without isolation, the noise introduced by contaminating populations impairs detection of the target-specific markers needed to inform clinical care. As assays continue to delve deeper into the interrogation of specific target populations – such as circulating fetal cells,[127,128] circulating tumor cells,[129,130] and stem cells[23] – cell isolation processes will become essential and drive development of commercial cell isolation products. Reflective of the ubiquitous nature of cell isolation in biologic studies, the current estimated market value (over 3.5 billion USD in 2016) is predicted to reach over 7.8 billion USD by 2021.[131]

Traditional approaches to tackle cell isolation, which purify or extract the intended population, have centered on filtration, centrifugation, sedimentation, and adherence. Filtration enables cell sorting based on size, largely performed by selecting or excluding populations using mesh filters of specific pore size.[132,133] Centrifugation and sedimentation enables sorting based on cell density, often aided by density gradients to subdivide subtle density differences across populations.[134,135] Adherence relies on differential cellular interactions with specific substrates over a specified timeframe.[136] While all are relatively simple and easy to scale, these methods are quickly limiting when cells lack significant, differential cell size, density, or adhesion, requiring new approaches to cell isolation.

Solving the limitations of density and size-based cell sorting is an emerging and quickly growing field, magnetic bead isolation. Magnetic bead isolation has found widespread use in biological assays and applications[137-139] utilizing small (nanometer or micrometer sized), magnetically responsive beads to

manipulate a biological target. A wide variety of magnetic beads with a diverse offering of surface chemistries are commercially available enabling easy manipulation of proteins,[140,141] nucleic acid,[142-144] and whole cells,[145-148] providing a powerful isolation tool.[149] For cell isolation, magnetic beads can be combined with a diverse offering of commercially available antibodies specific to cell surface proteins to enable the targeting of nearly any cell population.

While magnetic beads are widely developed with well-characterized physical traits and magnetic properties,[150,151] limited literature exists directly comparing multiple bead types within the same biological context to benchmark performance (i.e., capture efficacy, non-specific binding) and impact on common downstream endpoints (e.g., fluorescent staining of proteins to quantify localization, nucleic acid extraction, cell culture) across bead types. Here we evaluate five common cell isolation magnetic beads (**Fig. A.1**) – Dynabeads M-270 Epoxy, Dynabeads M-280 Streptavidin, CELLection Biotin Binder, FlowComp Dynabeads, and Sera-Mag SpeedBeads Streptavidin-Blocked Magnetic Particles – to highlight the tradeoffs and considerations in integrating cell isolation magnetic beads into biologic assays. These particular beads were selected to provide a range of capabilities that may be attractive to users, such as cell release – CELLection, FlowComp; biotin-based antibody conjugation for flexibility in cell capture – M-280, CELLection, FlowComp, SeraMag; batch conjugation of antibody to bead – M-270s; and advertised low non-specific binding – Seramag, M-270s. Based on these reported favorable cell capture attributes these commercially available magnetic beads were chosen for comparison. Beads were characterized in the context of EpCAM-specific (Epithelial Cell Adhesion Molecule) cell capture. EpCAM is a cell isolation and identification marker for epithelial cells including circulating tumor cells (CTCs).[6,8,152-154] CTCs are rare tumor cells, which are shed from a tumor lesion and enter the bloodstream. If captured, CTCs have the potential to provide insight into the cancer, and were thus selected as a representative rare cell population. To characterize capture of EpCAM-positive cells, we evaluated capture of cell lines with differential EpCAM expression, release of those cells following capture (for FlowComp, CELLection magnetic beads), and nonspecific capture of relevant background populations.

Furthermore, we assessed the impact of the beads in integrating with standard downstream assays, including: cell culture, fluorescent immunohistochemistry, and nucleic acid extraction. By evaluating a variety of magnetic bead types across a spectrum of molecular biologic assays, we aim to highlight the strengths, weaknesses, tradeoffs, and considerations when integrating beads into a cell isolation protocol.

2.2   Materials and Methods

2.2.1   Magnetic Beads, Antibody Conjugation, and Binding.

Capture experiments used a goat polyclonal anti-EpCAM antibody (Clone AF960) (AF960, R&D Systems) conjugated to the following magnetic beads: Dynabeads M-270 Epoxy beads (14311D, ThermoFisher), Dynabeads M-280 Streptavidin (11205D, ThermoFisher), CELLection Biotin Binder Kit (11533D, ThermoFisher), FlowComp Dynabeads (11061D, ThermoFisher), and Sera-Mag SpeedBeads Streptavidin-Blocked Magnetic Particles (21152104011150, GE Healthcare Life Sciences) (**see A.1 Extended Bead Information**). An overview of the magnetic beads evaluated is provided in **Fig. A.1**.

Antibody was batch conjugated to the M-270s, as per the manufacturer's instructions, using the Dynabeads^{TM} Antibody Coupling Kit (14311D, ThermoFisher). Due to the batch conjugation of M-270s, unlike the alternative magnetic bead types, an antibody to bead density could not be easily titrated. Thus, M-270s were conjugated following the manufacturer's recommendation, at a density of 6 μg antibody per milligram of beads. For all other beads, the antibody was first biotinylated following the manufacturer's instructions (DSB-X^{TM} Biotin Protein Labeling Kit D-20655, Thermo Fisher) to facilitate streptavidin-biotin binding of the antibody to the beads. Magnetic beads were washed by placing the beads on a magnetic tube rack (DynaMag Rack, ThermoFisher), removing the original buffer, and resuspending the beads in an identical volume of buffer (0.1% BSA in Ca^{+2} and Mg^{+2}–free PBS with 2 mM EDTA); the beads were again washed prior to use. Separately, the antibody was diluted into an identical volume of buffer, which was combined with the washed beads and tumbled for 30 minutes using a Labquake rotator (ThermoFisher) (set to approximately 6 rpm). Following binding, the fluid was removed, the beads washed, and resuspended in buffer for use.

2.2.2   EpCAM Expression.

To more robustly characterize EpCAM-based cell capture, EpCAM expression was assessed for each cell line to differentiate a high, medium, and low EpCAM expresser. Cells were stained with an anti-EpCAM antibody conjugated to phycoerythrin (PE) (clone VU-1D9) (ab112068, Abcam) (1:100) and Hoechst 33342 (H3570, ThermoFisher) (20 μg/mL), both diluted in 1x Phosphate Buffered Saline (PBS) supplemented with 2mM EDTA and 0.1% Bovine Serum Albumin (BSA). Once stained, cells were washed, resuspended in PBS, and imaged on glass. After imaging, the mean fluorescence (expression) of each cell line was calculated and normalized to the maximum EpCAM expressing cell line (LNCaP) to more easily compare relative expression (**Fig. A.3**). Based on the results, Du145 (low EpCAM), 22Rv1 (medium), and LNCaPs (high) were used in all subsequent experiments. Each of the three cell lines had a similar diameter ranging from ~15-25 μm, with Du145s being slightly smaller, generally ~15-20 μm. The cell lines screened for EpCAM expression were all prostate cell lines; prostate cancer has been one cancer type for which EpCAM-based capture has proven clinically relevant in capturing CTCs.[155,156] Additionally, many of the cell lines evaluated have been used in characterization of CTC capture platforms[6,8]; thus these cells represent a relevant target, spanning a wide range of EpCAM expression, in the context of EpCAM-based capture of prostate cancer cells.

2.2.3   Cell Isolation and Release

All cell isolation was performed using the Extractman (EM) (22100000, Gilson), a platform based on the Sliding Lid for Immobilized Droplet Extraction Technology.[157] EM allowed the simultaneous isolation of up to four samples. For *direct* isolation, 100 μg of antibody-coated magnetic beads (as described above) were incubated with cells (total volume of 475 μL) on a Labquake tumbler (ThermoFisher) rotating at ~6 rpm for 30 minutes at 4 °C. For *indirect* isolation, the anti-EpCAM antibody was first added to the cells solution (475 μL) and tumbled for 30 minutes at 4°C (as specified above); 100 μg of magnetic beads were then added and the solution tumbled for 10 minutes. After incubation, the entire volume was loaded into the input well of an Extractman plate (22100008, Gilson).

Using EM, the cells were then captured on the Extractman consumable strip (22100007, Gilson) as the built in EM magnets were held over the middle of the well for 30 seconds to enable collection of the beads and then the EM handle moved over a wash well (small wash well, 110 μL) where the lower magnets automatically pulled the beads into the well. Once dropped in the wash well, the EM collection strip was pulled back from the well, the lower magnet removed, and the cells mixed three times by pipette (set to 70 μL). For experiments characterizing non-specific capture of PBMCs, two additional washes (110 μL) were added to improve stringency in deciphering between nonspecific binding (i.e., cells and beads) and basic carryover (i.e., cells caught in residual fluid on the strip). Once mixed, the bead bound cells were recollected on the consumable strip by leaving the EM handle positioned over the well for 15 seconds and then moving the handle to the next well. The contents of all wells were then collected and imaged to ensure accurate cell counts.

For experiments involving release, a similar experimental design was followed except once washed, the beads and bound cells were dropped into a release well containing either FlowComp Release Buffer (FlowComp Flexi Kit 11061D, ThermoFisher) or CELLection release buffer (CELLection Biotin Binder Kit 11533D, ThermoFisher) prepared according to the manufacturer's instructions (release volume of 110 μL). Once dropped into the well, cells and beads were mixed by pipette (3 mixes; pipette set to 70 μL), allowed to incubate for the 20 minutes, mixed again, and collected. Any non-released population of cells was then collected by magnet and transferred using the EM to the final well. All wells were imaged to ensure accurate cell counts.

2.2.4 *Fluorescent Staining.*

Fluorescence characterization of the antibody-bead interaction and antibody-bead density was performed with either anti-goat Alexa488 (ab150129, Abcam) or anti-goat Alexa555 (ab150130, Abcam) secondary antibodies. In brief, following binding of the primary antibody to the beads, the diluted secondary (in buffer) was added to the beads for 30 minutes. The beads were then washed and resuspended in PBS prior to imaging. This fluorescence characterization was used to identify an optimal

anti-EpCAM antibody density for each bead type (with the exception of batch conjugated M-270s). Using the fluorescent secondary antibody, the fluorescent signal on the bead (due to bound anti-EpCAM antibody on the bead surface) was quantified across increasing amounts of primary antibody (**Fig. 2.1A**). The resultant intensity curve of fluorescence versus antibody density was then used to identify the 50% maximal binding capacity for each bead type (**Fig. A.2**). To standardize antibody function on the surface of the beads for all subsequent experiments unless noted (given differing surface areas and surface functionalities), the identified 50% maximal antibody density was used.

To first fluorescently characterize release, FlowComps (with bound primary and fluorescent secondary antibodies) were resuspended in FlowComp Release Buffer (110 μL). At set time points, the beads were collected by the EM handle and removed from the release buffer. The beads were then dropped in wash buffer to and imaged. The CELLection beads were similarly characterized (utilizing 110 μL CELLection Biotin Binder Kit Release Buffer). The measured bead fluorescence was corrected by subtracting the baseline autofluorescence of the blank bead incubation with the appropriate secondary as in the experimental conditions.

For all cell line-based capture experiments, cells were pre-stained with either Calcein, AM (C3100MP, ThermoFisher) or CellTracker[TM] Red CMTPX Dye (C34552, ThermoFisher) (CTR). CTR was utilized when background cells were present, which were concurrently stained with Calcein, AM to enable identification of each cell type. For viability experiments, a live-dead assay was performed on the cultured populations with Calcein, AM and Ethidium Homodimer-1 (E1169, ThermoFisher) at a final concentration of 1 μg/mL and 20 μg/mL respectively. Cells were allowed to incubate for 20 minutes and then imaged using a 10x objective.

To determine the impact of cell isolation beads on fluorescent immunohistochemistry, bead-captured LNCaPs were compared to an untouched population. Bead-bound cells were captured and washed using EM to ensure bead-free cells were removed from the population. Once isolated, the bead-bound cells and untouched population were incubated in buffer containing anti-EpCAM antibody

conjugated to PE (1:100) and Hoechst 33342 (20 μg/mL) for 30 minutes. The cells were then washed in buffer, fixed for 15 minutes in 4% PFA (P6148, Sigma), and washed. Following permeabilization (PBS with 1% Tween-20 and 0.05% Saponin for 30 minutes), cells were resuspended in buffer containing an anti-pan cytokeratin antibody (FITC) (35 μg/mL) (ab11214, Abcam) and anti-androgen receptor antibody (1:100) (5153S, Cell signaling) (incubated at 4°C overnight). After washing the cells, goat anti-rabbit AlexaFluor-488 (ab150073, Abcam), was added at 10 μg/mL for 1 hour in buffer. Samples were washed in buffer and resuspended in PBS prior to imaging on glass.

2.2.5   *Imaging and Image Analysis*

Samples were imaged on a Nikon Eclipse Ti at 10× magnification (0.33 μm/pixel) (Nikon, USA). Acquisition was collected with one of the following channels and filter sets: 390x440 (ex 390/22; em 440/40), 485x525 ((ex 485/25; em 525.30), 560x607 (ex 560/32; em 607/36), 648x684 (ex 648/20; em 684/24), 790x809 (ex 740/13; em 809/81). For capture and viability experiments, images were analyzed using the provided NIS-Elements AR Microscope Imaging Software.

Quantification of fluorescent intensity of beads and cells was performed using custom scripts written in MATLAB version R2016B (Mathworks, Natick, MA). All raw fluorescence images were corrected for background signal by subtracting the local median within a square-moving window with dimension at least 5 times the diameter of the cell type or bead of interest. Background-subtracted bead images were smoothed using a Gaussian filter ($\sigma = 0.66$ μm) and masks of the beads were generated by thresholding (Otsu method) off the autofluorescence of the unstained 485 nm channel. The relative density of EpCAM was quantified by calculating the mean intensity of the EpCAM channel (560) within the masks and dividing this value by the mean fluorescence.

For experiments investigating the relative EpCAM expression of cell lines, images were normalized to have zero local mean and unit local variance in order to improve the robustness of segmentation. Masks of the periphery of the cells were created by thresholding the normalized image at 1, and the relative EpCAM expression was quantified as the mean intensity within these masks.

For AR and GR localization experiments in LNCaPs, cell locations were manually marked in the cytokeratin channel (FITC), and masks of the cell were generated by thresholding (Otsu method) followed by morphological reconstruction using the manual markers. Nuclear masks were generated using the same method on the Hoechst channel, and cytoplasmic masks were calculating by subtracting the nuclear region from the cell masks. Relative nuclear and cytoplasmic expression of AR was quantified as the mean signal in these channels within each respective mask, and a nuclear localization metric was defined as the ratio of nuclear to cytoplasmic expression.

2.2.6   Cell Culture

Cells were cultured under sterile culture conditions at 37° C in 5% CO_2. VCaPs (courtesy of Dr. Scott Dehm, University of Minnesota) were cultured in DMEM (Gibco) supplemented with 10% Fetal Bovine Serum (FBS) (Gibco) and 1% Penicillin Streptomycin (PS) (Gibco). All other lines – LNCaPs (ATCC), Du145s (courtesy of Dr. Scott Dehm), 22Rv1s (courtesy of Dr. Douglas McNeel, University of Wisconsin-Madison), PC3s (courtesy of Dr. Scott Dehm), and PC3-MM2 (courtesy of Dr. C. Pettaway, MD Anderson Cancer Centre, TX, USA) were cultured in RPMI1640 media (#11875-093, Thermo Fisher Scientific) with 10% FBS and 1% PS. To maintain consistency across experiments, cells were counted, plated at 0.3 x 10^6 cells per well in a 6-well plate, and cultured for 48-hours prior to use.

2.2.7   Blood Processing

Peripheral Blood Mononuclear Cells (PBMCs) were isolated from whole blood for use as background cells. The whole blood – collected from healthy donors and treated with K3 EDTA (Biological Specialty Corporation) – was received within 24 hours of the blood draw and processed. Briefly, whole blood was mixed 1:1 with 1x PBS, overlaid on 15 mL of Ficoll Paque PLUS (17-1440-02, GE Healthcare), and centrifuged following the manufacturer's instructions. After centrifugation, the buffy coat was removed and diluted in 20 mL wash buffer (1x PBS supplemented with 0.5% BSA and 2mM EDTA). Cells were then centrifuged (200 rcf, 10 min), pelleted, and resuspended again in 20 mL wash

buffer and stored on ice until ready for use. Once ready for use, cells were centrifuged and resuspended as needed.

2.2.8   *Nucleic Acid Extraction and Quantification*

For RNA, cell samples (including 50 μg of beads) underwent either a spin column RNA extraction kit (AllPrep Spin Columns, Qiagen) or a magnetic bead-based extraction (Dynabeads mRNA Direct 61011, ThermoFisher). For the spin column, the manufacturer's protocol was followed, eluting into 15 μL of the provided elution buffer. For the magnetic bead-based RNA extraction, 200 μL of provided Lysis/Binding buffer and 20 μL of oligo(dt) beads were added to the cells. Using a magnetic rack (12321D, ThermoFisher), RNA was isolated following two 200 μL washes of Wash Buffer B (10 mM Tris-HCl (pH 7.5) (Sigma), 0.15 M LiCl (Sigma), 1 mM EDTA (Sigma)) followed by elution in 15 μL of elution buffer, 10 mM Tris-HCl. The eluted sample (including magnetic beads) was reverse transcribed using the High Capacity cDNA Reverse Transcriptase Kit (4387406, ThermoFisher) on a Techne TC-412 Thermal Cycler (37°C for 1 hour; 85°C for 5 minutes).

For DNA extraction (including 50 μg of magnetic beads), a spin column (QIAamp DNA Mini Kit 51304, Qiagen) and a silica magnetic bead-based approach were evaluated. For the spin column, the manufacturer's instructions were followed until elution where a modified elution volume of 15 μL was used. For the silica bead-based isolation approach, cells were lysed in 200 μL RLT Plus (1053393, Qiagen) with 5 μL of Magnesil KF magnetic beads (MD1471, Promega). Following lysis, beads and extracted DNA were washed in Wash Buffer B (above) and eluted in 15 μL of nuclease-free water.

To quantify, DNA or cDNA was mixed with LightCycler 480 Probes Master Mix (04535286001, Roche) and a Taqman assay for either GAPDH (Hs02786624_g1, LifeTech) (DNA) or HPRT (Hs11501003267_m1, LifeTech) (cDNA). The reaction underwent quantitative PCR on a LightCycler 480 (Roche) thermal cycler (pre-incubation 95°C for 5 minutes; 45 cycles of 95°C for 10 seconds, 60°C

for 30 seconds, 72°C for 1 second). The LightCycler software with the second derivative algorithm calculated cycle threshold.

2.2.9 *Statistics*

AR localization results were analyzed for difference by one-way ANOVA. Post hoc multiple comparisons were performed using a t-test with Bonferroni correction. Statistical significance was defined as $p \leq 0.05/15=0.0033$.Results

2.3 Results and Discussion

2.3.1 *Basic Magnetic Bead Characterization: Antibody Binding and Surface Density*

Cell capture exists as a balance between the frequency of antibody-antigen interactions (which can be problematic at low antibody densities), and steric hindrance (which emerges at high antibody densities). To first visualize antibody density on beads, bead bound capture antibody was fluorescently labeled. The fluorescence readout of antibody density on the magnetic beads generated of antibody-density curves for each bead type (**Fig. 2.1A**). Each magnetic bead demonstrated a saturation point, at which maximal binding was observed. Upon saturation, addition of more antibody resulted in no further increase in signal. In translating the generated antibody density curves (FlowComp, M-280) to cell capture (**Fig. 2.1B**), similar results are observed with poor capture at low antibody densities (too few antibody-cell interactions). Capture increased with increasing antibody density, until upon surpassing the maximal binding capacity densities identified in **Fig. 2.1A**, a decrease in capture was observed between capture at ~7 ng and 40 ng antibody per μg bead (FlowComp p=0.027; M-270 p =0.047, N=3). The subtle decrease was likely a result of steric hindrance due to the high, saturated density of the antibody. Notably, at least in the case of LNCaPs (high EpCAM expresser), too low of antibody density was much more detrimental to capture than too high of antibody density. Conceivably, the relationship between antibody density and capture is dependent on a number of factors including: antibody, cell type, antibody presentation, and the size of the magnetic bead and cell. Thus, understanding the balance between these metrics remains important for optimizing capture of a target population.
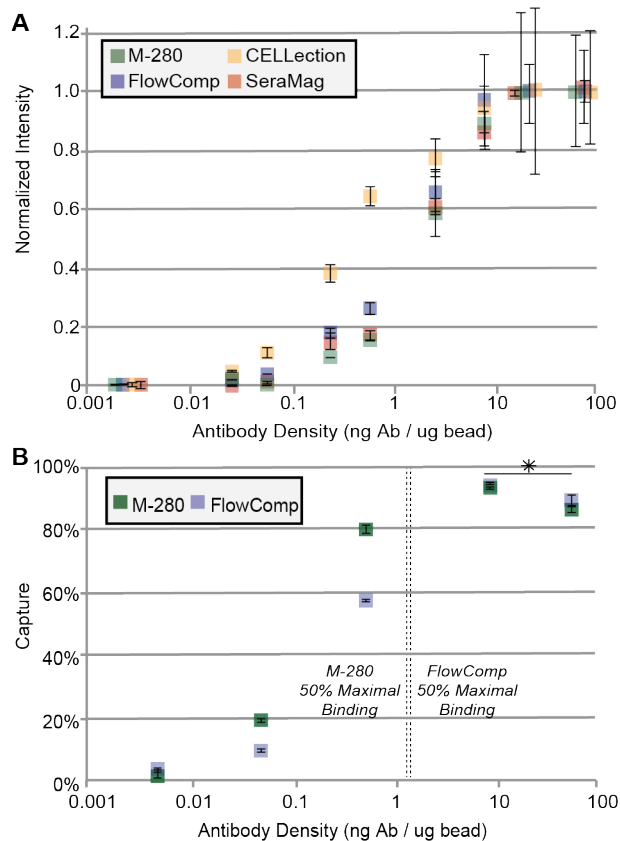
Fig. 2.1 Characterization of bead-antibody binding. (A) Magnetic bead fluorescence intensity curves generated by varying densities of a fluorescently labeled anti-EpCAM antibody on beads. A 50% maximal binding capacity for each bead type was then identified (Supplemental Table 2). Dots represent the average of three technical replicates consisting of 100 beads each (300 beads total) were averaged and presented on the graph; error bars represent standard deviation of the analyzed population. (B) Impact of the bead's antibody density on target capture. The 50% maximal antibody binding density for M-280s and FlowComps 50% maximal binding is denoted by their respective vertical dashed line. Points represent average of three technical replicates; error bars represent standard deviation; * denotes p < .05, which in this case applied to both M-280 and FlowComp.

Based on both the fluorescence titration curve of each antibody and the impact of antibody density on cell capture, an antibody density of fifty percent of maximal capacity was identified for each bead type (**Fig. A.2**). Due to the different sizes of beads used (**Fig. A.1**) as well as potential differences in surface roughness (not evaluated), the surface area of each bead type varies. Thus, rather than choosing a set concentration of antibody per milligram of beads (which ignores surface area discrepancies across bead types), the fifty percent maximal capacity was determined for each bead (**Fig. 2.1A**) and used in all capture experiments unless otherwise noted. Similarly, the number of beads per milligram was different across bead types. To determine the impact the quantity of beads had on cell capture, bead quantity per

sample was titrated using two different EpCAM cell lines, a high expressing EpCAM line (LNCaPs) combined with either a medium or low EpCAM expresser (22Rv1s or Du145s respectively) (**Fig. A.3**). Across the evaluated conditions, maximal capture was reached with ~100 μg of beads (e.g., upon addition of beads, little-to-no increase in capture was observed); thus for consistency, 100 μg of beads was used per sample (**Fig. A.4**). In applying beads to cell isolation, the combination of antibody density and bead number can greatly impact results; thus in ideal situations, antibody density and bead number introduced per isolation should be titrated for each application.

2.3.2   Cell Capture

Target cell capture efficiency and purity is one of the most important magnetic bead characteristics. To characterize target capture across each bead type, a low, medium, and high EpCAM expressing cell line (Du145s, 22Rv1s, and LNCaPs respectively) was tested (**Fig. 2.2A**). Despite the use of an identical antibody to capture with, capture varied greatly across magnetic bead types, especially in the low expressing Du145s; CELLection and FlowComp magnetic beads resulted in the lowest capture, while M-270s notably captured the largest population of Du145s. While the identical antibody lot (and conjugated stock with the exception of M-270 beads) was used across all capture experiments, differences in the bead surface (e.g., roughness, curvature due to size differences) or the functionality of the surface could impact how the antibody orientates on the surface of the bead. Antibody orientation would impact the antibody's potential for successful epitope binding, possibly explaining the variable capture observed. Similarly, how (and where) the antibody is biotinylated could impact antibody-bead performance, highlighting the need to optimize each component of the process for each bead type and new application.

While specific capture of the target is critical, for many endpoints (e.g., sequencing), high purity is also required as contaminating populations bias and mask target cell signatures. To determine the potential of contaminants to reduce purity for each magnetic bead type, the non-specific capture of each bead was estimated by incubating the beads with a mixed background population (PBMCs) (**Fig. 2.2B**). Sera-Mag and M-270 beads had the lowest rate of non-specific capture of PBMCs compared to the almost

10-fold increase in non-specific binding with M-280s and CELLection beads. The balance between specific capture (target cells) and non-specific capture (i.e., background cells) often determines the endpoints available as these metrics will determine both yield and purity, a consideration in identifying an assay-specific bead type.

Both capture efficacy and purity may also be impacted by the bead isolation method used: direct or indirect. Direct cell capture is the most common, wherein the pre-bound antibody-bead complex is incubated with the cells. In contrast, indirect capture involves incubating the antibody with the cells, followed by bead capture of the antibody-labeled cells. Typically, indirect capture results in higher capture efficacy, yet also results in increased contaminants. For indirect capture, the antibodies dispersed in the sample are free to interact, and may incidentally bind non-specifically with contaminant cells at higher frequencies than when attached to the bead in direct capture; the balance of captured target cells (increased target yields higher purity) and contaminants (increased contaminants yields lower purity) will ultimately determine the resultant impact of indirect capture on purity for each bead type. Indirect versus direct capture of low EpCAM expressing cells (Du145s) from PBMCs was evaluated for each bead type (except for M-270s, which are limited to direct capture) (**Fig. 2.2C**). Additionally, the resultant purity is reported (**Fig. 2.2D**). Interestingly, capture of Du145s was highly variable across bead types. Notably, if direct capture was low (average ~10% capture), switching to indirect capture had no impact as demonstrated by the CELLection and FlowComp magnetic beads. With indirect capture, the antibody is first added to the cells and is able to bind to the cells (little-to-no variation across conditions); then beads are added and the antibody-cell complex is bound to the beads. Based on the physical surface (e.g., roughness), the positioning of streptavidin on the surface of the bead, and the location of biotin(s) on the antibody, orientation of the bound antibody on the bead can be impacted. Binding of the biotin (or modified biotin) antibody to the streptavidin may result in an antibody orientation, which results in a decoupling of the antibody from the EpCAM, resulting in the release of the cell; this may explain why

some bead types saw little difference between direct and indirect. In contrast, M-280s and Sera-Mag magnetic beads improved capture efficacy when transitioning from direct to indirect capture.

When evaluating the impact of direct and indirect capture on purity (**Fig. 2.2D**), the results were surprisingly mixed. While M-280s gained in capture efficiency using indirect capture (from ~30% to ~50%), overall purity did not change, due to the increased contaminants captured in parallel. SeraMags, which also saw gains in capture efficiency (from ~30% to ~85%) with indirect, saw an increase in captured contaminants, but overall had an increase in purity with indirect capture. In contrast, CELLection and FlowComp saw little differences in capture efficiency and no notable changes in purity. While differences in indirect and direct capture are difficult to predict without experimentation and variable across beads, the gains in target capture efficacy for M-280 and Sera-Mag beads highlight the potential benefit of evaluating these metrics when evaluating bead types.
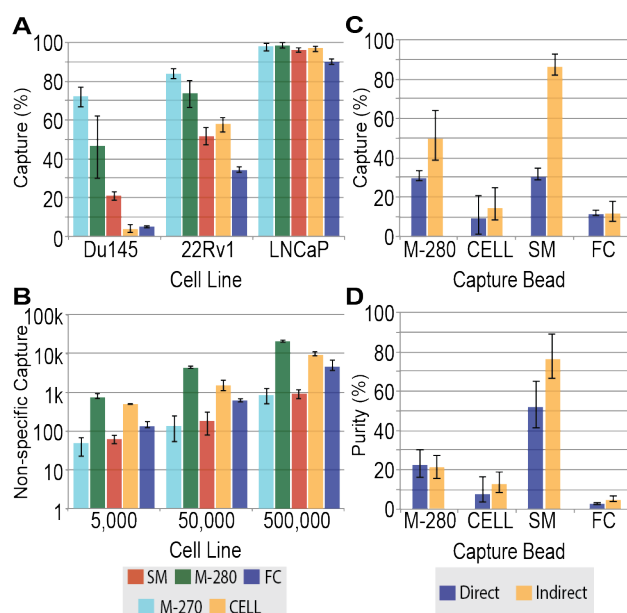
Fig. 2.2 Characterization of cell capture. (A) Capture of EpCAM-expressing cell lines (Du145 = low, 22Rv1 = medium, LNCaP = high) by each bead type. Beads are abbreviated as follows: SM = Sera-Mag, FC = FlowComp, CELL = CELLection. (B) Non-specific capture of PBMCs by each bead type across varying PBMC inputs. (C) Direct versus indirect capture of Du145s from a PBMC background. (D) Resultant purity of captured target cells from direct and indirect capture of Du145s. In all plots, the bars represent the technical replicate average (n=3) with error bars representing standard deviation.

2.3.3   *Magnetic Bead Release Characterization*

To characterize release, the release of fluorescently labeled EpCAM antibodies from each bead type was first characterized, followed by cell release from each bead type. For each of the bead types, release is accomplished through different mechanisms. FlowComp beads accomplish release by introducing D-biotin (B-1595 or B-20656, ThermoFisher) or D-desthiobiotin (D-20657, ThermoFisher). These molecules have a higher affinity for the (modified) streptavidin on the bead surface, and thus displace DSB-X biotin from the streptavidin, thereby releasing the antibody and attached cell from the bead. In contrast, CELLection beads utilize a DNA linker to connect the antibody to the bead; the provided DNase I can cleave the DNA linker attaching the bead to the antibody to release captured cells from the bead. To fluorescently characterize release, a fluorescent secondary system was used. Beads with different densities of primary antibody (22.5, 0.5, and 0.025 ng antibody per μg PMP) were placed in release buffer for varying durations, the magnetic beads removed from release buffer, and the remaining fluorescence on the magnetic bead measured (**Fig. 2.3A,B**). Using this system, a decrease in fluorescence intensity corresponds to a release of antibody. While CELLection (**Fig. 2.3A**) demonstrated the most rapid release, both chemistries demonstrated at least 50% release within 5 minutes. Notably, FlowComp bead release seemed hindered at higher antibody densities, resulting in incomplete release. The delayed release by FlowComp at a high, maximal antibody density could be due to limited access of the release buffer into the tightly packed antibody-bead complex. CELLection may not suffer from this issue as, by using a DNA linker between the bead and antibody, CELLection provides added space between the antibody and bead allowing easier access of the DNase.

Next, the release of captured cells across different time points was evaluated using 22Rv1s. Cells were captured on magnetic beads and allowed to incubate in release buffer for varying durations. The released and bead-bound fractions were then used to determine release efficiency across time (**Fig. 2.3C**) (**Fig. A.5**). Within 5 minutes, maximal release was obtained for each cell line. This data is comparable to the fluorescence data (**Fig. 2.3A,B**), where for medium and low antibody densities, substantial release

occurred by 1-minute, then diminishing gains were observed as time increased. In the context of cells however, the release was delayed – complete release occurred at 5 minutes instead of 1 minute. As cells are often bound to beads via multiple antibody linkages, multiple linkages must be broken to release the cell; this is likely slowing the process when compared to the fluorescent antibody characterization.

The release characteristics of the low, medium, and high EpCAM-positive cell lines were then characterized. Using the FlowComp (blue bars) and CELLection (yellow bars) beads, we evaluated the best-case capture efficacy of each bead (**Fig. 2.3D**). Next, cells were released for 20 minutes (**Fig. 2.3E**). CELLection beads were the most effective at releasing cells, releasing ~78-88%. Release from the FlowComp beads was considerably lower than the CELLection beads; yet, both bead types resulted in some cell loss during release due to an unreleased fraction remaining bound to the beads (**Fig. 2.3F**). The lack of complete release and discrepancies between release efficiencies of FlowComp and CELLection beads could be due to a number of differences in the release approaches. Release is dependent on ensuring the bead binds the cell through the antibody as the antibody is ultimately released from the bead. For instance, due to the close proximity of cells and beads, cells may non-specifically interact with the surface of the bead. As a result, antibody-based release methods become ineffective at release the beads as the cells are no longer solely bound via the specific antibody interaction. Additionally, CELLection use a spacer (DNA linker) between the cell and bead. This spacer may both place some additional distance between the target cell and bead (reducing direct bead interactions) as well as enable easier access of the releasing agent (DNase) to its target especially when a number of antibody-EpCAM interactions are likely occurring in a small area (e.g., high expressing cells, LNCaPs). Although difficult to determine the mechanism(s) impairing release, evaluating different approaches with a relevant target of interest is important for optimal results.

While magnetic beads are enabling in isolating a target population, bead removal may be required for optimal downstream techniques such as fluorescence microscopy. Releasable bead chemistries enable downstream separation and removal of the magnetic beads following capture, frequently by dissociation

of the antibody and bead. Thus, while release can be advantageous for an assay (i.e., imaging), the benefits may be counterbalanced by a decreased, resultant captured (and released) target population.
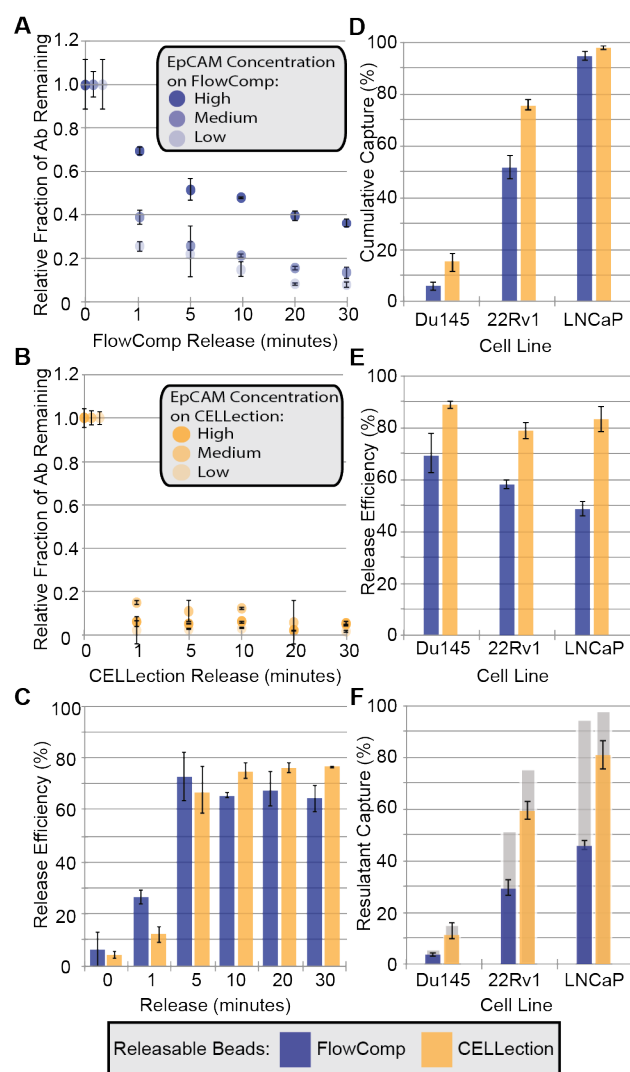
Fig. 2.3 Characterization of release from FlowComp and CELLection beads. (A & B) Release of a fluorescently labeled anti-EpCAM antibody from (A) FlowComp and (B) CELLection beads. Beads were labeled with low, medium, and high levels of antibody and released for the specified time intervals. (A&B) Dots represent the average of three technical replicates with each technical replicate representing the average of 100 beads (total of 300 beads); error bars represent standard deviation of the technical replicates. (C) Release of 22Rv1s from FlowComp and CELLection beads across time. (D) Capture efficiency of both FlowComp and CELLection beads when used to capture Du145, 22Rv1, and LNCaP cells. (E) Release efficiency of the three cell lines following bead-based capture. (F) Effective capture following release of the cells. Gray bars represent the population of cells lost during the release process due to inefficient release. In each plot, bars represents an average of n=3; error bars represent standard deviation.

2.3.4   *Impact of Magnetic Beads on Imaging and Analysis*

Once a cell population of interest is captured, many downstream applications involve fluorescent protein staining, either for verification and identification of the population or for protein localization and expression. In either application, fluorescent signal from the magnetic beads may ultimately limit the fluorescent stains or channels as well as impact the ability to discern localization or expression. Thus, the fluorescence of each bead type on glass was initially characterized across five different fluorescent channels (**Fig. 2.4A**), with the flowing excitation / emission filters (center(range)): 390(22) / 440(40), 485(25) / 525(30), 560(32) / 607(36), 648(20) / 684(24), and 740(13) / 809(81) (as highlighted in the methods) (**Fig A.6**). Each bead had some autofluorescence; while the intensity varied between bead types and channels, each bead peaked at an emission of (560 nm) (**Fig. 2.4A**).

To evaluate the potential impact of magnetic beads on evaluation of both protein expression (staining intensity) and localization (based on a nuclear and cytoplasmic staining), LNCaPs were captured with each magnetic bead type, fixed, and stained with a nuclear stain (Hoechst) as well as antibodies to pan-cytokeratin (Alexa790), EpCAM (PE), and androgen receptor (AR) (Alexa488). The captured cells were then compared to a bead-free population (**Fig. 2.4B**). Beads were found to have a variable impact on identified nuclear area, which conceivably would impact the ability to easily discern nuclear localization of proteins (**Fig. 2.4C**). Total calculated cellular AR resulted in statistical difference in every bead type compared to bead-free cells demonstrating the potential of magnetic beads to modify detected signal per cell, an issue when attempting to identify populations based on expression (or lack of) (**Fig. 2.4D**). In all bead types evaluated, cellular AR decreased, likely due to the beads attenuating fluorescent intensity; however, if a protein was lowly (or not) expressed, bead autofluorescence might lead to false quantification of positive signal. Using both the AR signal and localization based on determined nuclear area, the ratio of AR nuclear localization (nuclear AR to total AR) was calculated (**Fig. 2.4E**). While localization ratio seemed to correct for some noise observed in cellular AR signal, certain beads better maintained expression and localization patterns of bead-free cells (Sera-Mag and CELLection).

Magnetic beads impact evaluation of fluorescent staining for protein expression and localization as well as identification of nuclear area. One additional variable, bead coverage of the cells, is also likely to influence these results. All bead types appeared to variably cover cells ranging from only a few beads per cell to complete coverage within a single sample, highlighting cell-to-cell heterogeneity. As the number of beads, which bind to a cell is difficult to control and highly variable (as observed by the range of bead coverage of the cells within each bead type used), we assessed the impact of beads bound to cells using the entire population of captured cells (both highly and sparsely covered cells). Thus, assays relying on endpoint protein localization or cell identification through fluorescent staining should closely evaluate the impact of cell isolation beads, as beads can significantly distort population appearances; distortion likely impacted by the number of beads bound to a cell, a variable difficult to control.
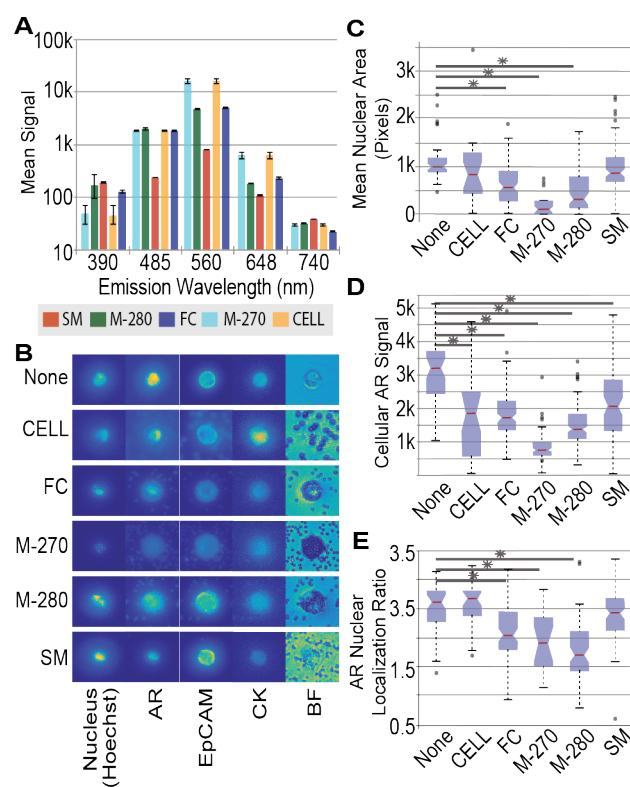
Fig. 2.4 Negative Selection. (A) Impact of magnetic beads on downstream fluorescence microscopy readouts. (A) Baseline autofluorescence of magnetic beads imaged on glass across five different filter sets with emission wavelengths listed. Bars represent the average of 200 beads; error bars represent standard deviation of the analyzed events. (B) Example images of single cells bound to each of the bead types (note: bead coverage of the cell greatly varied cell-to-cell for each bead type from a few beads to complete coverage), demonstrating variable staining patterns as influenced by the presence of cell isolation magnetic beads. (C) Impact of magnetic beads on identifying

LNCaP nuclear area based on Hoechst staining for a nucleus. (D) Identified cellular androgen receptor signal. (F) Ratio of nuclear to cytoplasmic androgen receptor identified in LNCaPs captured with each bead type and compared to bead-free cells. In the box plots, 50 cells were analyzed per condition; the notch represents the 95% confidence interval of the median and the circles are possible outliers. A statistically significant difference with respect to the no-bead group is indicated (*).

2.3.5   *Post-capture Culture of Cell Lines*

Following cell isolation, many assays require the user to culture the cells rather than perform a terminal endpoint assays such as intracellular staining. Thus, the viability of cells captured via magnetic beads was evaluated. Anti-EpCAM beads were incubated with ~5,000 22Rv1s or LNCaPs and isolated resulting in a captured bead-bound population. A total of 50 μg of magnetic beads were used for each bead type. Following isolation, the cells and beads were transferred into a 96-well plate and cultured for 3 days. After 3 days of culture, cells viability was assayed (**Fig. 2.5**).
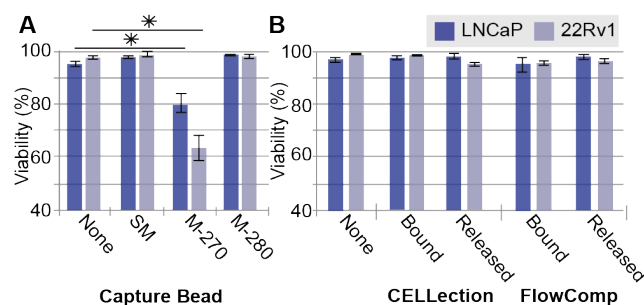
Fig. 2.5 Cell viability following capture and release. (A) The viability of cells (LNCaP, 22Rv1) bound to non-releaseable beads compared to untouched cells (underwent no magnetic bead isolation) following a 3-day culture (Abbreviations: SM = Sera-Mag). (B) The viability of cells bound to (bound) and released from (released) releaseable beads (CELLection, FlowComp) following a 3-day culture. Bars represent average of three technical replicates; error bars represent standard deviation.

For the non-releaseable beads, results indicated M-280 and Sera-Mag beads have no statistical impact on cell viability compared to cell only (no-bead) viability for both cell lines (**Fig. 2.5A**). M-270 beads resulted in a decrease in viability relative to the cell only control (p-value < 0.01 for both LNCaPs and 22Rv1s). For the releasable CELLection and FlowComp beads, viability is shown for 1) a no bead cell only control (None), 2) cultured bead-bound cells (Bound), and 3) cells cultured post capture and post release (Released) (**Fig. 2.5B**). Overall, the viability across conditions – including released and bead-bound cells –remained high in both LNCaPs and 22Rv1s. For LNCaP and 22Rv1 cells, no significant differences were seen, regardless of the bead type used or the culture condition (i.e., bound or release).

Ultimately, viability is likely an artifact of cell type, cell density, and bead density; nevertheless, these high viability results demonstrate both the promise and potential impact on post-capture culture.

2.3.6   *Integration of Magnetic Bead-based Cell Isolation with Standard Nucleic Acid Extraction Methods*

Downstream of cell isolation, many endpoints involve nucleic acid isolation. To characterize the potential impact of each cell isolation magnetic bead type on nucleic acid isolation, both RNA and DNA were evaluated. Ultimately each nucleic acid isolation protocol differs in buffers, nucleic acid binding mechanisms, and impact of the cell isolation beads; thus, to highlight the variable impact of cell isolation beads, two extraction methods were analyzed for completeness. For both RNA and DNA, a low cell number sample (~5,000 cells) was evaluated using a spin column and a bead-based technique.

For RNA, 50 µg of cell capture magnetic beads were added to each cell sample prior to addition of any lysis buffer to ensure the impact of cell isolation beads on the entire RNA isolation process was evaluated. RNA was then isolated with a magnetic bead-based method (Dynabeads mRNA Direct) as well as a spin column (Qiagen RNeasy Mini Kit). Following isolation, the eluted RNA (and beads) underwent reverse transcription (RT) into cDNA; the cDNA was quantified by real-time quantitative-PCR (qPCR) (no beads were loaded into the reaction). When cell isolation magnetic beads were integrated into a spin column isolation, little loss in RNA was detected compared to the cell only control (**Fig. 2.6A**); rather FlowComp magnetic beads resulted in a statistically significant (p-value = 0.032) increase in detected mRNA (**Fig. 2.6A**). Similarly, the bead-based mRNA extraction – the Dynabeads mRNA Direct Kit – resulted in no statistical difference in RNA quantified from the cell only condition or, in the case of CELLection and Sera-Mag (M-270s resulted in an average increase, but was not significant), a significant increase in RNA was detected (Sera-Mag p-value = 0.043; CELLection p-value = 0.012) (**Fig. 2.6B**).

Similarly for DNA, 50 µg of cell isolation magnetic beads were added to the cells prior to DNA isolation. Both a magnetic bead-based approach (DNA-binding silica bead) and a spin column approach (QIAamp DNA Mini Kit) were used to evaluate the potential impact of cell isolation beads. To quantify

the isolated DNA, qPCR was performed for a housekeeping gene (GAPDH). When DNA was isolated by

spin columns, no statistical differences were seen in detected DNA yields (**Fig. 2.6C**). In comparison,

when DNA was isolated by silica beads (**Fig. 2.6D**), DNA yield (via GAPDH) was comparable to the

control for M-280, Sera-mag, and CELLection beads. However, a decreased yield was observed when

either FlowComp or M-270s beads were present during the lysis step (p-value of 0.049 and <<0.01,

respectively). Notably, FlowComp bead resulted in some loss (approximately half the DNA yield

compared to control), but M-270s resulted in over a 90% decrease in detected DNA, a very different

result compared to the spin column DNA isolations. In this finite test of five bead types, M-270s and

FlowComp beads were the only beads that resulted in loss of DNA, specifically when DNA was isolated

using the bead-based DNA isolation protocol.
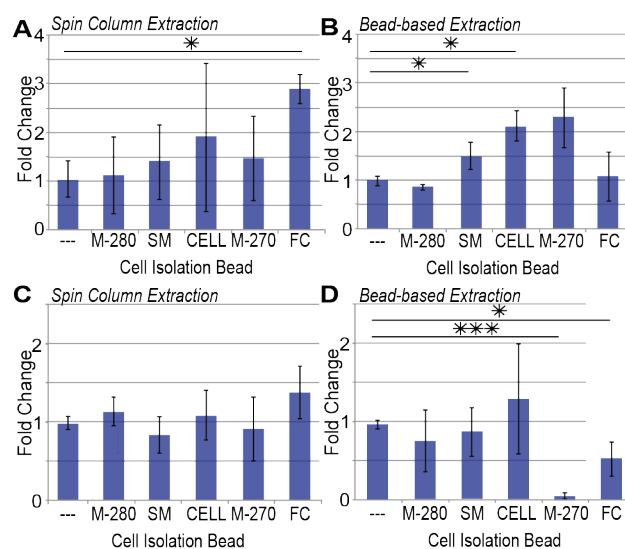
Fig. 2.6 Characterization of nucleic acid extraction with cell isolation magnetic beads present. (A&B) Relative fold change in mRNA transcript (HPRT) detected from LNCaPs. Isolations containing cell isolation beads were compared to no bead controls for two methods of RNA extraction: (A) spin columns and (B) bead-based extraction. (C&D) Similarly, relative fold change in GAPDH from DNA extracted via (C) spins columns or (D) bead-based extraction. Bars represent average of three technical replicates; error bars represent standard deviation; * denotes p < .05 and *** denotes p < .001; --- indicates the cell only, no bead control (Abbreviations: CELL = CELLection, SM = Sera-Mag, FC = FlowComp).

Across isolation methods in both RNA and DNA (i.e., spin columns versus magnetic bead-based

extraction), cell isolation magnetic beads had variable impacts on nucleic acid extraction based on the

nucleic acid approach used.   In bead-based DNA isolation, cell isolation magnetic beads could

significantly hinder yield, yet identical cell isolation beads had no statistical impact in a spin column isolation. Furthermore, bead types did not effect capture of all nucleic acids alike; a cell isolation bead that seemed to impact DNA yield, did not necessarily impact RNA yield (e.g., M-270). Overall, the variable impact of cell isolation beads across nucleic acid extraction methods highlights many of the potential nuances in integrating cell isolation beads into complex cell isolation protocols.

2.4   Conclusion

Cell isolation magnetic beads enable the rapid targeting of nearly any cell population, paired with a nearly endless offering of commercial antibodies. Yet, how the isolation magnetic beads perform and how they integrate into downstream endpoints impacts their utility to users. Different cell isolation magnetic-beads come with trade-offs in their ability to facilitate and integrate into different endpoints of cell isolation protocols. For baseline performance metrics, M-280s facilitated strong target capture enabling use of both direct and indirect capture approaches. For purity, Sera-Mag and M-270s paired strong capture with low non-specific binding of a complex background PBMC population. While FlowComp and CELLection did not perform as well, these beads enable release, which may be required for different culture applications as well as facilitate precise fluorescent immunohistochemistry endpoints such as protein localization. All cell isolation beads demonstrated compatibility with RNA and DNA extraction; yet results highlighted the method and buffer dependency of those results. This article aims to evaluate the beads in the presented context of EpCAM-specific cell capture, highlighting the range of results obtainable depending on the bead type utilized. While this paper attempted to ensure optimal performance across bead types, the attempts to standardize traits (e.g., fifty percent maximal binding capacity, bead number added) could all strongly influence resultant capture (and release). Similarly, while buffers were standardized across isolations, the buffers and additives (e.g., FBS, BSA, EDTA) could impact performance. Thus, this paper serves to introduce different bead types and provide insight to downstream users. Ultimately, further investigation is required to better understand the mechanisms behind the observed variation and direct the design of improved magnetic beads for cell applications.

Chapter 3   Adaptive Exclusion-based Sample Preparation Platform for Integrated Rare Cell Isolation and Analyte Extraction 3

Rare cell populations provide a patient-centric tool to monitor disease treatment, response, and resistance. However, understanding rare cells is a complex problem requiring cell isolation/purification and downstream molecular interrogation additionally challenged by populations, which vary patient-to-patient and change with disease. As such, cell isolation platforms must be amenable to a range of sample types, maintaining high efficiency and purity. The Multiplexed Technology for Automated Extraction (mTAE) is a versatile 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.

3.1　Introduction

Emerging discoveries have begun to highlight the biological and clinical significance of rare, discrete cell populations (e.g., minority 'stem' populations[23], circulating fetal cells[158,159], and circulating tumor cells[160]). Yet, rare cells are often masked within larger, more diverse backgrounds of cells (e.g., the bloodstream), complicating isolation[77,161] 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[162,163,164]. 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.[165]

There are two primary approaches in the growing field of antibody-based cell isolation: positive and negative selection[166,167]. 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[168,169], carbohydrate receptors[170], 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. Despite these advantages, negative selection typically results in incomplete background removal, yielding relatively low purity[171]. 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 through a 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[167,172]. However, it is now clear that enumeration alone is unlikely to revolutionize cancer monitoring and additional endpoints beyond enumeration will likely be necessary to expand the clinical utility of CTCs[165]. 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[173,155], CTCs undergoing epithelial-mesenchymal transition (EMT)[174-175], cancer stem cells[176]), emerging markers[177], and varying background populations[178,179]. 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.[8] 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.[6-180] 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.[172]

Building on a suite of exclusion-based sample preparation (ESP[TM]) technologies[157,181,182], 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. 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[TM]) technology – a low shear method for achieving high purity extraction of PMP-bound analytes.[157] 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.

3.2   Materials and Methods

3.2.1   Automated ESP Platform Overview

Cells and NA were isolated on a Gilson PIPETMAX automated liquid handler (Gilson Inc.) by integrating Exclusion-based Sample Preparation (ESP) technologies: SLIDE (Sliding Lid for Immobilized Droplet Extraction)[157] and a custom component that enables flexible manipulation of PMPs[183]. SLIDE is a technique for isolating and purifying PMP-bound analytes. SLIDE leverages convex droplets and a modified automated pipette head modified to house magnets, which is used to capture and transfer PMPs between wells. The SLIDE technology has been integrated into Gilson's EXTRACTMAN, thus we used EXTRACTMAN extraction plates (#22100008, Gilson) and collection strips (#22100007, Gilson) to achieve these isolations. 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. PMP-bound analytes 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 (**B.1 Automated ESP Platform Overview**).

3.2.2 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.

3.2.3 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. Samples were collected from patients with cancer (**Fig. B7-9**) who had a signed informed consent document under a University of Wisconsin-Madison IRB approved protocol. 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.

3.2.4   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.

3.2.5   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 (**Fig. B.2**): LNCaPs, HCC-2218, and PC3-MM2 (sub-clone of PC3 cell line). 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).

3.2.6   *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. [68] 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.[184] and Casavant et al.[182] 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.

3.2.7   *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**.**

3.2.8 *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[TM], 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[TM], 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 (Specified in Table 6, 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 (**Fig. B.6**).

3.3   Results

3.3.1   *Automated cell isolation platform design and optimization*

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 **(Fig. 3.1A,B)**. mTAE builds upon an automated liquid handler (PIPETMAX[TM], Gilson, Inc.), leveraging the SLIDE technology described by Casavant et al.[157], the magnetic system described by Guckenberger et al.[183], and EXTRACTMAN[TM] (Gilson) consumables (i.e.,

strips and plates). 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. 3.2A)** and non-target cells (**Fig. B.1**) 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 NA 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.
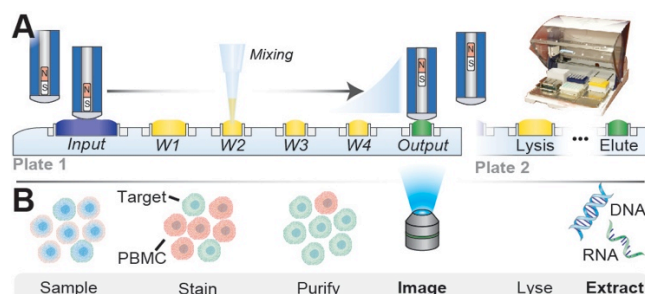
Fig. 3.1 mTAE system overview. (A) Cell extractions are performed on a modified Gilson PIPETMAX. The extraction plate contains six wells per sample with four samples per plate. The input is filled with sample (PBMC sample) and anti-target PMPs while the remaining wash wells (W1-W4) contain a combination 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 with Contrived Samples

3.3.2 *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) **(Fig. B.2)** from a background of 10 million PBMCs **(Fig. 3.2B)**. 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 **(Fig. B.2)**. 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 **(Fig. 3.2C,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.
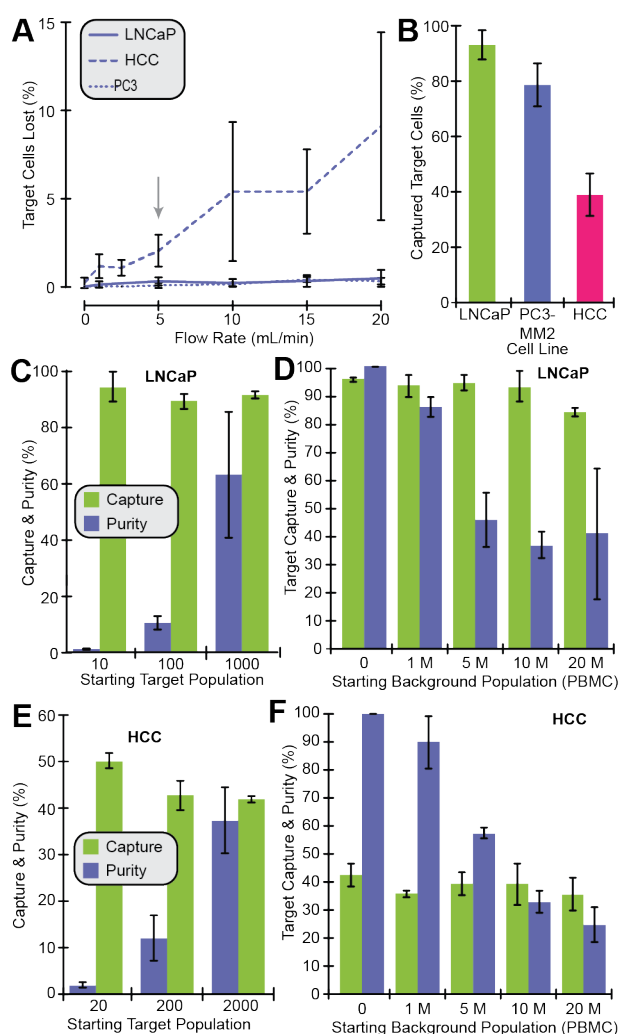
Fig. 3.2 Platform characterization and cell line validation. (A) Loss of target-cells as a result of flow introduced by pipette with flow rates ranging from 0 to 20 mL min⁻¹. A flow rate of 5 mL min⁻¹ was used for all subsequent experiments as noted by the arrow. (B) Capture efficacy of 500 target cells 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.

Next, to assess how background populations impact capture efficacy and purity, 1,000 target cells were isolated from 0-20 million background cells (PBMCs) **(Fig. 3.2D,F)**. For both LNCaPs and HCCs, capturing 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 (**Fig. B.3**). These results begin to characterize performance parameters and validate mTAE as a gentle, effective method of isolating rare cells from diverse background populations.

3.3.3   *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[185,186] and breast cancer[155,187,188]; within a subset of cancers, EpCAM-based CTC enumeration has provided predictive insight into prognosis and overall survival.[189,190] 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 **Fig. 3.3A,B**. While most CTCs were found as individual cells, clumps of CTCS **(Fig. 3.3C)** were identified in a subset of patients, a phenomenon previously observed in other CTC isolation platforms.[191,6] Capture was reported by quantifying the positively identified CTC events per milliliter of whole blood assayed **(Fig. 3.3D)**. 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.[152,192] 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 (**Fig. 3.3F**) and purity's relationship with CTC count was also evaluated (**Fig. 3.3E**). 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. 3.3G**). 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) (**Fig. B.4**). 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.
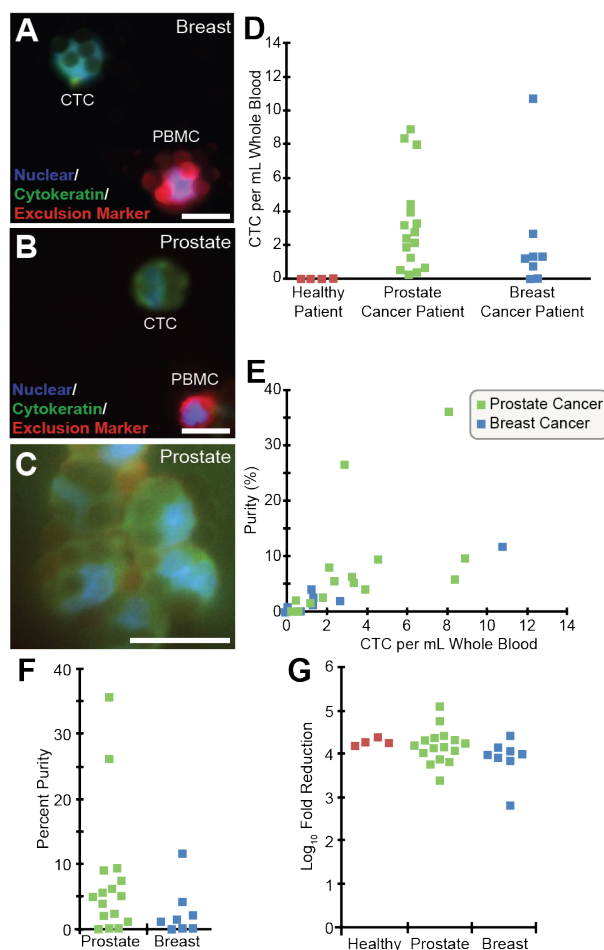
Fig. 3.3 Positive Selection. (A, B) Image of a CTC and PBMC captured from (A) a breast cancer sample and (B) a prostate cancer sample. (C) Image of a CTC clump captured from a prostate cancer sample. Small circular distortions in the stains are artifacts of the PMPs partially attenuating the fluorescent 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.

3.3.4 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. 3.4A)**. 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.[180] Using this approach, non-target cells, expressing low quantities of the selected depletion markers, may be been missed, contributing to the higher contaminant population.
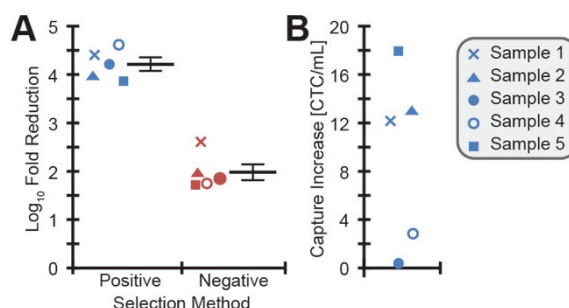
Fig. 3.4 Negative Selection. (A) A comparison between positive and negative selection demonstrating the efficacy in reducing the background PBMC population. Each point represents a single (breast cancer) patient sample (n=5). The large line represents the average and error bars represent standard error. (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.

Despite reduced background depletion, negative selection resulted in a greater number of CTCs identified (defined nuclear events positive for cytokeratin (CK), negative for exclusion markers) **(Fig. 3.4B)**. 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 poorly 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.

3.3.5 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.

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) (**Fig. 3.5A**). This result contrasted with the cell line characterization; characterization with cell lines identified little impact of background (0-20 million PBMCs) on specific capture (**Fig. 3.2F,G**) (i.e., capture efficiency was unchanged by increasing background populations). However contrived cell line samples are often inadequate representatives of patient samples. Interactions between CTCs and blood cells in circulation are relatively unknown, and

thus, as suggested by the improved CTC capture, there may be benefit in applying a depletion step prior to capture.

The impact of sequential selection on overall purity however, 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[193,194]).

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) (**Fig. 3.5B**). 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 (**Fig. 3.5C**). While it is unlikely that a sample with <1% purity would have much value beyond enumeration, 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.
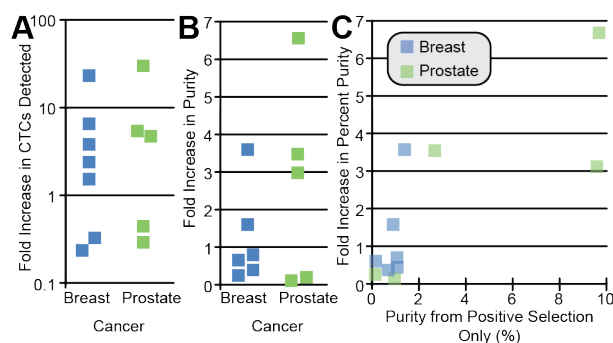
Fig. 3.5 Combinatorial selection. (A) Fold increase in CTCs detected when comparing CTCs captured with combinatorial to positive selection. The majority of samples from both breast and prostate saw an increase in CTCs

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.

3.3.6   *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., RNA, 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. 3.6A,B)**. 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 **(Fig. B.5)**; however, the presence of M-270s had an adverse impact on DNA, significantly impacting DNA yield **(Fig. B.5)**. 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.
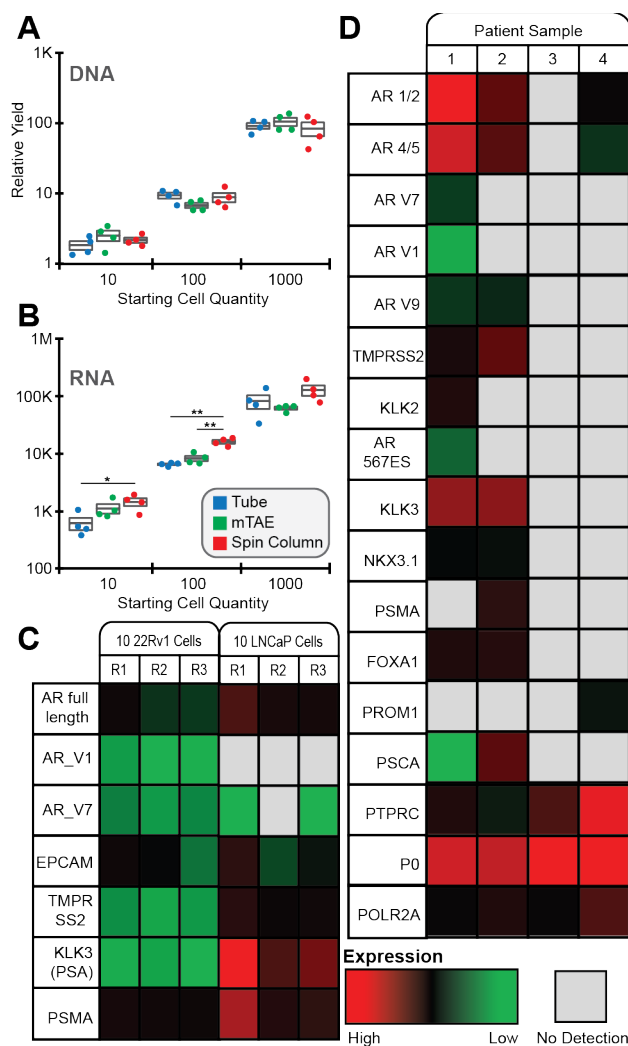
Fig. 3.6 Automated NA acid extraction on mTAE. (A) DNA extraction from 10, 100, and 1,000 cells compared across three methods: manual tube-based method (i.e., PMP-based capture followed by washing of NA bound PMPs performed in a micro-centrifuge tube using a magnetic tube rack), mTAE, and Qiagen QIAmp 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 (Ct values represented as a heat map). Cells were captured on mTAE followed by on-chip RNA extraction.

In CTCs from patients with prostate cancer, detection and quantification of AR gene expression, AR variants, 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).[195] 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 further validated specificity of the detected transcripts, demonstrating increased detection of the AR splice variants V1 and V7 in 22Rv1s in comparison to LNCaPs (**Fig. 3.6C**). Following cell line characterization, mRNA was extracted from prostate cancer CTCs (Fig. B.9) following positive selection via EpCAM. Detected transcripts include AR variants and downstream AR transcripts highlighting the involvement of these pathways in prostate cancer CTCs as well as confirming mTAE's ability to not only capture CTCs, but also detect therapy relevant transcripts (**Fig. 3.6D**).

3.4   Discussion

The isolation of specific cell populations has facilitated an understanding of the diverse biological roles and disease relevance of heterogeneous cell types. Despite the long history of cell isolation, growing interest in rare cell populations has pushed isolation technologies to their limits. Due to dynamically changing isolation criteria, rare cell capture has presented new challenges, creating demand for simple, yet versatile platforms able to further the understanding of rare populations. To assess cell isolation technologies, CTCs provide a relevant and challenging example of a rare population masked within a large, diverse background. A field complicated by heterogeneity (cancer heterogeneity, patient-to-patient heterogeneity, and intra-patient CTC variation), make CTCs a difficult population to isolate and purify. Isolated from the bloodstream, CTC isolation requires separation from a vast, diverse background, which varies with disease and treatment. Potential target antigens for CTC capture also vary with cancer type, resulting in a lack of a 'one size fits all' selection antigen. To add to the complexities, capture markers may differ in expression level between patients. Furthermore, within the CTC field, isolation demands and assay parameters are constantly changing as objectives of CTC technologies span from basic enumeration (i.e., where background is less impactful) to high purity isolations for transcriptomic or genomic analysis (i.e., where background populations can mask CTC-specific readouts). The stringent and variable demands governing CTC technologies make CTCs an ideal population with which to evaluate cell isolation platforms.

Like many emerging rare cell populations, the full clinical utility of CTCs has yet to be determined as clinical trials continue to match treatment response with molecular readouts. In the research field, diverse and dynamically changing approaches to cell isolation, both in CTCs and other rare populations, has necessitated that cell isolation technologies retain maximum flexibility. As isolation requirements vary experiment-to-experiment and researcher-to-researcher in their specific target population (i.e., CTC sub-populations), capture efficiency, and required purity, drive need for maintaining flexibility.

PMP-based isolation technologies provide users with flexibility, allowing the targeting of nearly any surface marker through readily available commercial products (e.g., antibodies). However, the majority of PMP technologies are limited in their selection methodology often perfecting either positive or negative selection. Positive selection yields a specific, relatively pure population. In order to maximize purity, the targeted marker should be specific to the target population. If non-target cells express the chosen marker or if the antibody nonspecifically interacts (e.g., cross reactivity) with contaminant cells, purity will be adversely impacted. In comparison, negative selection yields a larger contaminant population, resulting from incomplete depletion (e.g., background cells with little or no expression of the depletion markers). Despite generally low purity, negative selection may yield a larger population of target cells. 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. Instead, the markers used for depletion must not be expressed by the target population; while simple with CTCs, finding the right markers can be difficult with less differential populations. To maximize flexibility in cell isolation platforms and encompass the wide range of isolated populations and purities, platforms will need to move from solely positive or negative selection to performing both positive and negative selection, separately as well as in an integrated fashion.

To accommodate cell isolation flexibility in the lab and transition research-developed assays into a clinical setting, mTAE was developed. An automated PMP-based sample preparation platform, mTAE

facilitates flexibility in cell isolation on up to four samples at once. 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. When applied to patient samples, implementing negative selection prior to positive selection frequently resulted in higher CTC yields than positive selection of CTCs alone and higher purity than either positive or negative selection independently. This suggests that by removing contaminant cells, CTCs were more efficiently captured. Interestingly, however, in characterization with cell lines, this phenomenon was not observed – decreasing (or increasing) background had no discernable impact on target cell capture; 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.[196,197] Thus, a depletion step may help to reduce these potential interactions and improve interactions between the PMPs and CTCs.

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 and AR-driven transcripts.

For users, cell isolation is often only the first step towards endpoint analyses, which vary in input requirements. mTAE allows users to dynamically and flexibly approach cell isolation without being limited by selection methodology and able to adapt with each experiment, downstream analysis, and patient. While rigid isolation approaches have utility in well-studied populations and fully established assays, rare cell populations (e.g. CTCs) have yet to be fully characterized. Thus, maintaining isolation flexibility will allow researchers and clinicians to fully probe this emerging population and evaluate its full clinical potential.

3.5   Acknowledgements

Chapter 4 RNA-mediated TILDA for Improved Cell Capacity and Enhanced Detection of Multiply-spliced HIV RNA[4]

Quantification of the HIV viral reservoir is critical to understanding HIV latency, advancing patient care, and ultimately achieving a cure. To quantify the reservoir, a new metric was recently introduced, which quantified cells carrying multiply spliced HIV RNA. The developed assay, *Tat/rev* Induced Limiting Dilution Assay (TILDA), enables quantification of cells containing multiply-spliced HIV RNA events as an indicator of reservoir size. Due to TILDA's reliance on a limiting dilution format paired with the rarity of target events, numerous individual reactions are required to obtain a single endpoint. The current assay embodiment uses a whole cell input to detect target RNA sequences without the traditional preceding nucleic acid purification steps. Thus, while the direct measurement of target events from whole cells significantly streamlines the workflow, there is a cost in sensitivity and assay throughput. Here, we apply a new technique for rapid RNA isolation, Exclusion-Based Sample Preparation, to TILDA, with the goal of alleviating these limitations without significantly adding to the workflow. By combining TILDA with multiplexed RNA extraction enabled by Exclusion-Based Sample Preparation, assay sensitivity and capacity are improved while maintaining assay simplicity, advancements that could facilitate eventual clinical implementation in detecting rare events in patients.

4.1   Introduction

Recent HIV treatment strategies have largely succeeded in containing HIV progression while greatly reducing the risk of death from AIDS.[70,71,198,199] Yet, eradication of the virus is not yet possible as integration of HIV into the host cell genome throughout the course of infection leads to the establishment of a viral reservoir.[72,75,77,200] Termed the latent HIV reservoir, the reservoir serves as an 'underground spring' capable of replenishing virus production even after years of undetectable viral load measurements.[201] While the reservoir is widely accepted as a major barrier to HIV eradication, metrics to quantify and monitor the reservoir remain debated; furthermore, there is no FDA-approved assay for quantifying the reservoir, limiting clinical implementation of reservoir-based cure strategies.

Until recently, reservoir quantification centered around two targets: 1) HIV DNA and 2) host cells capable of producing replication competent virus. HIV DNA occurs in cells infected with HIV, as HIV reverse transcribes its RNA genome into DNA shortly after infection. Assays measuring HIV DNA provide a fast, relatively simple PCR-based reservoir quantification. For additional specificity, many have adopted Alu-PCR, designed to limit detection to HIV DNA that has integrated into the host genome, one step further in the infection process. Despite the assay's relative simplicity, the relevance of HIV DNA, either total or integrated (e.g., Alu-PCR), is unclear.[202] The detection of HIV DNA does not guarantee a cell will actively contribute to the reservoir (i.e., be able to produce HIV capable of infecting other cells). HIV DNA does not ensure a cell's downstream ability to produce virus (improper integration, ineffective integration sites, etc.) nor ensure the produced virus' infectiousness, two key components in contributing to the reservoir. Thus, HIV DNA is suspected to overestimate the size of the reservoir.[202] In contrast, the Quantitative Viral Outgrowth Assay (QVOA) aims to quantify only events that lead to the production of infectious virus, making it highly specific. However, QVOA requires extensive sample volume, time, and expertise to co-culture and propagate the virus to a detectable level, limiting widespread use. A recent iteration of this assay was able to decrease culture time;[99] yet significant hurdles remain. Despite QVOA's utility as a critical research tool and incremental improvements, QVOA's limited sensitivity and laborious protocols as well as shortcomings in reproducibility, and dynamic range continue to prohibit widespread

utilization. While both assays have expanded knowledge of the HIV reservoir, a niche exists between them for an assay, which maintains the simplicity of PCR-based approaches, but improves on the sensitivity of detected events.

Recently, Procopio et al. (2015) introduced a new metric and method to quantify the HIV viral reservoir: *Tat/rev* Induced Limiting Dilution Assay (TILDA). TILDA detects the presence of cellular *tat/rev* multiply-spliced HIV RNA (*tat/*rev msRNA) as an early indicator of HIV production.[91,93-95,97,98] As numerous defective HIV genomes often have modifications in the tat and rev genes,[86] the presence of *tat/rev* msRNA has been shown to be an early indicator of cells capable of producing functional, infectious virus. To detect *tat/rev* msRNA, TILDA begins with patient cell induction (e.g., PMA) throughout a 12-hour culture to reverse latency, and induce the cells to begin producing HIV-related mRNA transcripts, including the target *tat/rev* msRNA. To detect *tat/rev* msRNA, TILDA directly inputs the raw whole cells into the reverse transcription reaction, directly followed by a nested PCR reaction to amplify sequences specific to *tat/rev* msRNA. By detecting *tat/rev* msRNA, TILDA fills a gap between existing reservoir assays - integrated HIV DNA, considered as an overestimate of the viral reservoir, and QVOA, an assay that requires extensive sample volume, expertise, time, and cost. While TILDA remains yet untested in terms of its clinical implications and relevance in monitoring the reservoir in patients, TILDA is well formatted for clinical adoption. Yet, additional improvements are required to bring TILDA into the clinic.

While TILDA holds great potential, TILDA faces limitations, which hinder clinical adoption and widespread use. Currently, TILDA's utilization of whole cell inputs, rather than purified RNA, results in a simplified assay, without the need for traditionally laborious RNA isolations prior to downstream detection. However, while a whole cell input provides TILDA with simplicity, assay input material (e.g., purity, quantity, quality) ultimately impacts RT-PCR outcomes.[203-205] As a result, TILDA's whole cell input is detrimental to the RT-PCR reaction, extending assay time, impairing sensitivity, and ultimately limiting the number of cells screened. In quantifying the reservoir, the rarity of positive events, require an

assay best able to capitalize both on capacity and sensitivity. Furthermore, as treatments aiming to deplete the reservoir enter clinics, the assay will need to screen an increasing number of cells in each patient as the reservoir is depleted beyond the existing limit, ideally by simply increasing the number of events screened within the assay. In contrast to whole cells, isolated RNA provides a purified input material with limited background to amplify (i.e., RNases, inhibitors, proteins). However, the integration of RNA extractions into TILDA requires an easily multiplexed approach enabling parallel RNA extractions in order to fit into the existing assay, maintain the assay's simplicity, and support the large number of independent isolations. Manual magnetic bead approaches to RNA extraction require multiple laborious washing steps, in which the buffer is mixed with the RNA-specific paramagnetic particles (PMPs), the PMPs allowed to collect on a magnetic rack, then the fluid removed over the course of multiple washes. While automation can alleviate some of the hands-on time, these systems remain expensive and, due to the number of extractions required by TILDA, would likely be monopolized for a single patient sample making throughput a bottleneck in these systems. While many of these traditional RNA isolations remain laborious, time consuming, and limited in their ability to be performed in parallel (or cost prohibitive in the case of automated system), herein a bead-based technology termed Exclusion-Based Sample Preparation (ESP) is used to facilitate streamlined, multiplexed RNA extractions.[23] ESP allows RNA extractions to be performed simply, in sets of 8 parallel isolations, with no additional time. Thus ESP provides a well-suited technology to integrate the benefits of extracted RNA into the high volume demands of TILDA. By applying ESP-facilitated RNA extractions to modify the previously reported whole cell TILDA assay (ESP RNA TILDA), the potential gains in both assay capacity and sensitivity in detecting rare *tat/rev* msRNA are demonstrated.

4.2   Materials and Methods

4.2.1   Participants and Samples

Whole blood was either obtained from HIV-positive patients under IRB approval or from healthy donors (Biological Specialty Corporation, PA). HIV positive patients were on suppressive HIV medications at the time of blood draw. PBMCs were isolated by Ficoll paque (GE Healthcare) density

centrifugation according to the manufacturer's recommended protocol. Blood from HIV-positive donors was processed within 5 hours of the blood draw; healthy blood was shipped overnight and processed within 30 hours of blood draw. Cells were either utilized immediately or cryopreserved for later use. Prior to TILDA, HIV-positive donor samples were resuspended at $2x10^6$ cells/mL in complete RPMI (supplemented with 10% FBS and 1% Penicillin Streptomycin) for 3 hours at 37°C and 5% CO_2. After 3 hours, cells were stimulated for 12 hours with 100 ng/mL PMA (Sigma) and 1 µg/mL ionomycin (Sigma).[206]

4.2.2   Cell Culture

Background cells used in all experiments utilizing contrived samples consisted of Hut-78, a T lymphocyte line (ATCC), or PBMCs isolated from a healthy donor (Biological Specialty Corporation, PA). Cells were cultured in RPMI 1640 (Gibco, ThermoFisher) supplemented with 10% Fetal Bovine Serum (Gibco, ThermoFisher) and 1% Penicillin Streptomycin (Gibco, ThermoFisher). Cell lines were maintained in sterile conditions at 37 °C and 5% CO_2. J1.1 (generous gift from the NIH AIDS Reagent Program)[207] and ACH-2 (generous gift from the NIH AIDS Reagent Program)[208,209] cell lines were used for all contrived samples.

4.2.3   RNA Isolation

For the ESP RNA TILDA, RNA was isolated using the previously published technology termed Sliding Lid for Immobilized Droplets Extraction (SLIDE)[157] in either an automated (Gilson PipetteMax Automated Liquid Handler) or manual version designed for a 96-well plate (Figure 1B). For both methods, RNA extraction was multiplexed for 8 simultaneous RNA extractions. In brief, cells were counted and then loaded in the bottom of a 96-well plate (<5 µL volume). Lysis Binding Buffer (Dynabeads mRNA Direct, Thermo Fisher) with 15 µL oligo $(dT)_{25}$ paramagnetic particles (PMPs) (Dynabeads mRNA Direct, Thermo Fisher), and 3 µL of Magnesil KF PMPs (Promega) were then added to each well in a volume of 150 µL, mixed, and allowed to lyse for 10 minutes. After 10 minutes, the well volume was brought to 410 µL with Wash Buffer B (Dynabeads mRNA Direct, Thermo Fisher) and re-

mixed. Using a bar magnet (or 8 individual magnets in the automated platform) protected by a hydrophobic barrier (automated platform used EXTRACTMAN Bead Capture Strip (Gilson); manual RNA platform used a plastic film (Roche LightCycler® 480 Sealing Foils)). PMPs were collected at the surface of the well by hovering the collection surface backed by a bar magnet (or magnetic head in the automated platform) above the wells, allowing the PMPs to collect on the surface. The collection surface, now containing the PMPs and bound RNA were moved away from the lysis well and over a wash well pre-filled with 415 µL of Wash Buffer B. The wash buffer formed a convex drop at the top of the well allowing the collection surface containing the PMP pellets to come in direct contact with the fluid. The magnet was then removed from the strip and placed below the plate. Due to the new position of the magnet below the plate, the PMPs pulled off of the strip and into the wash fluid towards the magnet. Once the PMPs had released in the wash well, the magnet was removed from the bottom of the plate and the PMPs were mixed by multi-channel pipette. The washed PMPs and bound RNA were then collected on a new collection surface using a magnet again positioned at the top of the well. The collected washed PMPs and bound RNA were then resuspended in 5 µL nuclease-free water and added into a PCR plate filled with complete pre-mixed master mix.

4.2.4   TILDA

Either the whole cell or RNA input was distributed in a PCR plate along with one-step RT-PCR master mix (Taqman Fast Virus 1-Step Master Mix, ThermoFisher), and the previously reported TILDA primers. The sequences used were identical to those of the original TILDA assay, which were adapted from Pasternak et al. (2008) (tat1.4: 5'-TGG CAG GAA GAA GCG GAG A-3'; rev: 5'-GGA TCT GTC TCT GTC TCT CTC TCC ACC-3'). Based on the input, pre-amplification included: reverse transcription (50 °C for 15 min), denaturation (95 °C for 2 min), and 24 cycles of amplification. Unless otherwise noted in initial characterization experiments, when using whole cells, amplification was performed using 95 °C for 15 sec and 60 °C for 4 minutes; when using a RNA input, amplification was reduced to 95 °C for 15 sec and 60 °C for 30 sec. This was done in a real time Light Cycler 480 II (Roche). Following pre-

amplification, 40uL of TE buffer was added to each well and 1 μL was transferred into the next *tat/rev*

PCR reaction. For all whole cell versus ESP RNA experiments, the first reaction was performed on

separate plates due to differing thermocycling conditions. However, for the detection plate, both sample

sets were run on the same plate to better allow for comparisons. For the detection plate, identical to the

published TILDA, 5 μL of the Lightcycler Probe Master Mix (Roche), 0.2 μL of each 20 μM primer (tat2

and rev), 0.2 μL of the probe HIV FamZen at 5 μM and 3.4 μL of nuclease-free water for a 10 μL final

reaction volume. Primer sequences and probes were synthesized by IDT and taken from the original

TILDA publication which were adapted from Pasternak et al., 2008 (tat2: 5' ACA GTC AGA CTC ATC

AAG TTT CTC TAT CAA AGC A -3'; probe: 5'-/56-FAM/TTC CTT CGG /ZEN/GCC TGT CGG GTC

CC/3IABkFQ/-3'). Amplification and detection was performed in a real time Light Cycler 480 II (Roche)

(Preincubation 95 °C for 10 min, 45 cycles of 95 °C for 10 sec, 60 °C 30 sec, 72 °C for 1 sec followed by

cooling step of 40 °C for 30 sec). Fluorescent signal was measured every cycle. Positive events were

identified based on the second derivative maximum function using the Lightcycler 480 II. The cycle

threshold for each positive well was then used for comparison (cycle threshold called using the

LightlyCycler 480 Software, Second Derivative Maximum Function).

4.3 Results

4.3.1 *Integration of Exclusion-based Sample Preparation RNA Isolation*

Despite the simplicity of whole cells, a purified RNA input, if easily integrated, could reduce

background to improve amplification kinetics, remove potential inhibitors, and ultimately enhance

sensitivity and assay capacity (**Fig. 4.1A**).[204] However, due to TILDA's reliance on a large number of

individual reactions, traditional RNA extraction methods remain largely incompatible and unrealistic to

readily integrate. Previously, a PMP isolation approach termed Exclusion-Based Sample Preparation

(ESP) was developed as a means to perform simple PMP-based extractions, which deviated from

traditional tube-based approaches (i.e., remove the PMPs from solution rather than remove the fluid from

the PMPs). Building off of the ESP core technology, a multiplexed 96-well RNA extraction format was

adapted and developed for this application (**Fig. 4.1B**). In this ESP embodiment, RNA extractions were

performed in parallel sets of eight, requiring no additional time compared to a single extraction. To isolate RNA, cells were loaded into a 96-well plate while RNA lysis buffer and RNA binding PMPs were added by multichannel across the plate. Once lysed, the well volume was increased until the concave meniscus on the fluid became convex (~400 μL). A bar magnet, the same width of the plate, was then covered with a removable hydrophobic film and held directly over the first column of samples and positioned so the film just contacted the top of the convex droplet. The PMPs, attracted by the magnet, collected on the film. Once collected, the film, now containing a PMP pellet from each well (up to 8 RNA isolations), was then held over a new plate similarly filled with wash buffer. The magnet backing was removed and placed below the wash plate. Once the PMP pellets made contact with the convex fluid, the PMPs were pulled into the wash wells by the magnet below the plate. The PMPs in each wash well were then simultaneously mixed by multichannel pipette. After the PMPs and bound RNA were mixed, a new film was placed on the magnet and the PMPs collected onto its surface. The film was then removed and the magnet backing displaced. Each PMP pellet could then be resuspended with mastermix and loaded into a PCR plate.
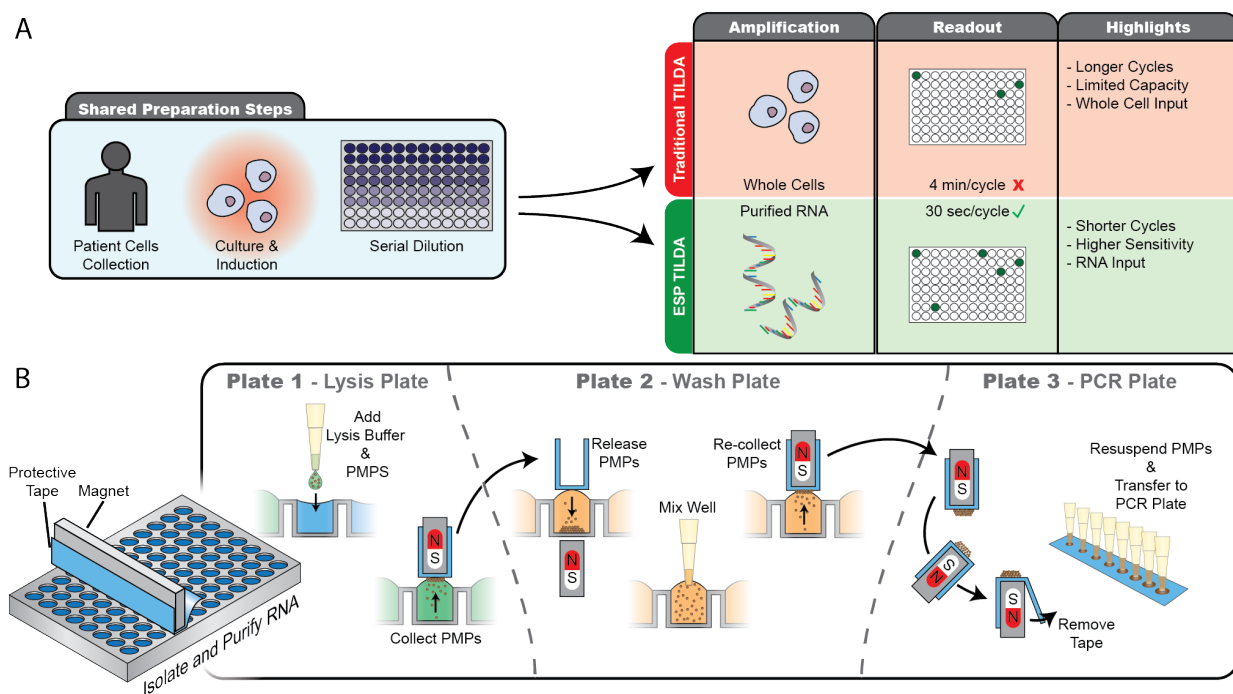

Fig. 4.1 Overview of whole cell and ESP RNA TILDA. (A) Schematic overview of whole cell TILDA vs. ESP RNA TILDA process. (B) Step-by-step schematic demonstrating RNA isolations using Exclusion-Based Sample

Preparation. The process can easily be performed on sets of 8-wells using a bar magnet, a standard 96-well plate, and multichannel pipette, with no additional time over a single RNA isolation in this format.

4.3.2   *Whole Cell Versus RNA*

TILDA provides a method for detection of *tat/rev* msRNA, as a new quantitative metric of the HIV viral reservoir. In the published assay, whole cells were directly utilized as the assay input to detect *tat/rev* msRNA, without prior purification or isolation of the target analyte, RNA. While the omission of RNA extraction results in a simple assay, we hypothesized whole cell inputs ultimately impair detection of *tat/rev* msRNA (i.e., due to limited access to target RNA within cells, presence of background material, inhibitors, RNases, etc.) (**Fig. 4.1A**). Using Exclusion-based sample preparation to perform multiplexed PMP-based RNA isolations for parallel RNA extractions (**Fig. 4.1B**), the potential gains in RNA rather than whole cell TILDA were evaluated. To compare inputs, low cell number samples of two HIV-infected cell lines (<10 cells per reaction) known to be positive for *tat/rev* msRNA were utilized as either whole cells or ESP-extracted RNA. Previously, these cell lines demonstrated constitutive production of HIV in culture, even at low cell numbers (as detected by viral RNA isolation from conditioned media – see **Fig C.1**).

Utilizing a highly idealized sample (consisting of low quantities of cell expressing *tat/rev* msRNA), both whole cell and ESP RNA inputs resulted in detection of *tat/rev* msRNA (**Fig. C.2**) under the published thermocycling conditions (4 minute elongation cycles). However, the published whole cell TILDA[206] ultimately employs atypically long elongation cycles (4 minutes) in contrast to the manufacturer's recommended conditions resulting in a longer assay with added equipment usage (manufacturer recommends elongation cycles of 30 seconds as opposed to 4 minutes); by reducing elongation time from 4 minutes to 30 seconds, 84 minutes of thermocycling time are eliminated from the process. We hypothesized an isolated RNA input would improved detection of rare events over a whole cell input as well as enable detection under standard PCR thermal cycling conditions. In reducing elongation cycling times to 30 seconds, signal was significantly reduced (>10,000 fold) in the whole cell inputs (**Fig. 4.2A**), resulting in missed detection in a number of replicates while ESP RNA inputs

maintained detection. Even a highly positive sample for *tat/rev* msRNA lacking background cells (HIV-negative cells), shortened elongation cycles create detection limitations for whole cell inputs, an issue which is only likely to worsen as (realistic) background is introduced. By implementing ESP-based RNA, the purified sample input enables a reduced elongation time (30 seconds) incompatible with a whole cells input, significantly accelerating time-to-answer (**Fig. 4.2A**).
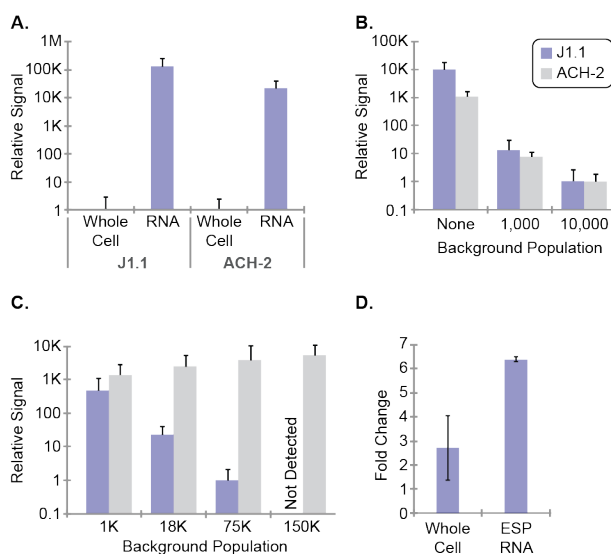
Fig. 4.2 Evaluation of whole cell versus RNA Input. (A) Comparison of a whole cell input to an ESP-isolated RNA input using shortened elongation cycling prior to detection (30-second elongation cycles). (B). Relative detection of *tat/rev* msRNA from whole cell inputs consisting of 10 HIV-infected cells (ACH-2 or J1.1) in various healthy PBMC background (4-minute elongation cycles). Relative signal normalized to lowest signal. (C) Relative expression of *tat/rev* msRNA from 10 ACH-2 cells in a varying background of healthy PBMCs using either a whole cell input (4 minute elongation cycles) or an ESP isolated RNA input (30 second elongation cycles). Positive signal was not detected from the whole cell sample in a background of 150,000 cells. (D) Sensitivity to detection of changes in *tat/rev* msRNA following induction of ACH-2 cells with PMA. All bars represent the average of three replicates while error bars represent standard deviation.

4.3.3 *Accelerating Assay Readout*

When evaluating a patient sample for *tat/rev* msRNA, TILDA relies on the detection of rare, target events from a large, varied background across range of cell quantities. To integrate and evaluate the impact of increasing background cell number on a whole cell TILDA, either 10 ACH-2 cells or 10 J1.1 cells were spiked into different quantities of background cells (e.g., healthy PBMCs) and detected using whole cell TILDA under the previously published assay (i.e., 4 minute elongation cycles). Results demonstrate, as background increases, the positive detection of *tat/rev* msRNA is delayed, despite the

equivalent number of detection events across samples (**Fig. 4.2B**). While *tat/rev* msRNA is readily detectable in the absence of background, detection sensitivity decays as more and more background cells are introduced.

Based on the improved signal ESP-isolated RNA provided in reducing thermocycling conditions, the impact of an ESP RNA input, rather than a whole cell input, was tested to evaluate whether a RNA input would recover the signal decay observed with increasing background with a whole cell input (**Fig. 4.2C**). Using ESP, an RNA isolation step was performed (in combination with a reduced 30-second elongation amplification step) for comparison to the whole cell input (in combination with the published TILDA 4-minute elongation amplification step) on samples of HIV-infected cell lines spiked into an increasing background population. While detection of *tat/rev* msRNA from whole cells saw a multiple-fold reduction across increasing background cells, an ESP RNA input eliminated the sensitivity decay in detection as background increased, in up to 150,000 background cells (maximum quantity tested). In comparison, the positive detection signal (as defined by calculation of a cycle threshold using the Lightcycler 480 Second Derivative Maximum function) from the whole cell inputs was delayed as background increased until, in a background of 150,000 cells, no positive events were detected (**Fig. 4.2C**).

As ACH-2 cells respond with increased HIV production following induction with PMA as measured by extracellular virus levels (**Fig. C.3**), we tested to determine 1) whether this change could be detected in intracellular tat/rev msRNA levels and 2) whether either whole cells or ESP RNA TILDA inputs might be more sensitive in detecting a response to induction. To evaluate, ACH-2 cells were either cultured for 12 hours in induction media containing 10 μM PMA or culture media. Ten cells from each population were then placed in a background of 1,000 cells for TILDA (RNA 30-second elongation cycles; whole cells 4-minute elongation cycles), either as whole cells or ESP isolated RNA. Compared to the non-induced sample, both sample inputs saw an increase in tat/rev msRNA (**Fig. 4.2D**). However, the ESP RNA input resulted in the detection of a much greater fold change in *tat/rev* msRNA detection as

well as less sample-to-sample variability compared to the whole cells improving the case for RNA as a more sensitive input for TILDA.

4.3.4   RNA TILDA Comparison and Validation with Contrived Samples

In order to validate ESP RNA TILDA (with 30 second elongation cycles) to whole cell TILDA (4 minute elongation cycles), a direct comparison was performed with contrived samples. Approximately 1-15 ACH-2 cells were spiked into a background of either Hut78 cells or healthy donor PBMCs (248,000 cells). These cells were then distributed across 32 total wells consisting of eight wells each of 18,000 cells, 9,000 cells, 3,000 cells, and 1,000 cells following the published TILDA dilutions; this enabled the creation of both ACH-2-positive wells and ACH-2-negative wells for analysis. For the ESP RNA TILDA, these cells were distributed (8 wells each of 18,000, 9,000, 3,000, or 1,000 cells) into a standard 96-well plate, RNA extracted with ESP, and the extracted RNA transferred to a PCR plate. Using the contrived samples of ACH-2 cells, the number of positive events detected in each method correlated well (Figure 3A). Thus this approach served to support RNA as an alternative input to whole cells. While both inputs resulted in similar events detected, on average, the RNA-based TILDA detected positive events ~5 cycles (corresponding to a 32-fold increase in relative RNA signal) earlier than the whole cell positive events (calculated based on the average cycle threshold of positive events obtained within each experiment across 10 experiments).

Due to the rarity of these cells in patients, ultimately an increased cell capacity would benefit the assay's clinical utility. Thus, the whole cell TILDA and ESP RNA TILDA were compared using high background samples consisting of 50,000 cells per well (**Fig. 4.3B**). Again, ACH-2 cells were spiked into identical samples and divided amongst the wells for either direct detection (whole cell TILDA) or ESP RNA TILDA. While the previous comparison resulted in nearly identical results across inputs (**Fig. 4.3A**), when background increased, ESP RNA TILDA shifted from the nearly 1-to-1 correlation observed in lower background, consistently detecting more positive events (**Fig. 4.3B**). With the exception of some of the lower input samples (i.e., 2 or fewer events detected per assay), ESP TILDA consistently detects

more events, suggesting improved sensitivity under increasing background. Notably, this is observed in samples likely high for *tat/rev* msRNA, as the cells constitutively produce HIV in the absence of induction (PMA), in contrast to truly latent events. As opposed to the whole cell samples, the ESP RNA TILDA detected more events in 9/13 samples (1 sample identical results; 3 samples less) when detecting from a background of 50,000 cells. Additionally, within the events detected, again ESP RNA TILDA, on average, detected positive events cycles earlier. While highly idealized samples consisting of constitutively HIV producing cells, RNA is already demonstrating potential gains in both sensitivity detecting events as well as improved detection in high background samples, even with greatly reduced thermocycling times. In patient samples, prevalence of *tat/rev* msRNA is only likely to decrease, highlighting the benefits of integrating RNA.
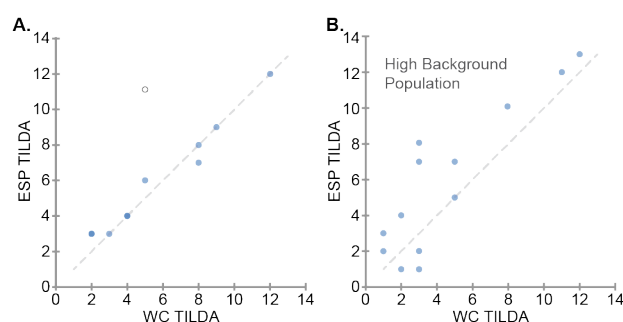
Fig. 4.3 Evaluation of whole cell versus RNA TILDA on contrived samples. (A) Head-to-head comparison of an ESP RNA input and a whole cell input for detection by TILDA using contrived samples of HIV-infected ACH-2 cells in a background population identical to the published TILDA (18,000 to 1,000 cells per well). Whole cell detection was done using 4-minute pre-amplification elongation cycles while the RNA was detected after 30-second pre-amplification elongation cycles. The number of positive events were then counted and compared across methods (n=10 plotted data points). The additional noted point (empty circle) was characterized as an outlier (fell outside the data's linear regression 95% confidence interval). (B) A direct comparison experiment was then performed in an increased background where ACH-2 cells were spiked at low number in a background population of 50,000 cells per well. Here, the ESP RNA samples consistently detected more positive events in the increased background (9/13 samples) (n=13). Dashed lines represent a hypothetical one-to-one correlation.

4.3.5 *Patient Sample Comparison*

The potential advantages beyond cell lines with ESP RNA TILDA were next evaluated on HIV-positive patients. The ESP RNA TILDA (30-second elongation cycles) was compared to the whole cell TILDA (4-minute elongation cycles) input on PBMCs isolated from an HIV-positive patient using the established TILDA cell quantities (i.e., 18,000 cells per well down to 1,000 cells per well). On average,

ESP RNA resulted in a higher frequency of *tat/rev* msRNA events detected from an identical number of inputted cells (average fold increase of 1.5) (**Fig. 4.4**). In samples obtained from healthy controls (n=3), both assays (whole cell and ESP RNA) resulted in no positive events across the dilution series (i.e., 8 wells each of 18,000, 9,000, 3,000, and 1,000 cells). While the contrived samples maintained very similar detection of events across methods within the cell quantities of the published assay (up to 18,000 cells), here we see ESP resulting in a larger quantitation of the reservoir consistently in patients, suggesting ESP RNA may increase sensitivity of the assay.
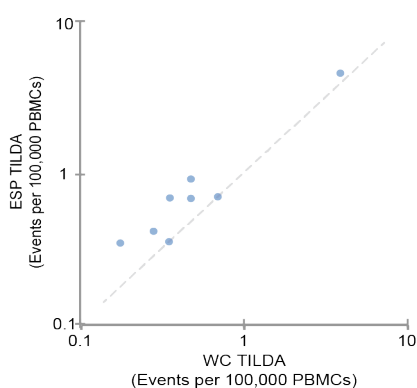
Fig. 4.4 Evaluation of whole cell versus RNA TILDA on patient samples. Comparison of detection of *tat/rev* msRNA from PBMCs collected from HIV-positive donors using both the published whole cell TILDA (4-minute elongation cycles) and the ESP RNA TILDA (30-second elongation cycles). In 6/8 patients, the ESP RNA TILDA resulted in an increased change in reservoir quantitation compared to the whole cell TILDA (two samples demonstrated no change). Dashed lines represent a one-to-one correlation.

4.4   Discussion

In order to facilitate the next generation of HIV monitoring and enable treatments targeting the viral reservoir, assays to quantify the reservoir are required to provide a benchmark of treatment success and patient eligibility for ART interruption. While numerous assays have been introduced (DNA-based assays, QVOA), variation across assays remains widespread, likely due to quantification of a range of analytes with varying relevance (e.g., HIV DNA, integrated DNA by alu-PCR, viral outgrowth). Recently, Procopio et al. (2015) introduced tat/rev induced limiting dilution assay (TILDA), which provides a new analyte, tat/rev msRNA, as a metric for reservoir quantification. A step beyond HIV DNA and a step short of detecting production of infectious virus, TILDA is poised to increase the specificity of

HIV DNA-based assays, detecting a further downstream target, while retaining their attractive simplicity and greater reproducibility compared to QVOA.

While TILDA is poised to provide a new metric for reservoir quantification, upstream sample processing and purification could improve the assay to increase assay sensitivity and capacity, two key factors in detecting rare events. ESP TILDA offers the following advantages: 1) Increased sensitivity in detection of rare events through use of a purified input material additionally enabling the interrogation of a larger number of patient cells; 2) Decreased RT-PCR thermocycler usage by 88%; 3) Maintained simplicity of original TILDA assay via implementation of an RNA isolation process, ESP, which is optimized for simplicity and speed. While the published TILDA advertised its simplicity in not requiring RNA extraction, we propose that the reliance on whole cells limits both TILDA's capacity and sensitivity along with increased assay duration. As TILDA requires a nested PCR reaction following reverse transcription, the assay faces limitations well known to this field, namely input sample impacts results. Traditional inputs for these reactions are purified RNA and/or DNA, as background material is often detrimental to reaction kinetics impacting sensitivity or detection altogether.[203-205] A whole cell input introduces significant background material, which impair reaction kinetics and detection capabilities compared to a reaction utilizing purified nucleic acid. Historically, RNA extractions have been tedious, requiring extensive time especially at the scale required by TILDA (i.e., numerous independent isolations). However, the development of bead-based RNA extraction kits combined with exclusion-based sample preparation, a bead isolation technology, allows users to rapidly preform RNA isolations, 8 isolations at a time, using either an automated or handheld system. The addition of RNA isolation requires less than 1 hour to perform 96 isolations. Additionally, by transitioning to a RNA input, pre-amplification elongation cycle time is reduced to 1/8 of the published time (96 minutes to 12 minutes), reducing thermocylcer usage by 84 minutes. While some hands-on time is required to perform RNA extractions (unless automated), the required RNA extraction time using the presented multiplexed

fashion, remains well below the thermocycling time saved (estimated at 40 minutes per 96-well plate). The timesaving, paired with improved sensitivity, presents an opportunity for RNA-based TILDA.

In comparing RNA to whole cell inputs in contrived samples, RNA presents significant advantages in stronger, often more consistent detection of *tat/rev* msRNA; thereby providing improved sensitivity in detecting either subtle changes or rare events. Significantly, a RNA input allows users to scale up the number of cells interrogated per well. Increasing assay capacity will become increasingly critical in the clinic as reservoir targeted treatments will deplete the reservoir, requiring the evaluation of larger cell quantities to monitor successful treatment. Ultimately, this may push the reservoir beyond the 744,000 cell events interrogated per plate. While conceivably TILDA could be performed using multiple plates or maximizing input to 18,000 cells in every well (maximum of 1.728 million events per plate), multiple plates would only increase assay time and equipment demands of the assay. Beyond this capacity, a whole cell input may suffer from reduced reaction sensitivity observed in contrived samples as background cell number increased. The phenomenon of decreasing signal with increasing background was reduced significantly by an ESP RNA input (Figure 2) speaking to the significant advantages RNA could provide in scaling up TILDA with expanding quantitative needs. Additionally, large single well inputs could have the potential in future TILDA applications if, to simplify processing, PBMCs rather than T-cells are evaluated; the switch to PBMCs would similarly require more events to be utilized. Additionally, wells of high cell number could be used to rapidly screen patients to get baseline readout for their reservoir (i.e., frequency of events is 1 in 100,000 vs. 1 in 10 million), before performing a full-scale assay to determine the sensitivity needed for each patient. As reservoir sizes vary greatly and currently are fairly unpredictable, screening patients to get a relative frequency would provide clinicians with the information necessary to scale the TILDA capacity to assay as many or as few cells as necessary (i.e., evaluate 500,000 events or 10 million or more events).

The published TILDA demonstrated a new analyte and assay on which to benchmark the HIV viral reservoir. We believe the presented modifications and method with which to integrate RNA inputs

provide substantial benefits on the published assay expanding capacity, which will be required as developing treatments continually diminish the reservoir size. RNA inputs demonstrate clear potential in expanding inputs, which could be used to both scale up the input without requiring multiple TILDA plates or enable qualitative screening, to predict reservoir size and ensure adequate cells are inputted into the quantitative assay.

4.5   Acknowledgements

Chapter 5   Dual Fluorescent Reporter for Quantification of Rare HIV Viral Reservoir Events Paired with Modified Well-based Culture for Improved Sensitivity

As HIV treatment approaches expand, monitoring of the HIV viral reservoir will be required to pair with developing clinical assays for combating and depleting the reservoir. While a number of analytes for reservoir quantitation have been proposed (e.g., unspliced HIV RNA, multiply-spliced HIV RNA, integrated HIV DNA), production of infectious, replication-competent virus remains the gold standard analyte for quantitation of the HIV viral reservoir. The existing assay targeting production of infectious, replication-competent virus, the Quantitative Viral Outgrowth Assay (QVOA) however, remains limited in clinical adoption due to high cost, labor, and time required for the assay (e.g., 7 day culture). Recently, a luciferase reporter-based assay was introduced, reducing culture time to 48-hours with a streamlined readout of infection. Building on the luciferase-based assay, here we propose a dual fluorescent reporter to enable real-time live reporting on HIV infection within the assay. The dual reporter enables highly specific identification of events through the integration of two independent infections to provide a visual, and visually monitorable readout of infection. In preliminary studies using contrived cell line-based studies the dual florescent reporter appears more sensitive at 48-hours. Furthermore, by controlling the culture media volume, we demonstrate the sensitivity of both reporter-based well plate assays can be enhanced through improved infection rates. Together, these approaches provide a new assay for applying to HIV viral reservoir quantitation.

5.1   Introduction

Due to successful treatment of HIV with Antiretroviral Therapy (ART), the risks associated with HIV progression, including development of AIDS, are largely reduced.[70,71,198,199] ART decreases the production of virus by infected cells by targeting multiple facets of the viral infection and production process (e.g., fusion/entry inhibitors, nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, integrase inhibitors, protease inhibitors). Paired with ART, clinical monitoring of HIV virus production is achieved with viral loads, which provide a quantitative measurement of free virus in the bloodstream. While on treatment, ART is able to drive HIV levels in the bloodstream to undetectable.[211-213] While highly successful in controlling viral production and thereby maintaining undetectable viral loads, if a patient were to lapse on treatment, HIV re-emerges, even following decades of successful ART and undetectable viral loads. Upon re-emergence of HIV, viral loads become detectable as virus production drives HIV progression. The inability to cure HIV, even in the face of decades of successful ART, is due to the establishment of an HIV viral reservoir.

The HIV viral reservoir is established early in HIV infection as HIV integrates into the host cell genome throughout infection generating a viral reservoir.[72,75,77,200] A latent reservoir, the HIV viral reservoir cells do not actively produce virus, allowing them to escape from ART detection, and enabling their persistence within the body. Largely residing in a population of long-lived cells (e.g., resting CD4+ T cells), the reservoir is capable of producing virus even after years of undetectable viral loads and successful ART treatment.[201] Despite the influence of the reservoir, quantitatively in the bloodstream, the viral reservoir is relatively small, consisting of approximately only 1-10 infectious units per million resting CD4+ T cells.[73,214,215] Yet, due to the presence of this small, inducible HIV viral reservoir capable of producing infectious virus, HIV is incurable with current HIV treatment approaches requiring a lifelong commitment to treatment.

However, interest in targeting the HIV viral reservoir for depletion remains, emerging as the next front of HIV treatment. By targeting the reservoir for treatment, clinicians may be able to reduce the

reservoir size, potentially enabling a cure in a subset of patients. While multiple clinical approaches to reservoir reduction have been hypothesized,[216,217] in order to validate these approaches, the reservoir needs to be concurrently quantified in patients to gage the effectiveness of reservoir-targeted treatments. While a number of potential analytes for quantifying the reservoir have been evaluated (e.g., HIV DNA[218-220], integrated HIV DNA[89,219,221], multiply-spliced RNA[91,93-95,97,98]), the efficacy of many of these in quantifying the reservoir remains debated.[202] The most relevant hypothesized metric by which to quantifying the HIV viral reservoir remains the detection of cells capable of reversing latency and producing infectious virus; this population of cells will play a significant role in driving viral rebound.

Until recently, quantifying the cells capable of reversing latency and producing infectious virus has been limited to one assay: Quantitative Viral Outgrowth Assay (QVOA).[99,100] QVOA detects the production of infectious virus by co-culturing patient cells with an HIV-susceptible amplifier cell line (gamma-irradiated PBMCs[26] or MOLT-4/CCR5[27]) until the virus has been propagated to a detectable level. While this approach has facilitated study and understanding of the HIV viral reservoir, the cost, labor (e.g., PCR[27] or ELISA-based[26] readout from each well), large volume of blood (120-180 mLs) and time (7-14 day culture) of the assay has prohibited its adoption in clinical settings. Furthermore, the QVOA may actually underestimate the size of the HIV viral reservoir as the assay only detects a fraction of the integrated, replication-competent events, likely in part due to incomplete reactivation of the latent reservoir.[86,222] More recently, the TZA assay was introduced. The TZA assay relies on an HIV-reporter cell line, TZM-bl, to report on the production of infectious virus by the co-cultured patient cells.[101] Using this approach, assay duration is reduced to only 48-hours and the readout is relatively simple luciferase assay. While extensively streamlined, to establish a baseline signal for the reporter, a culture plate from an HIV-negative donor must be assessed in parallel. Despite this addition, the TZA assay estimated a 70-fold increase in reservoir size compared to the QVOA[101], and due to its streamlined format, has the potential to advance detection of latency events into a more clinic-ready assay.

Building on the reporter concept of the TZA assay, here we present an assay designed to detect production of infectious virus in real-time through the use of a dual fluorescent reporter. The dual fluorescent reporter is a modified HeLa cell line expressing the receptor CD4 and co-receptors CXCR4 and CCR5, which are necessary for HIV infection. While single reporter systems may face challenges in specifically detecting infection events due to even low levels of noise in the reporter, the presented reporter uses two independent reporter mechanisms to determine a positive event: 1) the degradation of a constitutive expressed YFP-APOBEC3G by the HIV protein Vif (OFF signal) and 2) a Tat- and Rev-responsive gene cassette encoding HIV Gag-mCherry (ON signal). Using the dual fluorescent reporter in a well-based assay, we characterize infection response to a model latent HIV cell line ACH-2[209,210]. Furthermore, we then compare the dual fluorescent reporter against the TZA reporter as well as an adapted QVOA using the MOLT-4/CCR5 cell line with contrived samples and the same induction mechanism. Between assays, the dual fluorescent reporter closely compares to the existing platforms. In addition, we improved sensitivity of both reporter-based assays by moderating culture volume, beginning latency reversal (and subsequent virus production) at a low culture volume yielding a higher concentration or produced virus while then increasing volume to improve cell survival and signal. This approach demonstrated improved infection rates in both systems improving assay sensitivity. By integrating the dual fluorescent reporter and reduced media volume into a streamlined reporter-based viral outgrowth assay, the next generation of viral outgrowth assays may be better able to transition from the bench to the clinic enabling reservoir monitoring alongside existing HIV monitoring tools such as the HIV viral load.

5.2   Materials and Methods

5.2.1   *Cell Culture and Blood Processing*

Cells were cultured under sterile culture conditions at 37° C in 5% CO_2. ACH-2 (courtesy of NIH AIDS Reagent Program)[209,210] and PBMCs (obtained from healthy donors) were cultured in RPMI1640 (Gibco) supplemented with 10% Fetal Bovine Serum (FBS) (Gibco) and 1% Penicillin Streptomycin (PS) (Gibco). MOLT-4/CCR5 cells (courtesy of NIH AIDS Reagent Program)[223] were cultured in complete

RPMI supplemented with G418 (ThermoFisher) at a concentration of 1 mg/mL. The dual fluorescent reporter (courtesy of Dr. Nathan Sherer, UW-Madison) and TZM-bl (courtesy of NIH AIDS Reagent Program)[224-226] cell lines were each cultured in DMEM (Gibco) supplemented with 10% FBS and 1% PS.

For generation of contrived samples and HIV-negative donor samples, Peripheral Blood Mononuclear Cells (PBMCs) were isolated from whole blood. The whole blood – collected from healthy donors and treated with K3 EDTA (Biological Specialty Corporation) – was received within 24 hours of the blood draw and processed. In brief, whole blood was mixed with 1x PBS (1:1) and overlaid on 15 mL of Ficoll Paque PLUS (17-1440-02, GE Healthcare). The overlaid diluted blood was then centrifuged following the manufacturer's instructions. After centrifugation, the buffy coat was removed and diluted in 20 mL wash buffer (1x PBS supplemented with 0.5% BSA and 2mM EDTA). Cells were then centrifuged (200 rcf for 10 min), pelleted, and resuspended again in 20 mL wash buffer and stored on ice until ready for use. If freezing, cells were frozen to -80 °C in complete culture media supplemented with 10% DMSO (Sigma).

5.2.2 Contrived Sample Generation

For all experiments, contrived patient samples were generated by adding ACH-2s into a background of PBMCs isolated from a healthy donor. ACH-2 cells have been shown to activate and increase HIV production in response to external factors (e.g., PMA (**Fig. D.1**)). To create contrived patient samples for evaluation across the TZA, QVOA, and dual fluorescent reporter, all samples consisted of a series of wells containing 10,000 PBMCs in addition to <1 ACH-2 per well (resulting in some positive and some negative wells). Each sample consisted of 24-wells (240,000 total cells).

5.2.3 Dual Fluorescent Reporter Assay

To use the dual fluorescent reporter to detect cell events capable of producing infectious HIV, the cell line was seeded in a 96-well plate at 10,000 cells per well. This cell density enabled a nearly confluent covering of each well at 48-hours. Contrived samples were then added to the reporter cell line and cultured for 48-hours. After 48-hours cells were imaged at 10x on a Nikon TE Eclipse microscope

and the fluorescent signal from each reporter overlaid to determine infection events. Independent infection events were determined as events, which were completely surrounded by uninfected cells. Thus infection events could be a single infected cell as well as a cluster of infected cells. The number of infected events were frequently counted and reported for characterization of the cells line. When used in the context of the contrived patient samples in a viral outgrowth assay, wells were reported as positive if a single infection event (or more) occurred rather than reporting the number of events. In some wells, multiple infection events were observed; however, the well was just reported as positive.

5.2.4 TZA Assay

To perform the evaluate the performance of the TZA assay, the TZM-bl cell line, which stably expresses the receptors necessary for HIV infection (CD4, CXCR4, CCR5), also carries an integrated copy of the β-galactosidase (β-gal) gene under control of an HIV-1 long terminal repeat (LTR) promoter; paired with a luciferase assay, this cell line is capable of reporting HIV infection. TZM-bl cells were plated at a density of 10,000 cells per well; this is lower than the published TZA assay as the high plating densities published (i.e., 30,000 and 60,000 TZM-bl per well) appeared to decrease assay sensitivity in detecting infection relative to standard cell plating densities (e.g., 10,000 – 25,000 cells) (**Fig. D.2**); additionally, for consistency in the evaluation across reporters, identical reporter densities were used. Following plating along with the TZM-bl cell line, contrived patient samples were added and co-cultured for 48-hours. In parallel a second plate was set up containing only PBMCs, no ACH-2 to establish a baseline signal for the luciferase assay. Following co-culture, the media was removed and 30 µL of PBS was added to each well as well as 30 µL of Beta-Glo (Promega). The plate was incubated while protected from light for 20 minutes at room temperature. The luciferase signal was then read in a PHERAstar plate reader (BMG Labtech). To determine positive events, as published,[101] the no ACH-2 plate wells were averaged and two standard deviations added to the average. This was identified as the threshold for a positive event. Wells in the ACH-2 plate that exceeded the threshold were reported as positive events.

5.2.5 QVOA

To translate the QVOA into a 96-well plate to enable comparison with both the presented DFR-QVOA and TZA assay, 20,000 MOLT-4/CCR5 were plated per well (suspended cell line). Contrived sample was then added to each well, consisting of ~10,000 'patient' cells per well. The cells were allowed to culture for 7-days, with 50 μL of media volume exchanged at day 3 and day 5 (from the top so as not to remove any cells). On day 7, ~100 μL of conditioned media (taken from the top of the well to minimize number of collected cells) was collected and added to 200 μL PBS in a 96-well plate. To remove any cells (which could lead to detection of HIV DNA events), the well plate was then spun down at 800 x g for 5 minutes. Following centrifugation, 100 μL of the resultant diluted media (from the top) was then collected for downstream RNA extraction and detection.

5.2.6 RNA Extraction and Detection

For extraction of HIV RNA from conditioned media, RNA was isolated using the previously published technology termed Sliding Lid for Immobilized Droplets Extraction (SLIDE)[157] as demonstrated in Pezzi et al., 2017.[227] For RNA isolation, 100 μL of the desired sample was collected off the top of the well and added to Lysis Buffer (Abbott HIV-1 m2000sp Kit) containing 20 μL of magnetic beads in a well plate. The contents of the well plate were then heated to 50°C for 20 minutes. After 20 minutes, the well contents were brought to ~400 μL, the volume required to form a convex meniscus, and the isolation performed as previously described.[227] A single wash of Wash Buffer 2 was then performed. Following washing, the sample was eluted in 20 μL of provided elution buffer and heated for 20 minutes at 75°C. An additional 20 μL of Wash Buffer 2 was then added completing the elution process.

To quantify extracted HIV RNA, 5 μL of eluted RNA was added per reaction along with 2.5 μL of one-step RT-PCR master mix (Taqman Fast Virus 1-Step Master Mix, ThermoFisher), 2 μL of nuclease-free water, and 0.5 μL primers and probe specific to HIV-1. Specifically, the PCR primers and probe (FAM) used were specific to the LTR region of HIV (the primer and probe sequences were as follows:

forward primer: 5'-GCCTCAATAAAGCTTGCC-3'; reverse primer: 5'-GGCGCCACTGCTAGAGAT TTT-3'; probe: 5'-AAGTAGTGTGTGCCC-3') as specified by Rouet et al. and Veronique et al.[228,229] (synthesized by Life Technologies). Amplification and quantification was done a real time Light Cycler 480 II (Roche) under the following conditions: reverse transcription (50 °C for 15 min), denaturation (95 °C for 2 min), and 45 cycles of amplification (95 °C for 15 sec and 60 °C for 30 sec). Positive events were called based on reporting of a cycle threshold (Ct) value for that well. Ct values were included up to 40 while all amplification curves were assessed amplification and valid cycle threshold (Ct) calls to ensure detection of product.

5.3   Results and Discussion

5.3.1   Basic Performance Characterization of the Dual Fluorescent Reporter

Infection of the dual fluorescent reporter was first characterized with conditioned, cell-free media collected from an ACH-2 culture. To see if reporter response scaled with infectious virus, reporters were plated in a 96-well plate and serial dilutions of virus added. The infected area of each well was then calculated (**Fig. 5.1A**). Results indicate virus concentration scales with reporter response. Similarly, the response of the reporter to increasing numbers of HIV-producing cells, ACH-2s, was evaluated. As the number of ACH-2s added to the culture increased so did the number of infection events (**Fig. 5.1B**). At the high ACH-2 input, due to the higher frequency of infection events, the events reported may have underestimated the number of independent infection events. This is due to the chance two side-by-side events being reported as a single event due to their proximity. Thus, it is likely the number of infection events at the higher ACH-2 input is actually an underestimate of the total infection events observed.
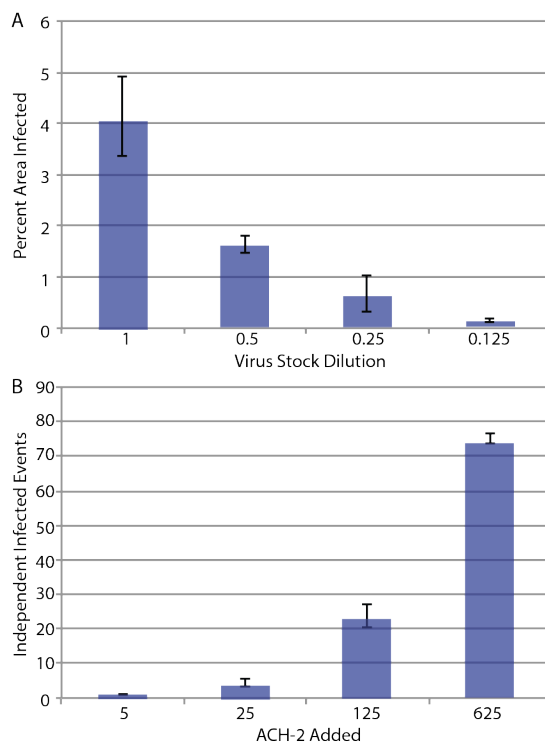
Fig. 5.1 Dual fluorescent reporter response to conditioned media and co-culture with ACH-2 (A) Reporter response to serial dilution of conditioned media collected from ACH-2s (following PMA induction at 5 ng/mL). (B) Number of independent infection events counted across increasing ACH-2 cells added to the culture. Bars represent the average of n=3 replicates and error bars represent standard deviation.

5.3.2   *Basic Performance Characterization of ACH-2 for use in Contrived Patient Samples*

To compare the three assays (QVOA, TZA, dual fluorescent reporter), a contrived patient sample was required. Thus ACH-2s,[209,210] a model latent HIV cell line, were used in combination with a PBMC background obtained from a healthy HIV-negative donor. While ACH-2 increase HIV production in response to PMA (**Fig. D.1**), we also observed that virus production from low numbers of ACH-2 (~25 cells) increased with increasing PBMCs present in co-culture as quantified by the number of infection events observed (**Fig. 5.2**). Despite identical numbers of ACH-2s across wells, more produced virus (infection events) was observed as the quantity of PBMCs in the well increased suggesting the presence of PBMCs themselves could reverse latency and increase HIV production in ACH-2s. Thus, rather than introduce PMA, a drug known to impact viability of various cell lines, the PBMCs were simply used in co-culture with the ACH-2 cells to generate contrived patient samples for the assay comparison.
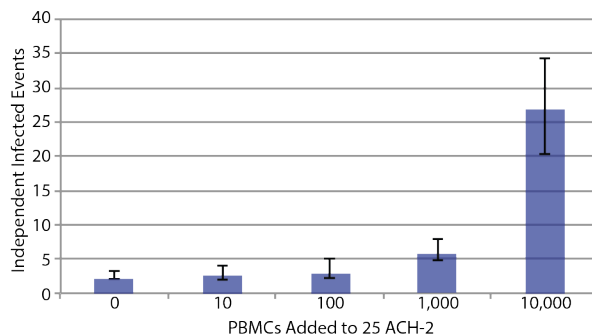
Fig. 5.2 Impact of PBMCs obtained from a healthy donor on a low number of identical ACH-2s (~25 cells) across wells in producing infectious HIV. Infection was measured based on infection events observed following a 48-hour culture with the dual fluorescent reporters. Bars represent the average of n=3 replicates and error bars represent standard deviation.

5.3.3 Evaluation of Contrived Samples across the QVOA, TZA, and Present Dual Fluorescent Reporter

To evaluate the QVOA, TZA, and dual fluorescent reporter assays, identical contrived patient samples (low numbers of ACH-2s spiked into a PBMC background from a healthy donor) were used. Though both published assays (TZA, QVOA) rely on different inducing agents for each assay (e.g., TZA uses CD3/CD28 beads, QVOA uses PHA), an identical induction mechanism was used across samples to ensure equivalent inductions across the assays. For induction of ACH-2, as shown in **Fig. 5.2**, PBMCs from a healthy donor were used to induce the ACH-2 in culture. Identical contrived patient samples each consisting of ~240,000 PBMCs evenly distributed across 24-wells in a 96-well plate were used. Each patient was set up across each of the three assays in parallel. Each assay was performed as described in the materials and methods; the number of infected wells per patient was then reported and plotted against the other assays for comparison. The comparison of the two reporter lines, TZM-bl (TZA assay) and the dual fluorescent reporter, appeared to correlate with a slight increase in event detection in the dual fluorescent reporter system (**Fig. 5.3A**). Example images of the dual fluorescent reporter positive events can be seen in (**Fig. 5.3C**), highlighting: the single infection event across the entire well including magnified a image of the event and images of the two independent reporters showing one increasing in signal while the other decreased. Similar results were seen when comparing either the TZA or dual

fluorescent reporter with the QVOA (**Fig. 5.3B**). Overall the assays appeared to report similar numbers of ACH-2 per sample. The lack of differences across assays may in part be due to the following reasons: 1) the ACH-2 were identically induced (by the presence of 10,000 PBMCs) unlike the evaluation in the published TZA assay where different induction protocols were used for the TZA and QVOA[101]; 2) while ACH-2 respond to induction with increased HIV production, these cells are not exact replicas of latent patient cells. Unlike true latent patient cells, ACH-2s are observed to produce low levels of HIV in the absence of any induction. Thus while this comparison served to highlight the potential of the dual fluorescent reporter as an alternative readout in quantifying the HIV viral reservoir, patient samples will need to be evaluated to fully evaluate the readout's potential; with patient samples, the reporter can better be assessed against both existing assays for sensitivity in detecting reservoir events. Additionally, in moving to patient samples, across assays, latency reversal agents (e.g., PMA, PHA, IL-2) should be kept consistent (as was done herein) to ensure comparison of each assay's performance rather than the effectiveness of the induction agent, a potential limitation in current published data comparing existing assays.
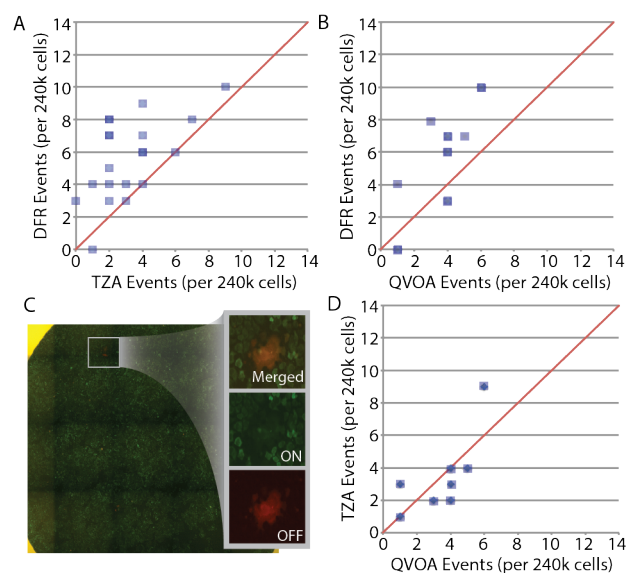
Fig. 5.3 Comparison of contrived samples across the dual fluorescent reporter assay (DFR), TZA, and QVOA. (A) Comparison of the TZA assay and DFR assay (n=20 samples). (B) Comparison of the QVOA to the DFR assay. (C) Image of a well containing a positive DFR event and the as well as showing the two independent reporters overlaid (merged), as well as the 'ON' and 'OFF' channel signals. (D) Comparison of the QVOA and TZA assays. The red

line represents a one-to-one correlation for reference. Each square represents as single sample consisting of 240,000 PBMCs (and low numbers of ACH-2) distributed across 24-wells in each assay.

5.3.4   *Micro – Macro Well-based Culture*

As subtle changes in culture are known to improve infection rates of cells, we hypothesized that manipulating culture volume could similarly impact infection rates. If a cell ends latency and begins producing virus, the volume the virus is released in ultimately determines the concentration of virus in the well; if the virus is produced and released into 50 µL, the concentration of virus will be higher than if released into 200 µL. Theoretically, this is likely to result in increased infection of cells cultured in the reduced media volume, due to a higher concentration of virus. In the context of reporter cell applications, as these assays aim to detect initial infection, enhancing early infection rates is likely to improve signal, enabling both more sensitive detection as well as clearer differentiation between infected and uninfected wells. To evaluate, we began by transitioning the 48-hour culture from a traditional media volume (~250 µL) to a reduced media volume in a well plate (~50 µL). Following 48-hour culture, while viability was not significantly impacted, both reporter cell lines (TZM-bl, dual fluorescent reporter) appeared be impacted by the prolonged culture in reduced media as cell morphology was rounded, with lower cell confluency at 48-hours compared to culture in standard media volumes (~250 µL). Additionally, the baseline signal of TZM-bl from a Beta Glow assay was reduced compared to the identical quantity of cells cultured in standard media volumes (~250 µL) (**Fig. 5.4A**). However, infection rates in reduced media volume were higher on average than in standard culture volume (**Fig. 5.4B**). Thus, preliminary testing demonstrated the potential of using lower volumes to enhance infection rates. However, given the reduced baseline signal (TZA assay) and poor cell morphology (potentially due to nutrient depletion or cellular waste buildup in the lower volume), there remained benefit in culturing in a higher media volume. Next, we hypothesized that an initial reduced media volume followed by an increase in media volume (at 24-hours) may be able to capitalize on both the enhanced infection kinetics and improved cell health. Thus, our protocol was adjusted so the co-culture of reporter and contrived patient sample was performed in 50 µL. After 24 hours, the media volume was raised to 250 µL. This resulted in a large increase in

relative infection compared to both a 48 hour reduced culture volume condition and a large culture volume (**Fig. 5.4B**). The enhanced infection signal suggests controlling the media volume could be an effective method to improve kinetics of initial infection benefiting both the DFR and TZA assay.
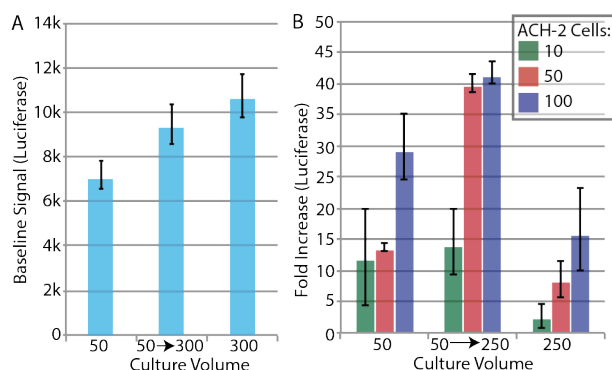
Fig. 5.4 Impact of culture volume on infection kinetics of the TZA assay using a luciferase assay. (A) Baseline signal from a TZM-bl cell only condition following 48-hours of culture in either 50 μL, adjusted media volume at 24-hous, or 250 μL. Increasing the media volume at 24-hours appears to recover some of the signal lost in prolonged 50 μL culture. (B) Evaluation of the impact of media volume on infection rates. Infection of TZM-bl with 10, 50, or 100 ACH-2 cells (in 10,000 HIV-negative PBMCs) in either 50 μL media, an adjusted media volume at 24-hours, or 250 μL of media. Infection signal is normalized to a no ACH-2 condition for each respective culture conditions. Bars represent the average of n=3 replicates and error bars represent standard deviation.

5.4 Conclusions

In order to continue development towards a cure for HIV and improve patient treatment, therapies towards targeted depletion of the reservoir are required. However, in order to enable use of reservoir depletion strategies, clinically compatible assays are also required to provide paired monitoring of the treatment efficacy. Despite the significance of the Quantitative Viral Outgrowth Assay (QVOA) in enhancing understanding of the viral reservoir, the assay remains poorly suited for clinical use, hindered by the cost, duration, and labor associated with the assay. The use of a reporter cell line to provide a direct readout of initial infection (e.g., 48-hours) provides an attractive alternative in maintaining the specificity detecting only infectious virus while greatly reducing assay length and streamlining the assay's readout. Using a dual fluorescent reporter cell line, we demonstrate the ability to detect low numbers of latent cells, evaluated through use of a model latent HIV cell line, ACH-2, from a relevant background population of healthy PBMCs (i.e., 240,000 cells). In detecting target cells in this system, the dual fluorescent reporter appears comparable to both the recently published TZA assay and the established

gold standard assay, QVOA. In addition, we demonstrate improved sensitivity from both reporter-based assays in detecting infectious events by controlling the culture volume. Specifically, by performing the initial induction and culture in a reduced media volume, concentrations of virus are higher within the well; this drives infection of the reporter. Then, to maintain cell line viability, the culture volume is raised at 24-hours and the result obtained after 48-hours of culture. This approach of initial culture in reduced volume, which was then raised to maintain cell viability demonstrated significant improvements in signal due to infection, a potential avenue for enhancing the sensitivity of reporter-based assays in detecting and quantifying the HIV viral reservoir. By integrating the dual fluorescent reporter and reduced media culture, a sensitive clinically adaptable HIV viral outgrowth assay may be possible following further validation with patient samples.

5.5   Acknowledgements

Chapter 6   Conclusions and Future Directions

A clear link to clinical progression, rare cell populations (RcPs) provide an underutilized tool with which to inform and augment clinical care. Despite the high potential clinical value (e.g., predictive, diagnostic, prognostic), these RcPs (e.g., circulating tumor cells (CTCs), and the HIV viral reservoir) face challenges in transitioning from research settings into the clinic where they can directly impact patient care. The rarity of these populations (frequently one in a million to one in a billion cells) present challenges, specifically in accessing the RcP, identifying the RcP, and adapting and utilizing heterogeneity in the RcP. In order to overcome the challenges presented by these rare, circulating populations, better tools and assays focused on adapting to these RcP challenges are required. Herein, existing tools are leveraged to develop technologies and assays enabling interrogation of the presented RcPs, specifically CTCs and the HIV viral reservoir.

6.1   Case Study 1: Circulating Tumor Cell Isolation Enabled by Magnetic Beads

One heavily utilized tool for cell isolation, magnetic beads, facilitates physical manipulation of bound cell populations. Due to the diverse offering of magnetic bead surface coatings, antibodies can be bound to their surface (e.g., through covalent binding, streptavidin-biotin) enabling the targeting of nearly any cell surface protein, and therefore any cell population of interest. The magnetic bead bound antibody, upon binding to the cell surface marker, allows the specific population targeted to then be physically manipulated from a mixed population; thus, magnetic bead antibody-based capture provides a powerful tool in capturing and assessing RcPs for continued interrogation.

Magnetic beads enable targeting of a wide number of analytes (e.g., whole cell, RNA, DNA) often spanning multiple isolation contexts. Thus, there exists a wide selection of commercially available options, which may or may not facilitate cell isolation. Despite availability, limited information exists comparing beads across similar biological contexts (e.g., capture of a specific cell line) impairing the user's ability to easily evaluate and decide between commercially available beads for cell isolation. The first chapter herein evaluates a selection of existing commercial magnetic beads (Dynabeads M-280, Dynabeads M-270, Dynabeads CeLLection, Dynabeads FlowComp, GE Sera-Mag SpeedBeads Streptavidin-Blocked) in the context of rare cell isolation, looking at each bead's performance in effectively capturing target populations and, importantly in the context of isolating RcPs, the nonspecific binding of background cells to the bead surface. Furthermore, as cell isolation processes often require downstream analytics, the impact of the beads on downstream applications were then evaluated, including: fluorescence imaging (i.e., identification, enumeration, localization), nucleic acid isolation assays (i.e., bead-based nucleic acid isolation approach and a spin column approach), and cell culture. Each bead type demonstrated variable results (e.g., capture efficiency, specificity, impact on imaging, nucleic acid extraction, protein localization, autofluorescence) across all evaluated analytic techniques, and is likely to further vary with differing target and background populations. For example in the application evaluated, M-270s demonstrated high capture efficiency and low nonspecific binding while

simultaneously demonstrating reduced viability following culture relative to the over bead types. The variable results across commercial beads highlight the importance bead selection has not only on the resultant population, but also the downstream results. In order for magnetic beads to most successfully integrate into cell isolation protocols, users must evaluate the influence of the beads specifically in the context of their desired application and endpoints. One specific RcP application we have compared, optimized, and evaluated magnetic beads for is in the isolation of CTCs.

To target CTCs, one of the evaluated bead types Dynabeads M-270s, which demonstrated high capture efficiency as well as low nonspecific binding of contaminant PBMCs was used for CTC isolation. To facilitate isolation of CTCs, an automated platform was developed around the bead-based sample preparation technology, exclusion-based sample preparation (ESP). ESP facilitates magnetic bead-based isolations by manipulating the magnetic beads out of the samples (rather than manipulating the fluid from the beads). Using ESP, we were able to develop an automated platform, which enabled CTC capture, downstream washing with integrated staining, as well as on platform nucleic acid isolation. Due to the integration of ESP in the platform, multiple cell selection methodologies were integrated allowing the user to select between positive, negative, and combinatorial selection of CTCs. To identify the captured population, on-platform staining of the bead-captured population was then performed. CTCs were identified based on identification of a nucleus (Hoechst) positive staining for cytokeratin and absence of any contaminant markers (i.e., CD45, CD14, CD11b, CD34).

Across an identical sample, the variable CTC results obtained between different selection methodologies demonstrate the existing tradeoffs in physical approaches to rare cell isolation from complex, diverse backgrounds. Positively selected CTCs from prostate cancer patients (EpCAM-based isolation) were isolated with high enough purity to enable detection of relevant downstream transcripts including androgen receptor (AR) and AR variants. While positive selection enabled isolation of a specific population with the potential for high purity (based on log-fold depletion of background populations), negative selection highlighted the opportunities available in a discovery approach to RcPs.

Overall more CTC events were detected following negative selection than positive selection (based on cytokeratin staining); while unsurprising due to the potential of EpCAM-negative CTCs (marker used for positive selection) within the sample, negative selection highlights the potential use of the platform in discovery-based approaches to rare cell isolation where a differential marker may be unknown. Building on independent positive and negative selection approaches, one of the advantages of ESP is the flexibility, enabling the integration of multiple isolation approaches (e.g., sequential selection, combinatorial selection). The ESP-enabled combinatorial selection approach highlights the potential use of this platform in integrating selection methodologies in sequence for a flexible approach to cell isolation, specifically following negative depletion with positive selection. The evaluated patient samples resulted in an overall improvement in both the number of CTCs identified and purity following combinatorial selection compared to positive selection alone.

Overall, the flexibility enabled by the developed cell isolation platform has broad potential both within and outside the CTC field. The flexibility facilitates numerous approaches to isolating cell populations and, due to the development in characterizing cell isolation magnetic beads, can be combined with cell isolation magnetic beads best tailored to the application and endpoints required. As CTC research moves forward, one of the emerging challenges is cell heterogeneity. While EpCAM is accepted as a CTC marker in a subset of cancers (e.g., prostate, breast), the emergence of EpCAM low or negative CTCs present challenges in limiting CTC isolation and analysis to only EpCAM-based capture approaches. Cell isolation platforms such as the presented automated platform could facilitate sequential selection, ensuring capture of not only the EpCAM population, but also other emerging markers. Thus, users could more easily build on CTC understanding by following selection of the well-characterized (and clinically proven) EpCAM population with capture and interrogation of additional emerging markers.

While the CTC field has made progress in highlighting the significance of CTCs as prognostic markers, more work is required to identify the potential of CTCs as predictive markers. To date, CTC enumeration has frequently been linked to prognosis, with increased CTCs proving indicative of a poor

prognosis. However, moving past identification for enumeration, a closer look at the intracellular components of CTCs (e.g., RNA transcripts, DNA modifications, protein expression, protein localization) may elucidate CTC's predictive value (e.g., treatment efficacy) and provide clinicians with valuable information in determining patient care. As CTC isolation approaches transition from enumeration to analysis of intracellular mechanisms (e.g., RNA transcripts, DNA modifications, protein expression, protein localization), cell isolation platforms will need to adapt in parallel to facilitate the study of each emerging marker.

In the context of CTCs, the developed platform demonstrates promise in the positive selection of CTCs paired with the downstream isolation and detection of rare RNA transcripts from CTCs. Preliminary data supports the successful CTC capture and downstream detection of clinically relevant RNA transcripts, most notably related to androgen receptor (AR) activity (protein, AR variants, downstream proteins). As the clinical relevance of AR is well known in prostate cancer and known to correspond to treatment decisions, the platform seems poised to facilitate EpCAM-based CTC capture followed by extraction of RNA transcripts. However, to transition this platform and assay into the clinic, additional work is required both in the context of target cell capture characterization and RNA detection characterization.

While EpCAM-based capture was well characterized in the initial platform studies (e.g., impact of target expression on target capture, quantity of target cells on capture, impact of background cell number), the impact of sample itself has not been extensively characterized. In order to implement clinically, increased understanding of the performance consistency (both in terms of cell capture and RNA transcript detection) across samples will be required to determine sample acceptance criteria. Much of the cell line performance characterization was done across a small selection of healthy donors. To better understand the impact of different patient background populations on target capture, cell line capture data should be repeated from additional blood samples; specifically, target capture should also be evaluated from blood samples drawn from patients with heightened or altered immune responses (e.g., chemotherapy patients)

to ensure understanding of the platform's performance (e.g., target cell recovery) from a broad spectrum of potential patient cohorts. Additionally, as blood ages, the sample is likely to become more challenging to process and extract the required populations from. Understanding our working window, to receive and process samples, will inform the opportunities for this assay in the CTC field either enabling the shipment of samples to central processing facilities or requiring onsite processing due to a short processing window. To ensure accuracy throughout the entire assay, not only should capture be assessed at different time points following blood draw, but also detection of RNA transcripts will need to be assessed as well. Similarly, due to the diversity in samples expected, the robustness in RNA transcript detection will need to be elucidated. Specifically in the context of both total CTC number and purity, characterization of the impact of background, non-target cells (e.g., PBMCs) on detection of CTC-related transcripts will have to be understood to determine an eligibility criteria for patients. For example, we will need to understand if we can detect target transcripts from 5 CTCs in a background of 100 cells as well as in a background of 10,000. By further characterizing sample limitations (e.g., time from draw to processing of a blood sample) as well as the required outputs (e.g., RNA quality, purity of the population, total CTCs) samples can be appropriately stratified as either compatible or incompatible with the platform. With the operating criteria established, healthy controls will then need to be evaluated to ensure specificity of the process and thresholds for specific transcripts identified if needed (e.g., use of healthy controls to set a baseline expression level for target transcripts). These characterizations will facilitate understanding of the limitations and criteria required for integration and compatibility of samples with the platform. While much retrospective data has been collected, ultimately in order to ensure clinical relevance and importantly predictive value to the patient, a full prospective clinical trial will be required.

6.2   Case Study 2: HIV Viral Reservoir Quantitation Through ESP-enabled TILDA

The HIV viral reservoir remains a major barrier to a cure for HIV. Quantitation of the HIV viral reservoir will be required to pair with emerging clinical treatments aiming to target and deplete the reservoir to monitor treatment effectiveness. The detection of rare tat/rev multiply-spliced HIV RNA from

induced patient cells using the Tat/rev Induced Limiting Dilution Assay (TILDA) provides a powerful new analyte for quantifying the HIV viral reservoir. While not yet clinically proven, the simplicity of the assay and short assay duration are two key aspects in the development of a clinically compatible assay. Building on the recent introduction of the assay and the known importance of sample preparation in PCR-based endpoints, we utilized a multiplexed (8-plex) ESP-based RNA extraction process to streamline the published assay and transition the whole cell input of the TILDA into RNA. By enabling the switch to RNA, we were able to ultimately reduce thermocycling time and improve the sensitivity of the assay in detecting target events (i.e., multiply-spliced HIV RNA). Furthermore, we were able to improve detection of target events in the presence of increased background cells enabling sampling of more patient cells per assay, a key obstacle due to the rarity of the HIV viral reservoir. While the integration of RNA extraction improved assay sensitivity, the clinical relevance of TILDA as a quantitative readout of the HIV viral reservoir remains relatively unknown. As the gold standard remains the detection of produced infectious virus measured via the Quantitative Viral Outgrowth Assay (QVOA), more work will have to be done to validate and assess how TILDA's newly introduced analyte, multiply-spliced HIV RNA, compares with either the QVOA or clinical progression to see broad adoption in clinical settings. If adopted, upstream RNA purification easily multiplexed with ESP appears likely to increase assay sensitivity and overall cell capacity of TILDA. While integrating a purified RNA input did improve sensitivity, limited detectable improvements were measured in quantified events in patient samples. Thus, in order for the added expense and labor associated with the presented RNA extraction method to have value, additional patients would need to be evaluated, which clearly demonstrate a meaningful gain in detected events (e.g., fold gains in event detected). Without clearer evidence demonstrating a marked improvement in the number of events detected by RNA TILDA over whole cell TILDA in patient samples, due to the inherent superior simplicity of the whole cell assay, the utility of a RNA extraction step is limited.

6.3   Case Study 3: HIV Viral Reservoir Quantitation Through use of a Dual Fluorescent Reporter Cell Line

The development of clinically compatible assays to quantify the HIV viral reservoir is required to support emerging reservoir depletion strategies. The gold standard approach to quantifying the HIV viral reservoir involves the detection of produced, infectious virus following latency reversal. Currently, the Quantitative Viral Outgrowth Assay (QVOA) remains the dominant assay in this space ensuring the production of infectious virus through the propagation of virus to a detectable level across a continual 7-day co-culture. Following culture, conditioned media is sampled for the propagated virus by extracting RNA for detection of HIV-specific sequences. While enabling, this assay remains prohibitive in clinical settings, due to its expense, duration, and labor-intensive protocol. As a result there exists an opportunity for an assay better able to facilitate clinical monitoring of the HIV viral reservoir, specifically in the context of measuring reservoir events capable of producing infectious virus. We propose the use of a cell-based readout with an integrated dual fluorescent reporter, which combines the specificity of two independent HIV-responsive reporters, with a simple fluorescent readout to detect a single infection event. Thus, we are able to detect infection after only 48-hours (as compared to 7-days) while requiring no RNA extraction, simplifying the readout. Additionally, unlike the recently introduced reporter-based TZA assay (uses a β-gal reporter for detection of infection via a luciferase assay after 48-hours of culture), we monitor infection in real-time as our readout is performed on live cells. Thus we can not only observe initial infection but also observe the number of cells infected. While the clinical relevance of this additional readout remains unknown, there is potential the size of the 'viral burst' (and therefore the number of cells infected in the well) could translate into a viral reservoir cell's potential for driving viral rebound (e.g., does activation result in a single infected cell or hundreds of cells). Regardless of the ability of the dual fluorescent reporter assay to quantify the reservoir, the visual readout of infection could provide additional information on existing infected cells (both actively infected and latently). By diluting the cells to less than one infectious cell per well, the extent of differential 'virus bursts' could be measured. Intra- and inter-patient heterogeneity in number of reporters infected following co-culture

could be assessed in patients. Conceivably, there could be connections between time to viral suppression on treatment and 'burst' size (e.g., do quickly suppressed patients have a lower 'viral burst' sizes). Thus, evaluation of the infection of the dual fluorescent reporters in patient samples may provide an added dimension of understanding of each patient's HIV infection.

Additionally, by moderating culture conditions within a standard well plate, we demonstrated the potential to increase assay sensitivity. Specifically, by reducing the initial culture volume (to ~50 μL) for the first 24-hours the initial virus concentration is higher facilitating infection. Then by raising the volume to standard culture volume for the second 24-hours (~250 μL), cell viability is maintained. Overall, this approach demonstrated improved infection rates of both the introduced dual fluorescent reporter and TZM-bl reporter (used by the TZA assay). While yet untested in patient samples, our presented method for improving infection signal highlights the potential to improve the sensitivity of both assays.

Due to the differences between true latent cells and the contrived samples of ACH-2, virus activation and production could greatly change and impact the measured response of the dual fluorescent reporters. In order to continue to develop this assay for eventual clinical use in measuring the HIV viral reservoir, patient samples will be required to assess performance against existing assays. Additionally, due to the poorly understood activation kinetics of the reservoir, readouts should be assessed at multiple intervals to ensure the chosen culture period for ACH-2 (48-hours) is adequate for detection of latent events in patient samples as well. However, due to the live reporting of infection by the dual fluorescent reporter, this evaluation will require no parallel experiments but can be assessed in the same culture by just re-imaging the culture at later time points. Additionally, the live cell reporting of infection provides an additional opportunity to explore activation of latent cell populations in patients. While many activation methods are utilized (e.g., PMA, PHA, CD3/CD28 beads), the efficacy of each across patients is relatively unknown. Due to the live cell reporting of infection, following activation with one inducing agent, a second inducing agent can be added to negative wells to determine if additional latent cells were present, but were simply not induced in the initial round of induction. While challenging to

mechanistically understand the implications of layering induction approaches, understanding the existence of additional events, which remained latently during the initial approach, are likely still clinically relevant with the potential to drive viral rebound in patients. Thus, using the dual fluorescent reporter readout to (re-) probe induction in latent populations provides an opportunity to understand and eventually identify, a robust activation approach. While much validation remains in evaluating the presented reporter-based assay in patient samples, these cells present an opportunity, demonstrating both high potential in quantifying the HIV viral reservoir through contrived samples (comparable to the TZA assay) as well as providing a new tool to assess HIV kinetics following latency reversal (e.g., viral burst size, layered induction).

Appendix A Integration of Magnetic Bead-Based Cell Selection into Complex Isolations [5]

A.1 Extended Bead Information

The beads evaluated for cell isolation differ in bead composition and surface coating. While not all bead composition information is available from the manufacturer, the available information for users on each bead type is compiled in Supplemental Table 1. Dynabeads M-270 Epoxy (14311D, ThermoFisher) are 2.8 μm superparamagnetic beads coated with glycidyl ether (epoxy) functional groups to faciliate antibody binding (binds to primary amino acids and sulfydral groups). The binding capacity of the beads is reported as 5 − 10 μg of IgG per milligram of M-270s. M-270s are described by the manufacturer as nonporous, hydrophillic beads with low non-specific binding. Beads are supplied lyophilized. Dynabeads M-280 Streptavidin (11205D, ThermoFisher) are also 2.8 μm superparamagnetic beads, but are hydrophobic. Due to their streptavidin coating, M-280s enable capture of biotin molecules (inlcuding biotinylated antibodies). The M-280s are coated with a monolayer of recombinant streptavidin covalently bound to the bead's surface, and are subsequently blocked with BSA. Beads are supplied in PBS (pH 7.4) with 0.1% BSA and 0.02% sodium azide. As per manufacturer specifications, the binding capacity of one milligram of M-280s is ~10 μg of biotinylated IgG. Sera-Mag SpeedBeads Streptavidin-Blocked Magnetic Particles (21152104011150, GE Healthcare Life Sciences) are 1 μm diameter streptavidin-coated beads. The beads consist of a polystyrene core encapsulated in two layers of magnetite coavalently bound to streptavidin. In addition, a proprietary blocking reagent is then used to reduce nonspecific binding to the surface of the bead (non-surfactant, and non-protein reagent). Particles are supplied in water containing 0.05% sodium azide.

The final two bead types are both releasable, enabling isolation of a bead-free cell population. Dynabeads FlowComp beads (11061D, ThermoFisher) are 2.8 μm superparamagnetic polymer

---

(unspecified) coated beads, with recominant streptavidin. The beads are supplied as part of a kit, which enables the user to label their own protein or antibody with a modified biotin, DSB-X biotin (derivative of desthiobiotin), for conjugation to the FlowComp Flexi beads (coated in a modified streptavidin as described by the manufacterer). By utilizing the DSB-X biotin-labeled antibody, the bead-antibody complex can be dissociated in release buffer, or by introducing D-biotin (B-1595 or B-20656, ThermoFisher) or D-desthiobiotin (D-20657, ThermoFisher) at a neutral pH. Beads are supplied in PBS (pH 7.4) supplemented with 0.1% BSA and 0.02% sodium azide. CELLection Biotin Binder beads (11533D, ThermoFisher) are 4.5 μm diameter superparamagnetic polystyrene beads coated with DNA linker containing a streptavidin on the end. Biotinylated antibodies are then bound to the DNA linker, which can then be cleaved (DNase I) after isolation. Beads are supplied in PBS (pH 7.4) supplemented with 0.1% Tween and 0.02% sodium azide.

PMP	Manufacturer	Releasable	Size (um)	Coating	Magnetic Properties
M-270 Epoxy	Dynabeads	No	2.8	Epoxy	Superparamagnetic
M-280	Dynabeads	No	2.8	Streptavidin	Superparamagnetic
Sera-Mag*	GE Healthcare	No	1.0	Streptavidin / proprietary blocking reagent	Magnetic
FlowComp	Dynabeads	Yes	2.8	Polymer coated beads with a Streptavidin coating (modified)	Superparamagnetic
CELLection	Dynabeads	Yes	4.5	Streptavidin-coated bead via a DNA linker	Superparamagnetic

*Sera-Mag SpeedBeads Streptavidin-Blocked Magnetic Particles

Fig. A.1 Overview of the cell isolation magnetic beads compared.

PMP	Calculated IC50 (ng Ab / ug PMP)
FlowComp	1.84
CELLection	0.35
M-280	1.33
M-270 Epoxy	N/A
Sera-Mag*	1.65

*Sera-Mag SpeedBeads Streptavidin-Blocked Magnetic Particles

Fig. A.2 Calculated fifty percent maximal anti-EpCAM antibody binding concentration values across M-280, FlowComp, CELLection, and Sera-Mag magnetic beads. For all experiments (unless noted), each bead type was conjugated to antibody at the identified density. M-270s were not analyzed in this way due to the provided

manufacturer's batch conjugation protocol.
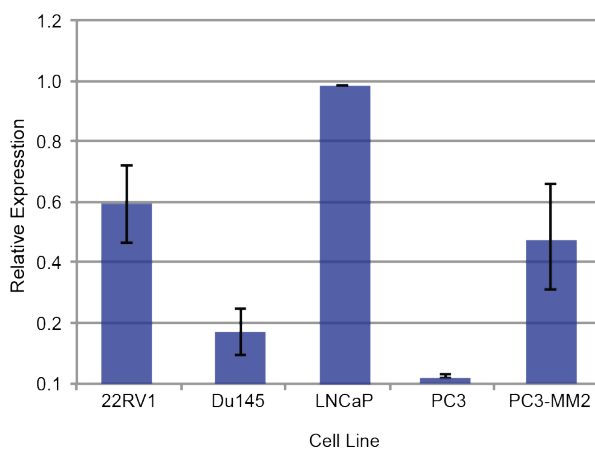
Fig. A.3 Panel of prostate cancer cell lines screened for relative EpCAM expression using a fluorescent anti-EpCAM antibody. Data normalized to average maximal EpCAM expression (LNCaP). Based on the data, a low, medium, and high EpCAM expresser were chosen for characterization with (Du145, 22Rv1, and LNCaPs respectively.
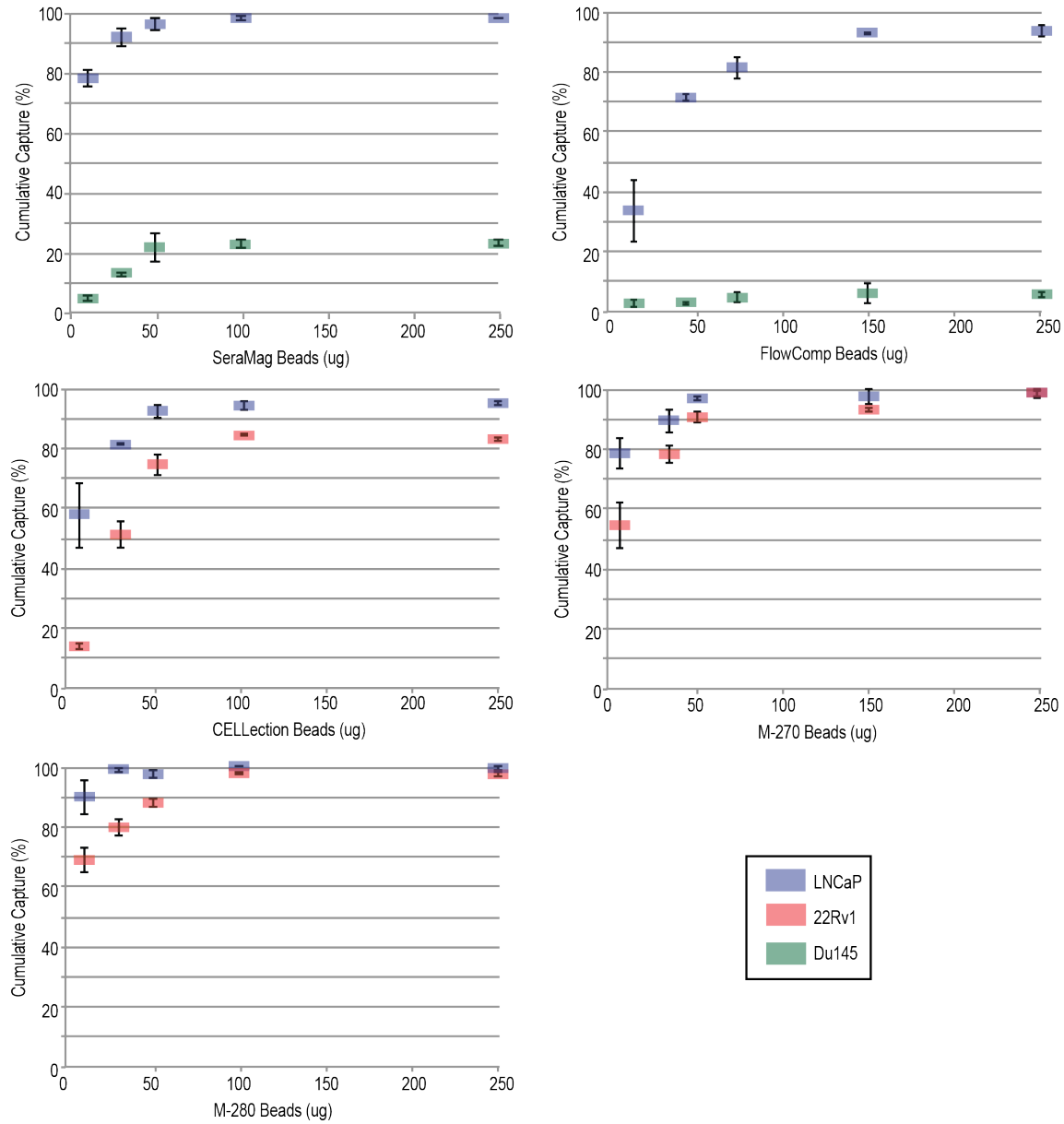
Fig. A.4 Quantity of cell isolation magnetic beads used per isolation titrated against two cell lines. Due to consistency in capture across bead types, 100 μg of magnetic beads was chosen for each bead type to move forward with in evaluated samples.
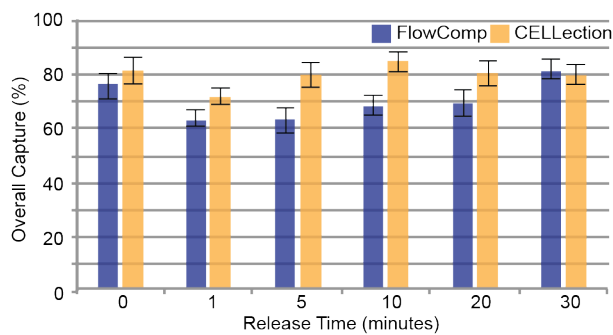
Fig. A.5 Overall capture (i.e., both released and bound captured cells compared to total population) across different release conditions. Capture remained consistent across conditions while the released population changed.
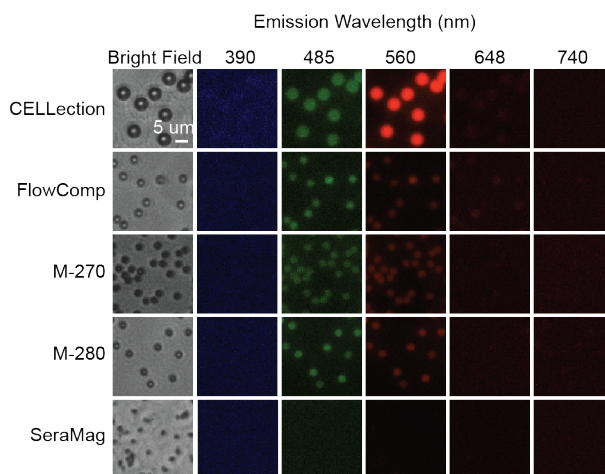
Fig. A.6 Representative multi-channel (bright field, 390, 485, 560, 648 and 740 nm) images of each magnetic bead type imaged under identical acquisition settings on glass. For visual comparison, fluorescence images within the same channel are displayed with the same contrast, and the lower limit is set at 3 standard deviations below the mean of the background intensity.

Appendix B Adaptive Exclusion-based Sample Preparation Platform for Integrated Rare Cell Isolation and Analyte Extraction [6]

B.1   Automated ESP Platform Overview

The PIPETMAX is equipped with: (1) a p200 pipette head, (2) a magnetic head, which is a p200 pipette head retrofitted with magnets (D36-N52, K&J Magnetics) and a rapid prototyped core (Midwest Prototyping) to fit the PMP collection strips, (3) a rapid prototyped rack for five the PMP collection strips (4) two magnet boxes[183], (5) a rack for 1.5 mL tubes, and (6) a rack of p200 tips (#DSF200ST, Gilson). Extraction plates (#22100008, Gilson), positioned on top of each magnetic box, contain four rows of wells, each row comprised of a sequence of six wells: one sample well (~475 µL), one large wash well (~250 µL), three small wash wells (110 µL each), and one output/elution well (110 µL for CTC samples or 15 µL for the NA samples). Buffy coat and PMPs are loaded into the sample well and allowed to bind for ~30 min and specific buffers (e.g., wash, stain, lysis buffer) are loaded into the remaining well at the predefined volumes. The magnetic head collects PMPs from the sample well by lowering over the well until the PMP collection strip contacts the fluid. The head then traverses to an adjacent well and the PMPs are released, mixed (either by pipette or magnetically), and recollected using the magnetic mechanism described by Guckenberger et al.[183]; this process is repeated until the PMPs reach the output well. Delays are added to specific wells (after the mixing step) to accommodate staining, fixing, and similar procedure. Once isolated, whole cells were either removed for microscopy or transferred to a second extraction plate for NA isolation. Cells were released directly in lysis buffer, allotted time for the NA to bind to PMPs, the PMPs carried through a series of washes and released into the elution buffer.
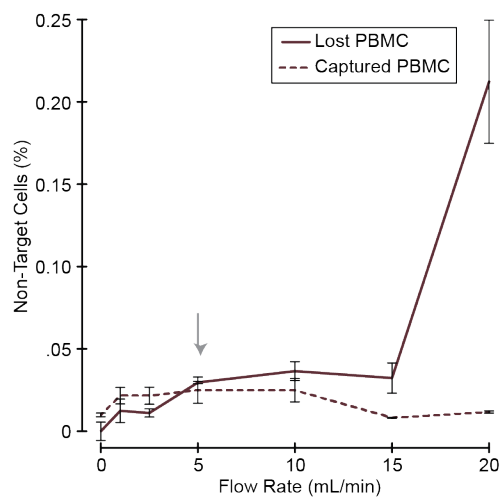
Fig. B.1 Impact of shear mixing by pipette on non-target PBMCs including both PBMCs released from the PMPs during mixing (lost PBMCs) and PBMCs that still remained bound to the PMPs after mixing (captured PBMCs).
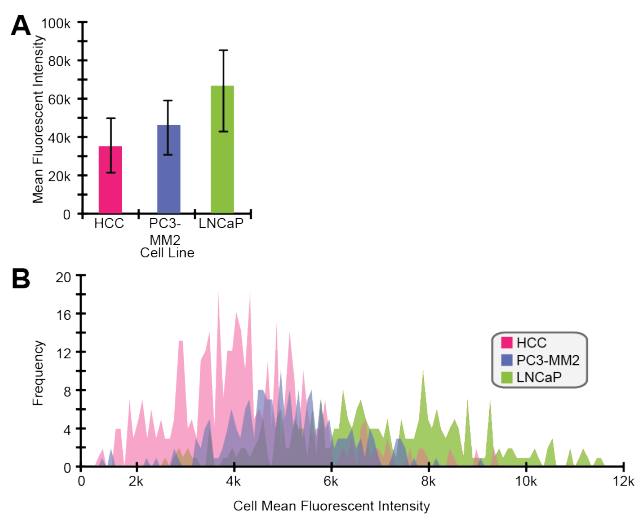
Fig. B.2 EpCAM protein expression (A) Average of each cell line's EpCAM mean fluorescent intensity. Error bars represent standard deviation (n=3) (A). Histogram of the each cell's means fluorescent EpCAM intensity (B).
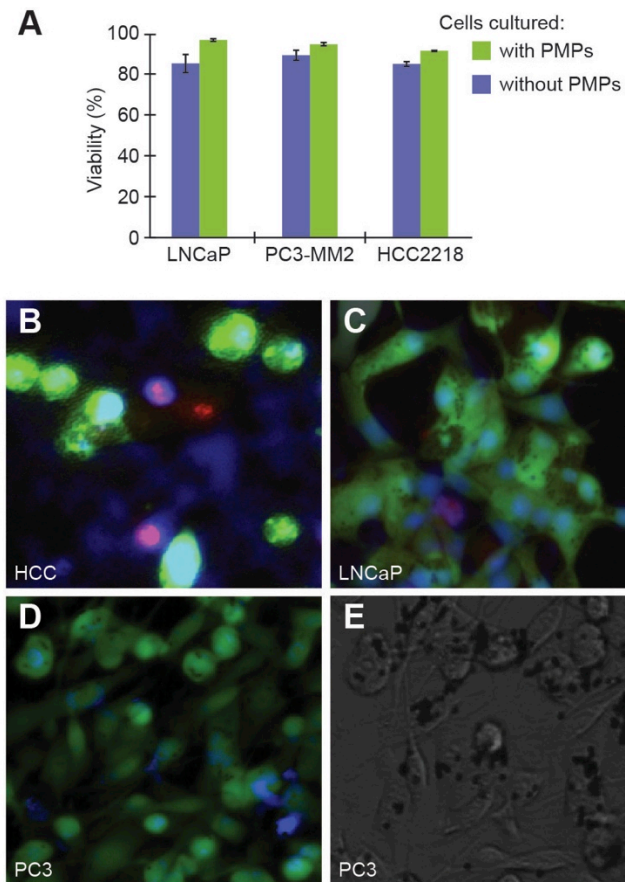
Fig. B.3 Impact of prolonged PMP exposure to cell viability. (A) Viability of cells cultured with PMPs for 5 days, following isolation from a PBMC background. (B-D) Representative images of Live/Dead staining of each of the cell lines following culture with PMPs (blue-Hoechst, green-live, red-dead) (B-HCC, C-LNCaP, D-PC3-MM2). (E) Corresponding brightfield image of PC3-MM2 showing PMPs still attached to cells (black dots).
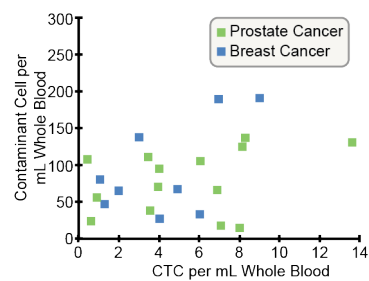
Fig. B.4 Captured contaminant cells normalized per mL of whole blood plotted against the number of CTCs captured per mL of whole blood in positive selection samples. Results indicate no relationship between contaminant cells and CTCs isolated from a sample.
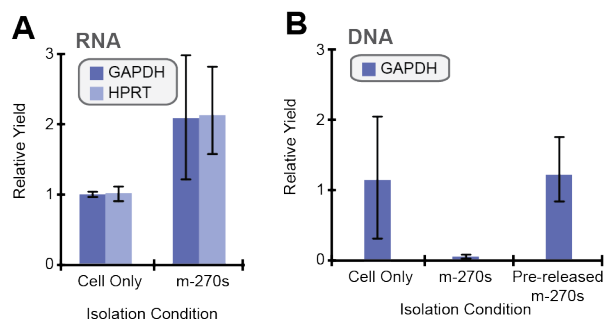
Fig. B.5 Impact of PMP on NA extraction. (A) Impact of M-270s on RNA extraction in the automated system. On average, M-270s improve RNA yields when included in with the RNA extraction PMPs. RNA quantification was done using two housekeeping genes, GAPDH and HPRT. (B) Impact of M-270s on DNA extraction in the automated system. Due to negative impact of M-270s on DNA extraction, a pre-lyse step was first performed, the M-270s removed, and DNA lysis buffer and PMPs added. Using this approach, DNA yields are consistent cell only DNA extraction yields.

	AR 1/2	Hs00907242_m1
	AR 4/5	Hs00171172_m1
AR	AR_V1	Custom (AI39R60)
	AR_V7	Hs04260217_m1
	AR_V9	Custom (AI6RPM5)
	KLK2	Hs00428383_m1
	KLK3 (PSA)	Hs02576345_m1
Prostate Cancer Specific Transcripts	TMPRSS2	Hs01120965_m1
	NKX3.1	Hs00171834_m1
	FOHL1 (PSMA)	Hs00379515_m1
	FOXA1	Hs04187555_m1
Epithelial	EpCAM	Hs00901885_m1
	KRT8 (cytokeratin 8)	Hs01595539_g1
Stem cell	PROM1 (CD133)	Hs01009250_m1
	PSCA	Hs04177224_g1
Blood	PTPRC (CD45)	Hs04189704_m1
Housekeeping	POLR2A	Hs00172187_m1
	P0	4333761F

Fig. B.6 Overview of the primers used for quantifying specific target transcripts in both cell lines and CTCs. All primers were Taqman (ThermoFisher).

Prostate CTC Cell Isolation Patients					Metastasis				
Patient	Age (yrs)	Gleason Score	Current Treatment	Time since diagnosis (Months)	Bone	Lymph Nodes	Liver	Brain	Lung parenchyma /lymphagitic spread/pleural based
1*	73	3+4	Surveillance off ADT	14	X				
2*	80	4+5	Enzalutamide	3	X	X			
3	76	9	Enzalutamide	16		X			

					Bone	Lymph Nodes	Liver	Brain	Lung
4	69	3+3	ADT	12		X			X
5	59	4+5	ADT	2	X	X			
6	71	4+5	ADT	2	X				
7*	75	4+5	Clinical trial with 52 cycles of ARN-509	8		X			
8	67	4+3	Between treatments after radium 223	7	X	X			
9	70	3+4	Enzalutamide	16	X	X			
10*	76	5+4	Abiraterone	1	X	X			
11	77	4+5	Enzalutamide	2	X	X			
12	67	Poorly differentiated	Palliative radiation	12	X				
13	66	4+5	Docetaxel	6	X	X			X
14	81	4+3	Enzalutamide	5	X	X			
15	80	5+4	Phase I trial	6	X	X	X		
16**									

Fig. B.7 Table of prostate cancer patient samples evaluated for CTC capture (* indicates multiple samples were obtained from the specified sample across multiple dates; ** patient data unknown).

Breast CTC Cell Isolation Patients						Metastasis				
Patient	Age (yrs)	Tumor Stage at Diagnosis	Primary Therapy (Primary site)	Current Therapy (Metastatic)	Time since diagnosis	Bone	Lymph Nodes	Liver	Brain	Lung
17	60	II	Docetaxel, pertuzumab/ trastuzumab,	Capecitabine (Xeloda), zoledronic acid (Zometa)	19	X	X			X
18*	60	N/A	Anastrozole plus pamidronate, multiple lines of endocrine therapy	Exemestane plus entinostat or placebo, radiation	11					
19**										
20	75	IV	Radiation	Faslodex and zoledronic acid	2	X		X		
21	75	IV	Arimidex	Nab/paclitaxel and zoledronic acid	13	X				
22	70	IA	Tamoxifen	Dexamethasone, Zoledronic acid, T-DM1	9	X		X	X	
23*	60	III	Doxorubicin and Cyclophosphamide, followed by Paclitaxel and Trastuzumab, tamoxifen	Trastuzumab (Herceptin), everolimus (Afinitor)	5	X	X	X	X	X
24*	51	IIA	AC, dd-paclitaxel	Eribulin and pembrolizumab	11		X			

Appendix C RNA-mediated TILDA for Improved Cell Capacity and Enhanced Detection of Multiply-spliced HIV RNA [7]
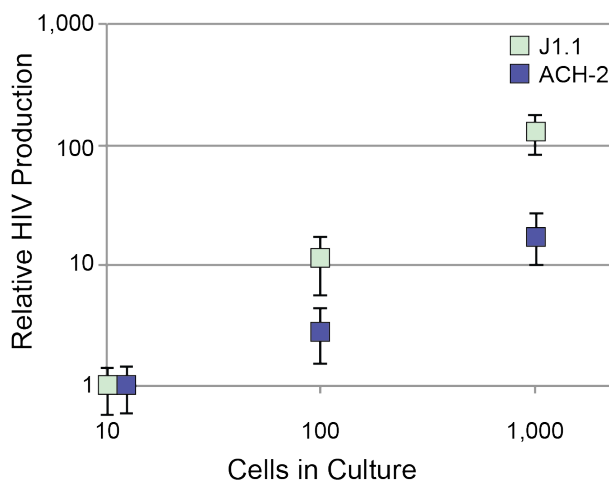
Fig. C.1 Relative detection of produced HIV from cell-free conditioned media from 10 to 1,000 inputted HIV producing cell lines, J1.1 and ACH-2. Conditioned media was collected from each culture after 24-hours of culture and assayed for HIV-specific RNA sequences using quantitative RT-PCR. HIV-specific RNA sequences were isolated from culture media 24 hours after plating in both cell lines. All samples were performed in triplicate with error bars representing standard deviation.
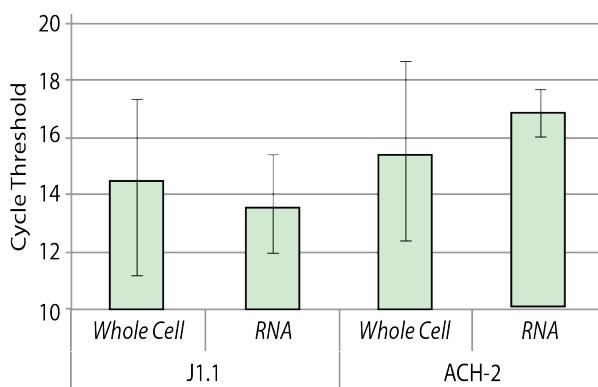
Fig. C.2 Reported cycle threshold from detection of *tat/rev* msRNA using either a whole cell input or ESP-isolated RNA. TILDA amplification and detection was done using the published 4-minute elongation cycles. Samples consisted of approximately 10 cells. Bars represent replicate averages (n=4) while error bars represent standard deviation. For both cell types, no statistically significant difference was observed (t-test, p>0.05).

---

7 This chapter has been adapted from Pezzi, H. M., Berry, S. M., Beebe, D. J., & Striker, R. (2017). RNA-mediated TILDA for improved cell capacity and enhanced detection of multiply-spliced HIV RNA. Integr Biol (Camb), 9(11), 876-884.
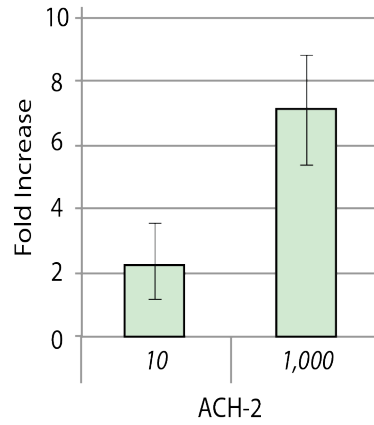
Fig. C.3 Fold increase in HIV production (quantified by RT-PCR for HIV-specific RNA sequences isolated from cell-free conditioned media) following a 24-hour induction period with 5 ng/mL PMA.

Appendix D Dual Fluorescent Reporter
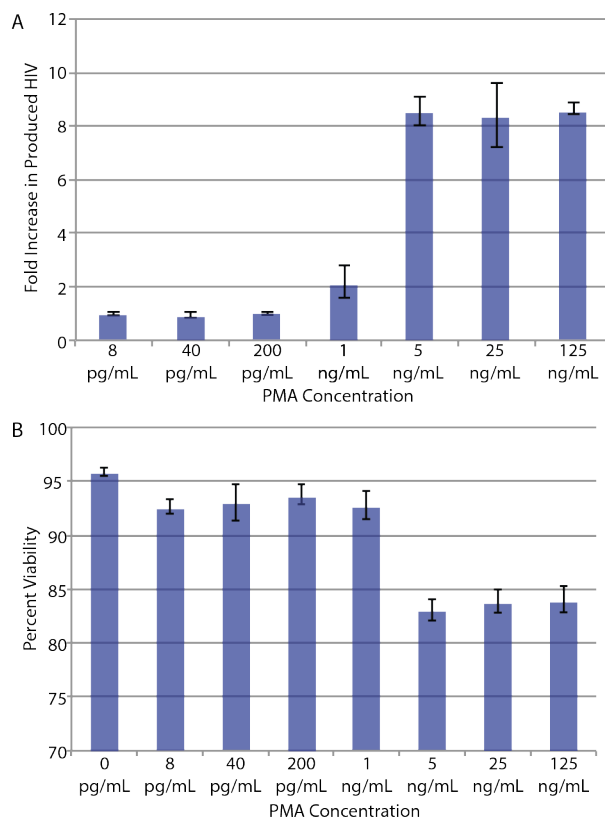
Fig. D.1 ACH-2 characterization in response to PMA induction. (A) Fold increase in HIV production (quantified by RT-PCR for HIV-specific RNA sequences isolated from cell-free conditioned media) over baseline HIV production (no ACH-2) of 10,000 ACH-2 following a 24-hour induction period. (B) Viability of ACH-2 populations after the 24-hour induction.
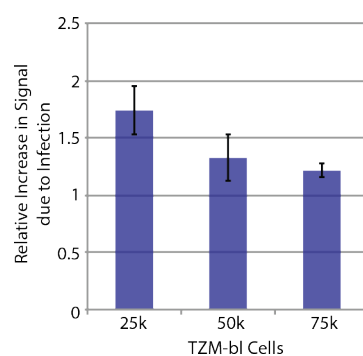
Fig. D.2 Impact of cell plating density on luciferase signal. Luciferase signal detected following a 48-hour infectious of different plating densities of TZM-bl with 100 ACH-2 in a 96-well plate (200 µL media). Signal was normalized to corresponding no ACH-2 density baseline. More relative infection signal is observed in this assay at lower densities as opposed to the published densities in a 96-well plate.

Appendix E Supplemental Publications

E.1 Overview

This appendix includes a list of publications that used the technologies described in this dissertation.

E.2 Technology Development to Facilitate Exclusion-based Sample Preparation

E.2.1 Air Jump: a new ESP technology for sample preparation

AirJump: Using Interfaces to Instantly Perform Simultaneous Extraction. Berry, S. M., **Pezzi, H. M.**, LaVanway, A. J., Guckenberger, D. J., Anderson, M., Beebe, D. J. *Applied Materials & Interfaces,* **2016.**

Analyte isolation is an important process that spans a range of biomedical disciplines, including diagnostics, research, and forensics. While downstream analytical techniques have advanced in terms of both capability and throughput, analyte isolation technology has lagged behind, increasingly becoming the bottleneck in these processes. Thus, there exists a need for simple, fast, and easy to integrate analyte separation protocols to alleviate this bottleneck. Recently, a new class of technologies has emerged that leverages the movement of paramagnetic particle (PMP)-bound analytes through phase barriers to achieve a high efficiency separation in a single or a few steps. Specifically, the passage of a PMP/analyte aggregate through a phase interface (aqueous/air in this case) acts to efficiently "exclude" unbound (contaminant) material from PMPbound analytes with higher efficiency than traditional washing-based solid-phase extraction (SPE) protocols (i.e., bind, wash several times, elute). Here, we describe for the first time a new type of "exclusion-based" sample preparation, which we term "AirJump". Upon realizing that much of the contaminant carryover stems from interactions with the sample vessel surface (e.g., pipetting residue, wetting), we aim to eliminate the influence of that factor. Thus, AirJump isolates PMP-bound analyte by "jumping" analyte directly out of a free liquid/air interface. Through careful characterization, we have demonstrated the validity of AirJump isolation through comparison to traditional washing-based isolations. Additionally, we have confirmed the suitability of AirJump in three important independent biological isolations, including protein immunoprecipitation, viral RNA isolation,

and cell culture gene expression analysis. Taken together, these data sets demonstrate that AirJump performs efficiently, with high analyte yield, high purity, no cross contamination, rapid time-to-isolation, and excellent reproducibility.

E.2.2   Magnetic technology to facilitate automation of ESP

A Magnetic System for Automated Manipulation of Paramagnetic Particles. Guckenberger D. J.*, **Pezzi H. M.***, Regier M. C., Berry S. M., Fawcett K, Barrett K, Beebe D. J. *Analytical Chemistry*. **2016**

* Co-authorship

The simple, rapid magnetic manipulation of paramagnetic particles (PMPs) paired with the wide range of available surface chemistries has strongly positioned PMPs in the field of analyte isolation. One recent technology, sliding lid for immobilized droplet extractions (SLIDE), presents a simple, rapid alternative to traditional PMP isolation protocols. Rather than remove fluid from PMP-bound analyte, SLIDE directly removes the PMPs from the fluid. SLIDE collects the PMPs on a hydrophobic, removable surface, which allows PMPs to be captured from one well and then transferred and released into a second well. Despite several key advantages, SLIDE remains limited by its passive magnetic manipulation that only allows for a one-time capture-and-release of PMPs, preventing wash steps and limiting purity. Furthermore, the strategy employed by SLIDE constrains the position of the wells, thereby limiting throughput and integration into automated systems. Here, we introduce a new, mechanically and operationally simplistic magnetic manipulation system for integration with the SLIDE technology to overcome the previously stated limitations. This magnetic system is compatible with nearly any plate design, can be integrated into automated workflows, enables high-throughput formats, simplifies mechanical requirements, and is amenable to a range of analytes. Using this magnetic system, PMPs can be collected, released, and resuspended throughout multiple wells regardless of proximity. We demonstrate this system's capabilities to isolate whole cells, mRNA, and DNA, demonstrating up to a 28-fold improvement of purity via the multiwash protocols enabled by this magnetic technology.

E.3 HIV Viral Load Development with ESP

E.3.1 Assay development for Low Cost HIV Viral Load

HIV Viral RNA Extraction in Wax Immiscible Filtration Assisted by Surface Tension (IFAST) devices. Berry, S. M., LaVanway, A. J., **Pezzi, H. M.**, Guckenberger, D. J., Anderson, M., Loeb, J. M., Beebe, D. J. *J Mol Diagn,* **2014.**

The monitoring of viral load is critical for proper management of antiretroviral therapy for HIV-positive patients. Unfortunately, in the developing world, significant economic and geographical barriers exist, limiting access to this test. The complexity of current viral load assays makes them expensive and their access limited to advanced facilities. We attempted to address these limitations by replacing conventional RNA extraction, one of the essential processes in viral load quantitation, with a simplified technique known as immiscible filtration assisted by surface tension (IFAST). Furthermore, these devices were produced via the embossing of wax, enabling local populations to produce and dispose of their own devices with minimal training or infrastructure, potentially reducing the total assay cost. In addition, IFAST can be used to reduce cold chain dependence during transportation. Viral RNA extracted from raw samples stored at 37C for 1 week exhibited nearly complete degradation. However, IFASTpurified RNA could be stored at 37C for 1 week without significant loss. These data suggest that RNA isolated at the point of care (eg, in a rural clinic) via IFAST could be shipped to a central laboratory for quantitative RT-PCR without a cold chain. Using this technology, we have demonstrated accurate and repeatable measurements of viral load on samples with as low as 50 copies per milliliter of sample.

E.3.2 Low Cost HIV Viral Load application and evaluation

Using Exclusion-based Sample Preparation (ESP) to Reduce Viral Load Assay Cost. Berry, S. M., **Pezzi, H. M**., Williams, E. D., Loeb, J. M., Guckenberger, D. J., Lavanway A. J., Puchalski, A. A., Kityo, C. M., Mugyenyi, P. N., Gaziano, F. M., Beebe, D. J. *Analytical Chemistry*. **2016**

Viral load (VL) measurements are critical to the proper management of HIV in developing countries. However, access to VL assays is limited by the high cost and complexity of existing assays. While there is a need for low cost VL assays, performance must not be compromised. Thus, new assays must be validated on metrics of limit of detection (LOD), accuracy, and dynamic range. Patient plasma samples from the Joint Clinical Research Centre in Uganda were de-identified and measured using both

an existing VL assay (Abbott RealTime HIV-1) and our assay, which combines low cost reagents with a simplified method of RNA isolation termed Exclusion-Based Sample Preparation (ESP).71 patient samples with VLs ranging from <40 to >3,000,000 copies/mL were used to compare the two methods. We demonstrated equivalent LOD (~50 copies/mL) and high accuracy (average difference between methods of 0.08 log, R2= 0.97). Using expenditures from this trial, we estimate that the cost of the reagents and consumables for this assay to be approximately $5 USD. As cost is a significant barrier to implementation of VL testing, we anticipate that our assay will enhance access to this critical monitoring test in developing countries.

E.4 Exclusion-based Sample Preparation-Facilitated Cell Capture Assays

E.4.1 *Evaluation of specificity required to assess Programmed Death-Ligand 1 in circulating tumor cells using an automated exclusion-based sample preparation platform*

High Specificity in Circulating Tumor Cell Identification is Required for Accurate Evaluation of Programmed Death-Ligand 1. Schehr, J. L., Schultz, Z. D., Warrick, J. W., Guckenberger, D. J., **Pezzi, H. M**., Sperger, J. M., Heninger, E., Saeed, A., Leal, T., Mattox, K., Traynor, A. M., Campbell, T. C., Berry, S. M., Beebe, D. J., Lang, J. M. *PloS one.* **2016.**

Expression of programmed-death ligand 1 (PD-L1) in non-small cell lung cancer (NSCLC) is typically evaluated through invasive biopsies; however, recent advances in the identification of circulating tumor cells (CTCs) may be a less invasive method to assay tumor cells for these purposes. These liquid biopsies rely on accurate identification of CTCs from the diverse populations in the blood, where some tumor cells share characteristics with normal blood cells. While many blood cells can be excluded by their high expression of CD45, neutrophils and other immature myeloid subsets have low to absent expression of CD45 and also express PD-L1. Furthermore, cytokeratin is typically used to identify CTCs, but neutrophils may stain non-specifically for intracellular antibodies, including cytokeratin, thus preventing accurate evaluation of PD-L1 expression on tumor cells. This holds even greater significance when evaluating PD-L1 in epithelial cell adhesion molecule (EpCAM) positive and EpCAM negative CTCs (as in epithelial-mesenchymal transition (EMT)).

E.4.2 Capture of CD4+ T cells from Whole Blood

Exclusion-based Capture and Enumeration of CD4+ T Cells from Whole Blood for Low-Resource Settings. Howard, A. L.[*], **Pezzi, H. M.*,** Beebe, D. J., & Berry, S. M. *J Lab Automation.* **2013.**

* Co-authorship

In developing countries, demand exists for a cost-effective method to evaluate human immunodeficiency virus patients' CD4[+] T-helper cell count. The T_H (CD4) cell count is the current marker used to identify when an HIV patient has progressed to acquired immunodeficiency syndrome, which results when the immune system can no longer prevent certain opportunistic infections. A system to perform T_H count that obviates the use of costly flow cytometry will enable physicians to more closely follow patients' disease progression and response to therapy in areas where such advanced equipment is unavailable. Our system of two serially-operated immiscible phase exclusion-based cell isolations coupled with a rapid fluorescent readout enables exclusion-based isolation and accurate counting of T-helper cells at lower cost and from a smaller volume of blood than previous methods. T_H cell isolation via immiscible filtration assisted by surface tension (IFAST) compares well against the established Dynal T4 Quant Kit and is sensitive at CD4 counts representative of immunocompromised patients (less than 200 T_H cells per microliter of blood). Our technique retains use of open, simple-to-operate devices that enable IFAST as a high-throughput, automatable sample preparation method, improving throughput over previous low-resource methods.

E.5   Cell-based Assays

E.5.1 Evaluation of neutrophil chemotaxis in dual gradients

Simple Microfluidic Device for Studying Chemotaxis in Response to Dual Gradients. Moussavi-Harami, S. F.[*], **Pezzi, H. M**.[*], Huttenlocher, A., & Beebe, D. J. *Biomedical microdevices, 17*(3), 51.

* Co-authorship

Chemotaxis is a fundamental biological process where complex chemotactic gradients are integrated and prioritized to guide cell migration toward specific locations. To understand the mechanisms of gradient dependent cell migration, it is important to develop *in vitro* models that recapitulate key attributes of the chemotactic cues present *in vivo*. Current *in vitro* tools for studying cell migration are not amenable to easily study the response of neutrophils to dual gradients. Many of these

systems require external pumps and complex setups to establish and maintain the gradients. Here we report a simple yet innovative microfluidic device for studying cell migration in the presence of dual chemotactic gradients through a 3-dimensional substrate. The device is tested and validated by studying the migration of the neutrophil-like cell line PLB-985 to gradients of fMLP. Furthermore, the device is expanded and used with heparinised whole blood, whereupon neutrophils were observed to migrate from whole blood towards gradients of fMLP eliminating the need for any neutrophil purification or capture steps.

E.5.2 *Enhanced sensitivity cell-based assay for detecting Botulinum Neurotoxin Type A*

Development of a Highly Sensitive Cell-Based Assay for Detecting Botulinum Neurotoxin Type A through Neural Culture Media Optimization. Hong, W. S., **Pezzi, H. M.**, Schuster, A. R., Berry, S. M., Sung, K. E., Beebe, D. J. *Journal of Biomolecular Screening.* **2015.**

Botulinum neurotoxin (BoNT) is the most lethal naturally produced neurotoxin. Due to the extreme toxicity, BoNTs are implicated in bioterrorism, while the specific mechanism of action and long-lasting effect was found to be medically applicable in treating various neurological disorders. Therefore, for both public and patient safety, a highly sensitive, physiologic, and specific assay is needed. In this paper, we show a method for achieving a highly sensitive cell-based assay for BoNT/A detection using the motor neuron–like continuous cell line NG108-15. To achieve high sensitivity, we performed a media optimization study evaluating three commercially available neural supplements in combination with retinoic acid, purmorphamine, transforming growth factor β1 (TGFβ1), and ganglioside GT1b. We found nonlinear combinatorial effects on BoNT/A detection sensitivity, achieving an EC_{50} of 7.4 U ± 1.5 SD (or ~7.9 pM). The achieved detection sensitivity is comparable to that of assays that used primary and stem cell–derived neurons as well as the mouse lethality assay.

References

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

2   Lima, J. *et al.* Characterization of B cells in healthy pregnant women from late pregnancy to post-partum: a prospective observational study. *BMC Pregnancy Childbirth* **16**, 139, doi:10.1186/s12884-016-0927-7 (2016).

3   Somerset, D. A., Zheng, Y., Kilby, M. D., Sansom, D. M. & Drayson, M. T. Normal human pregnancy is associated with an elevation in the immune suppressive CD25+ CD4+ regulatory T-cell subset. *Immunology* **112**, 38-43, doi:10.1111/j.1365-2567.2004.01869.x (2004).

4   Hazenberg, M. D. *et al.* Persistent immune activation in HIV-1 infection is associated with progression to AIDS. *AIDS* **17**, 1881-1888, doi:10.1097/01.aids.0000076311.76477.6e (2003).

5   Roederer, M. *et al.* CD8 naive T cell counts decrease progressively in HIV-infected adults. *J. Clin. Invest.* **95**, 2061-2066, doi:10.1172/JCI117892 (1995).

6   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).

7   Ozkumur, E. *et al.* Inertial focusing for tumor antigen-dependent and -independent sorting of rare circulating tumor cells. *Sci. Transl. Med.* **5**, 179ra147, doi:10.1126/scitranslmed.3005616 (2013).

8   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).

9   Yu, M., Stott, S., Toner, M., Maheswaran, S. & Haber, D. A. Circulating tumor cells: approaches to isolation and characterization. *The Journal of Cell Biology* **192**, 373-382, doi:10.1083/jcb.201010021 (2011).

10   Bolnick, J. M. *et al.* Trophoblast retrieval and isolation from the cervix (TRIC) for noninvasive prenatal screening at 5 to 20 weeks of gestation. *Fertil. Steril.* **102**, 135-142.e136, doi:10.1016/j.fertnstert.2014.04.008 (2014).

11   Ma, S. *et al.* Aldehyde dehydrogenase discriminates the CD133 liver cancer stem cell populations. *Mol. Cancer Res.* **6**, 1146-1153, doi:10.1158/1541-7786.MCR-08-0035 (2008).

12   Eramo, A. *et al.* Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ.* **15**, 504-514, doi:10.1038/sj.cdd.4402283 (2008).

13   Singh, S. K. *et al.* Identification of a cancer stem cell in human brain tumors. *Cancer Res.* **63**, 5821-5828 (2003).

14   Singh, S. K. *et al.* Identification of human brain tumour initiating cells. *Nature* **432**, 396-401, doi:10.1038/nature03128 (2004).

15   Pantaleo, G. *et al.* Lymphoid organs function as major reservoirs for human immunodeficiency virus. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 9838-9842 (1991).

16   Harper, M. E., Marselle, L. M., Gallo, R. C. & Wong-Staal, F. Detection of lymphocytes expressing human T-lymphotropic virus type III in lymph nodes and peripheral blood from infected individuals by in situ hybridization. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 772-776 (1986).

17   Racila, E. *et al.* Detection and characterization of carcinoma cells in the blood. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 4589-4594 (1998).

18   Hierlihy, A. M., Seale, P., Lobe, C. G., Rudnicki, M. A. & Megeney, L. A. The post-natal heart contains a myocardial stem cell population. *FEBS Lett.* **530**, 239-243 (2002).

19   Qian, C.-N. *et al.* Preparing the "Soil": The Primary Tumor Induces Vasculature Reorganization in the Sentinel Lymph Node before the Arrival of Metastatic Cancer Cells. *Cancer Res.* **66**, 10365-10376, doi:10.1158/0008-5472.CAN-06-2977 (2006).

20   Provenzano, P. P. *et al.* Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med.* **4**, 38, doi:10.1186/1741-7015-4-38 (2006).

21   Vona, G. *et al.* Enrichment, immunomorphological, and genetic characterization of fetal cells circulating in maternal blood. *Am. J. Pathol.* **160**, 51-58, doi:10.1016/S0002-9440(10)64348-9 (2002).

22   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. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 705-708 (1996).

23   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).

24   Lambotte, O. *et al.* Detection of infectious HIV in circulating monocytes from patients on prolonged highly active antiretroviral therapy. *J. Acquir. Immune Defic. Syndr.* **23**, 114-119 (2000).

25   Finzi, D. *et al.* Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* **278**, 1295-1300 (1997).

26   Cuckle, H. & Maymon, R. Development of prenatal screening--A historical overview. *Semin. Perinatol.* **40**, 12-22, doi:10.1053/j.semperi.2015.11.003 (2016).

27   Cunniff, C. Prenatal Screening and Diagnosis for Pediatricians. *Pediatrics* **114**, 889-894, doi:10.1542/peds.2004-1368 (2004).

28   Mujezinovic, F. & Alfirevic, Z. Procedure-related complications of amniocentesis and chorionic villous sampling: a systematic review. *Obstet. Gynecol.* **110**, 687-694, doi:10.1097/01.AOG.0000278820.54029.e3 (2007).

29   Akolekar, R., Beta, J., Picciarelli, G., Ogilvie, C. & D'Antonio, F. Procedure-related risk of miscarriage following amniocentesis and chorionic villus sampling: a systematic review and meta-analysis. *Ultrasound Obstet. Gynecol.* **45**, 16-26, doi:10.1002/uog.14636 (2015).

30   Norwitz, E. R. & Levy, B. Noninvasive prenatal testing: the future is now. *Rev. Obstet. Gynecol.* **6**, 48-62 (2013).

31   Lo, Y. M. *et al.* Presence of fetal DNA in maternal plasma and serum. *Lancet* **350**, 485-487, doi:10.1016/S0140-6736(97)02174-0 (1997).

32   Tounta, G. *et al.* Non-invasive prenatal diagnosis using cell-free fetal nucleic acids in maternal plasma: Progress overview beyond predictive and personalized diagnosis. *EPMA J.* **2**, 163-171, doi:10.1007/s13167-011-0085-y (2011).

33   Norton, M. E. *et al.* Cell-free DNA Analysis for Noninvasive Examination of Trisomy. *N. Engl. J. Med.* **372**, 1589-1597, doi:10.1056/NEJMoa1407349 (2015).

34   Bianchi, D. W. *et al.* DNA sequencing versus standard prenatal aneuploidy screening. *N. Engl. J. Med.* **370**, 799-808, doi:10.1056/NEJMoa1311037 (2014).

35   Gregg, A. R. *et al.* Noninvasive prenatal screening for fetal aneuploidy, 2016 update: a position statement of the American College of Medical Genetics and Genomics. *Genet. Med.* **18**, 1056-1065, doi:10.1038/gim.2016.97 (2016).

36   Lun, F. M. F. *et al.* Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin. Chem.* **54**, 1664-1672, doi:10.1373/clinchem.2008.111385 (2008).

37   Ashoor, G., Syngelaki, A., Poon, L. C. Y., Rezende, J. C. & Nicolaides, K. H. Fetal fraction in maternal plasma cell-free DNA at 11–13 weeks' gestation: relation to maternal and fetal characteristics. *Ultrasound Obstet. Gynecol.* **41**, 26-32, doi:10.1002/uog.12331 (2013).

38   Wright, C. F. & Chitty, L. S. Cell-free fetal DNA and RNA in maternal blood: implications for safer antenatal testing. *BMJ* **339**, b2451, doi:10.1136/bmj.b2451 (2009).

39   Taylor-Phillips, S. *et al.* Accuracy of non-invasive prenatal testing using cell-free DNA for detection of Down, Edwards and Patau syndromes: a systematic review and meta-analysis. *BMJ Open* **6**, e010002, doi:10.1136/bmjopen-2015-010002 (2016).

40   Chan, K. C. A. *et al.* Size distributions of maternal and fetal DNA in maternal plasma. *Clin. Chem.* **50**, 88-92, doi:10.1373/clinchem.2003.024893 (2004).

41   Poon, L. L. M., Leung, T. N., Lau, T. K., Chow, K. C. K. & Lo, Y. M. D. Differential DNA methylation between fetus and mother as a strategy for detecting fetal DNA in maternal plasma. *Clin. Chem.* **48**, 35-41 (2002).

42   Lapaire, O., Holzgreve, W., Oosterwijk, J. C., Brinkhaus, R. & Bianchi, D. W. Georg Schmorl on trophoblasts in the maternal circulation. *Placenta* **28**, 1-5, doi:10.1016/j.placenta.2006.02.004 (2007).

43   Herzenberg, L. A., Bianchi, D. W., Schröder, J., Cann, H. M. & Iverson, G. M. Fetal cells in the blood of pregnant women: detection and enrichment by fluorescence-activated cell sorting. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 1453-1455 (1979).

44   Walknowska, J., Conte, F. & Grumbach, M. PRACTICAL AND THEORETICAL IMPLICATIONS OF FETAL/MATERNAL LYMPHOCYTE TRANSFER. *The Lancet* **293**, 1119-1122, doi:https://doi.org/10.1016/S0140-6736(69)91642-0 (1969).

45   Mouawia, H. *et al.* Circulating trophoblastic cells provide genetic diagnosis in 63 fetuses at risk for cystic fibrosis or spinal muscular atrophy. *Reprod. Biomed. Online* **25**, 508-520, doi:10.1016/j.rbmo.2012.08.002 (2012).

46   Wessman, M., Ylinen, K. & Knuutila, S. Fetal granulocytes in maternal venous blood detected by in situ hybridization. *Prenat. Diagn.* **12**, 993-1000, doi:10.1002/pd.1970121204 (1992).

47   Coata, G. *et al.* Prenatal diagnosis of genetic abnormalities using fetal CD34+ stem cells in maternal circulation and evidence they do not affect diagnosis in later pregnancies. *Stem Cells* **19**, 534-542, doi:10.1634/stemcells.19-6-534 (2001).

48   Porra, V. *et al.* Identification and quantification of fetal red blood cells in maternal blood by a dual-color flow cytometric method: evaluation of the Fetal Cell Count kit. *Transfusion* **47**, 1281-1289, doi:10.1111/j.1537-2995.2007.01271.x (2007).

49   Fiddler, M. Fetal Cell Based Prenatal Diagnosis: Perspectives on the Present and Future. *J. Clin. Med. Res.* **3**, 972-985, doi:10.3390/jcm3030972 (2014).

50   Hamada, H., Arinami, T., Kubo, T., Hamaguchi, H. & Iwasaki, H. Fetal nucleated cells in maternal peripheral blood: frequency and relationship to gestational age. *Hum. Genet.* **91**, 427-432 (1993).

51   Little, M. T., Langlois, S., Wilson, R. D. & Lansdorp, P. M. Frequency of fetal cells in sorted subpopulations of nucleated erythroid and CD34+ hematopoietic progenitor cells from maternal peripheral blood. *Blood* **89**, 2347-2358 (1997).

52   Wachtel, S. S., Shulman, L. P. & Sammons, D. Fetal cells in maternal blood. *Clin. Genet.* **59**, 74-79 (2001).

53   Zheng, Y. L. *et al.* Prenatal diagnosis from maternal blood: simultaneous immunophenotyping and FISH of fetal nucleated erythrocytes isolated by negative magnetic cell sorting. *J. Med. Genet.* **30**, 1051-1056 (1993).

54   Mohamed, H., Turner, J. N. & Caggana, M. Biochip for separating fetal cells from maternal circulation. *J. Chromatogr. A* **1162**, 187-192, doi:10.1016/j.chroma.2007.06.025 (2007).

55   Bianchi, D. W. *et al.* Fetal gender and aneuploidy detection using fetal cells in maternal blood: analysis of NIFTY I data. National Institute of Child Health and Development Fetal Cell Isolation Study. *Prenat. Diagn.* **22**, 609-615, doi:10.1002/pd.347 (2002).

56   Jackson, L. Fetal cells and DNA in maternal blood. *Prenat. Diagn.* **23**, 837-846, doi:10.1002/pd.705 (2003).

57   Bulmer, J. N., Rodeck, C. & Adinolfi, M. Immunohistochemical characterization of cells retrieved by transcervical sampling in early pregnancy. *Prenat. Diagn.* **15**, 1143-1153 (1995).

58   Paraiso, M. F., Brady, K., Helmchen, R. & Roat, T. W. Evaluation of the endocervical Cytobrush and Cervex-Brush in pregnant women. *Obstet. Gynecol.* **84**, 539-543 (1994).

59   Imudia, A. N. *et al.* Retrieval of trophoblast cells from the cervical canal for prediction of abnormal pregnancy: a pilot study. *Hum. Reprod.* **24**, 2086-2092, doi:10.1093/humrep/dep206 (2009).

60   Bolnick, A. D. *et al.* Trophoblast Retrieval and Isolation From the Cervix for Noninvasive, First Trimester, Fetal Gender Determination in a Carrier of Congenital Adrenal Hyperplasia. *Reprod. Sci.* **23**, 717-722, doi:10.1177/1933719116632922 (2016).

61   Jain, C. V. *et al.* Fetal genome profiling at 5 weeks of gestation after noninvasive isolation of trophoblast cells from the endocervical canal. *Sci. Transl. Med.* **8**, 363re364-363re364, doi:10.1126/scitranslmed.aah4661 (2016).

62   Miltenyi, S., Müller, W., Weichel, W. & Radbruch, A. High gradient magnetic cell separation with MACS. *Cytometry* **11**, 231-238, doi:10.1002/cyto.990110203 (1990).

63   Allard, W. J. *et al.* Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. *Clin. Cancer Res.* **10**, 6897-6904, doi:10.1158/1078-0432.CCR-04-0378 (2004).

64   Penault-Llorca, F. *et al.* Optimization of immunohistochemical detection of ERBB2 in human breast cancer: Impact of fixation. *The Journal of Pathology* **173**, 65-75, doi:10.1002/path.1711730111 (1994).

65   Leong, A. S. & Gilham, P. N. The effects of progressive formaldehyde fixation on the preservation of tissue antigens. *Pathology* **21**, 266-268 (1989).

66   Srinivasan, M., Sedmak, D. & Jewell, S. Effect of Fixatives and Tissue Processing on the Content and Integrity of Nucleic Acids. *The American Journal of Pathology* **161**, 1961-1971 (2002).

67   Williams, C. *et al.* A high frequency of sequence alterations is due to formalin fixation of archival specimens. *Am J Pathol* **155**, 1467-1471, doi:10.1016/s0002-9440(10)65461-2 (1999).

68   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).

69   Perelson, A. S. *et al.* Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature* **387**, 188, doi:10.1038/387188a0 (1997).

70   Gulick, R. M. *et al.* Treatment with Indinavir, Zidovudine, and Lamivudine in Adults with Human Immunodeficiency Virus Infection and Prior Antiretroviral Therapy. *New England Journal of Medicine* **337**, 734-739, doi:doi:10.1056/NEJM199709113371102 (1997).

71   Palella, F. J. *et al.* Declining Morbidity and Mortality among Patients with Advanced Human Immunodeficiency Virus Infection. *New England Journal of Medicine* **338**, 853-860, doi:doi:10.1056/NEJM199803263381301 (1998).

72   Finzi, D. *et al.* Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* **278**, 1295-1300 (1997).

73   Finzi, D. *et al.* Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med* **5**, 512-517, doi:10.1038/8394 (1999).

74   Chun, T. W. *et al.* Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc Natl Acad Sci U S A* **94**, 13193-13197 (1997).

75   Wong, J. K. *et al.* Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science* **278**, 1291-1295 (1997).

76   Chun, T.-W. *et al.* In vivo fate of HIV-1-infected T cells: Quantitative analysis of the transition to stable latency. *Nat Med* **1**, 1284-1290 (1995).

77   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).

78   Spiegel, H., Herbst, H., Niedobitek, G., Foss, H. D. & Stein, H. Follicular dendritic cells are a major reservoir for human immunodeficiency virus type 1 in lymphoid tissues facilitating infection of CD4+ T-helper cells. *Am J Pathol* **140**, 15-22 (1992).

79   Burton, G. F., Keele, B. F., Estes, J. D., Thacker, T. C. & Gartner, S. Follicular dendritic cell contributions to HIV pathogenesis. *Seminars in immunology* **14**, 275-284 (2002).

80   Churchill, M. J. *et al.* Use of laser capture microdissection to detect integrated HIV-1 DNA in macrophages and astrocytes from autopsy brain tissues. *Journal of neurovirology* **12**, 146-152, doi:10.1080/13550280600748946 (2006).

81   Thompson, K. A., Cherry, C. L., Bell, J. E. & McLean, C. A. Brain cell reservoirs of latent virus in presymptomatic HIV-infected individuals. *Am J Pathol* **179**, 1623-1629, doi:10.1016/j.ajpath.2011.06.039 (2011).

82   Chun, T. W. *et al.* Persistence of HIV in gut-associated lymphoid tissue despite long-term antiretroviral therapy. *J Infect Dis* **197**, 714-720, doi:10.1086/527324 (2008).

83   Eisele, E. & Siliciano, R. F. Redefining the viral reservoirs that prevent HIV-1 eradication. *Immunity* **37**, 377-388, doi:10.1016/j.immuni.2012.08.010 (2012).

84   Archin, N. M. *et al.* Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature* **487**, 482, doi:10.1038/nature11286 (2012).

85   Shan, L. *et al.* Stimulation of HIV-1-specific cytolytic T lymphocytes facilitates elimination of latent viral reservoir after virus reactivation. *Immunity* **36**, 491-501, doi:10.1016/j.immuni.2012.01.014 (2012).

86   Ho, Y.-C. *et al.* Replication-Competent Noninduced Proviruses in the Latent Reservoir Increase Barrier to HIV-1 Cure. *Cell* **155**, 540-551, doi:http://dx.doi.org/10.1016/j.cell.2013.09.020 (2013).

87   Sanchez, G., Xu, X., Chermann, J. C. & Hirsch, I. Accumulation of defective viral genomes in peripheral blood mononuclear cells of human immunodeficiency virus type 1-infected individuals. *J Virol* **71**, 2233-2240 (1997).

88   Kieffer, T. L. *et al.* G→A Hypermutation in Protease and Reverse Transcriptase Regions of Human Immunodeficiency Virus Type 1 Residing in Resting CD4(+) T Cells In Vivo. *Journal of Virology* **79**, 1975-1980, doi:10.1128/JVI.79.3.1975-1980.2005 (2005).

89   Yu, J. J. *et al.* A more precise HIV integration assay designed to detect small differences finds lower levels of integrated DNA in HAART treated patients. *Virology* **379**, 78-86, doi:10.1016/j.virol.2008.05.030 (2008).

90   O'Doherty, U., Swiggard, W. J., Jeyakumar, D., McGain, D. & Malim, M. H. A sensitive, quantitative assay for human immunodeficiency virus type 1 integration. *J Virol* **76**, 10942-10950 (2002).

91   Fischer, M. *et al.* Residual cell-associated unspliced HIV-1 RNA in peripheral blood of patients on potent antiretroviral therapy represents intracellular transcripts. *Antiviral therapy* **7**, 91-104 (2002).

92   Furtado, M. R. *et al.* Persistence of HIV-1 transcription in peripheral-blood mononuclear cells in patients receiving potent antiretroviral therapy. *N Engl J Med* **340**, 1614-1622, doi:10.1056/nejm199905273402102 (1999).

93   Lewin, S. *et al.* Use of real-time PCR and molecular beacons to detect virus replication in human immunodeficiency virus type 1-infected individuals on prolonged effective antiretroviral therapy. *Journal of virology* **73**, 6099-6103 (1999).

94   Pasternak, A. O. *et al.* Cellular levels of HIV unspliced RNA from patients on combination antiretroviral therapy with undetectable plasma viremia predict the therapy outcome. *PloS one* **4**, e8490 (2009).

95   Vesanen, M., Markowitz, M., Cao, Y., Ho, D. D. & Saksela, K. Human immunodeficiency virus type-1 mRNA splicing pattern in infected persons is determined by the proportion of newly infected cells. *Virology* **236**, 104-109 (1997).

96   Sonza, S. *et al.* Selectively reduced tat mRNA heralds the decline in productive human immunodeficiency virus type 1 infection in monocyte-derived macrophages. *J Virol* **76**, 12611-12621 (2002).

97   Pasternak, A. O. *et al.* Highly sensitive methods based on seminested real-time reverse transcription-PCR for quantitation of human immunodeficiency virus type 1 unspliced and multiply spliced RNA and proviral DNA. *Journal of clinical microbiology* **46**, 2206-2211 (2008).

98   Fischer, M. *et al.* Attenuated and Nonproductive Viral Transcription in the Lymphatic Tissue of HIV-1–Infected Patients Receiving Potent Antiretroviral Therapy. *The Journal of Infectious Diseases* **189**, 273-285, doi:10.1086/380797 (2004).

99   Laird, G. M. *et al.* Rapid Quantification of the Latent Reservoir for HIV-1 Using a Viral Outgrowth Assay. *PLoS pathogens* **9**, e1003398, doi:10.1371/journal.ppat.1003398 (2013).

100   Siliciano, J. D. & Siliciano, R. F. in *Human Retrovirus Protocols: Virology and Molecular Biology* (ed Tuofu Zhu) 3-15 (Humana Press, 2005).

101   Sanyal, A. *et al.* Novel assay reveals a large, inducible, replication-competent HIV-1 reservoir in resting CD4+ T cells. *Nature Medicine* **23**, 885, doi:10.1038/nm.4347 https://www.nature.com/articles/nm.4347#supplementary-information (2017).

102   Gerlinger, M. *et al.* Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *N. Engl. J. Med.* **366**, 883-892, doi:10.1056/NEJMoa1113205 (2012).

103   Jiao, L. R. *et al.* Unique localization of circulating tumor cells in patients with hepatic metastases. *J. Clin. Oncol.* **27**, 6160-6165, doi:10.1200/JCO.2009.24.5837 (2009).

104   Canetti, E. F. D., Keane, J., McLellan, C. P. & Gray, A. B. Comparison of capillary and venous blood in the analysis of concentration and function of leucocyte sub-populations. *Eur. J. Appl. Physiol.* **116**, 1583-1593, doi:10.1007/s00421-016-3413-z (2016).

105   Tsavaris, N., Kosmas, C., Vadiaka, M., Kanelopoulos, P. & Boulamatsis, D. Immune changes in patients with advanced breast cancer undergoing chemotherapy with taxanes. *Br. J. Cancer* **87**, 21-27, doi:10.1038/sj.bjc.6600347 (2002).

106   Richman, C. M., Weiner, R. S. & Yankee, R. A. Increase in circulating stem cells following chemotherapy in man. *Blood* **47**, 1031-1039 (1976).

107   Campbell, J. J. *et al.* Chemokines and the arrest of lymphocytes rolling under flow conditions. *Science* **279**, 381-384 (1998).

108   Cristofanilli, M. *et al.* Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N. Engl. J. Med.* **351**, 781-791, doi:10.1056/NEJMoa040766 (2004).

109   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).

110   De Bono, J. S. *et al.* Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. *Clin. Cancer Res.* **14**, 6302-6309 (2008).

111   Munz, M., Fellinger, K., Hofmann, T., Schmitt, B. & Gires, O. Glycosylation is crucial for stability of tumour and cancer stem cell antigen EpCAM. *Front. Biosci.* **13**, 5195-5201 (2008).

112   Pauli, C. *et al.* Tumor-specific glycosylation of the carcinoma-associated epithelial cell adhesion molecule EpCAM in head and neck carcinomas. *Cancer Lett.* **193**, 25-32 (2003).

113   Wright, A. & Morrison, S. L. Effect of glycosylation on antibody function: implications for genetic engineering. *Trends Biotechnol.* **15**, 26-32, doi:10.1016/S0167-7799(96)10062-7 (1997).

114   Aktas, B. *et al.* Stem cell and epithelial-mesenchymal transition markers are frequently overexpressed in circulating tumor cells of metastatic breast cancer patients. *Breast Cancer Res.* **11**, R46, doi:10.1186/bcr2333 (2009).

115   Willipinski-Stapelfeldt, B. *et al.* Changes in cytoskeletal protein composition indicative of an epithelial-mesenchymal transition in human micrometastatic and primary breast carcinoma cells. *Clin. Cancer Res.* **11**, 8006-8014, doi:10.1158/1078-0432.CCR-05-0632 (2005).

116   Sophos, N. A. & Vasiliou, V. Aldehyde dehydrogenase gene superfamily: the 2002 update. *Chem. Biol. Interact.* **143-144**, 5-22 (2003).

117   Reyes, E. E. *et al.* Quantitative characterization of androgen receptor protein expression and cellular localization in circulating tumor cells from patients with metastatic castration-resistant prostate cancer. *Journal of translational medicine* **12**, 313, doi:10.1186/s12967-014-0313-z (2014).

118   de Bono , J. S. *et al.* Abiraterone and Increased Survival in Metastatic Prostate Cancer. *New England Journal of Medicine* **364**, 1995-2005, doi:10.1056/NEJMoa1014618 (2011).

119   Scher , H. I. *et al.* Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy. *New England Journal of Medicine* **367**, 1187-1197, doi:10.1056/NEJMoa1207506 (2012).

120   Antonarakis , E. S. *et al.* AR-V7 and Resistance to Enzalutamide and Abiraterone in Prostate Cancer. *New England Journal of Medicine* **371**, 1028-1038, doi:10.1056/NEJMoa1315815 (2014).

121   Tannock , I. F. *et al.* Docetaxel plus Prednisone or Mitoxantrone plus Prednisone for Advanced Prostate Cancer. *New England Journal of Medicine* **351**, 1502-1512, doi:10.1056/NEJMoa040720 (2004).

122   de Bono, J. S. *et al.* Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *Lancet* **376**, 1147-1154, doi:10.1016/s0140-6736(10)61389-x (2010).

123   Antonarakis, E. S. *et al.* Androgen Receptor Splice Variant 7 and Efficacy of Taxane Chemotherapy in Patients With Metastatic Castration-Resistant Prostate Cancer. *JAMA oncology* **1**, 582-591, doi:10.1001/jamaoncol.2015.1341 (2015).

124   Onstenk, W. *et al.* Efficacy of Cabazitaxel in Castration-resistant Prostate Cancer Is Independent of the Presence of AR-V7 in Circulating Tumor Cells. *European urology* **68**, 939-945, doi:https://doi.org/10.1016/j.eururo.2015.07.007 (2015).

125   Szaniszlo, P. *et al.* Getting the right cells to the array: Gene expression microarray analysis of cell mixtures and sorted cells. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* **59**, 191-202, doi:10.1002/cyto.a.20055 (2004).

126   Altelaar, A. F. M. & Heck, A. J. R. Trends in ultrasensitive proteomics. *Current Opinion in Chemical Biology* **16**, 206-213, doi:http://dx.doi.org/10.1016/j.cbpa.2011.12.011 (2012).

127   Bianchi, D. W. *et al.* Fetal gender and aneuploidy detection using fetal cells in maternal blood: analysis of NIFTY I data. *Prenatal diagnosis* **22**, 609-615, doi:10.1002/pd.347 (2002).

128   Vona, G. *et al.* Enrichment, Immunomorphological, and Genetic Characterization of Fetal Cells Circulating in Maternal Blood. *The American Journal of Pathology* **160**, 51-58, doi:http://dx.doi.org/10.1016/S0002-9440(10)64348-9 (2002).

129   Cristofanilli, M. *et al.* Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med* **2004**, 781-791 (2004).

130   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 (2008).

131   Markets, M. Cell isolation/cell separation market by product (regent, media, bead, centrifuge), cell type (human, stem cell, animal), technique (filtration, surface market), application (research, ivd) & end user (hospital, biotechnology)-forecast to 2019. 2014: Available at: http://www. marketsandmarkets. com. *Market–Reports/cell–isolation–market–103931479. html*.

132   Sethu, P., Sin, A. & Toner, M. Microfluidic diffusive filter for apheresis (leukapheresis). *Lab on a Chip* **6**, 83-89 (2006).

133   Ji, H. M. *et al.* Silicon-based microfilters for whole blood cell separation. *Biomedical microdevices* **10**, 251-257 (2008).

134   Campbell, G. L., Schachner, M. & Sharrow, S. O. Isolation of glial cell-enriched and -depleted populations from mouse cerebellum by density gradient centrifugation and electronic cell sorting. *Brain Research* **127**, 69-86, doi:http://dx.doi.org/10.1016/0006-8993(77)90380-8 (1977).

135   Friedman, S. L. & Roll, F. J. Isolation and culture of hepatic lipocytes, Kupffer cells, and sinusoidal endothelial cells by density gradient centrifugation with Stractan. *Analytical Biochemistry* **161**, 207-218, doi:http://dx.doi.org/10.1016/0003-2697(87)90673-7 (1987).

136   Dobbs, L. G., Gonzalez, R. & Williams, M. C. An Improved Method for Isolating Type II Cells in High Yield and Purity. *American Review of Respiratory Disease* **134**, 141-145, doi:10.1164/arrd.1986.134.1.141 (1986).

137   Pankhurst, Q. A., Connolly, J., Jones, S. K. & Dobson, J. Applications of magnetic nanoparticles in biomedicine. *Journal of Physics D: Applied Physics* **36**, R167 (2003).

138   Pamme, N. Magnetism and microfluidics. *Lab Chip* **6**, 24-38, doi:10.1039/b513005k (2006).

139   Gijs, M. A. M. Magnetic bead handling on-chip: new opportunities for analytical applications. *Microfluidics and Nanofluidics* **1**, 22-40, doi:10.1007/s10404-004-0010-y (2004).

140   Safarik, I. & Safarikova, M. Magnetic techniques for the isolation and purification of proteins and peptides. *BioMagnetic Research and Technology* **2**, 7 (2004).

141   Lin, A., Salvador, A. & Carter, J. M. Multiplexed Microsphere Suspension Array-Based Immunoassays. *ELISA: Methods and Protocols*, 107-118 (2015).

142   Price, C. W., Leslie, D. C. & Landers, J. P. Nucleic acid extraction techniques and application to the microchip. *Lab on a Chip* **9**, 2484-2494, doi:10.1039/B907652M (2009).

143   Rudi, K. *et al.* Rapid, universal method to isolate PCR-ready DNA using magnetic beads. *BioTechniques* **22**, 506-511 (1997).

144   Berensmeier, S. Magnetic particles for the separation and purification of nucleic acids. *Applied microbiology and biotechnology* **73**, 495-504 (2006).

145   Brian, D. P., Shashi, K. M. & Laura, H. L. Fundamentals and application of magnetic particles in cell isolation and enrichment: a review. *Reports on Progress in Physics* **78**, 016601 (2015).

146   Liu, C., Stakenborg, T., Peeters, S. & Lagae, L. Cell manipulation with magnetic particles toward microfluidic cytometry. *Journal of Applied Physics* **105**, 102014, doi:10.1063/1.3116091 (2009).

147   Avital, I. *et al.* Isolation, characterization, and transplantation of bone marrow-derived hepatocyte stem cells. *Biochemical and biophysical research communications* **288**, 156-164 (2001).

148   Asahara, T. *et al.* Isolation of putative progenitor endothelial cells for angiogenesis. *Science* **275**, 964-966 (1997).

149   Pamme, N. On-chip bioanalysis with magnetic particles. *Current Opinion in Chemical Biology* **16**, 436-443, doi:https://doi.org/10.1016/j.cbpa.2012.05.181 (2012).

150   Fonnum, G., Johansson, C., Molteberg, A., Mørup, S. & Aksnes, E. Characterisation of Dynabeads® by magnetization measurements and Mössbauer spectroscopy. *Journal of Magnetism and Magnetic Materials* **293**, 41-47, doi:https://doi.org/10.1016/j.jmmm.2005.01.041 (2005).

151   Gijs, M. A. M., Lacharme, F. & Lehmann, U. Microfluidic Applications of Magnetic Particles for Biological Analysis and Catalysis. *Chemical Reviews* **110**, 1518-1563, doi:10.1021/cr9001929 (2010).

152   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).

153   Nagrath, S. *et al.* Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* **450**, doi:10.1038/nature06385 (2007).

154   Riethdorf, S. *et al.* Detection of Circulating Tumor Cells in Peripheral Blood of Patients with Metastatic Breast Cancer: A Validation Study of the CellSearch System. *Clinical Cancer Research* **13**, 920-928, doi:10.1158/1078-0432.ccr-06-1695 (2007).

155   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).

156   de Bono, J. S. *et al.* Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **14**, 6302-6309, doi:10.1158/1078-0432.ccr-08-0872 (2008).

157   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).

158   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).

159   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).

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

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

162   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).

163   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).

164   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).

165   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).

166   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).

167   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).

168   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).

169   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).

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

171   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).

172   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).

173   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).

174   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).

175   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).

176   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).

177   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).

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

179 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).

180 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).

181 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).

182 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).

183 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).

184 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).

185 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).

186 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).

187 Königsberg, R. *et al.* Detection of EpCAM positive and negative circulating tumor cells in metastatic breast cancer patients. *Acta oncologica* **50**, 700-710 (2011).

188 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).

189 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).

190 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).

191 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).

192 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).

193 Cannon, G. J. & Swanson, J. A. The macrophage capacity for phagocytosis. *Journal of Cell Science* **101**, 907-913 (1992).

194 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.

195 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).

196 Leblanc, R. & Peyruchaud, O. Metastasis: new functional implications of platelets and megakaryocytes. *Blood* **128**, 24-31, doi:10.1182/blood-2016-01-636399 (2016).

197 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).

198 Ledergerber, B. *et al.* Clinical progression and virological failure on highly active antiretroviral therapy in HIV-1 patients: a prospective cohort study. *The Lancet* **353**, 863-868, doi:http://dx.doi.org/10.1016/S0140-6736(99)01122-8 (1999).

199 Zolopa, A. R. *et al.* Early antiretroviral therapy reduces AIDS progression/death in individuals with acute opportunistic infections: a multicenter randomized strategy trial. *PloS one* **4**, e5575 (2009).

200 Chun, T.-W. *et al.* HIV-infected individuals receiving effective antiviral therapy for extended periods of time continually replenish their viral reservoir. *The Journal of Clinical Investigation* **115**, 3250-3255, doi:10.1172/JCI26197 (2005).

201 Chun, T.-W. *et al.* Relationship between pre-existing viral reservoirs and the re-emergence of plasma viremia after discontinuation of highly active anti-retroviral therapy. *Nature medicine* **6**, 757-761 (2000).

202 Eriksson, S. *et al.* Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS pathogens* **9**, e1003174, doi:10.1371/journal.ppat.1003174 (2013).

203 Rådström, P., Knutsson, R., Wolffs, P., Lövenklev, M. & Löfström, C. Pre-PCR processing. *Molecular biotechnology* **26**, 133-146 (2004).

204 Al-Soud, W. A. & Rådström, P. Purification and characterization of PCR-inhibitory components in blood cells. *Journal of clinical microbiology* **39**, 485-493 (2001).

205 Wilson, I. G. Inhibition and facilitation of nucleic acid amplification. *Applied and Environmental Microbiology* **63**, 3741-3751 (1997).

206 Procopio, F. A. *et al.* A Novel Assay to Measure the Magnitude of the Inducible Viral Reservoir in HIV-infected Individuals. *EBioMedicine* **2**, 874-883, doi:10.1016/j.ebiom.2015.06.019 (2015).

207 Perez, V. L. *et al.* An HIV-1-infected T cell clone defective in IL-2 production and Ca2+ mobilization after CD3 stimulation. *Journal of immunology (Baltimore, Md. : 1950)* **147**, 3145-3148 (1991).

208 Folks, T. M. *et al.* Tumor necrosis factor alpha induces expression of human immunodeficiency virus in a chronically infected T-cell clone. *Proceedings of the National Academy of Sciences* **86**, 2365-2368 (1989).

209 Clouse, K. A. *et al.* Monokine regulation of human immunodeficiency virus-1 expression in a chronically infected human T cell clone. *Journal of immunology (Baltimore, Md. : 1950)* **142**, 431-438 (1989).

210 Folks, T. M. *et al.* Tumor necrosis factor alpha induces expression of human immunodeficiency virus in a chronically infected T-cell clone. *Proc Natl Acad Sci U S A* **86**, 2365-2368 (1989).

211 Autran, B. *et al.* Positive effects of combined antiretroviral therapy on CD4+ T cell homeostasis and function in advanced HIV disease. *Science* **277**, 112-116 (1997).

212 Komanduri, K. V. *et al.* Restoration of cytomegalovirus-specific CD4+ T-lymphocyte responses after ganciclovir and highly active antiretroviral therapy in individuals infected with HIV-1. *Nat Med* **4**, 953-956 (1998).

213 Lederman, M. M. *et al.* Immunologic responses associated with 12 weeks of combination antiretroviral therapy consisting of zidovudine, lamivudine, and ritonavir: results of AIDS Clinical Trials Group Protocol 315. *J Infect Dis* **178**, 70-79 (1998).

214 Bailey, J. R. *et al.* Residual Human Immunodeficiency Virus Type 1 Viremia in Some Patients on Antiretroviral Therapy Is Dominated by a Small Number of Invariant Clones Rarely Found in Circulating CD4(+) T Cells. *Journal of Virology* **80**, 6441-6457, doi:10.1128/JVI.00591-06 (2006).

215 Siliciano, J. D. *et al.* Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat Med* **9**, 727-728, doi:10.1038/nm880 (2003).

216   Shirakawa, K., Chavez, L., Hakre, S., Calvanese, V. & Verdin, E. Reactivation of latent HIV by histone deacetylase inhibitors. *Trends in Microbiology* **21**, 277-285, doi:https://doi.org/10.1016/j.tim.2013.02.005 (2013).

217   Archin, N. M. & Margolis, D. M. Emerging strategies to deplete the HIV reservoir. *Current opinion in infectious diseases* **27**, 29-35, doi:10.1097/QCO.0000000000000026 (2014).

218   Kabamba-Mukadi, B. *et al.* Human immunodeficiency virus type 1 (HIV-1) proviral DNA load in purified CD4+ cells by LightCycler® Real-time PCR. *BMC Infectious Diseases* **5**, 15, doi:10.1186/1471-2334-5-15 (2005).

219   Ibáñez, A. *et al.* Quantification of integrated and total HIV-1 DNA after long-term highly active antiretroviral therapy in HIV-1-infected patients. *AIDS* **13**, 1045-1049 (1999).

220   Jubault, V. *et al.* High rebound of plasma and cellular HIV load after discontinuation of triple combination therapy. *Aids* **12**, 2358-2359 (1998).

221   Carr, J. M. *et al.* Development of Methods for Coordinate Measurement of Total Cell-Associated and Integrated Human Immunodeficiency Virus Type 1 (HIV-1) DNA Forms in Routine Clinical Samples: Levels Are Not Associated with Clinical Parameters, but Low Levels of Integrated HIV-1 DNA May Be Prognostic for Continued Successful Therapy. *Journal of Clinical Microbiology* **45**, 1288-1297, doi:10.1128/JCM.01926-06 (2007).

222   Bruner, K. M. *et al.* Defective proviruses rapidly accumulate during acute HIV-1 infection. **22**, 1043-1049, doi:10.1038/nm.4156 (2016).

223   Baba, M., Miyake, H., Okamoto, M., Iizawa, Y. & Okonogi, K. Establishment of a CCR5-expressing T-lymphoblastoid cell line highly susceptible to R5 HIV type 1. *AIDS research and human retroviruses* **16**, 935-941, doi:10.1089/08892220050058344 (2000).

224   Platt, E. J., Bilska, M., Kozak, S. L., Kabat, D. & Montefiori, D. C. Evidence that ecotropic murine leukemia virus contamination in TZM-bl cells does not affect the outcome of neutralizing antibody assays with human immunodeficiency virus type 1. *J Virol* **83**, 8289-8292, doi:10.1128/jvi.00709-09 (2009).

225   Wei, X. *et al.* Emergence of Resistant Human Immunodeficiency Virus Type 1 in Patients Receiving Fusion Inhibitor (T-20) Monotherapy. *Antimicrobial Agents and Chemotherapy* **46**, 1896-1905, doi:10.1128/AAC.46.6.1896-1905.2002 (2002).

226   Platt, E. J., Wehrly, K., Kuhmann, S. E., Chesebro, B. & Kabat, D. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. *J Virol* **72**, 2855-2864 (1998).

227   Pezzi, H. M., Berry, S. M., Beebe, D. J. & Striker, R. RNA-mediated TILDA for improved cell capacity and enhanced detection of multiply-spliced HIV RNA. *Integrative biology : quantitative biosciences from nano to macro* **9**, 876-884, doi:10.1039/c7ib00112f (2017).

228   Rouet, F. *et al.* Impact of HIV-1 genetic diversity on plasma HIV-1 RNA Quantification: usefulness of the Agence Nationale de Recherches sur le SIDA second-generation long terminal repeat-based real-time reverse transcriptase polymerase chain reaction test. *Journal of acquired immune deficiency syndromes (1999)* **45**, 380-388, doi:10.1097/QAI.0b013e3180640cf5 (2007).

229   Avettand-Fenoel, V. *et al.* LTR real-time PCR for HIV-1 DNA quantitation in blood cells for early diagnosis in infants born to seropositive mothers treated in HAART area (ANRS CO 01). *Journal of medical virology* **81**, 217-223, doi:10.1002/jmv.21390 (2009).