A global live cell barcoding approach for multiplexed mass cytometry profiling of mouse tumors

Soren Charmsaz, …, Elizabeth Jaffee, Won Jin Ho


In-Press Preview

Technical Advance | In-Press Preview | Immunology | Oncology

Graphical abstract

Find the latest version:

https://jci.me/143283/pdf
A Global Live Cell Barcoding Approach for Multiplexed Mass Cytometry Profiling of Mouse Tumors

Authors: Soren Charmsaz*,1, Nicole Gross*,1, Elizabeth Jaffee1, Won Jin Ho1
*Co-first authors

Affiliation: 1Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

Corresponding Author:
Won Jin Ho, M.D.
1650 Orleans St, CRB1 488
Baltimore, MD 21287
Email: wjho@jhmi.edu

Running Title: A Live Cell Barcoding Approach for CyTOF Profiling of Mouse Tumors

Keywords: CyTOF, mass cytometry, barcoding, CD29, CD98

Acknowledgements: Funding was provided by Emerson Collective Cancer Research Fund (640183).

Conflicts of interest: W.H. reports co-inventorship of patents with potential for receiving royalties from Rodeo Therapeutics, receiving research grant from Sanofi, and consulting for Exelixis. E.M.J. reports receiving a commercial research grant from Bristol-Myers Squibb, Aduro Biotech, and Amgen, has ownership interest (including stock, patents, etc.) in Aduro Biotech, and is a consultant/advisory board member for CStone, Dragonfly, Genocea, and Adaptive Biotechnologies.
ABSTRACT

With the advent of cancer immunology, mass cytometry has been increasingly employed to characterize the responses to cancer therapies and the tumor microenvironment (TME). One of its most notable applications is efficient multiplexing of samples into batches by dedicating a number of metal isotope channels to barcodes, enabling robust data acquisition and analysis. Barcoding is most effective when markers are present in all cells of interest. While CD45 has been shown to be a reliable marker for barcoding all immune cells in a given sample, a strategy to reliably barcode mouse cancer cells has not been demonstrated. To this end, we identified CD29 and CD98 as markers widely expressed by commonly used mouse cancer cell lines. We conjugated anti-CD29 and anti-CD98 antibodies to cadmium or indium metals and validated their utility in 10-plex barcoding of live cells. Finally, we established a novel barcoding system incorporating the combination of CD29, CD98, and CD45 to multiplex ten tumors from subcutaneous MC38 and KPC tumor models, while successfully recapitulating the known contrast in the PD1-PDL1 axis between the two models. The ability to barcode tumor cells along with immune cells empowers the interrogation of the tumor-immune interactions in mouse TME studies.
INTRODUCTION

With remarkable progress in cancer immunology, mass cytometry or Cytometry by Time-of-Flight (CyTOF) has increasingly contributed to the understanding of the tumor microenvironment (TME)(1, 2). CyTOF is a technique that allows multiparametric analysis of single cells using metal tagged antibodies(3). Although similar to flow cytometry, CyTOF employs the use of antibodies conjugated to heavy metal ions rather than fluorophores. This strategy overcomes problematic spectra overlap inherent to fluorophores, allowing the parametrization of 40+ markers per cell(4). CyTOF also obviates technical vulnerabilities associated with tandem fluorophores and issues related to autofluorescence, which often complicate the flow cytometry applications in studying epithelial/cancer cells, myeloid cells, abundance of debris, or intracellular markers(4). Importantly, due to the wider range of channels available for assaying, additional channels may be assigned to barcoding to allow multiplexing of samples for batched data acquisition. Barcoding improves throughout of data acquisition, reduces sample-to-sample variations, and provides a much more robust downstream analysis(5).

To barcode each sample effectively, it is critical to select markers that are present in all of the cells of interest. The first demonstration of multiplexed CyTOF by barcoding utilized Maleimido-monoamide-DOTA (m-DOTA)(5). M-DOTA is a molecule that allows for the chelation of lanthanide metals to cellular thiol groups that is taken up by the cells(5). Naturally, this method requires cell permeabilization prior to barcoding, which was performed with methanol. As some epitopes are sensitive to alcohol(6), permeabilization with methanol requires surface staining prior to barcoding. Alternatively, barcoding with transient partial permeabilization using saponin was demonstrated, allowing barcoding to be performed prior to surface staining(6). However, methods that depend on cell permeabilization also necessitate preceding cell fixation with paraformaldehyde. Fixation methods also disrupt the assayability of some markers, which limit barcoding applications. Thus, more recently, “live cell” barcoding strategies, in which surface markers are utilized, have been employed. Given the prevalent application of CyTOF methods in understanding immunological processes, a leukocyte common antigen CD45 was shown to be a reliable barcoding antigen to mark all live immune cells in a given sample(7, 8). However, to obtain an expanded view of the TME (and for any other applications that do not focus on immune cells), it is important to be able to also barcode non-immune cells. To this end, it was found that beta-2-microglobulin (B2M) and CD298, which mark MHC class I molecules and a subunit of sodium-potassium ATPase, respectively, are suitable for barcoding live human cells(9). While B2M and CD298 are useful for human samples, there are no validated markers that are robustly expressed by commonly used mouse cancer cell lines, e.g. MC38, CT26, B16, KPC, and Panc02, which are critical components of many syngeneic immunocompetent studies. Furthermore, there are presently no reliable antibodies for mouse CD298. Herein, to address this problem, we present a newly developed live cell barcoding approach for the multiplexing of mouse cancer cell lines and singly dissociated cells from mouse tumors.

RESULTS

CD29 and CD98 identified as robust markers for labeling mouse cancer cell lines. In order to barcode all cells originating from tumors in addition to immune cells, we needed to identify markers that are widely expressed on the surface of the cancer cell lines. To find potential candidate markers, we utilized the Cell Surface Protein Atlas, which is an interactive database developed using Cell Surface Capture (CSC) technology and mass spectrometry to generate cellular surfaceome snapshots(10) (Supplementary Fig. 1). Using this resource, we found CD29 and CD98 as potential candidates as they are detectable in all of 31 different mouse cell lines (which includes immune cells, embryonal stem cells, glial cells, adipocytes, myocardial cells) within the database (Fig.1, Supplementary Table S1). As a point of reference, we also examined the results for CD45, and as expected, CD45 was detectable in only 7 of the 31 different mouse cell lines within the database (Fig. 1). CD29 is a surface marker involved in cell adhesion(11), while CD98 is a transmembrane protein involved in amino acid transport(12) and both are highly expressed also in mouse CD45- stromal cell types (ImmGen, Supplementary Table S2). CD29 and CD98 were particularly feasible markers to be tested given the commercial availability of purified monoclonal anti-mouse antibodies. Other notable top hits included CD298, CD54, and CD47 (Supplementary Table S1). In addition to CD29 and CD98, we also decided to test the following antigens...
due to their known functions that make them candidates to be widely expressed on the cancer cell lines: CD47, beta-2-microglobulin (B2M), and EPCAM. CD47 is a cell surface receptor that regulates macrophages from attacking healthy cells and is known to be overexpressed in cancers(13); B2M is a marker on MHC class 1 molecules(9); and EPCAM is a transmembrane receptor on epithelial cells that aids cell adhesion(14). CD45, Isotype hamster IgG, and isotype rat IgG2K were also tested as controls. In order to validate the utility of these candidates as barcoding markers, we tested their expression in the following commonly used cancer cell lines: MC38 (colon carcinoma)(15, 16), CT26 (colon carcinoma)(16, 17), B16 (melanoma)(16, 18, 19), Panc02 (pancreatic ductal adenocarcinoma)(16, 20), and KPC (pancreatic ductal adenocarcinoma)(21) (Supplementary Table S3). These antigens were also tested on splenocyte samples to examine their relative expression in immune cells. CD29 and CD98 were tested separately, with CD29 first being compared against the aforementioned candidate markers (Fig. 2A) and CD98 subsequently being tested in reference to CD29 (Fig. 2B). Of all of the antigens tested, CD29 and CD98 were the most broadly expressed antigens among cancer cell lines with CD98 having slightly higher expression than CD29 in MC38, B16, CT26, and Panc02 but almost equal expression in KPC (Fig. 2A, 2B). CD47 showed the next highest expression among cancer cell lines while the remaining antigens had relatively similar levels of expression with the exception of EPCAM having a considerable level of expression in Panc02 (Fig. 2A). CD29 and CD98 expression was reliably present in splenocytes but was less intense compared to the expression of CD45, demonstrating that CD29 and CD98 may not be as robust as CD45 for barcoding immune cells (Fig. 2A, 2B). To note, all cancer cell lines had much higher levels of autofluorescence compared to splenocytes. After finding that CD29 and CD98 are highly expressed in the tested mouse cancer cell lines, anti-CD29 and anti-CD98 antibodies were conjugated to three different cadmium isotopes (112Cd, 114Cd, and 116Cd) and two different indium isotopes (113In and 115In) (Fig. 2C). Finally, 116Cd-CD29 and 112Cd-CD98, in addition to 114Cd-CD45 as a control, were utilized to confirm that these antigens are robustly assayable in our cell lines via CyTOF (Supplementary Fig. 2).

**Validation of CD29 and CD98-based barcoding systems.** Having conjugated anti-CD29 and anti-CD98 antibodies and validated their stainability with CyTOF, we next tested the novel barcoding methods using the previously described 112Cd-, 113In-, 114Cd-, 115In-, and 116Cd-conjugated anti-CD29 and anti-CD98 antibodies. We employed the use of a 5-choose-3 barcoding strategy to yield ten unique barcodes for both CD29 and CD98 (Fig. 3A). 10 wells of MC38 and 10 wells of KPC cancer cell lines were stained with both anti-CD29 and anti-CD98 separately at a concentration of 0.25μg/100μl based on testing at several dilutions (Supplementary Fig. 3). After each well was stained with a unique barcode, MC38 and KPC cells were multiplexed into four 10-plex batches, one for MC38 with anti-CD29 barcodes, one for KPC with anti-CD29 barcodes, one for MC38 with anti-CD98 barcodes, and one for KPC with anti-CD98 barcodes. Upon data acquisition, the resulting cell events were manually debarcoded by hierarchal gating for cells that are positive for the three barcode channels and negative for the remaining two barcodes. This strategy was used for both anti-CD29 and anti-CD98 barcodes (representative results for MC38, Fig. 3B, and for KPC, Supplementary Fig. 4). Event counts revealed that the anti-CD29 barcoding strategy successfully captured 94.65% of all MC38 cells and 90.42% of all KPC cells, while the anti-CD98 barcoding strategy successfully captured 89.53% of all MC38 cells and 93.63% of all KPC cells. Furthermore, both anti-CD29 and anti-CD98 barcodes accounted for roughly 10% of the stained cells for both MC38 and KPC, demonstrating effective 10-plex runs based on CD29 and CD98 (Fig. 3C, 3D). We also employed a previously published algorithm(22) for debarcoding of the samples to see whether the barcoding strategy is robust to automation. Applying a separation threshold of 0.3, 77% and 84% of MC38 and KPC cells, respectively, could be reliably assigned to each barcode using anti-CD29, and 76% and 86% of MC38 and KPC cells, respectively, using anti-CD98. (Supplementary Fig. 5).

**Live cell barcoding system allows multiplexed mouse tumor profiling and recapitulates the presence of PD1-PDL1 axis in MC38 tumors.** Finally, to demonstrate the use of this barcoding system on mouse tumors, we used a subcutaneously inoculated syngeneic immunocompetent mouse model using KPC and MC38 cells. Tumors were harvested and enzymatically dissociated into single cells. To note, single cell dissociation of subcutaneous tumors involves Percoll separation of the enzymatically dissociated samples to enrich for immune cells by filtering out debris, dead cells, and red blood cells. We
multiplexed 10 single-cell dissociated tumor samples (5 KPC samples and 5 MC38 samples) using the combination of anti-CD29, anti-CD98, and anti-CD45 markers on the same channels to barcode based on the same 5-choose-3 scheme for each barcoding marker (Fig. 4A). We performed CyTOF using a panel of 20 markers to profile the tumors (Supplementary Table S4), and a clustering analysis of the resulting CyTOF data using FlowSOM algorithm identified 30 metaclusters that could be annotated into 21 final clusters (Fig. 4B, 4C, Supplementary Table S5). Using our barcoding strategy, we could observe several differences in the TME between MC38 and KPC tumors both as a percentage of live cells and as a percentage of CD45 cells, enabling analysis separately for the non-immune and immune compartments. Within the non-immune compartment, the CD45- or CD45dim cells, the vast majority of which are presumably cancer cells, belonged to four major phenotypically distinct clusters. The most abundant CD45- cell cluster was substantially represented in both KPC and MC38 tumors while the second most abundant CD45- cell cluster was primarily observed in MC38 tumors (Fig. 4D). Within the immune compartment, KPC tumors had significantly more infiltrating lymphoid cells (B cells, T cells, NK cells), and MC38 tumors in general had more tumor-associated macrophages (TAM) but also specifically TAM clusters 1 and 2. MC38 tumors also exhibited higher percentages of monocytic myeloid derived suppressor cells (M-MDSCs), whereas KPC tumors had greater abundance of granulocytic MDSC (G-MDSCs) (Fig. 4D, 4E, Supplementary Fig. 6). When we examined in detail the functional phenotypes of these clusters that were differentially abundant, we found MC38 tumors exhibited striking predominance of PDL1-expressing clusters in both immune and non-immune compartments (Tumor 2, TAM1, TAM2, M-MDSC1, M-MDSC2) (Fig. 4, 5A). While there was an abundance of PD1+ CD8 T cells in the KPC tumors also (Fig. 4D, 4E), abundance of PDL1-expressing cells were markedly lower (Fig. 5A), suggesting that the immune suppression within the subcutaneous KPC tumors may not be heavily dependent on PD1-PDL1 axis. Furthermore, when we compared the expression levels of PDL1 at the per-cell level within each cell type cluster, i.e. cell states instead of cell type abundances, the top two CD45- cell clusters, annotated as Tumor 1 and Tumor 2, showed substantial and statistically significant differences in PDL1 expression, again showing that MC38 tumors have higher PDL1 expression within the CD45- compartment (Fig. 5B). These results are consistent with previous observations including our own showing that subcutaneous MC38 tumors respond well to anti-PD1 therapy in both early- and late-treatment paradigms, whereas subcutaneous KPC tumors do not(23–26). In particular, our analysis of the PD1-PDL1 axis by simultaneously profiling both the non-immune and immune compartments further supports the previously established role of PDL1 expression in both the tumor cells and TAMs in the immune escape of MC38 tumors(25, 27–29).

**DISCUSSION**

In this study, we were able to establish both CD29 and CD98 as useful markers to barcode commonly used mouse cancer cell lines and demonstrate its utility in a 10-plex barcoding system in live mouse tumor samples. Whereas anti-CD45 alone only offers staining of immune cells, the combination of anti-CD29, anti-CD98, and anti-CD45 antibodies for barcoding is a novel strategy that allows for the staining of mouse immune and tumor cells in any given sample. Moreover, the dual staining of tumor cells with both anti-CD29 and anti-CD98 intuitively provides additional robustness for multiple reasons: (i) the expression of either of the two markers may change depending on the biological context, (ii) higher technical fidelity in the staining process, and (iii) having more than one option in cell lines that we did not test. We have also demonstrated the ability of the 5-choose-3 barcoding scheme to successfully debarcode multiplexed samples back into individual samples both in manual and automated methods. Our barcoding approach is especially valuable when profiling multiple different conditions and comparing across tumor models that are comprised of varying non-immune and immune components within the TME. It also provides the ability to look at cellular compositions as a percentage of total cells, which can be less predictable depending on the dissociation protocol used.

Importantly, our application of cadmium metals for barcoding, which has only been recently made available commercially, in conjunction with indium channels were particularly advantageous since despite the relatively lower mass-signal characteristics of these channels (112-116), they were used to target highly abundant markers and do not interfere with the signal of the channels dedicated to assaying the markers of interest (140+)(9). Also, by increasing the number of total metals used for barcoding with the
addition of another channel, e.g. 89Y or 194-198Pt, we could easily expand the current 10-plex scheme to 20-plex (6-choose-3) or theoretically more as needed. By increasing the number of samples multiplexed and leveraging the ability to multiplex samples upstream of any staining step without fixation or permeabilization for both immune and non-immune/tumor cells conserves reagents, decreases risk of batch variability, and allows us to rely on having fewer cells per sample. Lastly, CD29 and CD98 may be used as a barcoding target for proteogenomics platforms, e.g. CITE-seq, and membrane marker in emerging multiplexed imaging techniques to improve cell segmentation in mouse models. Although these applications will require further experimental validation, our study establishes the rationale.

Our study has the following limitations. First, the metal conjugated antibodies were conjugated across the span of several days, leading to variation in staining quality for each antibody. This lot-to-lot variation is expected and requires testing each time. The second limitation is that CD29 and CD98 expression is not inherently tumor-cell specific, and they may or may not be present in other non-immune cells. Thus, further extension of the application by including other phenotyping markers to identify less common cell types present in the CD45- compartment such as podoplanin (for stromal cells) and CD31 (endothelial cells) would be valuable. Also, though we have tested and shown that both CD29 and CD98 are robust in several cancer cell lines, barcoding cancer cell lines not tested in this study will require independent verification in those cell lines.

In summary, we have shown the utility of CD29 and CD98, in addition to CD45, as robust markers for barcoding in studies involving mouse tumors to permit interrogation of the TME beyond the immune compartment. This study expands the applicability of suspension mass cytometry and broadens the scope of inquiry into cancer-stroma interactions and therapy responses, especially in the context of targeting the TME.

METHODS

Cell Surface Protein Atlas: In order to identify CD29 and CD98 as pan-tumor markers, we utilized the Cell Surface Protein Atlas(10). First, we used the explore option for the downloadable “Matrix of all proteins and their detection in the different cell types” to open the relevant database. The cell types included in the database are primary cells from neural, myocardial, immune, pancreatic, glial among other types. From there, filters were used to select for the following: mouse, high confidence level, and CD markers. After sorting the results in descending order for “count of detection in different cell types,” the results in Figure 1 and Supplementary Table S1 could be generated.

ImmGen: To validate the CD29 and CD98 expression in non-immune cell types using another searchable external database, we used ImmGen(30) to explore at the RNAseq expression of Itgb1 (CD29) in stromal mouse cells available in the database. Among the reference populations, three were identified as CD45 negative. Expression of Itgb1 (CD29), Slc3a2 (CD98), Ptpc (CD45), and Pdpm in all reference populations were tabulated separately in Supplementary Table S2.

Antibodies: The mass cytometry antibodies and isotopes that were used for the duration of the experiment are shown in figure (Fig. 2C). Conjugation of the listed antibodies was carried out using Maxpar conjugation/metal labeling kits. The process was performed following the manufacturer’s protocol associated with the respective kits. To begin, 50Kda ultra filtration columns (Amicon) were used to perform a buffer exchange protocol on purified antibodies (Biolegend) that were then partially reduced with 4mM TCEP (Thermo Scientific). Concurrently, Polymers were loaded individually with isotopically enriched metals, 112Cd (Fluidigm), 113In (Trace Sciences), 114Cd (Fluidigm), 115In (Sigma), 116Cd (Fluidigm). Metal-loaded polymers yielded from this process were then conjugated to their respective antibodies. All conjugated antibodies were then subjected through a series of washing steps using a wash buffer before their concentrations in a wash buffer were quantified using Nanodrop. Final antibody concentrates were then diluted in their respective stabilization buffer. MCP9 polymer conjugated antibodies were diluted in HRP-Protector (Boca Scientific) while X8 polymer conjugated antibodies were diluted in a stabilization buffer (Candor) containing 0.3% sodium azide. Each unique antibody was then tested against 4 serial dilutions (0.0625μg/100μl, 0.125μg/100μl, 0.25μg/100μl, and 0.5μg/100μl) to ensure sufficient signal upon staining.
Cell Culture: All cell lines were thawed and in culture for over a week and 2 passages were conducted for each cell line before experimental use. MC38 Cells (Kerafast) were maintained in DMEM-based media (Gibco) containing 10% FBS (Gemini), 1% L-glutamine (Gibco), 1% penicillin/streptomycin (Gibco), 1% HEPES (Gibco), 1% sodium pyruvate (Sigma), and 1% non-essential amino acids (Gibco) in 5% CO2 at 37C. B16-F10 (ATCC) cells were maintained in DMEM-based media containing 10% FBS, 1% penicillin/streptomycin, 1% sodium pyruvate, and 1% non-essential amino acids in 5% CO2 at 37C. CT26 (ATCC) cells were maintained in RPMI 1640 with glutamine (Gibco) containing 10% FBS, 1% penicillin/streptomycin, 1% sodium pyruvate, 1% non-essential amino acids in 5% CO2 at 37C. Panc02 cells established and authenticated as previously described(20, 31) were maintained in DMEM-based media containing 10% FBS, 1% penicillin/streptomycin, and 1% L-glutamine (Gibco) in 10% CO2 at 37C.

MC38 and KPC tumor model: Experiment and euthanasia were performed in accordance with Johns Hopkins IACUC approved protocols. Mice were purchased from Jackson Laboratory and allowed to acclimate to facility environment for 1 week prior to experimental use. MC38 and KPC Cells were then subcutaneously injected in a syngeneic immunocompetent C57Bl/6J background. Cells were injected at a concentration of 5x10^6 cells per injection volume of 100µl into the right hind limb of 6-week old female C57Bl/6 mice and were left to grow for 15 days. On day 15, mice were euthanized, and tumors were harvested and placed in incomplete media on ice before being enzymatically dissociated. Tumors were transferred to gentleMACS C tubes (Miltenyi) to be enzymatically dissociated using a mouse Tumor Dissociation Kit (Miltenyi) enzymatic mix in combination with the use of a gentleMACS Octo Dissociator (Miltenyi), using the built-in heated mouse tumor protocol. The resulting homogenate was then transferred to 50ml conical tubes containing complete RPMI media through 100µm strainers in order to both filter and quench samples. Homogenates were then subject through another cleaning step that involved gradient centrifugation using Percoll (GE Healthcare). Homogenates were resuspended in 40% percoll and underlaid with 80% and then centrifuged to separate single cells from debris, dead cells, and red blood cells. Resulting single cells were then finally washed with complete RPMI and pellets of single cells were obtained.

Flow cytometry: Fluorescent flow cytometry on samples was conducted with the following 2 panels. Panel 1 consisted of: Anti-CD29 PE (HMb1-1, Biolegend, 1:100), Anti-CD47 PE (Miap301, Biolegend, 1:100), Anti-B2M PE (A16041A, Biolegend, 1:100), Anti-EPCAM PE (G8.8, Biolegend, 1:100), Anti-CD45 PE (30-F11, Biolegend, 1:100), Isotype Hamster IgG PE Antibody (HTK888, Biolegend, 1:100) and Isotype Rat IgG2a k PE (RTK2758, Biolegend, 1:100). Panel 2 consisted of: Anti-CD98 (RL388, Biolegend, 1:100), Anti-CD29 PE (HMb1-1, Biolegend, 1:100), Anti-CD45 PE (30-F11, Biolegend, 1:100), Isotype Rat IgG2a k PE (RTK2758, Biolegend, 1:100), and Isotype Hamster IgG PE Antibody (HTK888, Biolegend, 1:100). Control splenocytes were harvested from a healthy 8-week Balb/cJ female mouse (Jackson). Cancer cells in flasks were trypsinized and counted using a trypan blue dilution and a hemocytometer. 1 million cells per well were plated and washed. Following an initial wash, cells were then Fc blocked for 10 minutes at 4C. Samples were then stained at a concentration of 0.2µg/100µl and shielded from light for 30 minutes at 4C. Before data acquisition on a Beckman CytoFLEX, FACs buffer was used to wash the cells a total of two times. Supervised gating analysis was then performed on Cytobank.

Mass cytometry staining and data acquisition: 8 wells of MC38 cells, 8 wells of splenocytes, followed by 10 wells of MC38 cells, and finally 10 wells of KPC cells were plated on a 96 well plate at 1x10^5 cells per well. On a separate 96 well plate, 10 wells of MC38 and 10 wells of KPC were plated at 5x10^5 cells per well. Initially, plated cells were washed with PBS and 2mM EDTA. This is then followed by a 5-minute incubation in palladium chloride that is dissolved in DMSO (Sigma) and then diluted in PBS to a concentration of 500nM. This incubation at room temperature is used to mark viability. Following the 5
minutes, RPMI containing 10% FBS is added to the cells in order to quench any residual palladium(9). All wells were then stained with a unique metal barcode within their sample type for 25 minutes at room temperature followed by a series of washes with a cell staining buffer (Fluidigm). The 16 single samples (8 wells of MC38 and 8 wells of splenocytes) were then transferred into tubes via 40um filters while both sets of 10 wells of MC38 cells and both sets of 10 wells of KPC cells are 10-plexed (batched) and then transferred into tubes via 40um filters. After a series of washes, the cells were then stored in 1% methanol-free formaldehyde diluted in PBS (Thermo Scientific) until the day of data collection.

Single cell dissociated samples from mouse MC38 and KPC tumors were plated into 10 respective wells on a 96 well plate at 1.5x10^6 cells per well. Initially, plated cells were washed with PBS and 2mM EDTA. This is then followed by viability staining as above. The 10 wells were then stained with a unique metal barcode to differentiate between samples for 25 minutes at room temperature followed by a series of washes with a cell staining buffer (Fluidigm). The 5 samples were then batched and transferred into a 10-plexed tube using a 40um filter. The multiplexed tube was then blocked with 1ug anti-mouse Fc block (BD Biosciences) for 10 minutes at room temperature followed by a surface staining cocktail of antibodies (Supplementary Table S4) performed in cell staining buffer for 30 minutes at room temperature. After a series of washes, the cells were then stored in 1% methanol-free formaldehyde diluted in PBS (Thermo Scientific) until the day of data collection.

On the day of data collection, before samples are run, samples are stained with rhodium (Fluidigm) diluted at 1:500 in Maxpar Fix and Perm Buffer (Fluidigm) for 45 minutes at room temperature. Samples were run, events, and mass cytometry data were acquired on a Helios™ mass cytometer (Fluidigm) at the University of Maryland School of Medicine Center for Innovative Biomedical Resources (CIBR) Flow Cytometry and Mass Cytometry Core Facility, Baltimore, Maryland.

**Mass cytometry data processing and analysis:** All CyTOF analysis was performed as previously described(24, 32). Briefly, acquired data was randomized and beads were normalized then removed using CyTOF software (Fluidigm) v6.7. Gating based on Rh+ intensity (DNA-based cell ID) and the event length parameter was used to identify cell events, and whenever used, dead cells were gated out based on palladium intensity. Debarcoding of the preprocessed batches were manually performed by hierarchal gating using Cytobank (v7.3.0) or FlowJo (BD, v10.6.1). Automated debarcoding was performed using the Single Cell Debarcoder(22) within the `premessa` package (https://github.com/ParkerICI/premessa). Clustering analysis was performed using FlowSOM(33) algorithm and visualized by UMAP(34) in R (v3.6.2), based on a modified pipeline from a prior report by Nowicka et al(35). Dataset was clustered into 30 metaclusters that were annotated into major immune cell types based on canonical markers, e.g. CD45+CD3-CD19+B220+ representing B cells. Further subtyping of immune cell types were carried out based on key functional markers, e.g. PD-L1, MHC-II. Some metaclusters shared common marker expression profiles and were merged into one final annotated cluster (Supplementary Table S5).

**Statistics:** To analyze differential abundances of the clusters, we used edgeR as previously published(35, 36) to fit models and performs moderated tests at the cluster level while sharing information on variance across all samples. Raw p-values were adjusted by false-discovery rate.

**Figure Generation:** Figure 1 was screen captured from the user-interface Cell Surface Protein Atlas. Figure 2A, 2B, 3B, were acquired from Cytobank. Figure 2C was created on Biorender, Figure 3A was created in Microsoft Excel, Figure 3C and 3D were generated in GraphPad PRISM 8 (v8.3.1) using data originating from Cytobank, Figure 4A was created in Microsoft Excel, and Figure 4B, 4C, 4D, 4E were generated in R. Any figure arrangement was done in Inkscape (v1.0 for Windows or v1.0.0rc1 for MacOS).

**Study approval:** Experiments were performed in accordance with protocol approved by the institutional Animal Care and Use Committee.
AUTHOR CONTRIBUTIONS

SC and NG designed and performed experiments, analyzed data, wrote the manuscript, and revised the manuscript. EMJ analyzed data, revised manuscript, and provided funding. WJH conceived the study, designed experiments, analyzed data, wrote the manuscript, revised the manuscript, and provided funding. Order of co-first authors was determined based alphabetical order of last names.
REFERENCES


**Figure 1.** Potential surface markers for live barcoding of mouse cancer cell lines. The Cell Surface Protein Atlas utilizes Cell Surface Capture (CSC) technology to generate the above cellular surfaceome snapshot. The database shows protein expression in various cell types and has filters that allow results to be narrowed. The following filters led to the identification of CD29 and CD98 as top hits in mouse cell lines: mouse, high confidence level, CD markers, and sorted in the descending order by the number of cell types in which the antigen was detected. The results for CD45 are shown for comparison. The numbers listed under each cell line indicate the CD markers’ unique peptide counts in that cell type while the color intensity increases with the number of unique peptide counts.
Figure 2. Analysis of candidate pan-tumor markers in various cancer cell lines. (A) Flow cytometry was utilized to test the expression of the following candidate pan-tumor markers: CD29, CD47, B2M, and EPCAM in addition to CD45, isotype hamster IgG, and isotype rat IgG2K which serve as controls. (B) CD98 expression was tested separately via flow cytometry in addition to CD29, CD45, isotype hamster IgG, and isotype rat IgG2K which serve as controls. These markers were tested in several cancer cell lines: MC38, B16, KPC, CT26, and Panc02, in addition to splenocytes. Splenocytes were harvested from female BALB/cJ. CD29 and CD98 were identified as the most robustly expressed markers across cancer cell lines. (C) Anti-CD29 and anti-CD98 antibodies were covalently conjugated to the following isotopes: cadmium (112, 114, and 116 Cd) and indium (113 and 115 In).
Figure 3. Testing optimized CD29 barcoding system. (A) Using \textsuperscript{112}Cd-, \textsuperscript{113}In-, \textsuperscript{114}Cd-, \textsuperscript{115}In-, and \textsuperscript{116}Cd-conjugated anti-CD29 and anti-CD98 antibodies, ten unique barcodes were established for both anti-CD29 and anti-CD98 utilizing a 5-choose-3 barcoding strategy. MC38 and KPC cancer cell lines were stained using the aforementioned antibodies at a concentration of 0.25μg/100μl. (B) 10-plexed MC38 and 10-plexed KPC samples were debarcoded using their respective gating hierarchies back to their original individual samples. Representative gating schema for the MC38 batch stained with anti-CD29 barcodes are shown. (C) Distribution of individual anti-CD29 barcodes and (D) anti-CD98 barcodes.
yielded from the debarcoding process of MC38 and KPC samples. Both CD29 and CD98 barcoding systems resulted in a fairly even distribution of all ten barcodes in MC38 and KPC samples.

Figure 4

Figure 4. Demonstration of the barcoding strategy on mouse tumors. (A) Using 112Cd-, 113In-, 114Cd-, 115In-, and 116Cd-conjugated anti-CD45, anti-CD29, and anti-CD98 antibodies, ten unique barcodes were established based on a 5-choose-3 barcoding strategy. (B) Based on a dataset using 20 markers in a mouse immune profile mass cytometry panel, FlowSOM algorithm was then used to generate 30 metaclusters annotated into 21 final clusters. Displayed is the expression heatmap for all of the samples in the dataset. (C) Phenotype clusters identified by FlowSOM clustering shown as a UMAP plot. (D) Stacked bars represent the percentage of total live cells per mouse tumor within each cluster. (E) Stacked bars represent the percentage of CD45+ cells for each mouse tumor within each live cell cluster. Color legends on the right apply to panels B-D. Cell type abundances with significant difference between the KPC and MC38 are annotated. FDR adjusted p values by edgeR ****<0.0001, ***<0.005, **<0.01, *<0.05. Abbreviations: CD8 N, CD8+ naïve T cells; NK, natural killer cells; NKT, natural killer T cells; TAM, tumor-associated macrophages; DC, dendritic cells; G.MDSC/M.MDSC, granulocytic/monocytic myeloid-derived suppressor cells; UA, unassigned.
Figure 5. Comparison of PDL1-PD1 axis between KPC and MC38 tumor models. (A) Representative UMAPs for the tumor model, CD45, CD11b, F4/80, PDL1, and PD1 expression. (B) Violin plots comparing the expression levels of PDL1 and PD1 in specific cell type clusters with annotations of median fold difference (KPC/MC38) and FDR-adjusted p-values by edgeR: ***<0.005, *<0.05. Number of cells for each comparison of KPC vs. MC38: TAM1, 55607 vs. 519369; TAM2, 45595 vs. 122957; TAM3, 373020 vs. 26437; Tumor1, 511692 vs. 396897; Tumor2, 32692 vs. 107499; CD4, 30656 vs. 3229; CD8 N, 38463 vs. 6339; CD8 PD1+, 102213 vs. 17319; Treg 23766 vs. 7928. Abbreviations: CD8 N, naïve CD8+ T cells; TAM, tumor-associated macrophages; Treg, regulatory T cells.