PAPER

Multiplex protein detection on circulating tumor cells from liquid biopsies using imaging mass cytometry

Convergent Science Physical Oncology

PAPER

Multiplex protein detection on circulating tumor cells from liquid biopsies using imaging mass cytometry

RECEIVED
24 September 2017

REVISED
22 November 2017

ACCEPTED FOR PUBLICATION
8 December 2017

PUBLISHED
16 January 2018

Erik Gerdtsson[1], Milind Pore[1], Jana-Aletta Thiele[2], Anna Sandström Gerdtsson[1], Paymaneh D Malihi[1], Rafael Nevarez[1], Anand Kolatkar[1], Carmen Ruiz Velasco[1], Sophia Wix[1], Mohan Singh[5], Anders Carlsson[1], Amado J Zurita[3], Christopher Logothetis[3], Akil A Merchant[5], James Hicks[1,4] and Peter Kuhn[1,4,6,7,8]

[1] Bridge@USC, USC David and Dana Dornsife College of Letters, Arts, and Sciences, 3430 S Vermont Ave, TRF 114, MC3303, Los Angeles, CA, 90089-3303, United States of America
[2] Faculty of Medicine in Pilsen, Biomedical Center, Charles University in Prague, alej Svobody 76, 323 00 Pilsen, Czechia
[3] Department of Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Unit 207, Houston, TX, United States of America
[4] Biological Sciences, University of Southern California, Los Angeles, CA, United States of America
[5] Division of Hematology, Department of Medicine, USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, 1450 Biggy Street, Los Angeles, CA, United States of America
[6] Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, CA, United States of America
[7] Biomedical Engineering, University of Southern California, Los Angeles, CA, United States of America
[8] Medicine, University of Southern California, Los Angeles, CA, United States of America

E-mail: gerdtsso@usc.edu, pore@usc.edu, agerdtss@usc.edu, pmalihi@usc.edu, rnevarez@usc.edu, kolatkar@usc.edu, ruizvela@usc.edu, swix@usc.edu, ncarlsso@usc.edu, jameshic@usc.edu, pkuhn@usc.edu, Jana.A.Thiele@lfp.cuni.cz, azurita@mdanderson.org, clogothe@mdanderson.org, mohansin@usc.edu and akilmerc@usc.edu

Keywords: single cell analysis, imaging mass cytometry, CyTOF, protein biomarkers, HD-SCA, circulating tumor cells

Supplementary material for this article is available online

Abstract

Molecular analysis of circulating and disseminated tumor cells (CTCs/DTCs) has great potential as a means for continuous evaluation of prognosis and treatment efficacy in near-real time through minimally invasive liquid biopsies. To realize this potential, however, methods for molecular analysis of these rare cells must be developed and validated. Here, we describe the integration of imaging mass cytometry (IMC) using metal-labeled antibodies as implemented on the Fluidigm Hyperion Imaging System into the workflow of the previously established high definition single cell analysis (HD-SCA) assay for liquid biopsies, along with methods for image analysis and signal normalization. Using liquid biopsies from a metastatic prostate cancer case, we demonstrate that IMC can extend the reach of CTC characterization to include dozens of protein biomarkers, with the potential to understand a range of biological properties that could affect therapeutic response, metastasis and immune surveillance when coupled with simultaneous phenotyping of thousands of leukocytes.

Introduction

Enabling precision medicine in cancer care requires the ability to both deconvolute heterogeneity in the primary and metastatic tissues, and to characterize the liquid phase of the disease. In settings of chemotherapy, molecularly targeted and immune-system engaging treatment approaches, the high-resolution characterization of the disease along its evolutionary path at the time of the decision making is a primary challenge that can be addressed by the integration of quantitative single cell technologies within the solid and liquid biopsy workflows. The recognition of the complex interactions between the cancer and the immune system are particularly important drivers for the detailed analysis at the cellular, proteome and genome levels to both establish a treatment strategy and monitor its efficacy.

Since its commercial introduction in 2011, cytometry by time-of-flight (CyTOF) mass spectrometry using metal-labeled antibodies, has rapidly entered the basic and clinical research laboratory settings [1–3]. Its main distinction with respect to standard fluorescence-based cytometry is the capacity to assay the binding of 35 or more specific antibodies on each cell simultaneously. Bodenmiller and Günther *et al* have developed methods for coupling laser ablation

with CyTOF technology, enabling multiplexed image-based proteomic analysis of formalin-fixed paraffin-embedded (FFPE) tissue sections or cultured cells mounted on glass slides [4, 5]. The technology is now commercially available from Fluidigm Corporation (South San Francisco) as the Hyperion Imaging System for imaging mass cytometry, or IMC. In the IMC process, the target tissue or cell preparation is treated with a cocktail of antibodies, each labeled with a specific rare earth isotope. Regions of interest (ROI) on the slides are scanned *in situ* with a highly focused, pulsed laser, such that each pulse vaporizes a 1 μm^2 bloc of the sample and the resulting ions are introduced into the inductively coupled plasma time-of-flight mass spectrometer (ICP-TOF-MS) with helium as a carrier gas. The ion counts for each pulse can then be assembled into a protein expression image with a resolution of 1 μm^2 across the ROI allowing for limited characterization of sub-cellular localization.

In this report, we describe the integration of imaging mass cytometry (IMC) using the Hyperion instrument into the previously validated high definition single cell analysis (HD-SCA) workflow [6], with the goal to incorporate morphology, proteomics and genomics of rare single cells in a single streamlined process. The HD-SCA method was designed to identify and characterize ultra-rare (<0.0001%) cancer cells in liquid biopsies from blood and bone marrow aspirates, or high-complexity samples such as tissue 'touch preparations' on glass slides. In contrast to other CTC technologies, the HD-SCA workflow uses a direct analysis, 'no cell left behind' strategy in which the entire population of nucleated cells from a blood draw is spread on glass slides at 3 million cells per slide, stained, and imaged *in situ*. Fresh samples must be fractionated within 48 h following blood collection, which enables simple shipping from clinical sites. Candidate tumor cells are identified among the millions of leukocytes using immunofluorescent markers. In addition to the morphometric information obtained by imaging, the candidate circulating tumor cells (CTC) from blood or disseminated tumor cells (DTC) from bone marrow, or other cells of interest can subsequently be relocated and physically isolated for single cell genomic analysis [7–9]. The data derived from the HD-SCA workflow has been demonstrated to generate clinically actionable information [10]. Further, due to their minimally invasive nature, liquid biopsies can be repeated frequently during treatment to assess response and provide early indications of changes that could lead to disease progression. Despite its high sensitivity, the phenotyping of cells detected by HD-SCA has been largely limited to morphometry and four fluorescent markers. Targeted single cell proteomic analysis can complement genomic and whole plasma characterization to assign functional relevance to the molecular observations as part of an integrated signature. The longitudinal characterization of the disease using cancer cells from blood and bone marrow liquid biopsies provides insight into the evolution of the tumor, and the significance of sub-populations of circulating tumor cells present in specific environments.

Here, we present experimental procedures along with data analysis methods, specifically adapted for the HD-SCA workflow to incorporate multiplex targeted proteomic analysis on intact single cells using the Hyperion instrument. The functionality and performance of the workflow is demonstrated, including methods for validating metal conjugated antibodies for use specifically in the combined HD-SCA/IMC workflow. In addition, the robustness of the method allows for immunophenotyping of the leukocyte populations that are simultaneously imaged along with the CTCs, which is demonstrated here by gating the multi-parametric data from the HD-SCA/IMC workflow in a manner similar to cytometric sorting methods, yielding well separated sub-populations. Finally, we present combined proteo-genomic results relating the HD-SCA/IMC to the HD-SCA/genomics from the analysis of liquid biopsies from a metastatic prostate cancer patient, showing the distribution and subcellular location of characteristic prostate cancer protein markers in a clonally related population of CTCs and DTCs.

Methods

HD-SCA specimen preparation

Whole blood samples from primary and metastatic cancer patients were collected in cell-free DNA blood collection tubes (Streck, Omaha, USA) and shipped to the Kuhn laboratory at USC in temperature stabilized shippers (Standard 71, LLC, Los Angeles). Samples were processed within 24 h of collection using the HD-SCA protocol [6] and archived at −80 °C. The sample MDA-42109 was collected from a patient with metastatic prostate cancer in 2013 with informed consent and under IRB approval as previously described [8]. Normal control donor blood samples were obtained from The Scripps Research Institute, La Jolla, California, USA. All specimen collections were performed under IRB approval.

Cell lines spiked in blood

Cell lines were obtained from the American Type of Culture Collection, (ATCC, Manassas, VA, USA). Prostate cancer line LNCaP was cultured in RPMI-1640 medium and breast cancer line MDA-MB-231 was cultured in DMEM/F12 medium (Corning, Manassas, VA, USA). Cells were harvested using Versene solution (Gibco) and washed in PBS. For validation of the metal-labeled antibodies in the combined workflow, cultured cells were added at a ratio 1:10 000 to whole blood from normal blood donors and processed according to the standard HD-SCA protocol [6], including storage at −80 °C for a minimum of 24 h.

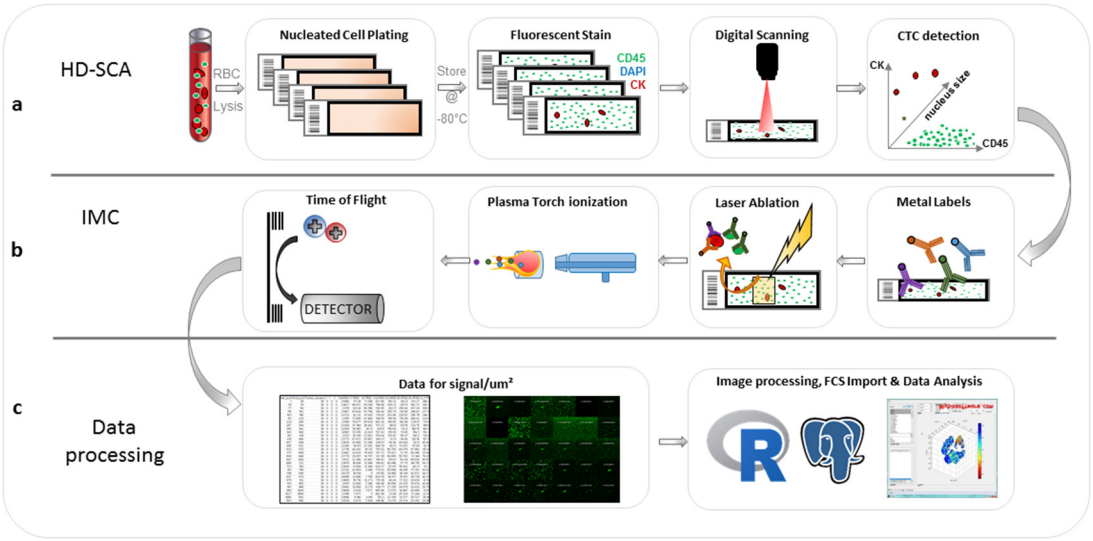

Figure 1. Schematic workflow for protein profiling of rare cells identified with the HD-SCA workflow. (a) Sample processing and fluorescent imaging; (b) IMC staining and data acquisition; (c) data processing.

HD-SCA fluorescent labelling, imaging and cell identification

To complete the HD-SCA procedure, slides were thawed, stained with DAPI (nuclear stain) and fluorescent antibodies against CD45 (leukocyte marker) and pan-cytokeratin (CK, epithelial marker), and imaged in three colors using a high-speed automated microscope scanner. As part of the normal HD-SCA image analysis, the candidate CK positive cells were identified and coordinates registered. High resolution 40× images of candidate cells were captured using a Nikon 80i microscope. Prior to staining with the metal-labeled antibody cocktail for IMC analysis, the slides were stored at 4 °C for several days to parallel normal operation of the HD-SCA protocol. Finally, cover slips were removed in PBS with minimal disruption of the cell layer. Slides were then washed with PBS to dissolve the imaging media in preparation for the metal-labeled antibody incubation. For the HD-SCA/Genomics workflow results presented in figure 3, samples were stained with 4-colors, adding AR to the DAPI, CD45 and CK panel described above.

MaxPar metal-labeled antibody staining

Metal-labeled antibodies were provided by Fluidigm, either from the standard CyTOF catalog (http://maxpar.fluidigm.com/product-catalog-metal.php) or as custom conjugates (supplementary table 1 (stacks.iop.org/CSPO/2/015002/mmedia)). Metal-labeled antibody cocktails were prepared in 0.1% Tween-20, 1% BSA in PBS as per the dilution scheme presented in supplementary table 1. All samples, prepared as described above, were first blocked with 1% BSA and 0.2 mg ml^{-1} mouse IgG Fc fragment (Thermo Scientific) in PBS for 30 min and then incubated with antibody cocktail for 1.5 h at RT, followed by washing with PBS and staining with DNA intercalator Ir-193 (Fluidigm) and cell membrane counter-stain In-115 (Fluidigm) for 30 min. Slides were again washed with PBS and rinsed with ddH20 for 5 s and dried overnight at room temperature prior to IMC analysis.

Statistical considerations and data analysis

Cell boundaries were determined by the segmentation method described in the results section, using the Bioconductor package EBImage in R [11]. The standard deviation over the mean (SDOM) per cell for each channel was defined as the difference between the average ion count for the cell and the mean signal for all the cells divided by the standard deviation of all cells within each ROI (400 × 400 μm). The noise level, used as a denominator for the S/N ratio was approximated as 1.5 × IQR (inter-quartile range) outside the upper quartile of all cells. When applicable, the limit of detection (LOD) for each marker was set as to mean + 3.3 × standard deviation (one sided 95% ∗ 2) or S/N = 3 [12]. For figure 5, the dataset was exported in a FCS format using the FlowCore package in R and the gating viSNE plots were generated with the CYT graphical software package [13] in MATLAB.

Results

The HD-SCA/IMC integrated workflow

Both the HD-SCA and the IMC were designed to operate on standard glass microscope slides in keeping with general procedures in the clinical pathology laboratory setting. The common slide format facilitated the integration of the two technologies into a single workflow viable for both basic and clinical research applications. The workflow, shown in figure 1(a), begins with processing the blood sample and spreading all nucleated cells on slides, followed by 3- or 4-color immunofluorescent staining and high-speed digital scanning to identify and image cells of

interest. The coordinates of each candidate cell are recorded and re-imaged at 40× for morphometric measurements, as previously reported [6, 8]. In preparation for the IMC (figure 1(b)), slides are re-incubated with MaxPar™ metal-labeled antibodies. The slide is placed in the IMC instrument and candidate cells are relocated using a previously determined offset between coordinates from the 40× imaging microscope and the IMC stage. A 400 × 400 μm ROI around the cell of interest undergoes laser ablation aerosolizing a 1 μm^2 area/pulse (200 Hz), followed by ionization and quantification in the CyTOF Helios instrument. Ion mass data is collected for each pulse and processed to render images for each individual channel at 1 μm resolution, where the intensity of each pixel corresponds to the ion count value. The images are analyzed using an in-house developed software application in R. Raw data and processed results are stored in a PostgreSQL database enabling export of datasets and integration with HD-SCA assay parameters (figure 1(c)).

Although the basic formats of the two assays are similar, adapting the IMC technology to the HD-SCA workflow required the development and validation of staining and analysis methods consistent with the previously validated HD-SCA protocols. The antibody preparation and staining steps were initially developed and optimized using cell lines spiked into normal donor blood samples (Methods).

For this study, a panel of metal-labeled antibodies was designed to include targets relevant to prostate cancer. The panel included PSA, PSMA and two epitopes for AR, along with epithelial markers EpCAM, E-cadherin, cancer stem cell marker CD44, EMT marker B-catenin, as well as a subset of markers for leukocyte characterization [14–16]. HLA-DR and CD14 were selected for monocytes identification along with CD3, CD8, CD4, CD45RA and CD38 to further characterize the T cell population [17]. CD66 was selected to identify mainly the neutrophil population, however, CD66 or CEA (carcinoembryonic antigen) is also expressed in in tumor cells and is known to have a tumor suppressive role in prostate cancer [18]. All catalog antibodies had been previously validated by Fluidigm for use in the CyTOF/Helios systems, but had not been validated in the IMC modality nor tested in the context of the HD-SCA workflow.

Each metal-labeled antibody was first tested for specificity and sensitivity against the prostate cancer line LNCaP, and metastatic breast cancer line MDA-MB231 (Materials and Methods and supplemental figure S1), using the leukocyte populations surrounding the candidate cells as an additional control. The unique challenge of determining the presence of the various antigens on the single candidate cell within each ROI in comparison to flow-based methods required the development of routines for quantification and scoring as described below.

Cell segmentation and statistical analysis

As the pulsed laser sequentially ablates subcellular areas across the candidate cell along with all of the ~300 leukocytes present in the defined ROI, an ion count value is reported for each isotope at each of the 160 000 acquired pulses. The resulting data matrix is stored in a text file from which false color 1 μm resolution images are generated from each array of ion count values. The range of ion count values is internally normalized within each image, usually so that the brightest pixel represents the 98th percentile of the cumulative ion count signal.

Figure 2 shows an example of the images obtained from the IMC analysis of the two spiked cell lines. In the top row, an aggregate of LNCaP cells is seen in the center of each frame. A single LNCaP cell is also captured within the ROI. The different localized signal for PSA (membrane, figure 2(e)), E-cadherin (cell junctions, figure 2(f)), and androgen receptor (AR) (translocated to nuclei upon activation, figure 2(g)), can clearly be seen. For segmentation, the DNA intercalator data (figure 2(d)) is used to identify the objects in each frame. Adaptive thresholding [11] is utilized to generate a binary image, and then applying a watershed algorithm [19] to separate objects that are touching. Finally, the membrane counterstain is used for Voronoi propagation [20]to define the cell boundaries of each object with the primary mask acting as seeds (figure 2(c)). The mask is used to quantify the signals associated with each cell for each channel, by applying the mask on the raw data array of ion count values, rather than the normalized image. The ion count signals within the boundaries of each segmented cell in the ablation field are used to generate biaxial plots showing the average signal per cell for each marker relative to a chosen leukocyte marker, typically CD45 (figures 2(m)–(o)). This quantitation leads to a heuristic scoring system for each marker as described below.

Design of a scoring system for marker classification

In practice, the quantitation of different markers is hampered by (i) subjectivity in adjusting the gain for normalization of images, (ii) a wide dynamic range of protein expression, (iii) inherent differences in detector sensitivity for the metal ions [21], (iv) experimental variation due to instrument fluctuation [22] and differences between staining batches etc, and (v) the difficulty of separating low abundant protein signals from background noise. As alluded to above, these challenges exacerbated in the setting of evaluating rare single cells, as compared to the common application of cytometry where larger populations are characterized.

Classification based on ion count alone was deemed unfeasible due to vast differences in dynamic range for each antibody-analyte pair, impeding normalization of the signals and their interpretation in heat maps. Thus, to establish an objective, semi-automated means for evaluating signals, a four level scoring system was created, with 0 being below the

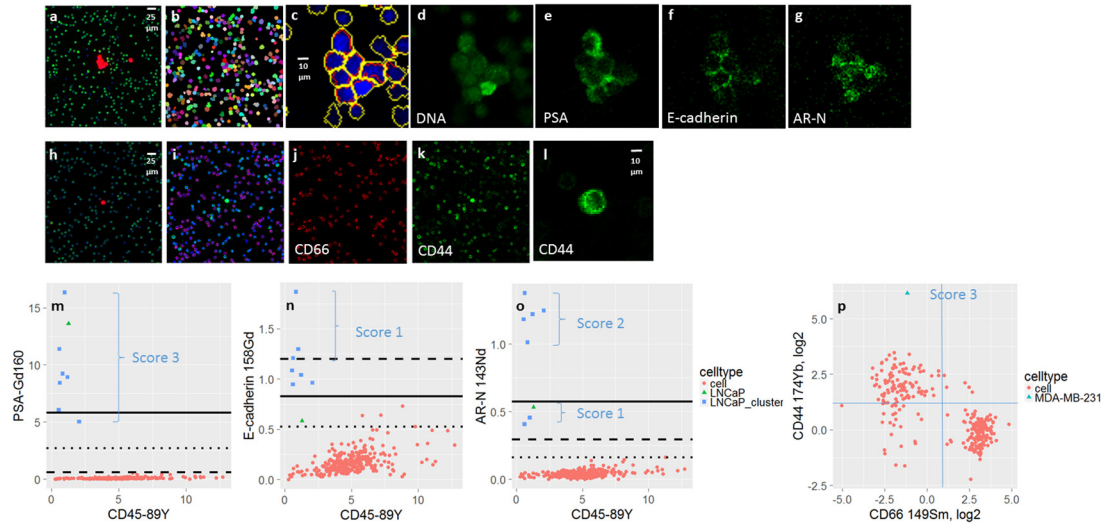

Figure 2. Illustrations of segmentation and scoring of individual cells and aggregates. (a) HD-SCA composite image of LNCaP cells spiked in normal blood containing a cluster of cells in the center of the frame along with a single cell to the right. (b) Colored segmented cells in the IMC ROI. (c) Zoomed in LNCaP cluster of cells segmented using the DNA-Intercalator signal (blue) to outline the nuclei (yellow) as primary masks acting as seeds for propagation onto the membrane image to identify the outer cell boundaries (red). (d)–(g) Images showing the signal for individual IMC channels as labeled. (h) HD-SCA composite image of a single MDA-MB-231 cell spiked in normal blood. (i) Composite image, showing CD66 (red), CD44 (green) and DNA (blue), demonstrating the mutually exclusive populations of leukocytes (j) and (k). (l) Enlarged image of the MDA-MB-231 cell stained with CD44. (m)–(o) Biaxial plots of the segmented cells from ROI in figure (b) with channels specified on the *y*-axis and CD45 on *x*-axis. The LNCaP cells are highlighted as green triangles (part of the cluster) or blue squares (single cell), with the associated score. The lines represents the estimated LOD based on either S/N (dashed), SDOM for all cells (solid) or SDOM based on the leukocytes only (dotted). (p) Biaxial plots of CD66 versus CD44 on a log2 scale of the cells present in the ROI in I along with gating lines in blue. The MDA-MB-231 cell is shown as a blue triangle, scored as a 3.

LOD, 1 exceeding LOD, and levels 2 to 3 set to discriminate the highest signal levels. The scoring system is based both on SDOM, and signal to noise (S/N) of the ion count value for the candidate cell to the surrounding leukocyte population, providing a relative measurement that is less sensitive to experimental variation than ion count alone. The LOD for each marker was set as equal to S/N = 3 or SDOM = 3.3 [17]. Cells exceeding the LOD were scored as 1+, signals with S/N of 7-20 or SDOM > 6 were scored 2+, and finally score 3+ was assigned to signals of S/N > 20 or SDOM > 12. In figures 2(m)–(p), the scores are marked in the plots along with the estimated LOD based on either S/N = 3 (dashed lines) or SDOM = 3.3.

Some of the markers tested, including CD44, CD66, and vimentin, are also present on subpopulations of leukocytes. As seen in figures 2(i)–(k), most of the cells in the ROI express either CD44 or CD66, comprising two mutually exclusive subpopulations. In such cases, manual or automated gating can be readily applied to score the cell of interest, as demonstrated for a spiked MDA-MB-231 cell shown in figure 2(p). In this case, the CD44 signal was above the gated CD44+ leukocyte population, and was given a score of 3. Additional scoring of the different markers for LNCaP and MDA-MB-231 are presented in supplemental figure 1.

Analysis of a liquid biopsy sample from a prostate cancer patient

To test the integration of IMC with the HD-SCA workflow and to demonstrate the added information provided by IMC over 4-color immunofluorescence, the combined HD-SCA/IMC workflow was performed on matched blood and bone marrow aspirate samples from a patient with metastatic castrate resistant prostate cancer (MDA-42109) collected as part of a larger study [8]. In the initial HD-SCA assay of MDA-42109 conducted in 2013, 9 CTCs ml^{-1} from blood and 207 DTCs ml^{-1} in bone marrow were identified, with 48% of the CTCs and 78% of the DTCs also scoring positive for AR expression [23]. High resolution (40×) images had been collected for all the CTCs and a subset of DTCs to verify the AR positivity and subcellular localization prior to isolating the cells for whole genome amplification using the HD-SCA genomics workflow [7]. The results showed that 41/42 of the successfully amplified cells yielded highly rearranged and closely related copy number profiles with common elements typical of metastatic prostate cancer, indicating a single cancer lineage with four distinct sub clones as shown in the heat map and representative profiles (figure 3(a)). The high degree of clonality in the CNV profiled cells provided assurance that the candidate CTCs and DTCs selected for subsequent IMC analysis were highly likely to be *bona fide* cancer cells from the same lineage. However, we note that despite sharing the same genotype, individual cells exhibit significant phenotypic variability in AR expression and subcellular location, consistent with previous results [7]. In most cases, amplification of the AR gene on chromosome X corresponds to a strong nuclear AR signal in the corresponding fluorescent

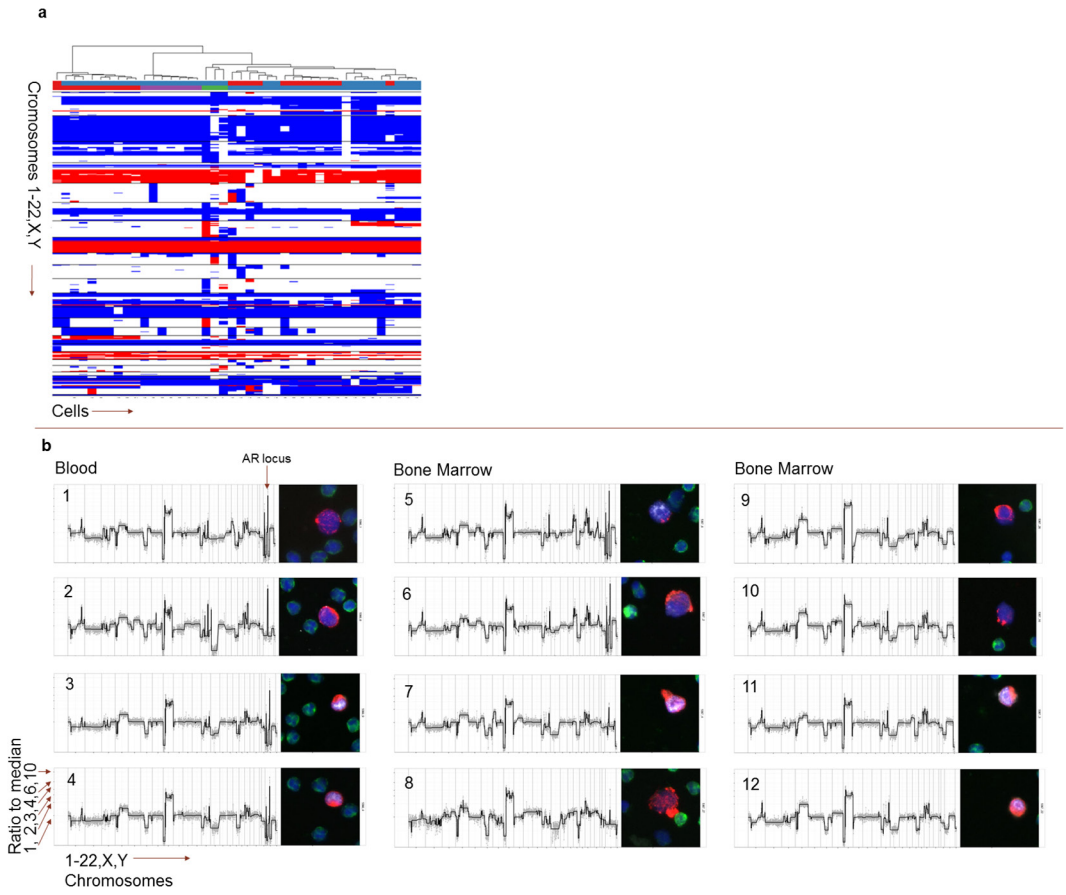

Figure 3. Genomic profiles of cells found in the blood and bone marrow from the prostate cancer patient. (a) Genomic heat map displays the CNV profiles (chromosomal deletions in blue and amplification in red) for the CTC/DTC analyzed from the prostate cancer patient. (b) Representative CNV profiles of a CTC (left) and DTC (center, right) including 40× immunofluorescent composite images (DAPI, blue; CK, red; CD45, green; and AR white). The AR loci on X chromosome is highlighted ((b), panel 1).

image (figure 3(b), panels 3, 4, 5 and 7). In other cases, the gene encoding AR was amplified even though the protein could not be detected by fluorescence (figure 3(b), panels 1, 2 and 6). Notably, there is also a genomic subclone comprising cases where AR amplification is not observed in the CNV profile (figure 3(b), panels 9–12) but AR is still detected in some of the fluorescent images (figure 3(b), panels 11 and 12). Morphological traits, such as cell and nuclear size and CK staining also varied among individual cells as is evident from the fluorescent images.

To evaluate the performance of the combined HD-SCA/IMC analysis in relation to the results from the HD-SCA fluorescent assay alone, we retrieved one unstained slide each from blood and bone marrow of MD42109 from the −80 °C archive where they had been stored for three years. The slides were stained and scanned using the standard HD-SCA 3-color protocol (DAPI, CD45 and pan-CK) to identify candidate cancer cells (figure 1). A total of 21 candidate DTCs and three CTCs were identified and their positions located on the respective slides. The stained slides were then stained with metal-labeled antibodies in preparation for imaging on the Hyperion instrument. The combined results of the HD-SCA immunofluorescence and IMC analysis of these cells are presented in figures 4 and 5.

Figures 4(a) and (b) show a selection of images generated in the IMC analysis of two representative cells, DTC #1271 from the bone marrow aspirate and CTC #1290 from blood, along with separate images of the markers being tested in the two modalities. The images from the fluorescence analysis of the representative cells (top panel) show that both cells were strongly CK positive and CD45 negative. The CD45 expression in leukocytes seen in the full ROI composite image generated by the IMC, show a high degree of correlation of CD45 expression ranging from low to high with the same cells in the corresponding ROI captured in the HD-SCA fluorescence assay image (figure 4(a), supplementary data, figure 1). The composite fluorescence images (figure 4(c)), demonstrate a phenotypic heterogeneity of both the morphology (size) and the variation in the CK signal among the cancer cells.

The heat map (figure 4(d)) provides a summary of the IMC signals in each channel for all 24 cells analyzed. The signal intensity for CK 8/18 varies even among the positive cells, and some of the cells that were positive for CK in the HD-SCA analysis were below the LOD in the IMC for CK 8/18, likely due to the fact that more CK epitopes are measured with the pan-CK antibody cocktail used in HD-SCA staining.

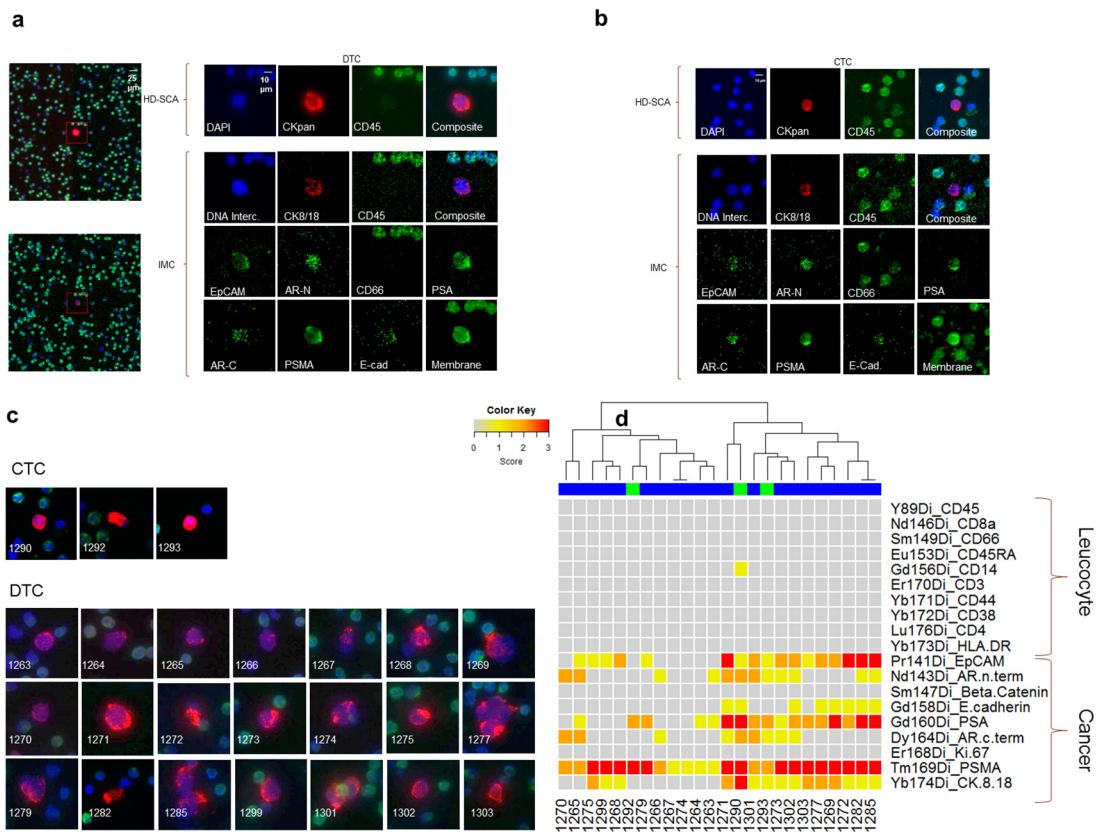

Figure 4. Combined IMC and HD-SCA characterization of circulating tumor cells from prostate cancer blood and bone marrow. (a) DTC (#1271) from a prostate cancer patient centered in the full frame ROI (400 × 400) together with cropped/zoomed in fluorescent images (top panel) and the IMC rendered images (bottom panels) as comparison for channel DAPI/Intercalator (blue), CK (red), CD45 (green) and as a 3 color composite image. The IMC generated results for EpCAM, AR-N, CD66, PSA, PSMA, E-cadherin and a lipid membrane counterstain are presented in green below. (b) Images of the CTC (#1290) with a similar expression profile from the same patient as in A. (c) 40× fluorescent images DAPI (blue), CD45 (green), and CK (red) of all the cells found and analyzed in the blood and bone marrow sample taken with optimal exposure for each cell. (d) Heat map of CTCs (green banner) and DTCs (blue banner) analyzed from the same patient colored according to the four step scoring scale; below detection (gray) detectable (yellow), strong (orange) and very strong (red) depending on the detected protein.

Fifty percent (12/24) of the cells showed AR positivity in the IMC, in agreement with the parallel HD-SCA assay. Of the AR N-terminal positive cells, the majority were also positive for the C-terminal specific antibody, indicating that the full-length receptor is present in these cells. The rendered images from the IMC show a clear nuclear localization of AR, which follows upon activation of the receptor [24] (figures 4(a) and (b)). A majority of the cells (18/24) were positive for the epithelial marker EpCAM and 6 cells were also positive for E-cadherin, including the cell clusters 1277 and 1285 (figures 4(c) and (d)). Importantly, all of the cells were PSMA positive and a majority were also PSA positive, providing independent evidence for the prostate origin of the cells (figures 4(a), (b) and (d)). Thus, the IMC analysis confirmed and complimented the fluorescent HD-SCA and genomic signatures by providing additional markers for phenotypic profiling. Further, these results demonstrate the feasibility of using archived samples for IMC analysis, allowing us to revisit and further characterize long stored and previously analyzed samples in the HD-SCA biobank.

Characterization of leukocyte populations

The HD-SCA assay is unique among the current CTC isolation methods in that it also assays and displays the full complement of leukocytes along with candidate CTC, enabling the analysis of a significant number of immune cells for multiple markers while analyzing CTCs. To demonstrate the feasibility of such an application, the surrounding leukocytes from the prostate cancer samples presented above were characterized by combining data aggregated from multiple ROIs on the slide into one dataset for each of the blood and bone marrow samples. To establish a reproducible analytical workflow, the first step was to remove all poorly segmented cells by size selection. Next, by gating the data (2089 cells) based on the average ion counts for each leukocyte on biaxial plots, the ratio compared to all nucleated cells of neutrophils ($CD66^{+}/CD44^{low}$, 67%) monocytes ($HLA\text{-}DR^{+}/CD14^{+}$, 7%) and T-cells ($CD3^{+}/HLA\text{-}DR^{-}$, 9%) could be determined (figure 5(a)). The ratio correlated well with the manual differential count performed by a trained clinical lab technician (69% neutrophils, 14% lymphocytes, 8% eosinophils, 8%

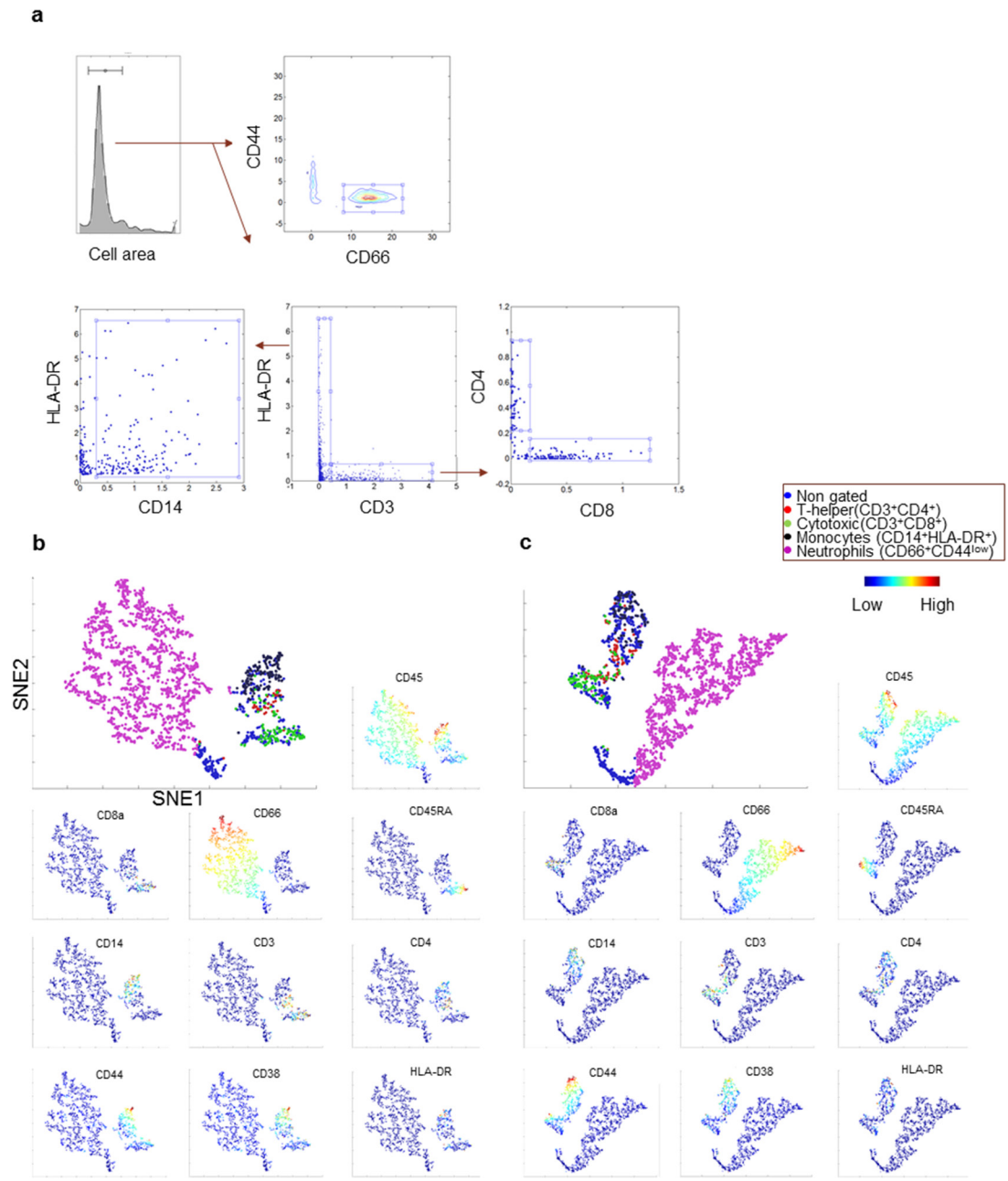

Figure 5. Immune cell phenotyping. Identification of subpopulations of leukocytes within the ROIs analyzed for sample MDA-42109 showing (a) conventional biaxial gating of cells (after initial size exclusion of non-single cells) to identify neutrophils ($CD66^+$), cytotoxic T-cells ($CD3^+$, $CD8^+$), T-helper cells ($CD3^+$, $CD4^+$) and monocyte populations ($HLA\text{-}DR^+$, $CD14^+$). (b) viSNE plots (SNE1 versus SNE2) of around 2000 cells captured with IMC in blood and (C) bone marrow, respectively. The larger image shows the same populations as gated above, overlaid by color. The smaller images show the individual markers with the expression represented as a gradient from low (blue) to high (red).

monocytes, 1% basophils). Furthermore, the T-helper $CD4^+$ cells (2%) and cytotoxic $CD8^+$ (5%) were also identified within the T-cell population. Amongst each of the T-helper and cytotoxic T-cell populations, we also detected small populations of activated ($CD38^+$, <1%) and naïve/effector cells ($CD45RA^+/CD4^+$, 1% and $CD45RA^+/CD8^+$, 3.8%). In figure 5(b), the signal intensity for different populations are shown overlaid in a 2D-viSNE map [13]. Similar data analysis for the bone marrow sample (2267 cells, figure 5(c)) demonstrated slightly lower ratio of neutrophils (55%), monocytes (4%) and lymphocytes (10%) compared to the blood. Thus, despite lacking markers for precursor cells in this particular analysis, the current results still indicate higher levels of immature cells present in the bone marrow compared to the blood, as expected [25].

In summary, we have developed an integrated approach of CTC and DTC characterization based on the HD-SCA workflow with the subsequent downstream multiplex proteomic analysis using IMC. The method was demonstrated using CTCs, DTCs as well as surrounding leukocytes in samples from a bone metastatic prostate cancer patient, showing added

value to the data generated by the HD-SCA assay alone. Results were compared to genomic profiling of matching cells from the same samples, showing a higher degree of heterogeneity on a protein level than on the genomic level.

Discussion

The results presented here provide proof of concept for the application of the combined HD-SCA/IMC workflow for targeted proteomics analysis on rare cancer cells and cell aggregates, as well as associated leukocyte populations, using blood and bone marrow aspirate from a metastatic prostate cancer patient. To achieve these results, we have coupled IMC, using the Fluidigm Hyperion instrument, with the existing HD-SCA method for CTC/DTC identification, morphometrics and single cell genomics. Significantly, this additional downstream high-content analysis was integrated without the need to modify any of the previously validated upstream processes of the HD-SCA. This continuous workflow enables a strategy of deep proteo-genomic analysis applicable for translational science studies within clinical trials. While the technology is currently applied as a discovery tool for markers or new combinations of markers, it has the potential for clinical validity either as a broader pathology platform, or during the development phase of a specific test. Typically, after discovery of a potential biomarker using such a high dimensional technology, a less complex, more focused assay based on a limited number of reagents and suitable for routine clinical implementation would likely be created, as demonstrated by Scher and colleagues [10].

The opportunity to utilize targeted proteomics for research on ultra-rare cells such as CTCs from liquid biopsies lies in the high level of multiplexing available to create 'smart panels' with a wide, purposeful spectrum of targets such as tissue specific markers and known markers of therapeutic importance, but also markers of discovery, such as mechanism of immune surveillance and escape [26] and markers for tumor cell states, including, dormancy, EMT, vascular mimicry and stemness. In practice, unique panels would be constructed for specific clinical studies. In this demonstration study, for example, we included markers for characterization of the surrounding leukocyte population. Patient specific immune signatures in addition to the cancer cells could potentially provide insight for classification and monitoring of cancer for precision medicine purposes [27, 28], and the possibility to do so simultaneously is a unique feature of the HD-SCA assay compared to other CTC technologies. The current panel contained four prostate specific markers, six epithelial and cell signaling or cancer-associated markers, and nine common leukocyte markers, plus a direct conjugated DNA intercalator and membrane counterstain. This combination of markers allowed us to both confirm the tissue of origin for the candidate prostate cancer cells and identify specific populations in the surrounding leukocytes, and further, to test the expression of additional proteins in the candidate cells.

In classifying rare cells it is important to establish normalization and quantification methods to enable cell to cell comparisons as well as comparisons between experiments. This is especially true for those proteins where expression approaches the LOD. Each individual antibody assay has to be evaluated in order to estimate the detection limit and linear range. Quantitative scoring of protein expression in individual candidate cells is an ongoing process. The scoring method presented in this study is rooted in that developed for the fluorescent measurements in the HD-SCA, wherein the signals from the surrounding white blood cells are used to establish a negative control baseline for most epithelial tissue markers and a means for normalizing signals from candidate cells. The SDOM measure as applied in the conventional HD-SCA assay is a straightforward means for establishing the background level based on the surrounding leukocytes. SDOM is however sensitive to outliers and is poorly suited for skewed populations, in particular for signals close to zero. SDOM is also limited in the case of CTC clusters, where multiple cells of interest contribute to the overall mean and standard deviation, as illustrated in figures 2(m)–(o) where the LOD (mean + 3.3σ) is estimated by excluding the seven cluster cells from the overall calculation (dotted lines) in comparison to estimate based on all cells (solid lines). The S/N measure based on the inter-quartile range was introduced because it is less sensitive to outliers, but on the other hand it gives only a rough estimate based on the background signal. Hence, both measures are currently taken into account to provide an objective score.

It should not be overlooked that each type of preparation, whether fresh frozen tissue, FFPE sections or the HD-SCA, requires separate validation of the antibody repertoire [29, 30]. In the process of integrating the two platforms, we noted that up to half of the antibody conjugates previously validated for suspension CyTOF did not work well or at all on the glass substrate HD-SCA preparations. We have therefore developed a standardized assay development process for the HD-SCA/IMC workflow using cell lines spiked into normal blood samples. Optimization of sensitivity and specificity for some problematic markers often required iterative testing of multiple antibody clones and metal-conjugate combinations. Expanding the repertoire of validated conjugates in the HD-SCA/IMC setting is a work in progress.

Rare cells from liquid biopsies show great potential for assessing the effectiveness of cancer therapy and detecting resistance or progression in near-real time through a minimally invasive procedure that can be repeated on a time scale of days or weeks, rather than months or years. New technologies such as cell free

DNA and single cell sequencing add volumes of new information to what was once limited to enumerating cancer cells in the blood. The HD-SCA workflow is a direct analysis platform that provides the opportunity for a complete characterization of the blood sample without enrichment and that accommodates multiple assay modes including both morphometric and molecular level analysis. The integration of the HD-SCA workflow with single cell targeted protein analysis creates a true multiplex technology for liquid biopsies that is accessible for both basic and translational clinical studies.

Conclusion

This study provides proof of concept that IMC analysis adds another dimension to the established HD-SCA workflow, enabling multiplex proteomic profiling in addition to morphometric and genomic characterization of ultra-rare circulating tumor cells. Importantly, this study provides evidence that samples stored for several years can be revisited and analyzed *de novo* as new protein targets are identified. With a multiplexing level of (currently) 40 proteins, highly informative biomarker panels can be readily tailored for the scientific question at hand, which gives the HD-SCA/IMC technology immense potential for deriving predictive and prognostic biomarkers from the liquid phase of cancer.

Acknowledgments

This work was based wholly or partially on research supported by Breast Cancer Research Foundation 004698-00003; National Cancer Institute and Leidos Biomedical Research, Inc HHSN261200800001E; Prostate Cancer Foundation 16CHAL04; Vassiliadis Research Fellowship, Polak Research Fellowship, the Vicky Joseph Research Fellowship, and the Charles University Research Fund (Progres Q39). USC is an early access partner of Fluidigm for the evaluation of the Hyperion System and Fluidigm provided both expertise and reagents to support this work. The content is solely the responsibility of the authors and does not necessarily represent the official views of these funding agencies and foundations. Competing Financial Interests are PK: Advisor to Epic Sciences, royalty recipient from Epic Sciences, shareholder at Epic Sciences; JBH: On the Clinical Advisory Boards of Epic Sciences, Inc., La Jolla, CA and CelMatix, Inc. of NY, NY.

ORCID iDs

Anna Sandström Gerdtsson https://orcid.org/0000-0003-1932-0365
Carmen Ruiz Velasco https://orcid.org/0000-0003-2370-3000
Peter Kuhn https://orcid.org/0000-0003-2629-4505

References

[1] Baca Q, Cosma A, Nolan G and Gaudilliere B 2017 The road ahead: implementing mass cytometry in clinical studies, one cell at a time *Cytometry* B **92** 10–1
[2] Spitzer M H and Nolan G P 2016 Mass cytometry: single cells, many features *Cell* **165** 780–91
[3] Bandura D R, Baranov V I, Ornatsky O I, Antonov A, Kinach R, Lou X, Pavlov S, Vorobiev S, Dick J E and Tanner S D 2009 Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry *Anal. Chem.* **81** 6813–22
[4] Wang H A, Grolimund D, Giesen C, Borca C N, Shaw-Stewart J R, Bodenmiller B and Günther D 2013 Fast chemical imaging at high spatial resolution by laser ablation inductively coupled plasma mass spectrometry *Anal. Chem.* **85** 10107–16
[5] Giesen C *et al* 2014 Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry *Nat. Methods* **11** 417–22
[6] Marrinucci D *et al* 2012 Fluid biopsy in patients with metastatic prostate, pancreatic and breast cancers *Phys. Biol.* **9** 016003
[7] Dago A E *et al* 2014 Rapid phenotypic and genomic change in response to therapeutic pressure in prostate cancer inferred by high content analysis of single circulating tumor cells *PloS One* **9** e101777
[8] Carlsson A, Kuhn P, Luttgen M S, Dizon K K, Troncoso P, Corn P G, Kolatkar A, Hicks J B, Logothetis C J and Zurita A J 2017 Paired high-content analysis of prostate cancer cells in bone marrow and blood characterizes increased androgen receptor expression in tumor cell clusters *Clin. Cancer Res.* **23** 1722–32
[9] Malihi P D *et al* 2017 Clonal diversity revealed by morphoproteomic and copy number profiles of single prostate cancer cells at diagnosis *Converg. Sci. Phys. Oncol.* **4** 015003
[10] Scher H I *et al* 2016 Association of AR-V7 on circulating tumor cells as a treatment-specific biomarker with outcomes and survival in castration-resistant prostate cancer *JAMA Oncol.* **2** 1441–9
[11] Pau G, Fuchs F, Sklyar O, Boutros M and Huber W 2010 EBImage—an R package for image processing with applications to cellular phenotypes *Bioinformatics* **26** 979–81
[12] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 1997 Validation of Analytical Procedures: Methodology Q2B *Fed. Reg.* **62** 27463–7 (www.ich.org/products/guidelines/quality/quality-single/article/validation-of-analytical-procedures-text-and-methodology.html)
[13] Amir E A D, Davis K L, Tadmor M D, Simonds E F, Levine J H, Bendall S C, Shenfeld D K, Krishnaswamy S, Nolan G P and Pe'er D 2013 viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia *Nat. Biotechnol.* **31** 545–52
[14] Ghosh A and Heston W D 2004 Tumor target prostate specific membrane antigen (PSMA) and its regulation in prostate cancer *J. Cell. Biochem.* **91** 528–39
[15] Jiang Y G, Luo Y, He D L, Li X, Zhang L L, Peng T, Li M C and Lin Y H 2007 Role of Wnt/beta-catenin signaling pathway in epithelial-mesenchymal transition of human prostate cancer induced by hypoxia-inducible factor-1alpha *Int. J. Urol.* **14** 1034–9
[16] Hurt E M, Kawasaki B T, Klarmann G J, Thomas S B and Farrar W L 2008 CD44 + CD24(-) prostate cells are early cancer progenitor/stem cells that provide a model for patients with poor prognosis *Br. J. Cancer* **98** 756–65
[17] Zola H *et al* 2005 CD molecules 2005: human cell differentiation molecules *Blood* **106** 3123–6
[18] Luo W, Tapolsky M, Earley K, Wood C G, Wilson D R, Logothetis C J and Lin S H 1999 Tumor-suppressive activity of CD66a in prostate cancer *Cancer Gene Ther.* **6** 313–21
[19] Beucher S 1994 Watershed, hierarchical segmentation and waterfall algorithm *Mathematical Morphology and Its*

Applications to Image Processing (Computational Imaging and Vision vol 2) ed J Serra and P Soille (Dordrecht: Springer)
[20] Jones T R, Carpenter A and Golland P 2005 Voronoi-based segmentation of cells on image manifolds *Lect. Notes Comput. Sci.* **3765** 535–43
[21] Tricot S *et al* 2015 Evaluating the efficiency of isotope transmission for improved panel design and a comparison of the detection sensitivities of mass cytometer instruments *Cytometry* A **87** 357–68
[22] Finck R, Simonds E F, Jager A, Krishnaswamy S, Sachs K, Fantl W, Pe'er D, Nolan G P and Bendall S C 2013 Normalization of mass cytometry data with bead standards *Cytometry* A **83** 483–94
[23] Lazar D C, Cho E H, Luttgen M S, Metzner T J, Uson M L, Torrey M, Gross M E and Kuhn P 2012 Cytometric comparisons between circulating tumor cells from prostate cancer patients and the prostate-tumor-derived LNCaP cell line *Phys. Biol.* **9** 016002
[24] Heinlein C A and Chang C 2004 Androgen receptor in prostate cancer *Endocr. Rev.* **25** 276–308
[25] Brooimans R A, Kraan J, van Putten W, Cornelissen J J, Lowenberg B and Gratama J W 2009 Flow cytometric differential of leukocyte populations in normal bone marrow: influence of peripheral blood contamination *Cytometry* B **76** 18–26
[26] Mohme M, Riethdorf S and Pantel K 2017 Circulating and disseminated tumour cells—mechanisms of immune surveillance and escape *Nat. Rev. Clin. Oncol.* **14** 155–67
[27] Gruber I, Landenberger N, Staebler A, Hahn M, Wallwiener D and Fehm T 2013 Relationship between circulating tumor cells and peripheral T-cells in patients with primary breast cancer *Anticancer Res.* **33** 2233–8
[28] Wistuba-Hamprecht K *et al* 2017 Establishing high dimensional immune signatures from peripheral blood via mass cytometry in a discovery cohort of stage IV melanoma patients *J. Immunol.* **198** 927–36
[29] Howat W J, Lewis A, Jones P, Kampf C, Ponten F, van der Loos C M, Gray N, Womack C and Warford A 2014 Antibody validation of immunohistochemistry for biomarker discovery: recommendations of a consortium of academic and pharmaceutical based histopathology researchers *Methods* **70** 34–8
[30] Uhlen M *et al* 2016 A proposal for validation of antibodies *Nat. Methods* **13** 823–7