RESEARCH ARTICLE

Whole Cell MALDI Fingerprinting Is a Robust Tool for Differential Profiling of Two-Component Mammalian Cell Mixtures

Valentina Z. Petukhova,[1] Alexandria N. Young,[1] Jian Wang,[2] Mingxun Wang,[2] Andras Ladanyi,[3] Rajul Kothari,[4] Joanna E. Burdette,[1] Laura M. Sanchez[1]

[1]Department of Medicinal Chemistry and Pharmacognosy, University of Illinois at Chicago, 833 S Wood St., MC 781, Room 539, Chicago, IL 60612, USA
[2]Ometa Labs, 3210 Merryfield Row, San Diego, CA 92121, USA
[3]Department of Obstetrics & Gynecology, Rush University Medical Center, 1653 W Congress Pkwy, Chicago, IL 60612, USA
[4]Department of Obstetrics & Gynecology–Division of Gynecologic Oncology, University of Illinois at Chicago, 820 S Wood St., Chicago, IL 60612, USA
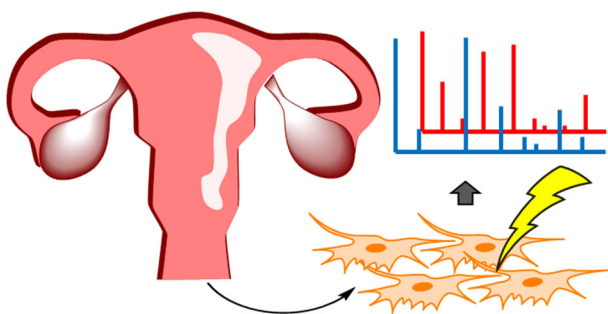
Abstract. MALDI fingerprinting was first described two decades ago as a technique to identify microbial cell lines. Microbial fingerprinting has since evolved into an automated platform for microorganism identification and classification, which is now routinely used in clinical and environmental sectors. The extension of fingerprinting to mammalian cells has yet to progress partly due to compartmentalization of eukaryotic cells and overall higher cellular complexity. A number of publications on mammalian whole cell fingerprinting suggest that the method could be useful for classification of different cell types, cell states, and monitoring cell differentiation. We report the optimization of MALDI fingerprinting workflow parameters for mammalian cells and its application for differential profiling of mammalian cell lines and two-component cell line mixtures. Murine fallopian tube cells and high-grade ovarian carcinoma cell lines and their mixtures are used as model mammalian cell lines. Two-component cell mixtures serve to determine the method's feasibility for complex biological samples as the ability to detect cancer cells in a mixed cell population. The level of detection of cancer cells in the two-component mixture by principle component analysis (PCA) starts to deteriorate at 5% but with application of a different statistical approach, Wilcoxon rank sum test, the level of detection was determined to be 1%. The ability to differentiate heterogeneous cell mixtures will help further extend whole cell MALDI fingerprinting to complex biological systems.
Keywords: Whole cell MALDI fingerprinting, Mammalian cell lines, Two-component cell line populations

Introduction

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI -TOF MS) is a powerful analytical technique widely used for analysis of proteins in a broad range of biological samples, which vary from lysed cells extracts [1–10] to tissue sections [11–18]. Whole cell MALDI fingerprinting is based on the mass spectral analysis of a whole cell without additional preparatory steps, such as fractionation or extraction. Minimal sample processing circumvents complex and time-consuming experimental steps necessary for cell proteomic analyses. MALDI protein spectral profiles generated during analyses are species specific and serve as unique cellular

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13361-018-2088-6) contains supplementary material, which is available to authorized users.

Correspondence to: Laura Sanchez; *e-mail:* sanchelm@uic.edu

ID fingerprints [19–21]. Speed, simplicity, sensitivity, and specificity for large biomolecules in complex matrices confer whole cell MALDI fingerprinting as a robust, facile, and inexpensive technique. The most advanced application—identification and classification of microorganisms—is an automated platform, which is routinely used in clinical [22–26], biodefense [27], and environmental [28–30] laboratories.

Mammalian fingerprinting has yet to be broadly deployed in clinical applications in part because of the overall complexity of eukaryotic cells and a lack of standardized workflows for mammalian fingerprinting. Mammalian cell compartmentalization results in numerous regulatory pathways for cellular networking and multiple cell cycle states, which affect the resulting mammalian cellular fingerprint. The lack of standardized workflows in mammalian fingerprinting was recently addressed by Munteanu et al. [31]. The review comprehensively covers all aspects of the fingerprinting workflow necessary for method standardization, such as sample preparation (type of matrix, matrix/solvent composition/additives, sample application and cell density), instrument parameters (type of MS instrument, laser wavelength/number of shots, acquisition parameters), and statistical assessment models [principal component analysis (PCA), partial least square analysis (PLSA), hierarchical clustering analysis (HCA), Pearson coefficient]. For example, it is known that in bacterial cells, different matrices and crystallization methods produce different MALDI spectra of the same sample [32, 33], and given the complexity of eukaryotic cells, there is a high need for rigorous standardization of sample preparation/instrument parameters. Also noted is a lack of extensive databases for mammalian cellular fingerprints that would require universal automated workflows and classification algorithms.

Despite the obvious challenges of mammalian fingerprinting, numerous reports on mammalian cells characterization have been published. Zhang et al. demonstrated the ability to differentiate three mammalian cell lines after cell lines were lysed in 2,5-dihydroxybenzoic acid matrix solution [34]. Another study identified 66 cell lines representing 34 species from insects to primates based on MALDI analysis of formic acid/acetonitrile extractions of cultured cells [35]. More advanced applications were aimed at distinguishing different cell types originating from the same cell lineage: identification of two different pancreatic hormone-secreting cell lines [36], the comparison of primary human blood cells and blood cell lines [37, 38], molecular phenotyping of central nervous system (CNS) glial cells (astroglial, microglial, and oligodendroglial) [39], and MALDI-MS fingerprinting of different melanoma cell lines [40]. Further applications of mammalian fingerprinting have focused on physiological changes of a single cell, reflecting its specific cell states or cell transformations, such as differentiation of human colon carcinoma [41] or leukemia [38] cell lines, multifaceted activation of human macrophages [42], identification of resting and activated human monocyte subsets [43], rapid detection of apoptosis/necrosis signature [44], and monitoring of histone deacetylase drug target engagement [45]. Regardless of the scope of the aforementioned studies, no consistency in

method parameters was observed (such as matrix, cell density, cell media, sample application technique, and laser frequency/number of shots) for either cell authentication [35–40] or close monitoring of a single cell change application [41–45].

Based on literature, mammalian fingerprinting has the potential to discriminate between different cell types and cell states, but there is a scarcity of publications on application of this technique towards heterogeneous samples. Heterogeneity is a hallmark of cancer, and there is unmet medical need for sensitive detection of cancerous cells in a complex environment of biological samples or detection of relevant biomarkers using emerging liquid biopsy techniques. As a cancer diagnostic tool, MALDI fingerprinting has been successfully applied towards clinical fine-needle aspirates of lung cancer cells [46] and oral mucosa brush biopsies [47] to obtain cancer cell specific protein profiles, which differentiate tumor samples and non-tumor controls. Unfortunately, the sample heterogeneity or the percentage of cancer cells detected was not reported, limiting the utility of these reported methods. We used a reduced model system for workflow optimization consisting of two-component cell line mixtures with known concentrations of cancer cells. Method parameters were optimized for whole cell MALDI fingerprinting workflow and validated using defined cell line mixtures. Optimized method parameters allowed for the discrimination between non-cancerous and cancer mammalian cell lines as well as between two-component cell line mixtures with the minimum threshold for cancer cells to be 1% in an otherwise non-cancerous "healthy" cellular background.

Methods

Reagents

Acetonitrile (ACN, HPLC grade) and trifluoroacetic acid (TFA, LC-MS grade) was purchased from Sigma-Aldrich (St. Louis, MO, USA). Milli-Q water (ddH2O; Millipore) was prepared in-house. Sinapinic acid (SA) (matrix substance for MALDI-MS, ≥ 99%) was obtained from Sigma-Aldrich (St. Louis, MO, USA). α-Cyano-4-hydroxycinnamic acid (CHCA), 2,5-dihydroxybenzoic acid (DHB), and protein calibration standards I and II were purchased from Bruker Daltonics (Bremen, Germany).

Cell Culture

Cell Lines The human ovarian cancer cell line (OVCAR3) was purchased from the American Type Culture Collection. OVCAR8 cells expressing red fluorescent protein (OVCAR8-RFP) were a gift from Sharon Stack at the University of Notre Dame. OVCAR4 and OVCAR8 were obtained from the NCI 60 Cell Panel Cell Bank Repository. OVCAR4-RFP was generated with lentiviral transduction according to the manufacturer's protocol (GenTarget #LVP023). OVCAR3 was grown in RPMI 1640 supplemented with 20% fetal bovine serum (FBS), 1% penicillin/streptomycin (P/S), and 10 μg/mL insulin. OVCAR8 and OVCAR8-RFP were grown in DMEM with 10% FBS and 1% P/S. OVCAR4-RFP was grown in RPMI

1640 supplemented with 10% FBS, 1% P/S, and 1% L-glutamine. Our non-cancerous cell line was murine oviductal cells (MOE) obtained from Dr. Barbara Vanderhyden at the University of Ottawa and were maintained in α-modified Eagle's medium supplemented with 10% FBS, 1% P/S, 1% L-glutamine, 2 μg/ml epithelial growth factor, 5 μg/ml insulin, 5 μg/ml transferrin, 5 ng/ml sodium selenite, 1 mg/ml gentamycin, and 20 ng/ml β-estradiol. Cultured cells were maintained in a humidified incubator at 37 °C in 5% CO_2. Cells were passaged a maximum of 30 times. Cell lines were validated by short tandem repeat analysis and tested mycoplasma-free in 2017.

Primary Human Cells   Human fallopian tube specimens were obtained from benign gynecologic surgeries performed at the University of Illinois Hospital after consent and de-identification. Fallopian tube epithelium was mechanically isolated and incubated in MOE media with 1 mg/mL collagenase and 0.05% trypsin for 20 min at 37 °C with occasional vortexing. Dissociated cells were allowed to attach and grow in culture plates for 5 days before FBS concentration was reduced to 1% to enrich for epithelial cells. Cells were passaged a maximum for five times thereafter. Tissue collection was approved by the UIC Institutional Review Board (IRB #2012-0539) and performed in accordance with NIH guidelines on human subject research.

Aqueous Cell Suspensions   Cells were collected by trypsinization and washed with PBS. Cell densities were measured via hemocytometer and cells were re-suspended in ddH_2O to obtain a cell concentration of 10,000 cells/μL and stored as frozen suspensions at − 20 °C until use.

Live Cells in PBS   For live cell applications, live cells were collected by trypsinization and washed with PBS. Cell densities were measured via hemocytometer after which cells were re-suspended in PBS to obtain a cell concentration of 10,000 cells/μL, followed by storage on ice prior to sample preparation.

Sample Preparation for MALDI-MS Analysis

Matrices containing 10 mg/mL of SA (SA10), CHCA (CHCA10), or DHB (DHB10) were prepared in the solution of ACN/water = 70/30 containing 0.1% TFA. Matrix solution of SA with a higher aqueous composition (SA20) was prepared by dissolving 20 mg of SA in 1 mL of ACN/water/TFA = 30/70 containing 0.1% TFA.

Dried Droplet and Double Layer Sample Applications from Aqueous Cell Suspensions   Frozen cell suspensions were thawed on ice and diluted with ddH_2O to desired concentrations. For dried droplet application, equal volume amounts of a cell suspension and matrix solution were mixed together and left on ice for 10 min. To avoid the problem of non-specific interactions

of cellular proteins with the pipette surface, each pipette tip was saturated with matrix-analyte mixture (by pipetting the mixture three times) prior to application. Then 1–1.5 μL of the resulting mixture was spotted on 384-well ground steel target (Bruker Daltonics) in n technical replicates ($n = 3, 4, 16$, or 24) and air dried. For double layer application, first, 1–1.5 μL of SA in acetonitrile (1 mg/mL of ACN) was applied to 384-well ground steel plate and air dried. Then, it was followed by dried droplet application of matrix-analyte mixture as described above.

Whole Cell Applications of Live Cells from PBS Media

Method A   Live cells in PBS media were diluted with PBS to desired concentrations. The samples were kept on ice in between dilution steps and sample preparation. Before sample application, cell mixture was gently re-suspended by light agitation. Each pipette tip was saturated with cell mixture (by gentle pipetting the mixture three times) prior to application. For sample preparation, 1.5 μL of cell mixture was applied on 384-well ground steel plate in triplicates and dried at room temperature. Then, 1.5 μL of SA20 matrix (20 mg/mL of ACN/water = 30/70 + 0.1%TFA) was applied over the dried cell culture and air dried.

Dried Droplet Applications of Live Cells from PBS Media

Method B   Live cells in PBS media were diluted with PBS to desired concentrations. The samples were kept on ice in between dilution steps and sample preparation. Equal volume amounts of a cell suspension and matrix solution were mixed together and left on ice for 10 min. To avoid the problem of non-specific interactions of cellular proteins with the pipette surface, each pipette tip was saturated with matrix-analyte mixture (by pipetting the mixture three times) prior to application. Then, 1.5 μL of the resulting mixture was spotted on 384-well ground steel target (Bruker Daltonics) in triplicate and air dried. Dried droplet applications that were washed with water to remove PBS salts were prepared following all the steps described above. Then, 1.5 μL of ice-cold water was gently applied to the top of a dry crystalline preparation and briefly removed with Kimwipe trying not to disturb the sample integrity. After washing, step crystalline samples were dried at room temperature.

Mixed Cell Populations from Aqueous Cell Suspensions

Frozen MOE and OVCAR8 cell suspensions at 1×10^4 cells/μL were thawed on ice and diluted with cold ddH_2O to 2500 cells/μL. Then, MOE and OVCAR8 suspensions were mixed to obtain a range of two-component cell line populations containing 50 to 10% (with 10% increment) or 9 to 1% (with 1% increment) of OVCAR8 cancer cells in the mixture. All cell suspensions

were kept on ice during dilution and mixing steps. For dried droplet application, equal volume amounts of a cell suspension and SA20 matrix solution were mixed together and left on ice for 10 min. To avoid the problem of non-specific interactions of cellular proteins with the pipette surface, each pipette tip was saturated with matrix-analyte mixture (by pipetting the mixture three times) prior to application. Then, 1.5 μL of the resulting mixture was spotted on 384-well ground steel target (Bruker Daltonics) in 16 or 24 replicates per population and air dried.

MALDI-MS and Data Analysis

All mass spectra were acquired on an Autoflex Speed LRF MALDI-TOF mass spectrometer (Bruker Daltonics) equipped with Smartbeam™-II laser (355 nm) in positive linear mode within the mass range of m/z 5000 to 20,000 Da with a low mass gate at 3300 Da. Protein spectra can be acquired manually or automatically using the AutoXecute function of the FlexControl 3.4 software with 2-kHz Smartbeam laser, acceleration voltage setting at 19.5 kV, and delayed extraction at 170 ns. Four thousand laser shots were accumulated from 20 different positions of the MALDI target chosen by random walk for manual acquisition or by "spiral_large" measuring raster for automated runs. All spectra were externally calibrated from calibration spots adjacent to the sample spots and processed by FlexAnalysis 4.2 (Bruker Daltonics) software using Top Hat algorithm for baseline subtraction. PCA analysis was performed using ClinProTools 3.0 software. The following CPT processing parameters (adapted from [38]) were used: resolution, 800; top hat baseline subtraction with 10% minimal baseline width; mass range, 5000–20,000 Da. All spectra were automatically normalized by their total ion count by CPT software. Mass spectra were further smoothed using the Savitsky-Golay algorithms (5 cycles with $m/z = 6$ Da). Null spectra and noise spectra exclusion with a noise threshold of 1.00 were enabled. Spectra were recalibrated, allowing a maximal shift tolerance of 1000 ppm match on 25% of the peaks. Peak picking was performed on average spectra using S/N > 5, and peak calculation was performed using peak intensities.

Statistical Analysis

For statistical evaluation profile, mass spectra were converted to mzML format.

To detect differentially expressed features between cell lines, mass spectra in the mzML format were pre-processed using the MALDIQuant software [48] following the workflow suggested in the software documentation (https://cran.r-project.org/web/packages/MALDIquant/MALDIquant.pdf). Protein features across all samples were assembled into a matrix and queried for features, which were differentially expressed in the mixed cell samples versus the non-cancerous cell line (MOE). Utilizing the abundances for each feature, a Wilcoxson rank-sum test [49] was applied to assess whether the abundance differences between the mixture and controls were statistically significant. The Wilcoxson rank-sum test was chosen because it makes very few assumptions about the distribution of the data, e.g., normality in the case of t

tests. Bonferroni-correction [50] was applied to the p value of each test to account for multiple hypothesis testing and a threshold of 0.01 was applied on the corrected p value for each test; other methods were also considered and the results are reported in Figure S21. For each feature, we determined its minimum level of detection L, such that all mixed samples with cancer cell population $\geq L\%$ show statistically different abundance for that feature. For example, if a feature is deemed as differentially expressed in sample with 2 and 5–50% of cancer cells samples, its limit of detection was determined at 5%.

Results and Discussion

Optimization of Matrix/Solvent Composition and Cell Density Using Aqueous Cell Suspensions

The basis of whole cell MALDI fingerprinting can be defined as the cellular protein profile with the best ionization efficiency (highest number of characteristic and/or intense features). Sample ionization is dependent on many factors and therefore required optimization to identify the combination of matrix/solvent composition and plating cell density for generating unique cellular fingerprints. Two to twenty kilodaltons has historically been popular for mammalian fingerprinting and SA is widely used [38–41, 43, 46, 51–54], whereas consensus on the matrix/solvent formulation and sample application technique has varied across the literature. For the initial round of optimization, we chose three different matrices (CHCA, DHB, and SA) for simultaneous optimization of matrix/solvent composition and plating cell density using dry droplet sample application. For side to side comparison, all three matrices were prepared at 10 mg/mL in 70% ACN$_{(aq)}$ (CHCA10, DHB10, and SA10). For the matrix with higher aqueous composition, we used the combination of 20 mg/mL of SA in 30% of ACN$_{(aq)}$ (SA20). The matrix solvent acidity was kept low (0.1% of TFA), since higher TFA concentrations were reported to produce fewer characteristic signals [38].

For cell plating density, 5000-, 2500-, 1250-, 625-, 312-, 156-, and 78-cells/μL concentrations were tested. As seen from Figure 1a, b, the best ionization efficiency for OVCAR3 cell line was observed upon sample dilution with the highest number of peaks appearing around 1250 cells/μL. Less signals were seen upon further dilution, and the cellular signature was almost depleted at 78 cells/μL. The same trend of strong ionization at lower cell counts was observed for MOE cells (Figure S1a and S1b). Interestingly, DHB10 and CHCA10 matrices also showed better ionization upon sample dilution (Figures S2–S4). We also checked higher cell densities (50,000, 25,000, and 10,000 cells/μL), which were previously reported for mammalian cell lines at 10,000 or 100,000 cells/spot [34, 36, 37, 39, 52, 53]. High cellular densities resulted in fewer observed peaks for OVCAR8 cell line, which can be attributed to ion suppression effects (Figure S5). Out of four different matrix/solvent combinations tested, SA20 matrix provided rich and uniform cellular fingerprints as seen for OVCAR3 (Figure 2) or MOE (Figure S6) cells. The other
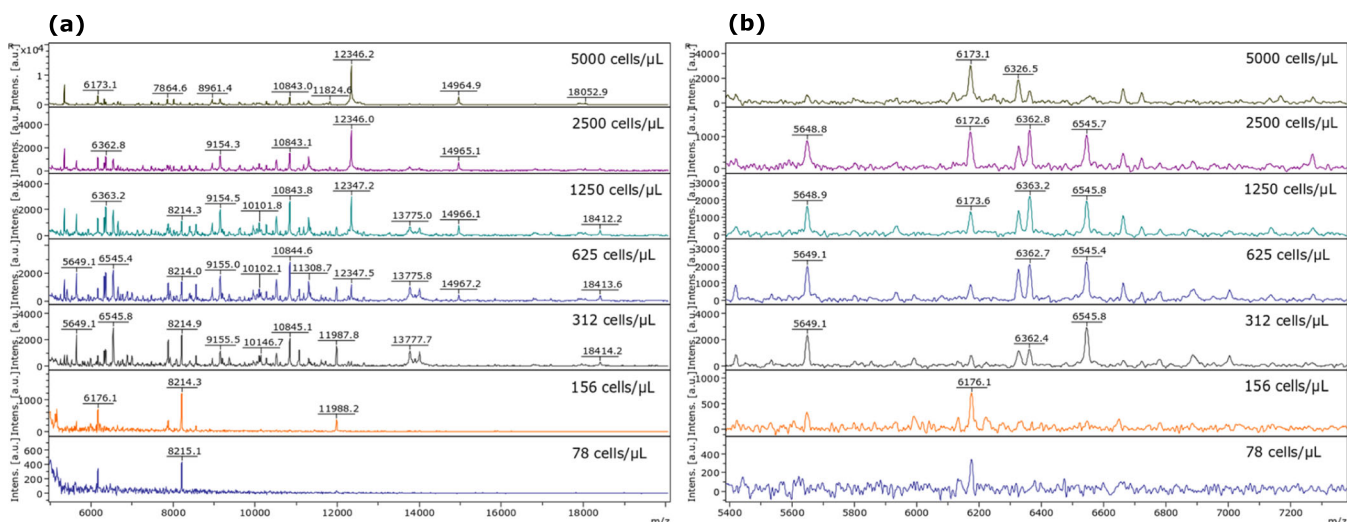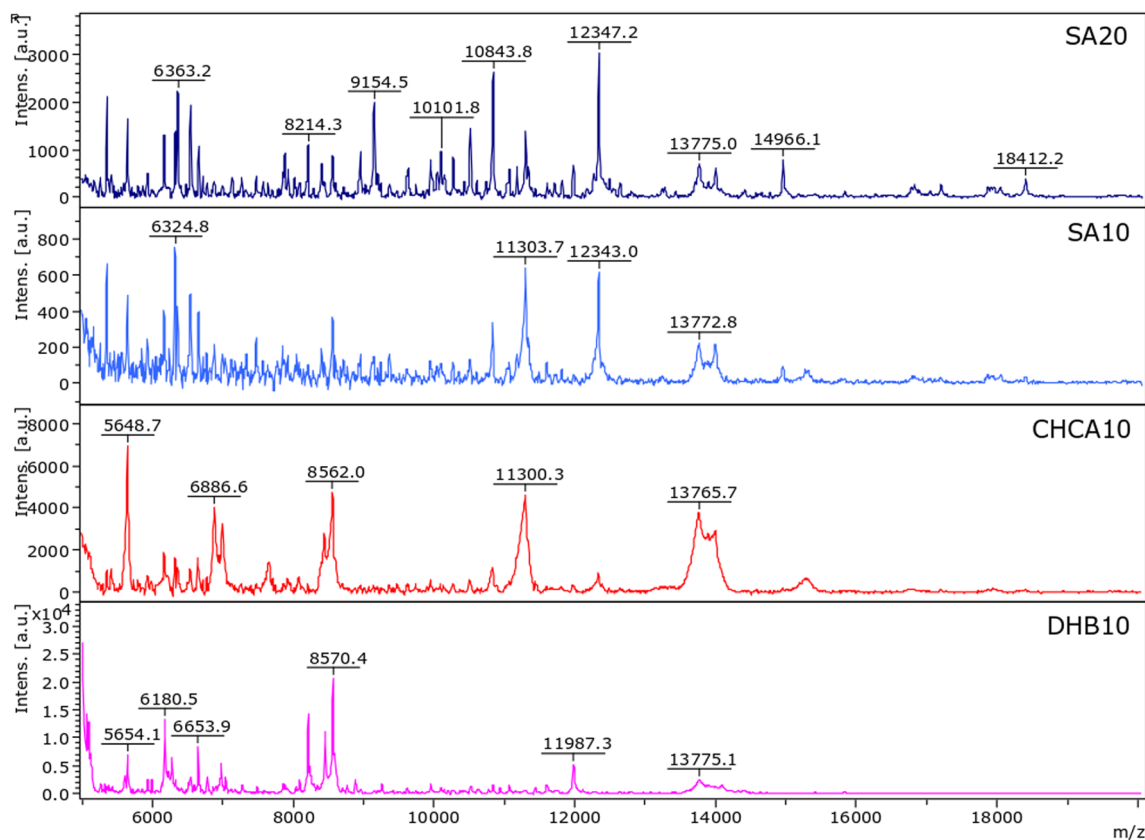
Figure 1. Representative MALDI-TOF spectra of OVCAR3 cell line at different cell densities obtained with SA20 matrix/dried droplet application. (**a**) *m/z* range 5000–20,000 and (**b**) *m/z* range 5400–7400

matrix/solvent combinations resulted in lower ionization efficiency (SA10 and DHB10), the presence of poorly resolved clusters (CHCA10), or poor ionization of proteins with molecular weights higher than 15 kDa (all three matrices: SA10, CHCA10, DHB10). Literature reports also confirm better quality spectra obtained from SA matrix comparing to DHB or

CHCA [36, 40, 53] despite the difference in cell media/sample application techniques (whole cells from PBS applied on MALDI plate, dried, and covered with matrix [36, 53] or cells fixed on aluminum foil, dried, and covered with matrix [40]). Based on these results, all further experiments were conducted with SA20 matrix at 1250 cells/µL plating density.

Figure 2. Representative MALDI-TOF spectra of OVCAR3 cell line at 1250 cell/µL obtained with SA20, SA10, CHCA10, and DHB10 matrices/dried droplet application

Optimization of Mass Detection Range Using Aqueous Cell Suspensions

In general, published methods for mammalian fingerprinting rely on medium protein range, i.e., 5–20 kDa for cell authentication. We decided to check the high protein range (20–50 kDa) to verify the choice of mass detection range for mammalian cell lines. More characteristic peaks were observed in 5–20-kDa range for three different cell lines (OVCAR3, OVCAR8, and MOE, Figure 3a, b) when compared to the 20–50-kDa range (Figure S7). The differences in MOE protein signatures compared to the OVCAR3/OVCAR8 fingerprints are immediately apparent from appearance of very distinctive peaks (marked by green stars) for MOE cells (Figure 3a). Further examination of spectra shows additional peak features and differential peak ionization pattern characteristic to each of the cell lines (Figure 3b). All the above data confirm that 5–20 kDa is most suitable for the generation of characteristic cellular fingerprints.

Sample Application and Cell Media

Previous publications reported better ionization efficiency using the double layer application technique [38, 39, 55] for sample preparation. To further check this method's sensitivity, dried droplet application was compared to double layer application using aqueous cell suspensions under optimized matrix/cell density conditions (SA20, 1250 cells/μL). With our SA20 matrix combination for dried droplet and double layer applications, no differences in ionization efficiency between the two applications for MOE (Figure 4) or OVCAR3 (Figure S8A and S8B) cells were observed. Further, we confirmed previous experimental observation [37] that freezing aqueous cell suspensions obtained from harvested cells would not alter the cellular fingerprint (data not shown). We next tested whether live cells (cells suspended in PBS) would provide a better characteristic cellular fingerprint as opposed to cells suspended in water or freshly frozen in water prior to analysis. For this purpose, MALDI spectra were also acquired from live cells suspended in PBS media using method A or method B and SA20 matrix. Less uniform ionization and significant reduction of protein signals above 14 kDa were observed from whole cell applications (method A) from OVCAR3 cells (Figure S9). Dried droplet application of OVCAR3 cells (method B) suspended in PBS media did not improve cellular fingerprint and removal of PBS salts from dried droplet applications by a quick aqueous wash resulted only in slight improvement of peak intensities (Figure S10). Suppression of protein signals from PBS media due to phosphate anions competing for protons was reported previously [38, 39, 53] and likely depends on sample plating density, i.e., proteins/PBS salts ratio.

Previous reports [52, 53] using SA matrix for dried droplet preparation relied on samples with high cellular densities (10,000 cells/spot). In our case, higher cell densities (10,000 or 5000 cells/μL) yielded modest improvement for intensities/additional peaks in the 6000–12,000 region for whole cell and dried droplet applications from PBS media for OVCAR3 cells (Figure S11A and S11B), but suppression of protein signals above 14 kDa occurred. Aqueous cell suspensions provided the most efficient and uniform protein ionization coupled with the dried droplet application. All further sample preparations utilized this combination.

Optimization of Plating Volume

For plating volume optimization, the statistical distribution of different cell densities (5000, 2500, 1250, 625, and 312 cell/μL) at 1- and 1.5 μL plating volumes was compared. As seen from PCA plots and ClinProTool Peak statistics data for MOE cells (Figure S12, Tables S1 and S2), 1.5 μL volume resulted in more ordered distribution and tighter clustering of different MOE cell density data, which could be the result of overall improved values
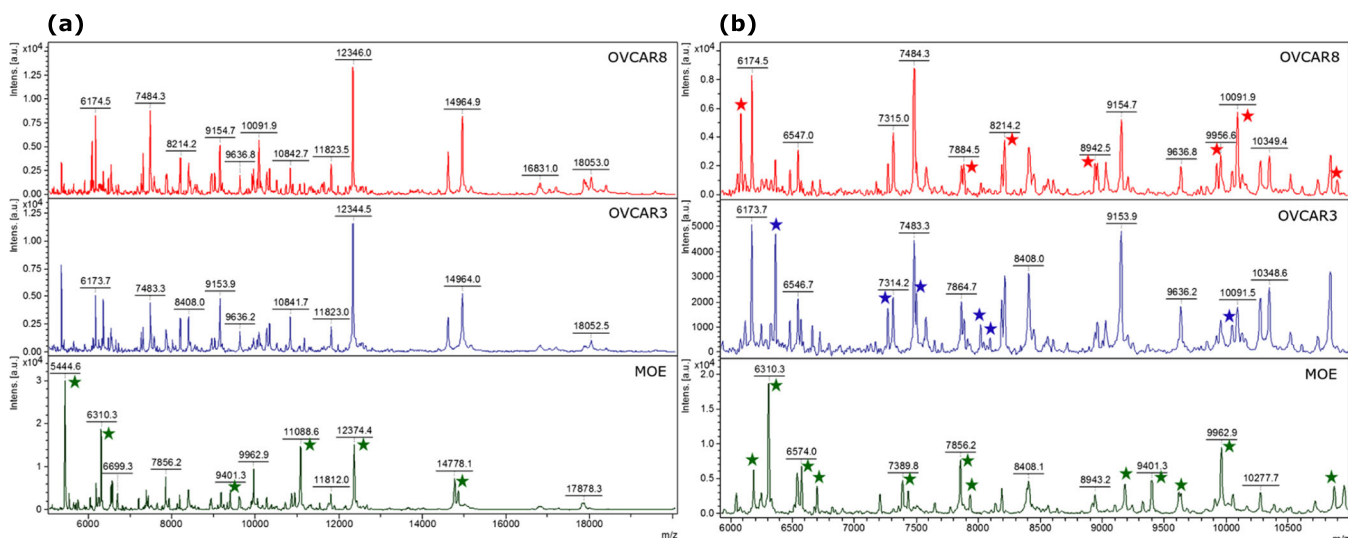
Figure 3. Comparison of OVCAR8, OVCAR3 and MOE cell lines fingerprints at 1250 cells/μL obtained with SA20 matrix/dried droplet application. (**a**) *m/z* range 5000–20,000 range and (**b**) *m/z* range 6000–11,000. Unique signals and ionization patterns characteristic to each cell line are marked by a star: red for OVCAR8, blue for OVCAR3, and green for MOE
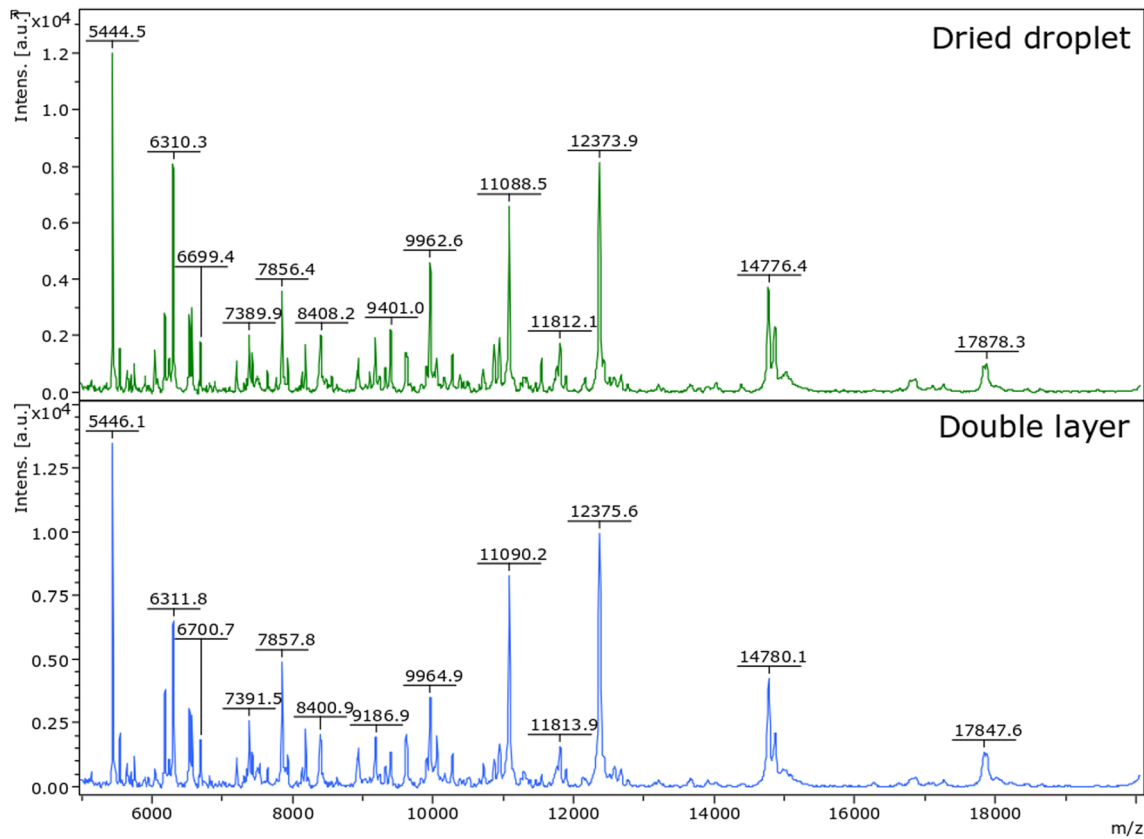
Figure 4. MOE ionization efficiency for dried droplet (upper green) and double layer (bottom blue) applications at 1250 cells/µL

for statistical parameters: p values, standard deviations, and coefficients of variation for peak intensities. Similar results were obtained for OVCAR3 cells as well (data not shown).

Comparison of Cancer and Normal Mammalian Cells Lines Fingerprints

After optimization of method parameters for sample preparation, the workflow was applied to different non-

cancerous and patient-derived cells (MOE, patient 1, patient 2 fallopian tube cells) and cancer (OVCAR3, OVCAR4-RFP, OVCAR8, and OVCAR8-RFP high-grade ovarian carcinomas) cell lines to check the method's ability to distinguish between different mammalian cell lines. The cell lines were tested with $n = 3$ technical replicates per population at 1250 cells/µL plating density and 1.5 µL plating volume. Differences in
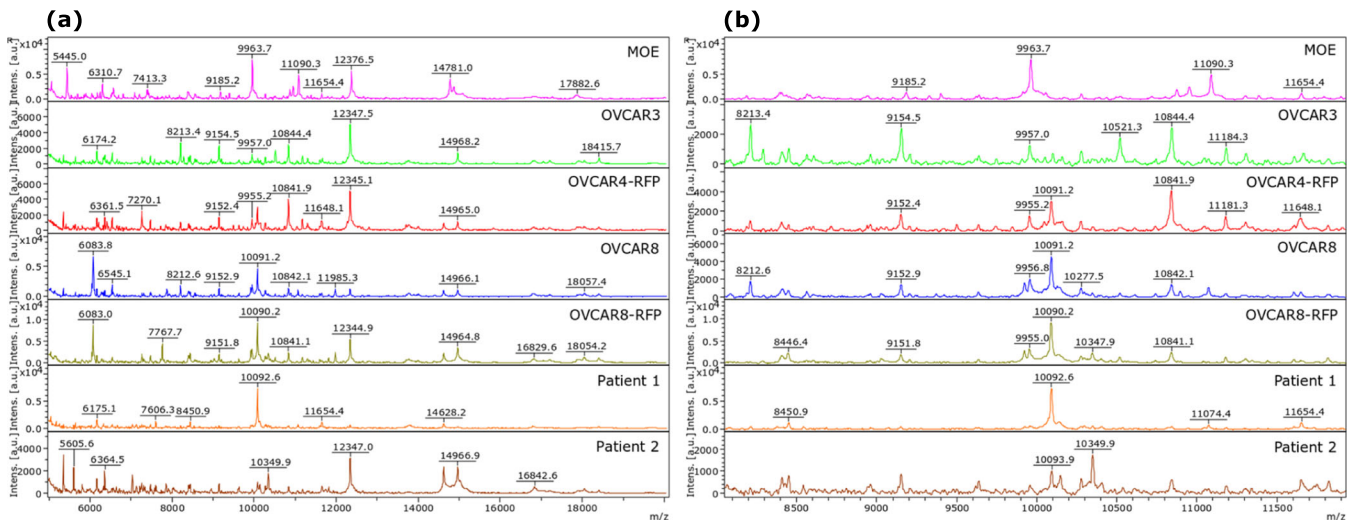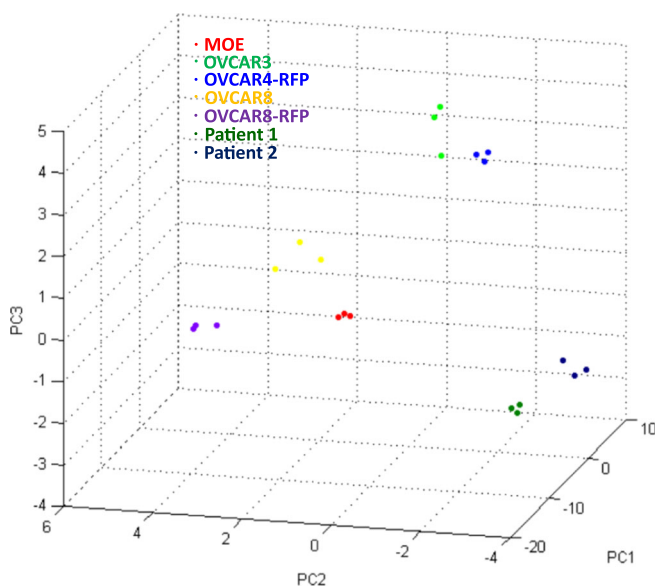

Figure 5. Representative cellular fingerprints of normal and cancer mammalian cell lines at 1250 cells/µL obtained with SA20 matrix/dried droplet application. (**a**) m/z range 5000–20,000. (**b**) m/z range 7000–11,900
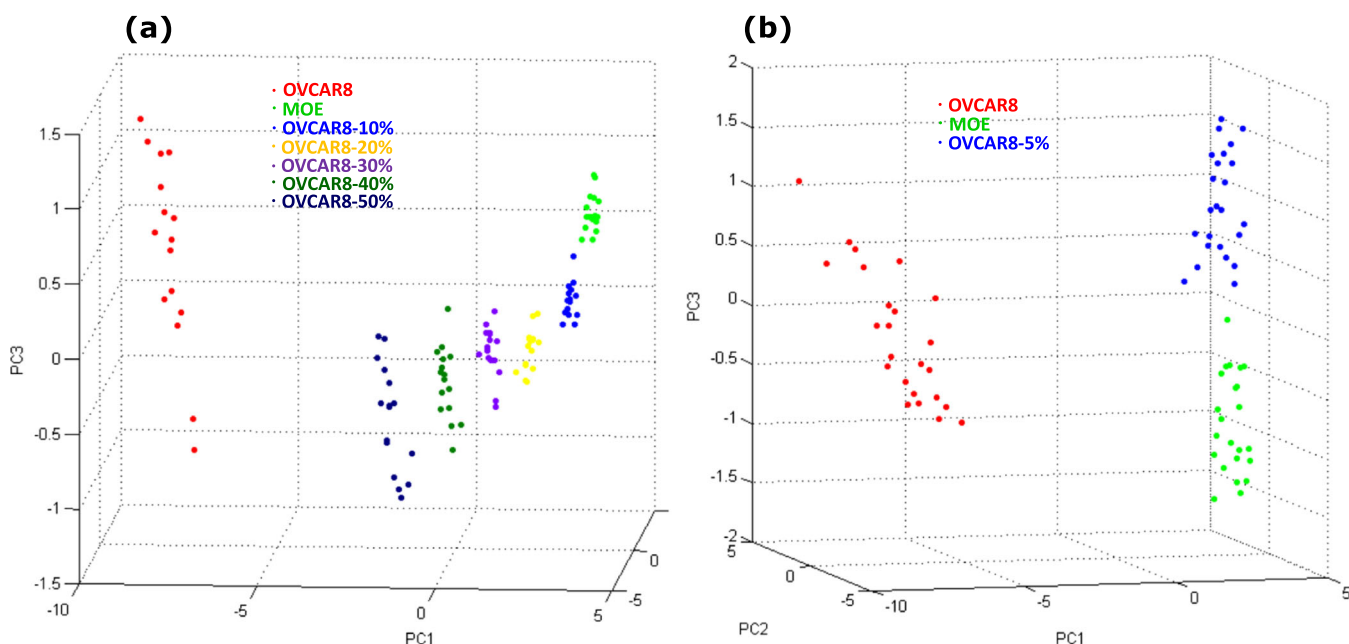
Figure 6. PCA distribution plot for non-cancerous and cancer cellular fingerprints at 1250 cells/μL obtained with SA20 matrix/dried droplet application

cellular fingerprints are obvious from quick spectral overview for all seven cell lines (Figures 5a and S13) and more differential peaks can be revealed upon closer inspection of their spectra (Figure 5b). As confirmed by PCA analysis (Figure 6), non-cancerous and cancer cell populations are unique and are cleanly separated from one another (cluster independently from one another) under optimized method parameters.

## Mixed Two-Component Cell Line Populations

For further method validation, two-component mixtures of normal MOE and high-grade ovarian carcinoma OVCAR8 cell lines were prepared, where cancer cell concentrations varied from 50 to 10% (with 10% increment) of OVCAR8 cells in the mixture. The limit of detection was queried using a range of lower concentrations of cancer cell line such as 9 to 1% (in 1% increments) of OVCAR8 cells per population. Pure and mixed cell populations were reproducibly tested at $n = 16$ (Figure 7a) and $n = 24$ (Figure S14) replicates per population to compensate for statistical errors and improve statistical outcomes. As seen from Figure 7a, mixed cell populations can be delineated from original cell lines by PCA with as low as 10% of cancer cells in the mixture. Further separation of mixed cell populations is detectable at 9–5% of cancer cells in the mixture (Figure S15a) but deteriorates at 5% cancer cells in the mixture (Figure 7b), precluding identification of lower cancer cell levels in mixtures (Figure S15b).

Below 5%, we used the Wilcoxon rank-sum tests (see "Methods"—"Statistical Analysis"), which showed that cancer specific features were determined to be significantly differentially expressed in mixtures down to 1% cancer cells. As the percentage of cancer cells increased (from 1 to 10%), the number of significantly differentially expressed features in the MALDI spectra increased (Figure 8). Under 5% cancer cells in the two-component mixtures, the spectra were dominated by signals from the non-cancerous (MOE) cells. Thus, it is not surprising that methods based on the global profile of the MALDI spectra, e.g., PCA, were not effective at differentiating cancer cells from a non-cancerous background. However, there

Figure 7. (a) PCA distribution plot for two-component MOE/OVCAR8 cell mixtures containing 50 to 10% cancer cells in the mixture at $n = 16$ replicates per population. (b) 5% of OVCAR8 cancer cells in the mixture is the limit of detection for PCA analysis at $n = 24$ replicates per population
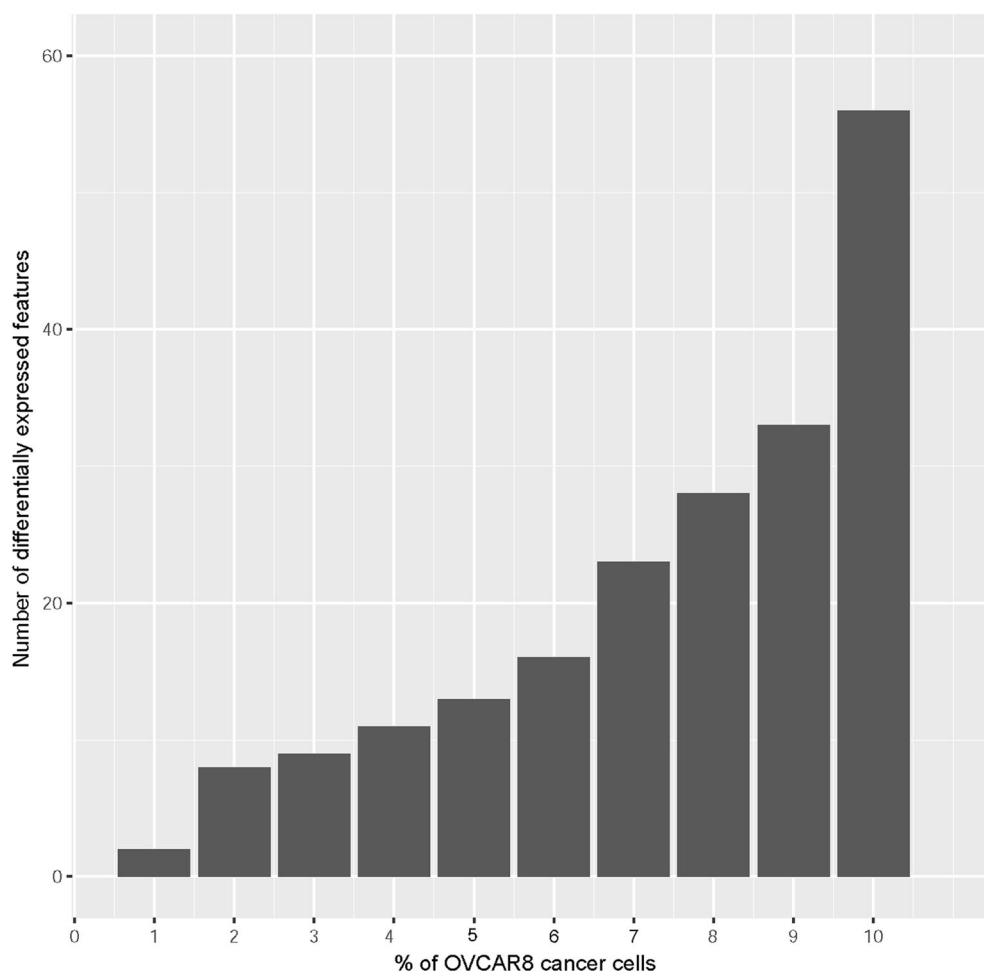
Figure 8. The number of features in the MALDI spectra shown to be statistically significant between 0% OVCAR8 cancer cells and 1–10% OVCAR8 cancer cells in mixed cell populations. Significance was calculated upon 24 replicates of each cell population and tested using the Wilcoxon sum-rank test with Bonferroni-correction multiple hypothesis correction. A significance threshold of 0.01 was applied to the corrected *p* values

may still be cancer specific features that are detectable in the MALDI spectra. Therefore, we determined which features were differentially expressed in the mixed cell population when compared to the non-cancerous cells. For each such feature, we calculated the minimum percent of cancer cells in the mixed sample at which the feature is still differentially expressed. We summarized these findings and show the total number of features, which are differentially expressed at each cancer cell percentage below 10% (Figure 8 and Figure S21). Even at 1% cancer cells, we observe two features that are shown to be differentially expressed. As expected, as the percentage of cancer cells in the mixed cell population increases, we can detect more cancer-specific features.

Biological and Technical Reproducibility

Stability of a cellular fingerprint is defined by its biological and technical reproducibility over a period of time. During the course of method optimization and experiments with mixed cell populations, we acquired multiple protein spectra of cells harvested at different dates or spectra were repeatedly recorded

on different dates from the same frozen stocks. These data became a small in-house data set and spectra acquired on different days from different batches were compared for assessment of biological and technical reproducibility of cellular fingerprints. Both MOE (Figure S16a and S16b) and OVCAR3 (Figure S17a and S17b) cells offer good biological reproducibility for three biological replicates. While some variations for peak patterns and intensities are observed in individual spectra, the unique signatures of cellular fingerprints are recognizable from peak values and peak patterns. Technical reproducibility is exemplified by four replicate spectra acquired on different days from the same frozen stocks of OVCAR3 (Figure S18a and S18b) and MOE (Figure S19a and S19b) cells. It also should be noted that cell growth media does not influence protein cellular fingerprints (Figure S20a–c).

Conclusion

Mammalian whole cell MALDI fingerprinting is a developing area of mass spectrometry but still requires attention in the

standardization of workflows and creation of universal databases to become a valuable tool for industrial and clinical applications. The goal of this study was to approach workflow standardization based on the whole cell MALDI fingerprinting techniques published in the literature and identification of workflow parameters that provide the most robust mammalian cellular fingerprint. We report the optimization of sample preparation parameters, such as cell media for cell lines under study, sample application technique, matrix/solvent composition, matrix/analyte density, and plating volume. Additionally, our work supports the use of the medium protein range (5–20 kDa) for providing robust protein spectra for mammalian fingerprinting. In our experiments on the AutoflexSpeed LRF model (Bruker), we found that 500 Hz laser frequency with 3-medium or 4-large SmartBeam II parameter settings resulted in efficient protein ionization with appropriate laser power and detector gain adjustments and those settings can be used for both manual and automated acquisitions. We also demonstrated the extension of mammalian fingerprints to heterogeneous cell mixtures with the ability to statistically discriminate 1% of cancer cells in an otherwise non-cancerous "healthy" cell background. We aim to further expand the utility of this method to detection of aberrant cells in host organisms in the future.

Acknowledgments

References

1. Wang, H., Kachman, M.T., Schwartz, D.R., Cho, K.R., Lubman, D.M.: Comprehensive proteome analysis of ovarian cancers using liquid phase separation, mass mapping and tandem mass spectrometry: a strategy for identification of candidate cancer biomarkers. Proteomics. **4**, 2476–2495 (2004)

2. Feldmann Jr., R.E., Bieback, K., Maurer, M.H., Kalenka, A., Burgers, H.F., Gross, B., et al.: Stem cell proteomes: a profile of human mesenchymal stem cells derived from umbilical cord blood. Electrophoresis. **26**, 2749–2758 (2005)

3. Rouleau, A., El Osta, M., Lucchi, G., Ducoroy, P., Boireau, W.: Immuno-MALDI-MS in human plasma and on-chip biomarker characterizations at the femtomole level. Sensors (Basel). **12**, 15119–15132 (2012)

4. Altuntas, E.G., Ayhan, K., Peker, S., Ayhan, B., Demiralp, D.O.: Purification and mass spectrometry based characterization of a pediocin produced by *Pediococcus acidilactici* 13. Mol. Biol. Rep. **41**, 6879–6885 (2014)

5. Agrawal, H., Joshi, R., Gupta, M.: Isolation, purification and characterization of antioxidative peptide of pearl millet (*Pennisetum glaucum*) protein hydrolysate. Food Chem. **204**, 365–372 (2016)

6. Anklesaria, J.H., Pandya, R.R., Pathak, B.R., Mahale, S.D.: Purification and characterization of CRISP-3 from human seminal plasma and its real-time binding kinetics with PSP94. J Chromatogr B Analyt Technol Biomed Life Sci. **1039**, 59–65 (2016)

7. Alcolea, P.J., Alonso, A., Garcia-Tabares, F., Mena, M.D.C., Ciordia, S., Larraga, V.: Proteome profiling of the growth phases of *Leishmania pifanoi* promastigotes in axenic culture reveals differential abundance of immunostimulatory proteins. Acta Trop. **158**, 240–247 (2016)

8. Anders, U., Schaefer, J.V., Hibti, F.E., Frydman, C., Suckau, D., Pluckthun, A., et al.: SPRi-MALDI MS: characterization and identification of a kinase from cell lysate by specific interaction with different designed ankyrin repeat proteins. Anal. Bioanal. Chem. **409**, 1827–1836 (2017)

9. Kelley, A.R., Perry, G., Bach, S.B.H.: Characterization of proteins present in isolated senile plaques from Alzheimer's diseased brains by MALDI-TOF MS with MS/MS. ACS Chem. Neurosci. **9**, 708–714 (2018)

10. Yang, J., Liu, Y., Xu, S., Lin, H., Meng, C., Lin, D.: Expression, purification and characterization of the full-length SmpB protein from *Mycobacterium tuberculosis*. Protein Expr. Purif. **151**, 9–17 (2018)

11. Ait-Belkacem, R., Berenguer, C., Villard, C., Ouafik, L., Figarella-Branger, D., Chinot, O., et al.: MALDI imaging and in-source decay for top-down characterization of glioblastoma. Proteomics. **14**, 1290–1301 (2014)

12. Scott, A.J., Jones, J.W., Orschell, C.M., MacVittie, T.J., Kane, M.A., Ernst, R.K.: Mass spectrometry imaging enriches biomarker discovery approaches with candidate mapping. Health Phys. **106**, 120–128 (2014)

13. Calligaris, D., Feldman, D.R., Norton, I., Olubiyi, O., Changelian, A.N., Machaidze, R., et al.: MALDI mass spectrometry imaging analysis of pituitary adenomas for near-real-time tumor delineation. Proc. Natl. Acad. Sci. U. S. A. **112**, 9978–9983 (2015)

14. Mourino-Alvarez, L., Iloro, I., de la Cuesta, F., Azkargorta, M., Sastre-Oliva, T., Escobes, I., et al.: MALDI-imaging mass spectrometry: a step forward in the anatomopathological characterization of stenotic aortic valve tissue. Sci. Rep. **6**, 27106 (2016)

15. Negrao, F., De, O.R.D.F., Jaeeger, C.F., Rocha, F.J.S., Eberlin, M.N., Giorgio, S.: Murine cutaneous leishmaniasis investigated by MALDI mass spectrometry imaging. Mol. BioSyst. **13**, 2036–2043 (2017)

16. Delcourt, V., Franck, J., Quanico, J., Gimeno, J.P., Wisztorski, M., Raffo-Romero, A., et al.: Spatially-resolved top-down proteomics bridged to MALDI MS imaging reveals the molecular physiome of brain regions. Mol. Cell. Proteomics. **17**, 357–372 (2018)

17. Alberts, D., Pottier, C., Smargiasso, N., Baiwir, D., Mazzucchelli, G., Delvenne, P., et al.: MALDI imaging-guided microproteomic analyses of heterogeneous breast tumors-a pilot study. Proteomics Clin Appl. 12, (2018)

18. Klein, O., Hanke, T., Nebrich, G., Yan, J., Schubert, B., Giavalisco, P., et al.: Imaging mass spectrometry for characterization of atrial fibrillation subtypes. Proteomics Clin Appl. e1700155 (2018)

19. Demirev, P.A., Ho, Y.P., Ryzhov, V., Fenselau, C.: Microorganism identification by mass spectrometry and protein database searches. Anal. Chem. **71**, 2732–2738 (1999)

20. Dalluge, J.J.: Mass spectrometry for direct determination of proteins in cells: applications in biotechnology and microbiology. Fresenius J. Anal. Chem. **366**, 701–711 (2000)

21. Clark, C.M., Costa, M.S., Sanchez, L.M., Murphy, B.T.: Coupling MALDI-TOF mass spectrometry protein and specialized metabolite analyses to rapidly discriminate bacterial function. Proc. Natl. Acad. Sci. U. S. A. **115**, 4981–4986 (2018)

22. Fenselau, C., Demirev, P.A.: Characterization of intact microorganisms by MALDI mass spectrometry. Mass Spectrom. Rev. **20**, 157–171 (2001)

23. Walker, J., Fox, A.J., Edwards-Jones, V., Gordon, D.B.: Intact cell mass spectrometry (ICMS) used to type methicillin-resistant *Staphylococcus aureus*: media effects and inter-laboratory reproducibility. J. Microbiol. Methods. **48**, 117–126 (2002)

24. Donohue, M.J., Smallwood, A.W., Pfaller, S., Rodgers, M., Shoemaker, J.A.: The development of a matrix-assisted laser desorption/ionization mass spectrometry-based method for the protein fingerprinting and identification of *Aeromonas* species using whole cells. J. Microbiol. Methods. **65**, 380–389 (2006)

25. Demirev, P.A., Fenselau, C.: Mass spectrometry for rapid characterization of microorganisms. Annu Rev Anal Chem (Palo Alto, Calif). **1**, 71–93 (2008)

26. Croxatto, A., Prod'hom, G., Greub, G.: Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. FEMS Microbiol. Rev. **36**, 380–407 (2012)

27. Demirev, P.A., Fenselau, C.: Mass spectrometry in biodefense. J. Mass Spectrom. **43**, 1441–1457 (2008)

28. Karlsson, R., Gonzales-Siles, L., Boulund, F., Svensson-Stadler, L., Skovbjerg, S., Karlsson, A., et al.: Proteotyping: proteomic characterization, classification and identification of microorganisms—a prospectus. Syst. Appl. Microbiol. **38**, 246–257 (2015)

29. Santos, I.C., Hildenbrand, Z.L., Schug, K.A.: Applications of MALDI-TOF MS in environmental microbiology. Analyst. **141**, 2827–2837 (2016)

30. Sandrin, T.R., Demirev, P.A.: Characterization of microbial mixtures by mass spectrometry. Mass Spectrom. Rev. **37**, 321–349 (2018)

31. Munteanu, B., Hopf, C.: Emergence of whole-cell MALDI-MS biotyping for high-throughput bioanalysis of mammalian cells? Bioanalysis. **5**, 885–893 (2013)

32. Vaidyanathan, S., Winder, C.L., Wade, S.C., Kell, D.B., Goodacre, R.: Sample preparation in matrix-assisted laser desorption/ionization mass spectrometry of whole bacterial cells and the detection of high mass (> 20 kDa) proteins. Rapid Commun. Mass Spectrom. **16**, 1276–1286 (2002)

33. Williams, T.L., Andrzejewski, D., Lay, J.O., Musser, S.M.: Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells. J. Am. Soc. Mass Spectrom. **14**, 342–351 (2003)

34. Zhang, X., Scalf, M., Berggren, T.W., Westphall, M.S., Smith, L.M.: Identification of mammalian cell lines using MALDI-TOF and LC-ESI-MS/MS mass spectrometry. J. Am. Soc. Mass Spectrom. **17**, 490–499 (2006)

35. Karger, A., Bettin, B., Lenk, M., Mettenleiter, T.C.: Rapid characterisation of cell cultures by matrix-assisted laser desorption/ionisation mass spectrometric typing. J. Virol. Methods. **164**, 116–121 (2010)

36. Buchanan, C.M., Malik, A.S., Cooper, G.J.: Direct visualisation of peptide hormones in cultured pancreatic islet alpha- and beta-cells by intact-cell mass spectrometry. Rapid Commun. Mass Spectrom. **21**, 3452–3458 (2007)

37. Ouedraogo, R., Flaudrops, C., Ben Amara, A., Capo, C., Raoult, D., Mege, J.L.: Global analysis of circulating immune cells by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. PLoS One. **5**, e13691 (2010)

38. Munteanu, B., von Reitzenstein, C., Hansch, G.M., Meyer, B., Hopf, C.: Sensitive, robust and automated protein analysis of cell differentiation and of primary human blood cells by intact cell MALDI mass spectrometry biotyping. Anal. Bioanal. Chem. **404**, 2277–2286 (2012)

39. Hanrieder, J., Wicher, G., Bergquist, J., Andersson, M., Fex-Svenningsen, A.: MALDI mass spectrometry based molecular phenotyping of CNS glial cells for prediction in mammalian brain tissue. Anal. Bioanal. Chem. **401**, 135–147 (2011)

40. Bondarenko, A., Zhu, Y., Qiao, L., Cortes Salazar, F., Pick, H., Girault, H.H.: Aluminium foil as a single-use substrate for MALDI-MS fingerprinting of different melanoma cell lines. Analyst. **141**, 3403–3410 (2016)

41. Marvin-Guy, L.F., Duncan, P., Wagniere, S., Antille, N., Porta, N., Affolter, M., et al.: Rapid identification of differentiation markers from whole epithelial cells by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry and statistical analysis. Rapid Commun. Mass Spectrom. **22**, 1099–1108 (2008)

42. Ouedraogo, R., Daumas, A., Ghigo, E., Capo, C., Mege, J.L., Textoris, J.: Whole-cell MALDI-TOF MS: a new tool to assess the multifaceted activation of macrophages. J. Proteome. **75**, 5523–5532 (2012)

43. Portevin, D., Pfluger, V., Otieno, P., Brunisholz, R., Vogel, G., Daubenberger, C.: Quantitative whole-cell MALDI-TOF MS fingerprints distinguishes human monocyte sub-populations activated by distinct microbial ligands. BMC Biotechnol. **15**, 24 (2015)

44. Dong, H., Shen, W., Cheung, M.T., Liang, Y., Cheung, H.Y., Allmaier, G., et al.: Rapid detection of apoptosis in mammalian cells by using intact cell MALDI mass spectrometry. Analyst. **136**, 5181–5189 (2011)

45. Munteanu, B., Meyer, B., von Reitzenstein, C., Burgermeister, E., Bog, S., Pahl, A., et al.: Label-free in situ monitoring of histone deacetylase drug target engagement by matrix-assisted laser desorption ionization-mass spectrometry biotyping and imaging. Anal. Chem. **86**, 4642–4647 (2014)

46. Amann, J.M., Chaurand, P., Gonzalez, A., Mobley, J.A., Massion, P.P., Carbone, D.P., et al.: Selective profiling of proteins in lung cancer cells from fine-needle aspirates by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin. Cancer Res. **12**, 5142–5150 (2006)

47. Maurer, K., Eschrich, K., Schellenberger, W., Bertolini, J., Rupf, S., Remmerbach, T.W.: Oral brush biopsy analysis by MALDI-ToF mass spectrometry for early cancer diagnosis. Oral Oncol. **49**, 152–156 (2013)

48. Gibb, S., Strimmer, K.: MALDIquant: a versatile R package for the analysis of mass spectrometry data. Bioinformatics. **28**, 2270–2271 (2012)

49. Henry, B., Mann, D.R.W.: On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Stat. **18**, 50–60 (1947)

50. Dunn, O.J.: Multiple comparisons among means. Journal of American Statistical Association. 52–64 (1961)

51. Valletta, E., Kucera, L., Prokes, L., Amato, F., Pivetta, T., Hampl, A., et al.: Multivariate calibration approach for quantitative determination of cell-line cross contamination by intact cell mass spectrometry and artificial neural networks. PLoS One. **11**, e0147414 (2016)

52. Feng, H.T., Sim, L.C., Wan, C., Wong, N.S., Yang, Y.: Rapid characterization of protein productivity and production stability of CHO cells by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. **25**, 1407–1412 (2011)

53. Feng, H.T., Wong, N.S., Sim, L.C., Wati, L., Ho, Y., Lee, M.M.: Rapid characterization of high/low producer CHO cells using matrix-assisted laser desorption/ionization time-of-flight. Rapid Commun. Mass Spectrom. **24**, 1226–1230 (2010)

54. Povey, J.F., O'Malley, C.J., Root, T., Martin, E.B., Montague, G.A., Feary, M., et al.: Rapid high-throughput characterisation, classification and selection of recombinant mammalian cell line phenotypes using intact cell MALDI-ToF mass spectrometry fingerprinting and PLS-DA modelling. J. Biotechnol. **184**, 84–93 (2014)

55. Schwamb, S., Munteanu, B., Meyer, B., Hopf, C., Hafner, M., Wiedemann, P.: Monitoring CHO cell cultures: cell stress and early apoptosis assessment by mass spectrometry. J. Biotechnol. **168**, 452–461 (2013)