Live single cell mass spectrometry reveals cancer-specific metabolic profiles of circulating tumor cells

Yasmine Abouleila[1,2,3,12], Kaoru Onidani[4,5,12], Ahmed Ali[1,2,3], Hirokazu Shoji[4,6], Takayuki Kawai[1,7,8], Chwee Teck Lim[9,10], Vipin Kumar[1], Shinobu Okaya[4], Ken Kato[6], Eiso Hiyama[2], Toshio Yanagida[1], Tsutomu Masujima[1], *Yoshihiro Shimizu[1], *Kazufumi Honda[4, 11]

[1]RIKEN Center for Biosystems Dynamics research (BDR), Osaka, Japan.
[2]Natural Science for Basic Research and Development, Hiroshima University, Hiroshima, Japan.
[3] Misr International University Research Center (MIU-RC), Cairo, Egypt.
[4]Department of Biomarkers for Early Detection of Cancer, National Cancer Center Research Institute, Tokyo, Japan.
[5]Department of Oral and Maxillofacial Surgery, Tokyo Dental College, Tokyo, Japan.
[6]Gastrointestinal Medical Oncology Division, National Cancer Center Hospital, Tokyo, Japan.
[7]Japan Science and Technology Agency, PRESTO, Saitama, Japan.
[8]Graduate School of Frontier Biosciences, Osaka University, Osaka, Japan.
[9]Department of Biomedical Engineering, National University of Singapore, Singapore.
[10]Biomedical Institute for Global Health Research and Technology, National University of Singapore, Singapore.
[11]Japan Agency for Medical Research and Development (AMED) CREST, Tokyo, Japan.
[12]These authors contributed equally to this work.

*Correspondence should be addressed to:
Kazufumi Honda, D. D. S., and Ph.D.
Department of Biomarkers for Early Detection of Cancer, National Cancer Center Research Institute, 5-1-1 Tsukiji Chuo-ku, Tokyo 104-0045, Japan
Tel: +81-3-3542-2511
Fax: +81-3-3542-3013
Email: khonda@ncc.go.jp

Yoshihiro Shimizu, Ph.D.
Laboratory for cell-free protein synthesis, RIKEN Center for Biosystems Dynamics (BDR), 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan

Tel: +81-6-6872-1885

Fax: +81-6-6872-1885

Email: yshimizu@riken.jp

Word count: 4566 words

Number of tables/figures: 5 figures and 1 table

Quantity of supporting information: 7 supporting figures, and 3 supporting tables

Summary

Recently, there has been an increased attention to circulating tumor cells (CTCs) analysis, also known as liquid biopsy, owing to its potential benefits in cancer diagnosis and treatment. CTCs are released from primary tumor lesions into the blood stream and eventually metastasize to distant body organs. However, a major hurdle with CTC analysis is their natural scarcity. Existing methods lack sensitivity, specificity or reproducibility required in CTC characterization and detection. Here, we report untargeted molecular profiling of single CTCs obtained from gastric cancer (GC) and colorectal cancer (CRC) patients, using live single cell mass spectrometry (LSC-MS) integrated with microfluidics-based cell enrichment technique. Using this approach, we demonstrated the difference in the metabolomic profile between CTCs originating from different cancer groups. Moreover, potential biomarkers were putatively annotated to be specific to each cancer type.

Keywords: Cancer biomarker; Circulating tumor cells (CTCs); Liquid biopsy; Mass spectrometry; Single cell analysis;

Abbreviations:

CTCs, circulating tumor cells; GC, gastric cancer; CRC, colorectal cancer; LSC-MS, live single-cell mass spectrometry; MS, mass spectrometry; EMT, epithelial-mesenchymal transition; FACS, fluorescent assisted cell sorting; DFF, dean flow fractionation; FWHM, full width half maximum; TIC, total ion count; PCA, principle component analysis; DA, discriminant analysis; KEGG, Kyoto encyclopedia of genes and genomics; HMDB, human metabolome database; %RSD, percent relative standard deviation; GPLs, glycerophospholipids; SLs, sterol lipids; FAs, fatty acyls; PLs, prenol lipids; TCA cycle, tricarboxylic acid cycle; MUFAs, monounsaturated fatty acids; PUFAs, polyunsaturated fatty acids; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PS, phosphatidylserine

Introduction

Recent studies have shown that cancer cells are highly heterogeneous on the single cell level, which might be one of the factors contributing to tumor relapse and increased incidence of metastasis[1-3]. However, studying cancer on the single cell level poses several challenges such as the difficulty of getting tumor samples in large scale clinical studies. Aside from imaging techniques for diagnostic purposes, the only reliable way of studying cancer cells in a clinical setting is to perform a surgical biopsy on the tumor site, a process that is inconvenient for the patient because of its invasive nature. Moreover, such invasive examination is unsuitable for large scale studies to monitor the biological behavior in real time, which can be altered under the pressure of various therapeutic agents.

An alternative procedure termed 'liquid biopsy' has been introduced as a possible method that can allow access to molecular information of cancer cells without the need for invasive procedures[4-8]. Liquid biopsy is a simple, non-invasive technique that targets circulating free nucleic acids (cfDNA and cfRNA) or circulating tumor cells (CTCs) which are often found in the peripheral blood of cancer patients. CTCs in particular, present wide-ranging features of phenotypic and genotypic variation depending on their primary tumor source[9]. CTCs are cells that shed from the primary tumor into the vasculature after undergoing epithelial-mesenchymal transition (EMT)[10], which is a phenotypic conversion of epithelial cells to gain more mesenchymal features allowing them to circulate through the blood stream and potentially metastasize to various body organs. Accordingly, they ought to have insightful and essential information on the primary tumor which will be of great importance for differential cancer diagnoses.

Previous CTC studies have mainly focused on enumerating their presence in peripheral blood, which can act as a predictive biomarker for early detection of tumor metastasis as well as monitoring the therapeutic efficacy and response of anticancer drugs[11]. Despite the prognostic value of CTC enumeration, it is still not enough to gain a comprehensive understanding of the tumor nature and its characteristics. Recent studies have reported genome and transcriptome analysis of single CTCs using next generation sequencer[12]. CTC sequencing proved to be an efficient liquid biopsy tool that can monitor the variation in gene expression among different cancer stages which could be used to investigate tumor origin, evolution and tumor progression during treatment. However, there's an untapped potential that lies in CTC metabolomic profiling which can shed light on the heterogeneity of cancer cells as well as the possible role played by CTCs in cancer progression and metastasis. In addition, analysis of CTCs provides the capability of getting a snapshot on the functional state of endogenous metabolites, thus, playing an essential role in filling the "genotype-phenotype gap". Furthermore, the metabolic profile

as a whole is considered as a complex biomarker which can be of great value in the field of oncology[9,13-18].

However, CTC metabolomic profiling is often a difficult task. The number of CTCs obtained per patient could greatly vary depending on the cancer type and the clinical stage. In general, the number of CTCs usually varies between zero to few hundreds or even more in approximately 7.5 mLs of peripheral blood[18,19]. Therefore, obtaining sufficient numbers of CTCs suitable for carrying out analysis still remains a great challenge. On this background, achieving efficient CTC enrichment with minimal sample loss is an issue for CTC metabolic profiling. Among the methods used for CTC enrichment are size-based filtration, immune-mediated capturing[20] and fluorescence assisted cell sorting (FACS)[21]. However, multi-step processes with harsh conditions could result in low throughput, CTC recovery, and cell viability. To overcome such problems, a microfluidics device that relies on differences in size and deformability of CTCs and blood cells for CTC separation was developed, which ensures accurate and effective label-free approach that maintains cell viability for further downstream analysis[22].

Another issue is selection of suitable analytical techniques. Conventional metabolomic techniques such as liquid/gas chromatography are unsuitable due to their insufficient sensitivity and inapplicability to small volumes associated with single cells. Live single-cell mass spectrometry (LSC-MS)[23] has been developed as a promising technique that has enough sensitivity for single cell metabolic profiling. In LSC-MS, a single cell is collected into a tapered glass micro-capillary under video-microscopy, which is then ionized and directly introduced to a mass spectrometer. LSC-MS was successfully applied to plant cells[24], mammalian cells[25], and CTCs[21], albeit with focus on targeted analysis for a limited number of compounds. To achieve molecular characterization of CTCs for future diagnostics, untargeted analysis must be achieved to gain a comprehensive metabolomic information about the primary tumor.

In this study, by integrating LSC-MS and microfluidics-based CTC enrichment technique, untargeted analysis was performed for CTCs obtained from two cancer types, gastric cancer (GC) and colorectal cancer (CRC) (**Figure 1**). We explored the possibility of discriminating between CTCs and lymphocytes obtained from the same patients, as well as discriminating between CTCs obtained from different cancer types and patients on the single cell level.

Materials and Methods
Patients and peripheral blood samples. Participants comprised 10 patients with advanced gastric cancer (GC) and colorectal cancer (CRC). The patients' information is

summarized in **Table 1**. Once informed consent was secured from these patients, their blood samples were processed for CTCs analysis. Peripheral blood was collected in 5 mL EDTA vacutainers or Streck tubes (TERUMO, Tokyo, Japan) and processed within 24 hours. This study was approved by the ethics committee of the National Cancer Center (2013-001; the term is during 2013-2022) and RIKEN (Kobe1 2017-07; the term is during 2017-2022).

CTCs enrichment. The ClearCell® FX system (ClearBridge Biomedics, Singapore) was used to capture and enrich CTCs from peripheral blood samples according to the manufacturer's protocol. Five mL of blood was mixed with 15 mL of red blood cell lysis buffer (G-Biosciences, St. Louis, MO, USA) at room temperature for 10 min. After incubation, the samples were centrifuged at 500 g for 10 min followed by aspiration of supernatant, and finally resuspended in 4.3 mL of suspension reagent supplied by the manufacturer. The samples were then processed through the ClearCell® FX system. The ClearCell® FX system is an automated CTCs enrichment system driven by the CTChip® FR1, a microfluidic biochip to isolate CTCs based on size, deformability and inertia. The isolation principle takes advantage of the inherent Dean vortex flows present in curvilinear channels for CTCs enrichment, termed dean flow fractionation (DFF)[22].

Single-cell sampling. The enriched CTCs samples were centrifuged at 500 g for 10 min followed by aspiration of supernatant and resuspended in 300 μL PBS. After reconstitution, CTC sample solution was transferred to a Cell Imaging Dish 145 μm, 35 mm × 10 mm (Eppendorf, Germany) for microscopic visualization. The samples were stained with fluorescent antibody, mouse anti-human CD45-FITC (130-080-202, Miltenyi, Germany, **Figure S1**). A single CD45 negative CTC was chosen under microscope and sucked into a Cellomics tip (CT-2, Humanix, Japan) using a micromanipulator and piston syringe. Similarly, a single CD45 positive lymphocytes were also sucked into a tip as a control. Single CTCs and lymphocytes were selected by morphology. Afterwards, the samples were frozen at -80°C until subsequent mass spectrometry analysis.

Sample preparation. The collected samples were thawed, and the organic solvent was introduced from the rear end of the Cellomics tip. The organic solvent consisted of 80% methanol, 10% dimethyl sulfoxide, and 0.1% formic acid. All the reagents used in the organic solvent were of LC-MS grade and were obtained from Sigma-Aldrich, USA. Ultra-sonication was then applied to the Cellomics tips containing the cells using a

homogenizer rod (UR-20P, Seiko CO., Japan) outfitted with an in-house attachment for the tips. Sonicating the tips before mass spectrometry measurement enhances the extraction of metabolites and lipids in addition to improving robustness by reducing tip plugging[21].

Mass spectrometry measurement. Mass spectrometry measurement was done using LTQ orbitrap Velos pro instrument (Thermo Fisher Scientific Inc., USA) equipped with a nanospray source (Nanospray Flex, Thermo Fisher Scientific Inc., USA). The distance between the Cellomics tip and the inlet of the instrument was set to 2 mm and the inlet capillary temperature was set to 200°C. The spray voltage was chosen to be 1 ~ 1.5 kV maintaining a spray current between 100 nA ~ 150 nA. The resolution was set to 100,000 FWHM. Since the mass spectrometer used is LTQ orbitrap Velos pro, which requires 90 minutes stabilization time after switching the polarity between positive and negative modes, positive mode was selected in this study to ensure wider metabolite and lipid coverage. Generally, positive ion mode exhibits overall more exhaustiveness than negative mode due to the higher efficiency of protonation compared to deprotonation process[26,27]. Despite using positive mode, several lipids that are usually detected in negative mode (i.e., Fatty acyls), could still be detected by our method as cationic ion-conjugates (potassium and sodium adducts)[28,29]. These adducts are readily formed and observed in electrospray ionization analyses[30]. For untargeted analysis, the instrument was set to selective ion monitoring (SIM) mode and it scanned from 100 m/z to 2,000 m/z in 50 m/z increments. This method of 'SIM stitching' allows for higher dynamic range as well as lower overall signal to noise ratios without compromising mass accuracy, thus improving the number of metabolites and lipids detected[31].

Data handling. The data generated from the mass spectrometer was converted from Thermo's raw proprietary format to text files using an in-house script, peak alignment was done using MarkerView® software (AB SCIEX, USA). Afterwards, text files were imported to R statistical software for further processing. Peaks with signal to noise ratio of less than 3 were eliminated. In addition, peaks appearing in less than 4 samples (10% of the total number of samples) have been removed as a pre-processing step to eliminate any noise or insignificant peaks. Furthermore, log transformation was applied to the spectra, followed by total ion count (TIC) normalization by using the MALDI-quant package[32].

Statistical analysis. To visualize the metabolomic differences across multiple samples in

a reduced dimensional space, a supervised approach utilizing principle component analysis followed by discriminant analysis (PCA-DA) was performed using MarkerView® software. The loading plots for all PCA-DA performed are shown in **Figure S2-S4**. Furthermore, to discern the possible unique peaks to the different groups, Welch's t-test was done between CTCs and lymphocytes (control) and between GC CTCs and CRC CTCs, the test was done on R statistical software. Peaks with more than 1 log2 foldchange or less than -1 log2 foldchange and with p-value less than 0.05 were selected (**Figure S5**). For peak identification, we followed a two-pronged approach depending on the accurate mass (*m/z* less than 5 ppm) and isotopic pattern following the metabolomics standard initiative[33]. The significant peaks were run through an in-house script that matches possible peaks against Kyoto encyclopedia of genes and genomics (KEGG)[34], human metabolome database (HMDB)[35], and LIPID MAPS structure database[36]. All the annotation of putatively identified lipids was done by using the shorthand lipid notation system suggested by Liegbisch *et al.*[37], except for eicosanoids lipids, where common names were used. Lipid candidates were annotated on the fatty acyl/alkyl position level, since exact mass measurements are incapable of determining the stereochemistry, and double bond geometry.

Results

Identification of unique metabolic profile between CTCs and lymphocytes.

   We enriched the CTCs using microfluidics-based enrichment technique and cytosolic metabolite was harvested from single cell CTCs according to the processes which was described in Materials and Methods (**Figure 1** and **Figure S1**). To examine whether LSC-MS can detect the metabolomic profile difference between a single CTC and other control cells, we analyzed CTCs and lymphocytes collected from the same patient. After peak alignment and normalization, we performed supervised PCA followed by DA to visualize the samples in a reduced dimensional space. (**Figure 2** and **Figure S6**). The supervised PCA-DA showed clear clustering behavior between three groups; CTCs, lymphocytes and the organic solvent blank per cancer (**Figure 2**) and per patient (**Figure S6**). Due to the limited number of cells obtained per patient, supervised PCA-DA was only performed on patients' samples, in which the number of CTCs and lymphocytes analyzed is more than 2. Detailed information about the patients and obtained samples is shown in **Table 1**.

Possible biomarkers for CTCs.

   It is known that there are aberrant profiles with the presence of unique metabolites or

lipids in malignant cells compared with normal cells[38-40]. Accordingly, we investigated the biomarkers found in all CTCs compared to the control lymphocyte cells. Since the sample numbers have unequal size and variance, Welch's t-test was used to determine which metabolites are present only in CTCs across all patients. The comparison was performed between CTCs (n = 22) and lymphocytes collected from all patients regardless of a cancer type (n = 16), after subtracting solvent peaks from both groups. The resulting peaks were filtered by selecting statistically significant peaks (p-value < 0.05).

In total, 119 peaks were putatively identified that were specific to CTCs by using their respective exact mass values with mean error of less than 5 ppm (**Figure 3**). Interestingly enough, out of the 119 peaks found, 75 were identified as glycerophospholipids. Considering their function in the cell as a structural component in biological membranes, the possible unique membrane profile common to CTCs can be shown. A table with the putatively identified metabolites and lipids unique to CTCs with p-value, percent relative standard deviation (%RSD), chemical formula and class for each compound is shown in **Table S1**.

Single cell profiling of gastric cancer and colorectal cancer CTCs.

Several studies have suggested that unique metabolomic profiles were observed in the primary site of different cancer types[41-43]. However, to the best of our knowledge, the metabolomic profiles of single CTCs in several cancer types has never been reported. In addition, CTCs are easily accessible from peripheral blood and their molecular characterization may have significant prognostic and diagnostic values[44-46], especially if metabolic differences can be discerned between CTCs originating from different cancer types. To investigate this further, we compared the CTCs obtained from GC patients to those obtained from CRC patients to visualize the unique cellular profiles of each group.

To this end, supervised PCA-DA was performed on GC and CRC CTCs as well as the blank. Despite the heterogeneity caused by samples obtained from different patients, significant clustering could still be observed where CTCs appear to cluster in two distinct groups corresponding to their cancer type as shown in **Figure 4a**, suggesting the potential of CTC metabolome characterization as a future tool for cancer diagnosis.

Upon further inspection of the data, a trend was noticed in the frequency of peaks distributed along the *m/z* scale, especially in the case of GC, in which a higher incidence of high *m/z* peaks was detected. This was demonstrated by comparing the histograms of the average spectra of GC CTCs versus those of CRC, as shown in **Figure 4b**. Since most metabolites have relatively low molecular weight, the increased incidence of relatively high molecular weight peaks in GC CTCs suggest a distinctive metabolic "finger print"

of this cancer that most likely involves a higher distribution of lipids which can be used in the future as a biomarker for GC.

Possible biomarkers for colorectal cancer and gastric cancer CTCs.
To identify possible metabolites or lipids that are unique to GC or CRC CTCs, a Welch's t-test was done on the two groups (n = 9 and n = 13, respectively). In total, 155 significant peaks were extracted according to their log2 based foldchange (more than 1 or less than -1 log2 foldchange) and p-value less than 0.05 (**Figure 5a, S5**). Among those peaks, 69 were present in GC CTCs, while 86 peaks were suggested to be specific to CRC CTCs. A histogram of the *m/z* distribution of potential biomarkers to each cancer type is shown in **Figure 5b**. A summary with the putatively identified metabolites and lipids unique to GC CTCs and CRC CTCs with p-value, %RSD, chemical formula and class for each compound is shown in **Table S2** and **Table S3**, respectively.

Among the statistically significant peaks found, acyl carnitine metabolites as well as sterol lipids were elevated in CRC. Furthermore, eicosanoids were also observed to be more abundant in CRC CTCs, which is further corroborated with other studies done on this cancer type[47]. On the other hand, glycerophospholipids were noticed to be elevated in GC CTCs which matches with recent literature[48].

Discussion

Previous studies in genomics[12], transcriptomics[49] and proteomics[50] succeeded in depicting the effective role of CTC analyses in monitoring tumor progression and prognosis as well as its clinical impact. Since metabolomics is the final frontier of omics, it is essential to study the metabolic phenotypes of CTCs to gain a compressive understanding on the primary and metastatic tumor biology. However, metabolomics still faces several challenges arising from its inherent complexity and variability especially on the single cell level. As a result, dimensional reduction techniques followed by discriminant analysis are required to simplify obtained data and reveal significant biological meanings. This could possibly be achieved by using PCA followed by discriminant analysis (DA) a process we term PCA-DA. In which, PCA provides unsupervised dimensionality reduction that simplifies the data and enhances its visualization[51], while discriminant analysis highlights the metabolomic differences (i.e., metabolic fingerprint) among different cell types originating from different cancer phenotypes[52]. Nonetheless, there are limitations associated with discriminant analysis as it may provide overly optimistic results, mainly due to the lack of suitable statistical validation and use by non-experts without considering its potential pitfalls[53].

In this paper, our results demonstrated the metabolic differences between CTCs and lymphocytes derived from the same patients (**Figure 2**). In which, PCA-DA showed clear clustering behavior between CTCs and lymphocytes per cancer, which gives credence to the theory that despite cellular heterogeneity, overall metabolomic differences between cancer cells and control cells are evident even on a single cell scale. Furthermore, the metabolic difference among different cancer types (GC and CRC) was depicted in **Figure 4a**, where CTCs clustered into two distinct groups corresponding to their cancer type, suggesting the potential of CTC metabolome characterization as a future tool for cancer diagnosis. The difference in phenotype was further confirmed by a histogram, showing the discrepancy in the frequency of peaks distribution across the *m/z* scale of different cancer types (**Figure 4b**).

The important role of lipids' synthesis and metabolism in tumorigenesis and tumor progression has been showcased in previous studies[40,54–56]. Hence, targeting specific lipids or blocking certain pathways is a promising therapeutic strategy for cancer treatment[57]. Accordingly, we focused in our study on detecting mainly lipids to explore their role in both cancer types (GC and CRC). This was done by optimizing and tuning our mass spectrometry method to focus on the higher *m/z* ranges and including DMSO in our ionization solvent. DMSO is known to improve the overall ionization performance of lipids in mass spectrometry[58]. In addition, applying ultra-sonication on the capillaries containing single cells proved to improve exhaustiveness and lipid coverage[21]. Consequently, most of the potential biomarkers detected in this study were mainly lipids such as: glycerophospholipids (GPLs), fatty acyls (FAs), acyl carnitines, sterol lipids (SLs) and prenol lipids (PLs). In addition, different abundance and distribution of lipid classes took place in each cancer type (**Figure S7**). As for the case of CRC CTCs, SLs were highly elevated compared to other lipid classes. Previous studies showed the correlation between high levels of SLs and distant metastasis in CRC patients[59], which is consistent with our study subjects. Moreover, since chronic inflammation is associated with the development of CRC, it was expected to detect eicosanoids in relatively high levels. In which, eicosanoids are believed to affect CRC development and progression by inflammation induction, regulation of cellular oxidative stress, and alteration of membrane dynamics[60]. Furthermore, low levels of FAs and GPLs were detected in CRC CTCs. This could be due to the increased catabolism of GPLs to FAs followed by the subsequent degradation of FAs during the β -oxidation process[61,62]. Since β-oxidation usually takes place in the mitochondria, FAs then bind to acylcarnitines (carriers) that transport FAs to the mitochondria. Therefore, higher levels of acylcarnitines are thought to be specific to CRC (in contrast to GC)[63], which is supported by our results.

In GC CTCs, FAs and GPLs were preeminent. This is probably due to the increased de novo synthesis and elongation of FAs and GPLs[64–67]. In which, GPLs synthesis was proved to be a key factor in cancer proliferation as its responsible mainly for membrane and energy production[68]. Since GPLs synthesis requires acetyl-CoA, and citrate is an acetyl-CoA donor for this process, higher lipid metabolism can be associated with elevated tricarboxylic acid cycle (TCA cycle) activity in GC[69,70]. As previously mentioned, GC exhibited overall higher levels of FAs than CRC. Specifically, saturated fatty acids (SFAs) were found to be relatively elevated in comparison to monounsaturated fatty acids (MUFAs) and polyunsaturated fatty acids (PUFAs)[71]. Recent studies demonstrated that SFAs can act as a source of energy in tumor cells, enhancing tumor proliferation and survival and also might affect tumor resistance to treatment[72,73]. Withal, numerous metabolites/lipids were detected to be common in all CTC samples vs lymphocytes. Chief among them are several GPLs that are believed to have biological significance in cancer physiology including, PC(32:1), PC(34:1), PS(38:5), PE(38:6), PC(32:3) and PC(34:2)[5,74]. Detailed information about identified lipids and metabolites for each cancer group are shown in Supplementary **Tables ST1-3**. As a result, lipid profiling of CTCs may be used as biomarkers for the diagnosis and prognosis of CRC and GC as well as novel targets for their treatment. However, the specific impact of lipids on tumor development, progression and metastasis is not yet fully understood[75], which highlights the need for precise monitoring of any alterations in lipid metabolism in cancer cells. Along with our results, this presents a new avenue for the diagnostics and treatment of cancer.

Previous studies utilized human serum and plasma to perform untargeted analysis and discriminate between different cancer types[76–79]. However, to the best of our knowledge, our study is the first to achieve this utilizing CTCs at the single cell level. Besides the previously mentioned advantages of CTC molecular profiling, our method can be used to investigate new potential biomarkers on the single cell level that could not be detected using the conventional methods, highlighting the novelty of this method in cancer research. Nevertheless, it is worth noting that the aim of this paper is not to provide a be all and end all definitive list of biomarkers that can be used in clinical diagnosis today. Instead, we aim to highlight the untapped potential of CTC molecular characterization in both clinical and research settings using LSC-MS. Moreover, despite using exact mass with a high degree of accuracy (less than 5 ppm) in this paper for annotation, additional verification techniques must be used such as tandem mass spectrometry (MS/MS), capillary electrophoresis or HPLC for positive compound identification. However, performing MS/MS on signals obtained from a single cell in a robust manner is still a

challenging task that needs to be addressed with improvements in the instrumentation itself or by incorporating an enrichment step before analysis.

In summary, untargeted analysis of human derived CTCs at the single cell level was performed for the first time utilizing LSC-MS. CTCs and lymphocytes obtained from the same patient could be successfully discriminated as well as CTCs of different cancer types. In addition, the possible role played by lipids and higher molecular weight compounds distribution on the single cell level in classifying different cancer types based on their "metabolic fingerprint" was highlighted. Several promising biomarkers were putatively annotated that are mostly specific to GC CTCs, CRC CTCs, and CTCs in general. Finally, due to the scarcity of human derived samples, especially in inherently rare cells such as CTCs, we hope that this paper and its results spur the much-needed collaborative efforts to upscale CTC characterization to established large scale studies that will possibly aid in the clinical applications in the near future.

Disclosure Statement

The authors declare no financial nor non-financial conflict of interests.

References

(1)   Fisher R, Pusztai L, Swanton C. Cancer Heterogeneity: Implications for Targeted Therapeutics. Br J Cancer. 2013;108:479-485.

(2)   Dagogo-Jack I, Shaw A T. Tumour Heterogeneity and Resistance to Cancer Therapies. Nat Rev Clin Oncol. 2018;15:81-94.

(3)   Alizadeh A A, Aranda V, Bardelli A, et al. Toward Understanding and Exploiting Tumor Heterogeneity. Nat Med. 2015;21:846-853.

(4)   Karachaliou N, Mayo-de-Las-Casas C, Molina-Vila, et al. Real-Time Liquid Biopsies Become a Reality in Cancer Treatment. Ann Transl Med. 2015;3:36.

(5)   Hofman P. Liquid Biopsy for Early Detection of Lung Cancer. Curr Opin Oncol.

2017;29:73-78.

(6)   Pantel K. Blood-Based Analysis of Circulating Cell-Free DNA and Tumor Cells for Early Cancer Detection. PLoS Med. 2016;13:e1002205.

(7)   Yőrűker E E, Holdenrieder S, Gezer U, Blood-Based Biomarkers for Diagnosis, Prognosis and Treatment of Colorectal Cancer. Clin Chim Acta. 2016;455:26-32.

(8)   Ilié M, Hofman P. Pros: Can Tissue Biopsy Be Replaced by Liquid Biopsy? Transl Lung Cancer Res. 2016;5:420-423.

(9)   Zhou L, Dicker D T, Matthew E, et al. Circulating Tumor Cells: Silent Predictors of Metastasis. F1000Res. 2017;6:F1000 Faculty Rev-1445.

(10)   Jie X X, Zhang X Y, Xu C J. Epithelial-to-Mesenchymal Transition, Circulating Tumor Cells and Cancer Metastasis: Mechanisms and Clinical Applications. Oncotarget. 2017;8:81558-81571.

(11)   Toss A, Mu Z, Fernandez S, et al. CTC Enumeration and Characterization: Moving toward Personalized Medicine. Ann Transl Med. 2014;2:108.

(12)   Zhu Z, Qiu S, Shao K. et al. Progress and Challenges of Sequencing and Analyzing Circulating Tumor Cells. Cell Biol Toxicol. 2018;34:405-415.

(13)   Krebs M G, Hou J M, Ward T H, et al. Circulating Tumour Cells: Their Utility in Cancer Management and Predicting Outcomes. Ther Adv Med Oncol. 2010;2:351-365.

(14)   Maas M, Hegemann M, Rausch S, et al. Circulating Tumor Cells and Their Role in Prostate Cancer. Asian J Androl. 2017;19:1-8

(15)   Gallo M, De Luca A, Maiello M R, et al. Clinical Utility of Circulating Tumor Cells in Patients with Non-Small-Cell Lung Cancer. Transl Lung Cancer Res. 2017;6:486-498.

(16)   Polasik A, Tzschaschel M, Schochter F, et al. Circulating Tumour Cells, Circulating Tumour DNA and Circulating MicroRNA in Metastatic Breast Carcinoma - What Is the Role of Liquid Biopsy in Breast Cancer? Geburtshilfe Frauenheilkd. 2017;77:1291-1298.

(17)   Kong S L, Liu X, Suhaimi N M, et al. Molecular Characterization of Circulating Colorectal Tumor Cells Defines Genetic Signatures for Individualized Cancer Care. Oncotarget. 2017;8:68026-68037.

(18)   Gold B, Cankovic M, Furtado L V, et al. Do Circulating Tumor Cells, Exosomes, and Circulating Tumor Nucleic Acids Have Clinical Utility? A Report of the Association for Molecular Pathology. J Mol Diagn. 2015;17:209?224.

(19)   Galletti G, Portella L, Tagawa S T, et al. Circulating Tumor Cells in Prostate Cancer Diagnosis and Monitoring: An Appraisal of Clinical Potential. Mol

Diagn Ther. 2014;18:389-402.

(20)   Millner L M, Linder M W, Valdes R. Circulating Tumor Cells: A Review of Present Methods and the Need to Identify Heterogeneous Phenotypes. Ann Clin Lab Sci. 2013;43:295-304.

(21)   Hiyama E, Ali A, Amer S, et al. Direct Lipido-Metabolomics of Single Floating Cells for Analysis of Circulating Tumor Cells by Live Single-Cell Mass Spectrometry. Anal Sci. 2015;31:1215-1217.

(22)   Hou H W, Warkiani M E, Khoo B L, et al. Isolation and Retrieval of Circulating Tumor Cells Using Centrifugal Forces. Sci Rep. 2013;3:1259.

(23)   Masujima T. Live Single-Cell Mass Spectrometry. Anal Sci. 2009;25:953-960.

(24)   Fujii T, Matsuda S, Tejedor M L, et al. Direct Metabolomics for Plant Cells by Live Single-Cell Mass Spectrometry. Nat Protoc. 2015;10:1445-1456.

(25)   Ali A, Abouleila Y, Amer S, et al. Quantitative Live Single-Cell Mass Spectrometry with Spatial Evaluation by Three-Dimensional Holographic and Tomographic Laser Microscopy. Anal Sci. 2016;32:125-127.

(26)   Zhong H, Fang C, Fan Y, et al. Lipidomic Profiling Reveals Distinct Differences in Plasma Lipid Composition in Healthy, Prediabetic, and Type 2 Diabetic Individuals. Gigascience. 2017;6:1–12.

(27)   Yanes O, Tautenhahn R, Patti G J, et al. Expanding Coverage of the Metabolome for Global Metabolite Profiling. Anal Chem. 2011;83:2152–2161.

(28)   Murphy R C, Axelsen P H. Mass Spectrometric Analysis of Long-Chain Lipids. Mass Spectrom Rev. 2011;30:579–599.

(29)   Wang C, Wang M, Han X. Applications of Mass Spectrometry for Cellular Lipid Analysis. Mol Biosyst. 2015;11:698–713.

(30)   Kind T, Fiehn O. Advances in Structure Elucidation of Small Molecules Using Mass Spectrometry. Bioanal Rev. 2010;2:23–60.

(31)   Southam A D, Payne T G, Cooper H J, et al. Dynamic Range and Mass Accuracy of Wide-Scan Direct Infusion Nanoelectrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry-Based Metabolomics Increased by the Spectral Stitching Method. Anal Chem. 2007;79:4595-4602.

(32)   Gibb S, Strimmer K. MALDIquant: A Versatile R Package for the Analysis of Mass Spectrometry Data. Bioinformatics. 2012;28:2270-2271.

(33)   Goodacre R, Broadhurst D, Smilde A. K, et al. Proposed Minimum Reporting Standards for Data Analysis in Metabolomics. Metabolomics. 2007;3:231–241.

(34)   Kanehisa M, Goto S, Hattori M, et al. From Genomics to Chemical Genomics: New Developments in KEGG. Nucleic Acids Res. 2006;34:D354-357.Wishart

(35)   D S, Feunang Y D, Marcu A, et al. HMDB 4.0: The Human Metabolome Database for 2018. Nucleic Acids Res. 2018;46:D608-617.

(36)   Sud M, Fahy E, Cotter D, et al. LMSD: LIPID MAPS Structure Database. Nucleic Acids Res. 2007;35:D527-532.

(37)   Liebisch G, Vizcaino J A, Kofeler H, et al. Shorthand Notation for Lipid Structures Derived from Mass Spectrometry. J Lipid Res. 2013;54:1523-1530.

(38)   Griffin J L, Shockcor J P. Metabolic Profiles of Cancer Cells. Nat Rev Cancer. 2004;4:551-561.

(39)   Koppenol W H, Bounds P L, Dang C V. Otto Warburg's Contributions to Current Concepts of Cancer Metabolism. Nat Rev Cancer. 2011;11:325-337.

(40)   Santos C R, Schulze A. Lipid Metabolism in Cancer. FEBS J. 2012;279:2610-2623.

(41)   Kim H Y, Lee K M, Kim S H, et al. Comparative Metabolic and Lipidomic Profiling of Human Breast Cancer Cells with Different Metastatic Potentials. Oncotarget. 2016;7:67111-67128.

(42)   Boroughs L K, DeBerardinis R J. Metabolic Pathways Promoting Cancer Cell Survival and Growth. Nat Cell Biol. 2015;17:351-359.

(43)   Brown D G, Rao S, Weir T L, et al. Metabolomics and Metabolic Pathway Networks from Human Colorectal Cancers, Adjacent Mucosa, and Stool. Cancer Metab. 2016;4:11.

(44)   Jacob K, Sollier C, Jabado N. Circulating Tumor Cells: Detection, Molecular Profiling and Future Prospects. Expert Rev Proteomics. 2007;4:741-756.

(45)   Polzer B, Medoro G, Pasch S, et al. Molecular Profiling of Single Circulating Tumor Cells with Diagnostic Intention. EMBO Mol Med. 2014;6:1371-1386.

(46)   Aktas B, Kasimir-Bauer S, Heubner M, et al. Molecular Profiling and Prognostic Relevance of Circulating Tumor Cells in the Blood of Ovarian Cancer Patients at Primary Diagnosis and after Platinum-Based Chemotherapy. Int J Gynecol Cancer. 2011;21:822-830.

(47)   Xu J, Chen Y, Zhang R, et al. Global Metabolomics Reveals Potential Urinary Biomarkers of Esophageal Squamous Cell Carcinoma for Diagnosis and Staging. Sci Rep. 2016;6:35010.

(48)   Lario S, Ramírez-Lázaro M J, Sanjuan-Herráez D, et al. Plasma Sample Based Analysis of Gastric Cancer Progression Using Targeted Metabolomics. Sci Rep. 2017;7:17774.

(49)   Welty C J, Coleman I, Coleman R, et al. Single Cell Transcriptomic Analysis of Prostate Cancer Cells. BMC Mol Biol. 2013;14:6.

(50) Sinkala E, Sollier-Christen E, Renier C, et al. Profiling Protein Expression in Circulating Tumour Cells Using Microfluidic Western Blotting. Nat Commun. 2017;8:14622.

(51) Worley B, Powers R. Multivariate Analysis in Metabolomics. Curr Metabolomics. 2013;1:92-107.

(52) Jombart T, Devillard S, Balloux F. Discriminant Analysis of Principal Components: A New Method for the Analysis of Genetically Structured Populations. BMC Genet. 2010;11:94.

(53) Gromski P S, Muhamadali H, Ellis D I, et al. A Tutorial Review: Metabolomics and Partial Least Squares-Discriminant Analysis--a Marriage of Convenience or a Shotgun Wedding. Anal Chim Acta. 2015;879:10-23.

(54) Beloribi-Djefaflia S, Vasseur S, Guillaumond F. Lipid Metabolic Reprogramming in Cancer Cells. Oncogenesis. 2016;5:e189.

(55) Corbet C, Feron O. Emerging Roles of Lipid Metabolism in Cancer Progression. Curr Opin Clin Nutr Metab Care. 2017;20:254-260.

(56) Ray U, Roy S S. Aberrant Lipid Metabolism in Cancer Cells - the Role of Oncolipid-Activated Signaling. FEBS J. 2018;285:432-443.

(57) Liu Q, Luo Q, Halim A, et al. Targeting Lipid Metabolism of Cancer Cells: A Promising Therapeutic Strategy for Cancer. Cancer Lett. 2017;401:39-45.

(58) Szabó P T, Kele Z. Electrospray Mass Spectrometry of Hydrophobic Compounds Using Dimethyl Sulfoxide and Dimethylformamide as Solvents. Rapid Commun. Mass Spectrom. 2001;15:2415–2419.

(59) Yan G, Li L, Zhu B, et al. Lipidome in Colorectal Cancer. Oncotarget. 2016;7:33429-33439.

(60) Tuncer S, Banerjee S. Eicosanoid Pathway in Colorectal Cancer: Recent Updates. World J Gastroenterol. 2015;21:11748-11766.

(61) Halama A, Guerrouahen BS, Pasquier J, Diboun I, et al. A. Metabolic Signatures Differentiate Ovarian from Colon Cancer Cell Lines. J Transl Med. 2015;13.

(62) Schooneman MG, Vaz FM, Houten M, et al. Acylcarnitines. Diabetes. 2013; 62:1–8.

(63) Jung J, Jung Y, Bang EJ, et al. Noninvasive Diagnosis and Evaluation of Curative Surgery for Gastric Cancer by Using NMR-Based Metabolomic Profiling. Ann Surg Oncol. 2014;21:S736-742.

(64) Huang C, Freter C. Lipid Metabolism, Apoptosis and Cancer Therapy. Int J Mol Sci. 2015;16:924–949.

(65) Xiao S, Zhou L. Gastric Cancer: Metabolic and Metabolomics Perspectives

(Review). International Journal of Oncology. 2017;51:5–17.

(66) Kusakabe, T.; Nashimoto, A.; Honma, K.; Suzuki, T. Fatty Acid Synthase Is Highly Expressed in Carcinoma, Adenoma and in Regenerative Epithelium and Intestinal Metaplasia of the Stomach. Histopathology. 2002;40:71–79.

(67) Perrotti F, Rosa C, Cicalini I, et al. Advances in Lipidomics for Cancer Biomarkers Discovery. Int J Mol Sci. 2016;17.

(68) Dolce V, Cappello AR, Lappano R, et al. Glycerophospholipid Synthesis as a Novel Drug Target against Cancer. Curr Mol Pharmacol. 2011;4:167–175.

(69) Aa J, Yu L, Sun M, et al. Metabolic Features of the Tumor Microenvironment of Gastric Cancer and the Link to the Systemic Macroenvironment. Metabolomics. 2012;8:164–173.

(70) Wise DR, Ward PS, Shay JES, et al. Hypoxia Promotes Isocitrate Dehydrogenase-Dependent Carboxylation of α-Ketoglutarate to Citrate to Support Cell Growth and Viability. Proc. Natl Acad Sci USA. 2011;108:19611–19616.

(71) Song H, Wang L, Liu H-L, et al. Tissue Metabolomic Fingerprinting Reveals Metabolic Disorders Associated with Human Gastric Cancer Morbidity. Oncology Reports 2011;26: 431–438.

(72) Forman MR, Mahabir S. Saturated Fatty Acids and Cancer. In: Milner JA, Romagnolo DF, eds. Bioactive Compounds and Cancer. Nutrition and Health. Totowa, NJ:Humana Press;2010:213–233.

(73) Ventura R, Mordec K, Waszczuk J, et al. Inhibition of de Novo Palmitate Synthesis by Fatty Acid Synthase Induces Apoptosis in Tumor Cells by Remodeling Cell Membranes, Inhibiting Signaling Pathways, and Reprogramming Gene Expression. EBioMedicine. 2015;2:808–824.

(74) Chen X, Chen H, Dai M, et al. Plasma Lipidomics Profiling Identified Lipid Biomarkers in Distinguishing Early-Stage Breast Cancer from Benign Lesions. Oncotarget. 2016;7:36622-36631.

(75) Baenke F, Peck B, Miess H, et al. Hooked on Fat: The Role of Lipid Synthesis in Cancer Metabolism and Tumour Development. Dis Model Mech. 2013;6:1353-1363.

(76) Yang L, Cui X, Zhang N, et al. Comprehensive lipid profiling of plasma in patients with benign breast tumor and breast cancer reveals novel biomarkers. Anal Bioanal Chem. 2015; 407:5065-77.

(77) Shen S, Yang L, Li L, et al. A plasma lipidomics strategy reveals perturbed lipid metabolic pathways and potential lipid biomarkers of human colorectal cancer. J

Chromatogr B. 2017;1068-1069:41-48.

(78)   Del Boccio P, Perrotti F, Rossi C, et al. Serum lipidomic study reveals potential early biomarkers for predicting response to chemoradiation therapy in advanced rectal cancer: A pilot study. Adv Radiat Oncol. 2017;2:118-124.

(79)   Jinping Gu, Xiaomin Hu, Wei Shao, et al. Metabolomic analysis reveals altered metabolic pathways in a rat model of gastric carcinogenesis. Oncotarget. 2016;7:60053-60073.

Figure Legends

Figure 1. Schematic of single cell analysis of CTCs using LSC-MS. Blood samples were collected from GC and CRC patients. CTCs were isolated and enriched using microfluidics technique. Single CTCs were sampled and analyzed using the LSC-MS system. Finally, data processing and statistical analysis (t-test and PCA) was done.

Figure 2. PCA-DA of CTCs and lymphocytes. (a) The difference in the metabolic profile between CTCs and lymphocytes collected from GC patients. (b) CRC patients CTCs and lymphocytes metabolic profile. Each dot corresponds to a single cell.

Figure 3 Heatmap of significant peaks found in all collected CTCs in comparison with lymphocytes. The p-value of each annotated peak is shown above the figure.

Figure 4 Single cell profiling of GC and CRC CTCs. (a) PCA-DA discriminating between GC CTCs, CRC CTCs and blank. Each dot corresponds to a single cell. (b) Histogram of the frequency of peaks distribution across the *m/z* scale of different cancer types.

Figure 5. Characterization of significant peaks found in both GC and CRC CTCs. (a) Heatmap of significant peaks found in both GC CTCs and CRC CTCs. The p-value of each annotated peak is shown above the figure. (b) Histogram showing the frequency of the unique peaks (potential biomarkers) to each cancer type distribution across the *m/z* scale of the two cancer types.

Table 1. Clinical characteristics for study subjects.

Supporting information

Figure S1 Live single cell mass spectrometry. Microscopic image of the sampling processes of (a) single CD45 negative CTC and (b) single CD45 positive lymphocytes are shown.

Figure S2. PCA-DA loadings plot for PCA-DA discriminating between GC CTCs and CRC CTCs.

Figure S3. PCA-DA loadings plot for (a) PCA-DA of GC CTCs vs lymphocytes and (b) CRC CTCs vs lymphocytes.

Figure S4. PCA-DA loadings plot for per patient PCA-DA of CTCs and lymphocytes of CRC and GC patients.

Figure S5. Volcano plot of significant and insignificant *m/z* peaks. Values with threshold more than 1 log2 fold change or less than -1 log2 foldchange and p-value less than 0.05 are shown.

Figure S6. Per-patient PCA-DA of CTCs and lymphocytes of CRC and GC patients.

Figure S7. Bar plot showing the unique distribution of lipid classes and lipid pathways related classes across CTC samples of GC and CRC.

Table S1 Putatively identified metabolites and lipids unique to CTCs. Statistically significant p = < 0.05 peaks of CTCs when compared to lymphocytes are shown.

Table S2 Putatively identified metabolites and lipids unique to GC CTCs. Statistically significant p = < 0.05 peaks of GC CTCs in comparison to CRC CTCs are shown.

Table S3 Putatively identified metabolites and lipids unique to CRC CTCs. Statistically significant p = < 0.05 peaks of CRC CTCs in comparison to GC CTCs are shown.
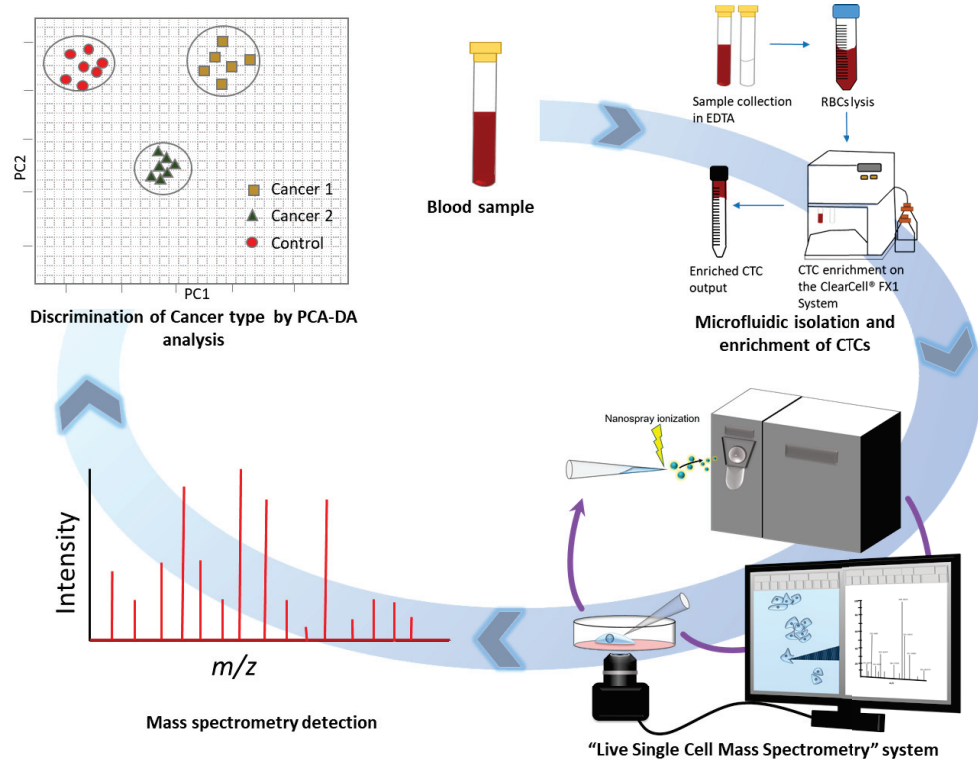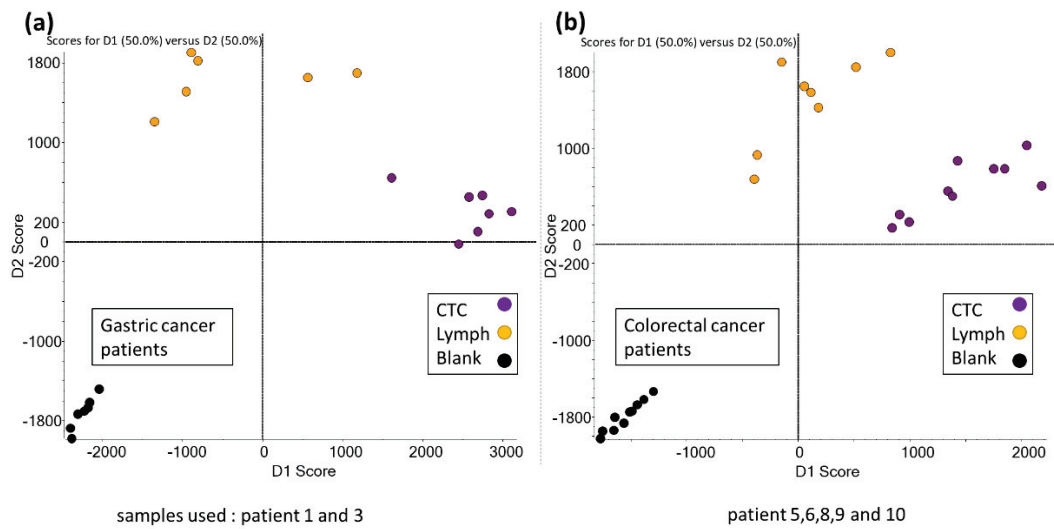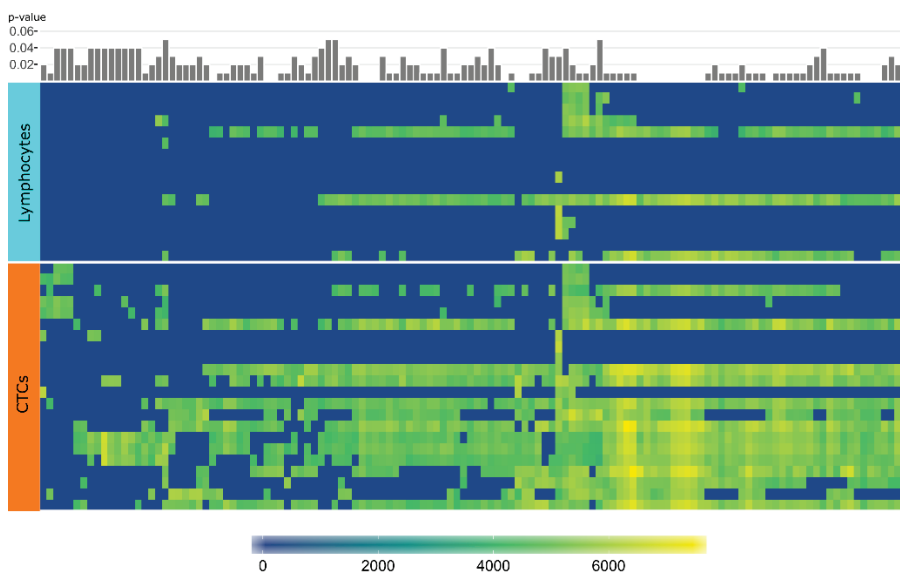
Figure 1:

Discrimination of Cancer type by PCA-DA analysis

Blood sample

Sample collection in EDTA

RBCs lysis

CTC enrichment on the ClearCell® FX1 System

Enriched CTC output

Microfluidic isolation and enrichment of CTCs

Nanospray ionization

"Live Single Cell Mass Spectrometry" system

Intensity

m/z

Mass spectrometry detection

Figure 2:

(a) Scores for D1 (50.0%) versus D2 (50.0%)

Gastric cancer patients

CTC
Lymph
Blank

D2 Score

D1 Score

samples used : patient 1 and 3

(b) Scores for D1 (50.0%) versus D2 (50.0%)

Colorectal cancer patients

CTC
Lymph
Blank

D2 Score

D1 Score

patient 5,6,8,9 and 10

Figure 3:

Figure 4:

Figure 5:
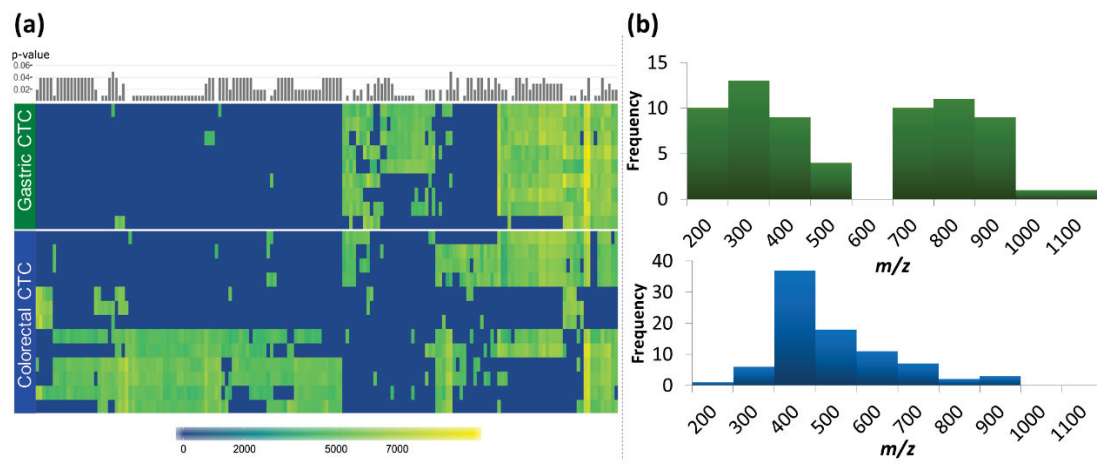
Figure S1:
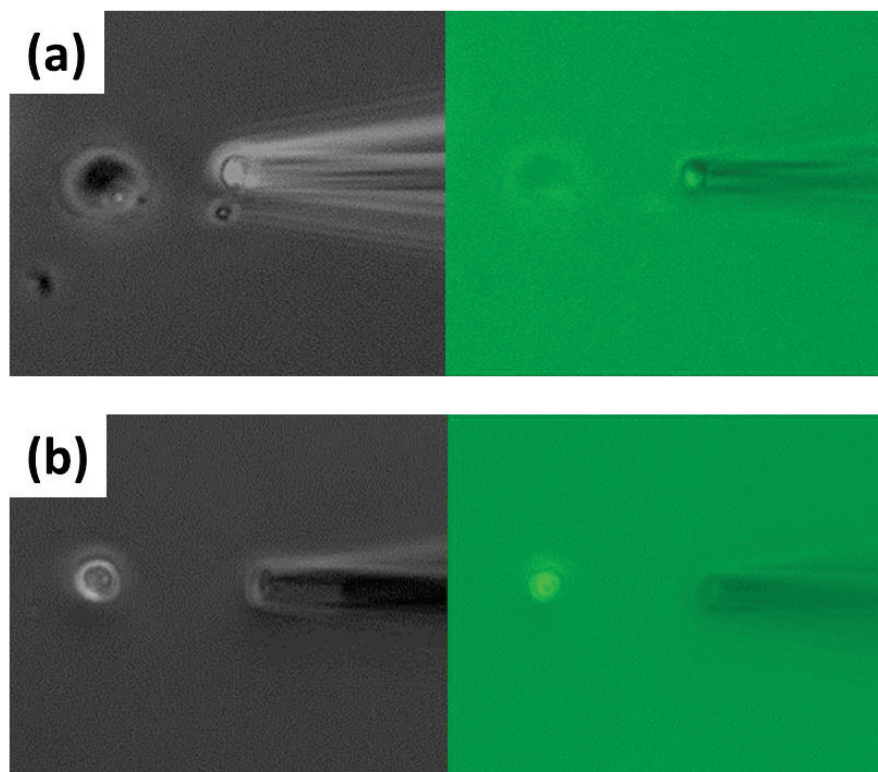
Figure S2:
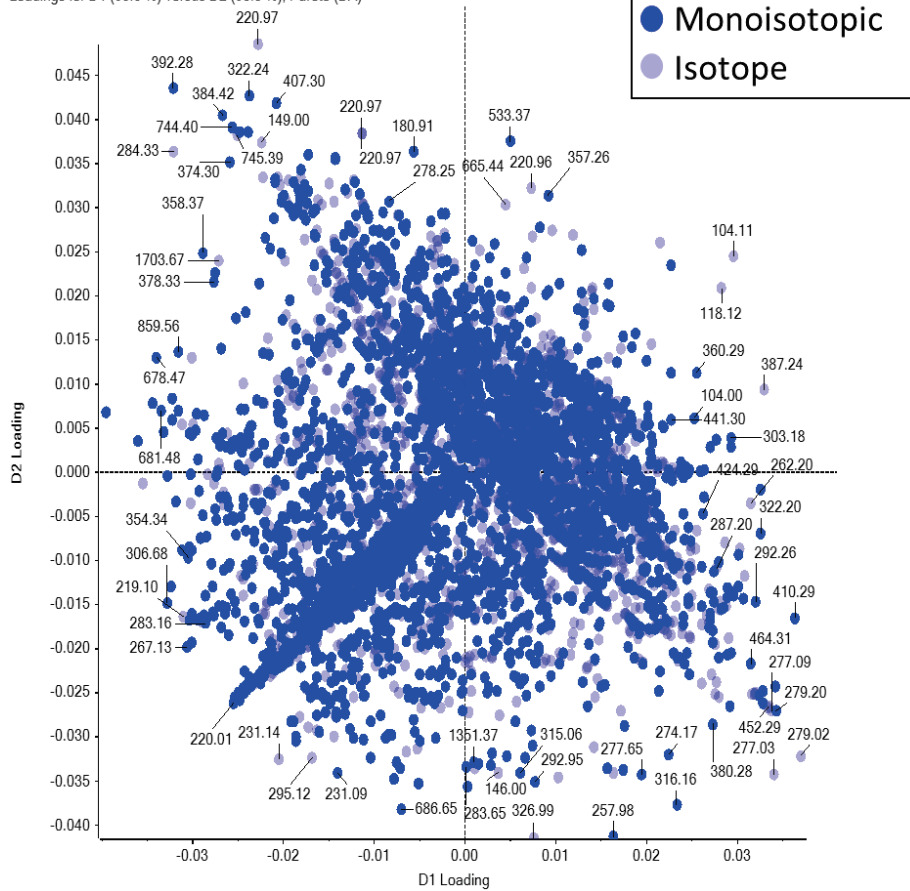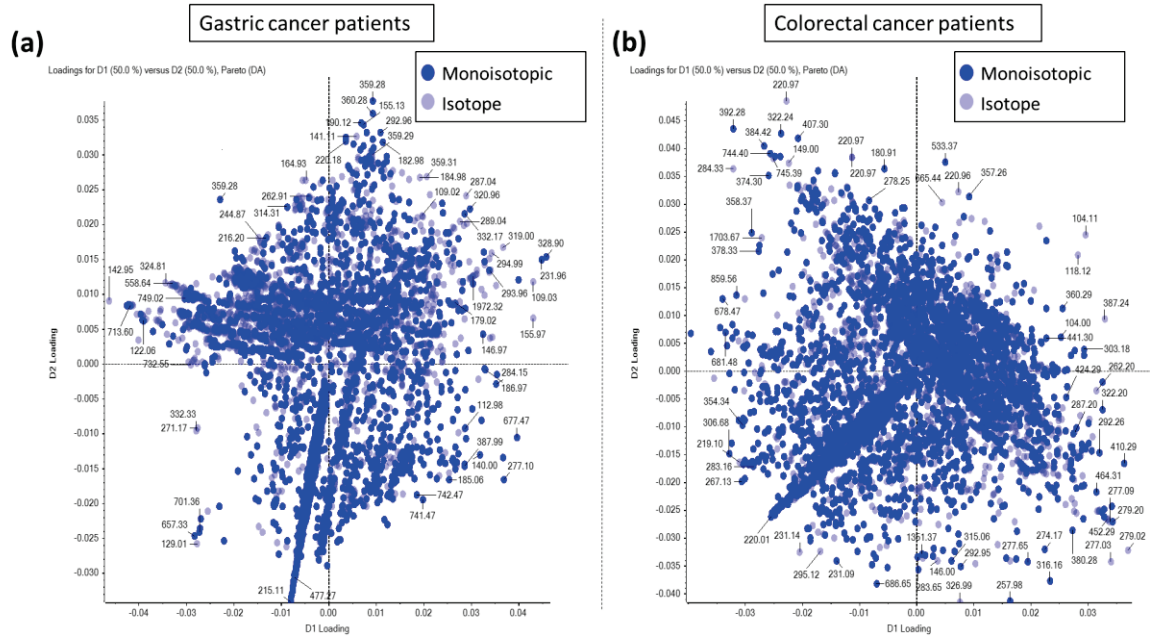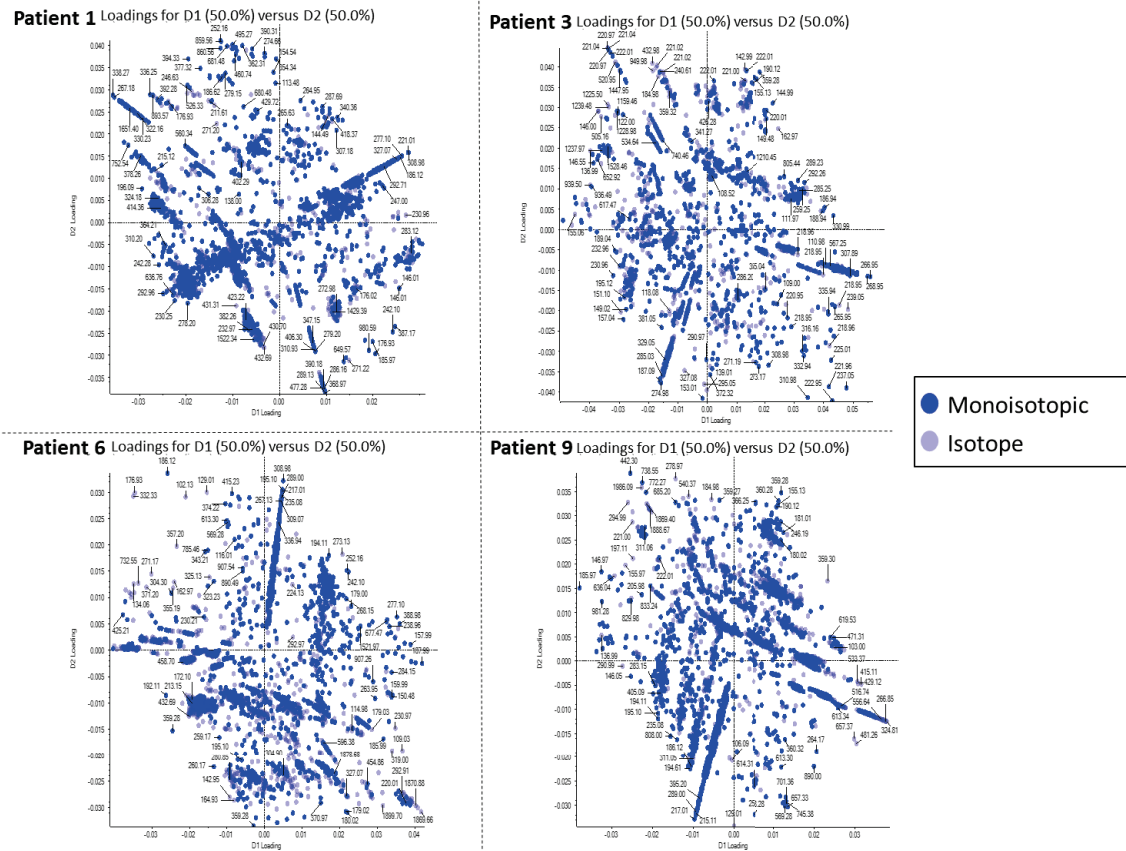
Figure S3:

Figure S4:

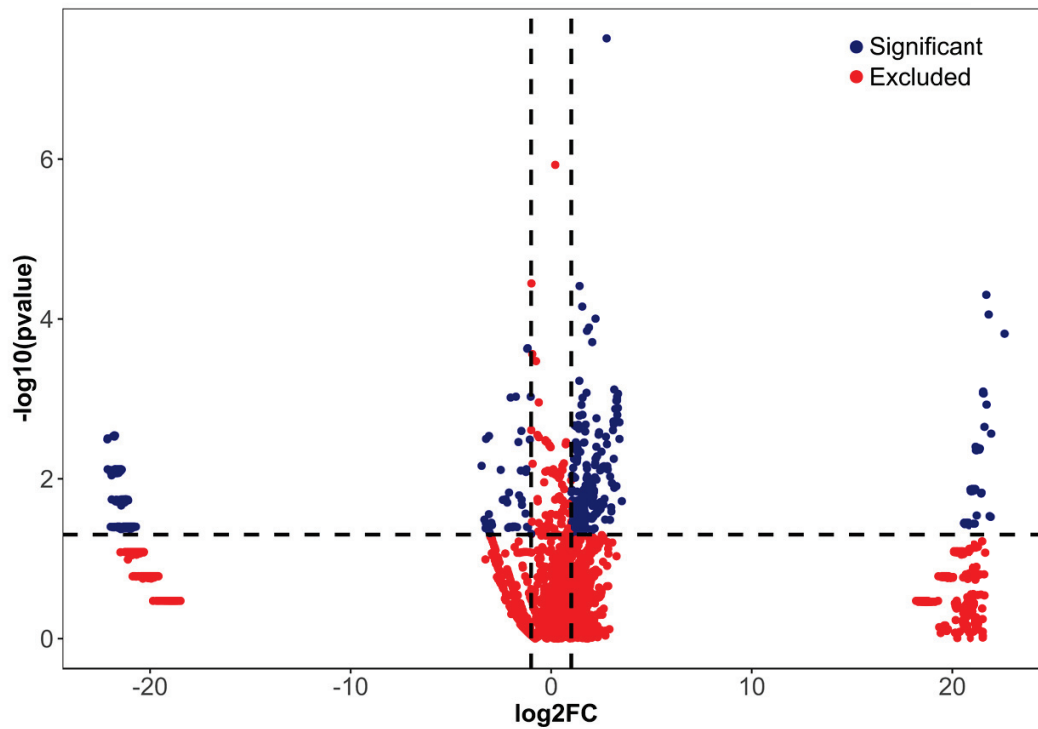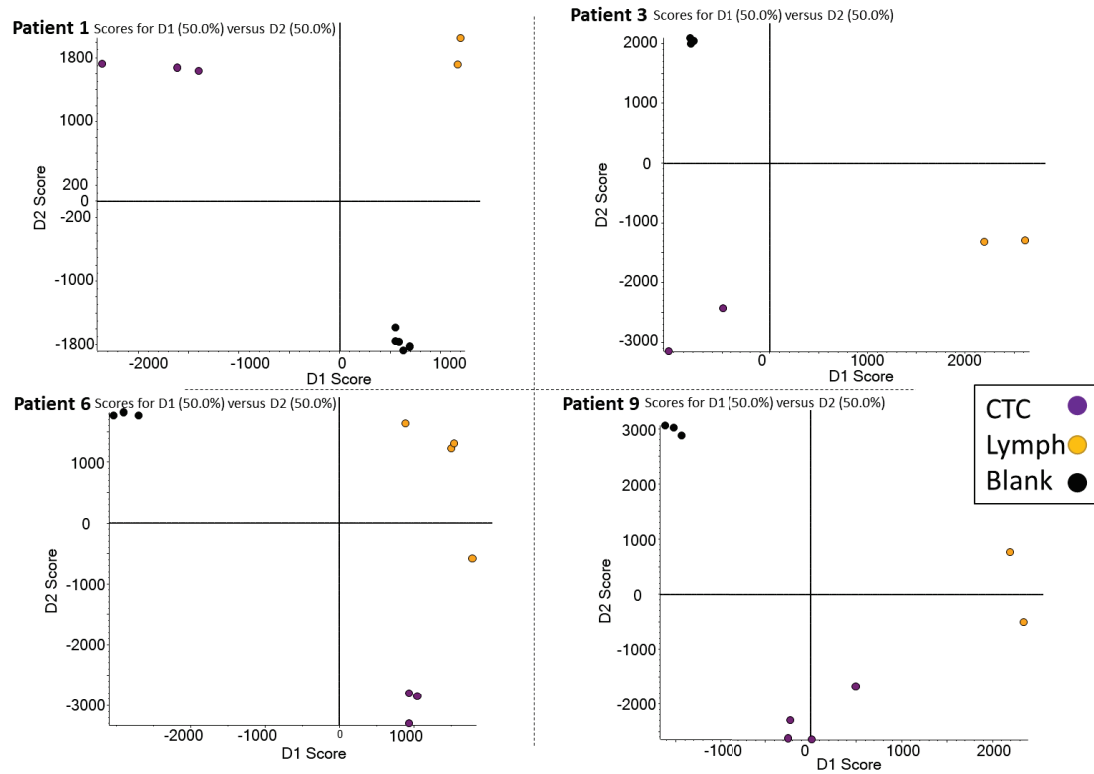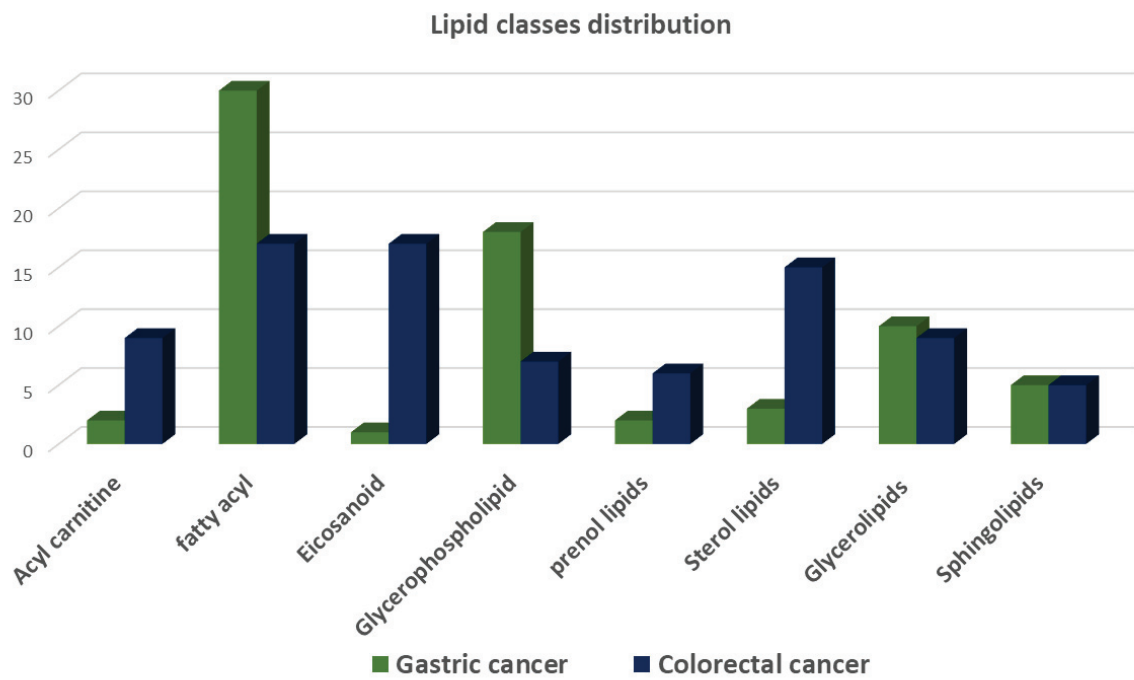**Figure S5:**

Figure S6:

Figure S7:

Table 1:

Patient Number	Gender	Age	Primary site	Histology	Histological differentiation	Disease status	Distant metastatic site
1	male	60	stomach	adenocarcinoma	poorly	recurrence	liver, bone
2	female	71	colon	adenocarcinoma	well to moderately	recurrence	small intestinal, uterus
3	male	63	stomach	adenocarcinoma	moderately to poorly	stage IV	liver, adrenal gland, para-aortic lymh nodes
4	male	80	stomach	adenocarcinoma	moderately to poorly	stage IV	liver, peritoneum
5	male	51	colon	adenocarcinoma	poorly	stage IV	liver
6	male	70	colon	adenocarcinoma	unknown	recurrence	lung
7	male	70	colon	adenocarcinoma	moderately	recurrence	liver, lung
8	male	68	colon	adenocarcinoma	poorly	recurrence	non-regional lymph nodes
9	male	56	colon	adenocarcinoma	moderately	recurrence	liver, lung
10	male	72	colon	adenocarcinoma	well	recurrence	liver, lung