Opportunities and challenges for sample preparation and enrichment in mass spectrometry for single-cell metabolomics

Wevers, D.; Ramautar, R.; Clark, C.P.; Hankemeier, T.; Ali, A.M.A.M.

Received: 11 May 2023 | Revised: 8 August 2023 | Accepted: 19 August 2023

DOI: 10.1002/elps.202300105

ELECTROPHORESIS

REVIEW

Opportunities and challenges for sample preparation and enrichment in mass spectrometry for single-cell metabolomics

Dirk Wevers[1,2] | **Rawi Ramautar**[2] | **Charlie Clark**[2] | **Thomas Hankemeier**[2] | **Ahmed Ali**[2]

[1]Wageningen University and Research, Wageningen, The Netherlands

[2]Metabolomics and Analytics Centre, Leiden Academic Centre for Drug Research, Leiden, The Netherlands

Correspondence
Ahmed Ali, Leiden Academic Centre for Drug Research, Einsteinweg 55, 2333 CC Leiden, The Netherlands.
Email: ali@lacdr.leidenuniv.nl

Colour online: See the article online to view Figures 1–7 in colour.

Abstract

Single-cell heterogeneity in metabolism, drug resistance and disease type poses the need for analytical techniques for single-cell analysis. As the metabolome provides the closest view of the status quo in the cell, studying the metabolome at single-cell resolution may unravel said heterogeneity. A challenge in single-cell metabolome analysis is that metabolites cannot be amplified, so one needs to deal with picolitre volumes and a wide range of analyte concentrations. Due to high sensitivity and resolution, MS is preferred in single-cell metabolomics. Large numbers of cells need to be analysed for proper statistics; this requires high-throughput analysis, and hence automation of the analytical workflow. Significant advances in (micro)sampling methods, CE and ion mobility spectrometry have been made, some of which have been applied in high-throughput analyses. Microfluidics has enabled an automation of cell picking and metabolite extraction; image recognition has enabled automated cell identification. Many techniques have been used for data analysis, varying from conventional techniques to novel combinations of advanced chemometric approaches. Steps have been set in making data more findable, accessible, interoperable and reusable, but significant opportunities for improvement remain. Herein, advances in single-cell analysis workflows and data analysis are discussed, and recommendations are made based on the experimental goal.

Abbreviations: CCS, collision cross-section; CNN, convolutional neural network; CTC, circulating tumour cell; CyESI–MS, cytometry ESI–MS; DBDI, dielectric barrier discharge ionization; DI, direct infusion; DIA, data-independent acquisition; DL, deep learning; DTIMS, drift time ion mobility spectrometry; FAIR, findable, accessible, interoperable and reusable; GC × GC–MS, comprehensive GC–MS; GGM, Gaussian graphical model; HCA, hierarchical cluster analysis; HR, high-resolution; IMS, ion mobility spectrometry; KDE, kernel density estimation; LAESI, laser ablation electrospray ionization; LV, latent variable; LVC, liquid vortex capturing; ML, machine learning; MRM, multiple reaction monitoring; NBC, naïve Bayes classifier; OPLS-DA, orthogonal partial least squares discriminant analysis; PCA-DA, principal component analysis discriminant analysis; PLS-DA, partial least squares discriminant analysis; RF, random forest; RS-ESI, remote sampling electrospray ionization; SHAP, Shapley additive explanation; SIMS, secondary ion mass spectrometry; SRM, selected reaction monitoring; SWATH-MS, sequential window acquisition of all theoretical mass spectra; TIMS, trapped ion mobility spectrometry; *t*-SNE, *t*-distributed stochastic neighbour embedding; TWIMS, travelling-wave ion mobility spectrometry.

KEYWORDS
data processing, experimental design, mass spectrometry, single-cell heterogeneity, single-cell metabolomics

1 | INTRODUCTION

Cells are the fundamental units of life, and each cell is unique, even when the type and genome are the same [1]. This is true for healthy cells, but also, for example cancer cells [2, 3]. An infamous example is breast cancer: Its heterogeneity makes it complicated to treat it effectively [4], which is one of the reasons that breast cancer is one of the leading causes of death among women [5]. Furthermore, circulating tumour cells (CTCs) that have detached from the tumour may be present in the bloodstream, and these can provide information about the type of tumour present and its metastasis [6]. Next to heterogeneity in disease (sub)type, cells may also be heterogeneous in their resistance to drugs [7]. When cells are analysed in bulk, as is done in the analysis of a biopsy, the differences between subpopulations of cells are averaged out, as is shown in Figure 1. Therefore, there is an interest in analytical techniques capable of studying the chemical composition of a single cell, so that the heterogeneity within the population can be studied as well.

In cells, an overwhelmingly complex network of chemical reactions takes place. The reactions in a cell are ultimately controlled by its genome; part of the genome is transcribed to mRNA, and the mRNA is translated to proteins and enzymes, the proteome of the cell. Enzymes catalyse biochemical reactions that allow the cell to grow, divide, communicate and so forth. In these reactions, a wide variety of small molecules – metabolites – take part, the entirety of which is referred to as the metabolome. The metabolome gives the closest view of the status quo in a cell: Although genes may not be expressed, mRNA can be degraded before it is translated and proteins may counteract each other, metabolites are (intermediate) products of reactions occurring or just finished in the cell [8, 9]. Therefore, the metabolome is of interest when studying the (heterogeneity in) behaviour of single cells. A complicating factor in the analysis of single cells is that metabolites, as opposed to for instance DNA and mRNA, cannot be amplified; hence, one is bound to analyses at very low concentrations and in very small volumes.

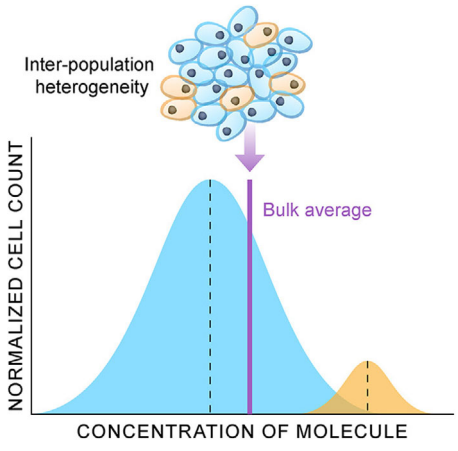

FIGURE 1 Relevance of studying single cells: When an entire population is considered, differences between single cells are averaged out, thereby masking their heterogeneity. *Source*: Reproduced with permission from Taylor et al. [81].

Due to its high sensitivity and high-resolution (HR), MS is the technique of choice when analysing the molecular composition of single cells [10–14]. In general, (single-cell) MS workflows in metabolomics comprise sample enrichment to unlock the full potential of MS. Sample enrichment can be done by endashing MS to LC (LC–MS), GC (GC–MS) or CE (CE–MS) [15–17]. These separation steps, however, decrease the throughput and reproducibility. A different widely applied technique in metabolomics is NMR spectroscopy: It has sensitivities down to the nM range, is non-destructive, can be applied *in vivo* and requires much less sample preparation than MS, therefore allowing for higher throughput and lower costs [18]. The employment of higher field strengths and 2D-NMR experiments as well as techniques to hyperpolarise nuclei can improve NMR sensitivity even further [19]. Nonetheless, GC–MS and LC–MS still outperform NMR in sensitivity by one-to-two orders of magnitude [19]. Next to that, CE–MS gets more and more attention due to the extraordinarily high sensitivities that can be reached [20–23]. Furthermore, the risk of overlapping signals is higher in NMR than in, for example CE–MS, limiting the number of analytes that can be identified in a sample. Therefore, MS techniques are preferred over NMR in the analysis of single cells.

An important consideration in experimental design is that the turnover rate of some human metabolites is in the order of seconds [24]. Single-cell isolation may lead to disturbance of the cell, which may affect its metabolome. Therefore, halting the metabolism and cell isolation should be done fast and as non-intrusively as possible. After that, a sample is taken from the cell; this can be done by chemical extraction or using micro-sampling probes (Section 2.1). Next to that, cells can be fixated after

which spatially resolved MS analysis can be executed (Section 2.2), which does not require metabolite extraction. As the volumes and concentrations are minute, sample enrichment (Section 2.3) utilizing LC, GC or CE can be applied to increase the MS sensitivity, but this also leads to longer analysis times. Direct-infusion MS (DI–MS) does not entail a chromatographic step, which reduces the analysis time. The absence of separation, however, leads to the simultaneous injection of all analytes into the ionization chamber, which may lead to ionization suppression [25]. Nonetheless, in 2008, the first MS workflow was published that could be used for the real time, comprehensive MS analysis of single plant cells [26, 27] and human tumour cells [28]; this technique is also referred to as live MS. Additional separation can be achieved using ion mobility spectrometry (IMS), which separates ions based on their shape (cf. Section 2.3.4) using a neutral buffer gas flow. Lastly, one can selectively detect ions, which can increase MS sensitivity. Vast numbers of cells are required to do proper statistics, which poses the need for high-throughput analytical workflows. A variety of techniques is available for single-cell isolation, metabolite extraction and sample enrichment, among which techniques that allow for high throughput. Isolation of single cells can be accelerated using image recognition, but microfluidics has been employed for cell picking and sample preparation as well [29, 30]. Next to that, CE can also be executed on microfluidic chips coupled to an MS instrument [31]. Besides minimizing the analysis time, increasing the throughput necessitates as much automation of the analysis steps as possible. The transition to high-throughput analysis will be discussed further in Section 2.4. The prerequisites and opportunities for higher throughput in single-cell MS have also been reviewed recently by Xu et al. [32].

Lastly, the combination of increased throughput and increased sensitivity poses the need for data analysis techniques that can handle the large amount of high-dimensional data produced. Besides conventional data analysis techniques such as *t*-tests, principal component analysis discriminant analysis (PCA-DA) and partial least squares discriminant analysis (PLS-DA) [14, 33, 34], machine learning (ML) [35, 36] and deep learning (DL) [37–39] are employed more and more. ML models such as random forest (RF) classification and logistic regression have been used successfully for the prediction of drug resistance of single cells based on their metabolic profile [7]. DL algorithms are comparably novel and require vast amounts of data to be of merit but have been shown to perform especially well in handling large amounts of data while requiring less user intervention than ML algorithms; DL algorithms have already been used successfully in the field of proteomics [40, 41]. In addition to increasing the throughput of the analysis, making use of findable, accessible, interoperable and reusable (FAIR) data [42] also increases the amount of data from which conclusions can be drawn. This also demands, however, that experimentalists adhere to certain standards and guidelines regarding experimental set-up and data handling. Examples of different types of data analysis techniques and the use of FAIR data will be evaluated in Section 3.

This review aims to provide analytical scientists with an overview of the opportunities and challenges in the field of single-cell MS. Furthermore, the community is provided with recommendations based on the literature studied. The field of single-cell metabolomics may be a relatively young one, but it is developing quickly. Herein, experimental design (Section 2) and data analysis and handling (Section 3) are discussed, after which some concluding remarks (Section 4) are shared. Scientific literature was explored using Google Scholar using keywords such as 'single-cell metabolomics' and 'single-cell mass spectrometry', together with more specific terms such as '*in vivo*', 'untargeted', 'data analysis' and 'CE–MS'. Next to that, it was checked if and where publications were cited, as a means to find follow-up research.

2 | EXPERIMENTAL DESIGN

Single-cell isolation and subsequent metabolite extraction (Section 2.1) is an essential step in the analysis of single cells [43, 44] and starts with the halting of the cell metabolism. This is a crucial step as subsequent analysis steps may inflict stress on the cell, which may affect the cell's metabolome. Metabolite extraction can be done using an extraction solvent as well as using micro-sampling probes. On the other hand, the employment of spatially resolved MS techniques (Section 2.2) such as secondary ion MS (SIMS) and MALDI-MS will be discussed: Occasionally, these require the embedding of the sample in a matrix, but these analyses can be executed *in situ* and in some cases also under ambient conditions; there is no possibility nor the need for metabolite extraction in this type of analyses. When doing (single-cell) MS one must be aware of ionization suppression [45]. When low-volatile analytes are present, these may affect droplet formation and evaporation in ESI for co-eluting analytes, which may in turn influence which ions reach the mass analyser and in which relative intensities. Compounds that tend to have this influence include salts [46], which are abundantly present in the samples discussed herein. Ways to reduce the risk of ionization suppression will also be discussed. After metabolite extraction, the sample may need to be enriched as metabolites are present in vast numbers, and concentrations may range over multiple orders of magnitude; sample

enrichment is discussed further in Section 2.3. Lastly, to be able to do proper statistics, the metabolomes of vast numbers of cells are needed. To this end, high-throughput single-cell analyses are needed, discussed in Section 2.4.

2.1 | Single-cell isolation and extraction

2.1.1 | Halting the cell metabolism

When it comes to halting the metabolism, one can distinguish between cells in suspension and cells in tissues. When cells are in suspension, one can remove the medium by filtration and apply the quenching agent onto the residue. Washing away the medium can be done with warm PBS, but this is only applied when necessary as it may lead to metabolite leakage. For adherent cells, the quenching agent can be applied directly to the sample. A discussion of multiple quenching protocols can be found in the work by Dietmair et al. [47].

A widely employed technique for halting cell metabolism is the immersion of cells in a hot or cold organic solvent such as acetonitrile [48] or methanol; mixtures of organic solvents, water and buffers are also used frequently [47, 49]. Boiling ethanol and cold methanol are used most: Boiling ethanol carries the risk of degrading thermally labile metabolites, whereas cold organic solvents carry the risk of being slow at quenching the metabolism [50].

The extent of quenching can be studied by investigating the conversion of isotopically labelled compounds, as is done in the work by Wang et al. [51]: The conversion of these compounds after quenching can be used to quantify the remaining metabolism. Their work showed that the use of 100% methanol at −80°C is most effective at quenching metabolism; 30% methanol at −24°C is slightly less effective but allows for higher throughput. On the other hand, in the work by Onjiko et al. [52] blastomeres from frog embryos were transferred to cooled methanol (4°C) to denature enzymes and prevent metabolites from degrading. Subsequently, the solvent was removed, after which an aqueous solution of 50% methanol with 0.5% acetic acid was added. Samples were sonicated in ice-cold water and vortexed to promote extraction. The authors were able to identify several metabolites that played a role in the embryonic phenotype and cleavage. The same approach has been used by Nemes et al. [53] for metabolite extraction from sea hare cells as a way to classify cells based on their metabolome.

Another possibility to halt cell metabolism is freezing the cells using liquid nitrogen [54]. However, as the heat transfer between biological samples and liquid nitrogen is slow due to vapour formation between the surface and the liquid, cooling is preferably done using precooled metal plates [55]. However, this way of halting cell metabolism is more difficult to implement for large numbers of single cells. Therefore, methanol-based quenching methods seem to be the most effective quenching methods; the percentage of methanol used may differ [49].

2.1.2 | Metabolite extraction using extraction solvents

After the metabolism has been halted, one can proceed with the analysis of the metabolome, the first step being metabolite extraction from single cells. Extraction techniques are employed to remove matrix interferents from the sample as these might complicate subsequent MS analysis. In this way, matrix effects can be reduced. The optimum choice of extraction method varies strongly per sample [56].

In the work by Zhang et al. [57], droplets of acetonitrile are extruded from a capillary to selectively extract the immerged cell's metabolites. Subsequently, the droplet is sucked back, most of the solvent is evaporated, and the metabolites are re-dissolved. This allowed for the removal of matrix interferents and reliable and stable detection of different metabolites. In a later publication of their group [58], the aforementioned microdroplet extraction is combined with an in-house developed pico-ESI source [59] to study the differences between human astrocyte cells and glioblastoma cells. This pico-ESI source makes use of a constant, high DC voltage and a pulsed ESI source. The small flowrate ejected by the pico-ESI source allows for a lasting and stable flow, so MS/MS spectra can be acquired as well; this is needed for the distinction among different fatty acids and enabled the identification of hundreds of metabolites. Furthermore, this approach enabled the classification of diseased and healthy cells. The throughput of this method is, however, limited as cells needed to be selected manually using a microscope.

Adjustment of the extraction solvent composition can allow for the extraction of more hydrophobic analytes. In the work by Onjiko et al. [60], three extraction solvents, varying in pH and polarity, were used: Comparison of the results showed that there was overlap as well as variation in the metabolites identified using the different extraction solvents, providing wider coverage of the metabolome. Furthermore, varying the ESI mode can also widen metabolome coverage. Wang et al. [11] made use of multiple microextraction steps to study the metabolic profile of breast cancer cells to identify different subtypes. The use of multiple extraction solvents as well as ESI detection in both positive and negative mode provided a broad view of the different metabolites present, which significantly

improved classification accuracy. It does, however, require a careful combination of the mass spectra obtained for the different samples.

2.1.3 | Metabolite extraction using microfluidic devices

Besides the use of an extraction solvent, various microfluidic approaches can be employed for metabolite extraction. Traditional benefits of microfluidics, such as reduced sample and reagent consumption [61], lower limits of detection [62], and the ability to create automated, high-throughput workflows [63] are also seen in single-cell work. An additional advantage of microfluidic extraction is that the risk of ionization suppression can be reduced, as micro-sampling leads to the introduction of fewer matrix interferents than whole-cell dissection [64]. Due to the small volumes used, however, also smaller amounts of matrix interferents can have a pronounced effect. In the work by Gong et al. [65], a tungsten probe was inserted into live cells, to which the metabolites were adsorbed. Subsequently, the metabolites are desorbed from the probe and injected into the MS instrument. In this way, a wide variety of metabolites was analysed with minimal sample preparation. After use, the probe could easily be cleaned and re-used. Furthermore, this technique was shown to work well also at the subcellular level; this requires, however, staining of the nucleus which may affect the cell's metabolism. Another limitation of this approach is that one needs to ensure that all metabolites adsorb to the same extent to the tungsten tip.

Li et al. [66] successfully extracted metabolites from cells using a pulled borosilicate capillary, with a stainless steel electrode inserted to supply DC voltage. A single yeast cell in 0.5 μL aqueous solution was driven through the capillary by electromigration, after which it was lysed by electroporation. Subsequent MS analysis was performed with a pico-ESI source (Section 2.1.2). In this study, however, yeast cells were used, the cell wall of which might influence electroporation efficiency. Fang et al. [67] applied this same technique to human cells: They supplied bone samples with mannose to differentiate between osteosarcoma cells and osteoblasts; metabolomes of these cells were studied using, that is the aforementioned pico-ESI source. Wu et al. [68] improved the concept of the pico-ESI source by amongst others lowering the voltage. A decrease in the voltage lowers the chance of electrochemical reactions of the analytes. They managed to decrease the ESI voltage to around 250 V while maintaining sensitivity. Furthermore, in their work, ionization was done on-demand, using a discontinuous ambient pressure interface, which facilitates the MS and MS/MS analysis of minute volumes.

2.1.4 | Micro-sampling probes to combine sampling, extraction and measurement

There have been major contributions to single-cell analysis by developing microfluidic devices which can combine single-cell sampling, metabolite extraction and coupling to MS for measurement. These 'micro-sampling probes' vary in their size (intracellular vs. single-cell analysis), throughput and capability for integrated sample preparation. Pan et al. [69] developed the single-probe that can be applied for *in situ*, real-time analysis of single cells and has found a wide variety of applications. The tip is small enough to be inserted into a eukaryotic cell and consists of two channels: One channel ejects an ionization solvent, and the other sucks it up together with the cell's contents. The entire analysis process takes roughly 3 min and does not require sample preparation, only positioning of the cells and rinsing with PBS are sufficient. A possible limitation of the single-probe is that it requires the cells to be taken from their natural environment. In some circumstances, this is not a detriment, however, as cellular activity may be more representative. Nonetheless, it has been used to study heterogeneity in responses to *Trypanosoma cruzi* infection [70] and irinotecan treatment for cancer cells [71]. Standke et al. [72] used the single-probe to study leukemic cells almost in their native environment, the only sample preparation being the suspension of cells. In this way, they managed to distinguish between leukemic cells that had and that had not been treated with the cancer drug Taxol. Chen et al. [34] used the single-probe to study the effects of different treatments on irinotecan-resistant cancer cell metabolomes. This resulted in the identification of different metabolites and lipids associated with the effectiveness of different treatments, as well as insights into the underlying mechanism.

Yoshimura et al. [73, 74] developed an ESI-based probe that allowed for remote sampling electrospray ionization (RS–ESI) suitable for real-time and *in situ* analysis of biological samples and high-throughput analysis. A schematic picture of this probe is shown in Figure 2. For *in vivo* analyses, first an extraction solvent was applied to the sample; the droplet size of this extraction solvent was also the limiting factor in the spatial resolution that could be obtained. A motorized micromanipulator was used for the positioning of the probe, which acquired samples via negative pressure produced by the Venturi effect [75]. In the case of high-throughput analyses, samples were loaded in a micro-well plate, and results were obtained within seconds. This set-up has less precise sampling than some other probes which can target specific individual cells, for example. However the high throughput (one sample per minute) is a major advantage over other approaches, and

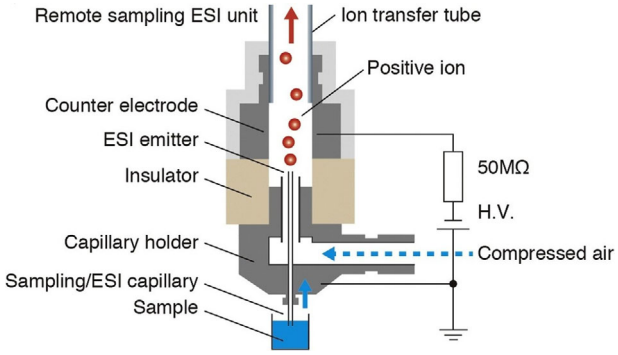

FIGURE 2 Schematic of remote sampling electrospray ionization (RS–ESI) source used by Yoshimura et al. In the ESI emitter, the Venturi effect is used to generate a local reduced pressure needed for ionization. *Source*: Reproduced with permission from Yoshimura et al. [73].

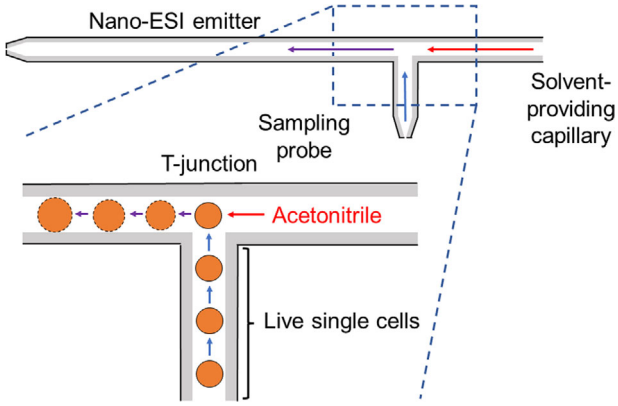

FIGURE 3 Schematic of the T-probe, incorporating online cell lysis. The flow of ionization solvent causes negative pressure, causing the cellular contents to be sucked in. *Source*: Adapted from Zhu et al. [12].

the motorized well plate for automated sampling further reduces the need for operator input.

Liu et al. [12] developed the T-probe, which operates on a similar principle to the single-probe. The major difference is that the T-probe does not eject ionization solvent into the cell, but rather samples the entire cell into a continuous flow of ionization solvent (50/50 water/methanol with 0.1% formic acid). This solvent flow provides enough suction to aspirate live cells through a glass capillary and into DI–MS. The schematic of the T-probe is shown in Figure 3. Using the T-probe, *in situ* real-time studies of cellular metabolomics are possible. Furthermore, in contrast to the single-probe, the capillary diameter can be decreased to sub-1 μm, thereby enabling the analysis of subcellular volumes. The T-probe can be re-used for multiple cells, which facilitates high-throughput analysis. However, there is an added risk of unwanted carryover between samples, as cells can adhere to the glass sampling capillary, if not treated with surface coating [76]. The T-probe demonstrated limits of detection similar to the single-probe; however, additional separation of analytes using CE or IMS is needed for more accurate identification. Zhu et al. [77] redesigned the T-probe, incorporating online cell lysis with acetonitrile as the sampling flow. This enabled a distinction between cancer cells that had and that had not undergone irinotecan treatment. One potential drawback of the T-probe approach is that there is no precise control over the sampling volume. This is a challenge for quantitation, as the dilution factor during lysis is not necessarily consistent.

Another challenge in single-cell analyses is that some metabolites are difficult to ionize in ESI due to their hydrophobicity. Cao et al. [78] developed a workflow with on-probe derivatization that enabled the analysis of fatty alcohols and sterol metabolites and showed that these can be used to differentiate among different cell types. In their work, the cell contents were sucked out and injected into the derivatization solution. Subsequently, alcohol groups were quaternized using excess $SOCl_2$ and pyridine. After derivatization, the reaction mixture was subjected to MS analysis. To perform ionization of this apolar mixture, a variant of nano-carbon fibre ionization [79] was used. The nano-fibres this probe consists of can be used to apply the analytes to, they can be coupled with LC, GC and supercritical fluid chromatography (SFC), or they can be placed against the sample. Subsequently, a high voltage is used to ionize the analytes, after which they are transferred to the MS instrument. Selective detection of derivatized compounds is easily done in selected reaction monitoring (SRM), making use of pyridine as an easily detachable group. Using this method, MS detection of fatty alcohols and sterols was strongly enhanced.

A complicating factor when dealing with ionization suppression is that analyte concentrations may vary among different parts of the cell, so variations in measured analyte concentrations are not necessarily a consequence of ionization suppression. This can be compensated for by using an isotopically labelled internal standard. One needs to make sure, however, that this internal standard is not present in excessively high concentrations, as this might also lead to ionization suppression [80]. Working with internal standards in live single cells, however, presents a challenge in itself: When using an extraction solvent, the internal standard can be a constituent thereof, but when the sample is not extracted before MS analysis, the internal standard would have to be added to the cell itself.

The individual steps of metabolite extraction for single-cell MS can be executed in a wide variety of ways. Whether extracting from cells in tissues or cells in suspension, the cell's metabolism needs to be halted to avoid interference from subsequent analysis steps. Cold aqueous methanol

solutions yield high quenching efficiencies. Cells in suspension can easily be extracted using micro-droplet extraction and micro-sampling using, for example a T-probe can be applied to both cells in suspension and tissues. An additional advantage of micro-sampling is that less stress is inflicted on the cells. The use of multiple extraction solvents or a hybrid ionization source gives a wider coverage of the metabolome. The analysis of non-polar analytes in ESI–MS is challenging; this can be overcome using derivatization.

2.2 | Spatially resolved MS ionization methods

Instead of extracting the metabolites from the cell, their spatial distribution over the cell can be measured as well. Spatially resolved MS enables imaging of tissues at the cellular level, providing the chemical information corresponding to the individual pixels; this may also be useful in the identification of biomarkers. Advances in spatially resolved MS analysis in single-cell metabolomics have recently been reviewed by Taylor et al. [81] and Hansen [82]. Subcellular resolutions have been achieved in spatially resolved MS of single-cell metabolomes [83]. MALDI–MS requires the sample to be embedded and fixated in a matrix, which is a clear limitation. Furthermore, the spatial resolution that can be obtained is limited by the laser spot size. The use of MALDI–MS for single-cell metabolomics was recently reviewed by Krismer et al. [84].

SIMS makes use of an ion beam for ionization and has a higher resolution than MALDI, but its ionization is quite hard [81]. Therefore, SIMS is suitable for the detection of, for example, trace elements, but it is limited in the ionization of biomolecules. Therefore, when metabolism is studied, cells are generally grown on ^{13}C- or ^{15}N-rich media, after which single cells are probed and small fragments containing these heavy isotopes are annotated with metabolites [85]. This limits the use of SIMS for, for example cancer cell characterization or analysis of other tissue samples. Furthermore, the applicability of SIMS is limited due to the need for fixation and vacuum. Nonetheless, SIMS-TOF has been used successfully for the identification and HR imaging of α-tocopherol [86]. Furthermore, SIMS has been coupled to TOF and Orbitrap mass analysers for the identification and HR and high-speed localization of phospholipids and other metabolites in mouse brains [83]. The high speed and high spatial resolution of TOF-SIMS and the high mass resolution, the high mass accuracy and the possibility to acquire MS/MS spectra of the Orbitrap combined can be a great tool in biomarker identification. Subcellular resolutions could be reached for the distribution of various metabolites. *In vivo* analysis may not be feasible, but this technique might provide valuable insights into cancer metabolomics, as it allows for subcellular analysis and analysis of distributions within organelles [87].

Spatial resolution can also be obtained using laser ablation ionization. Lu et al. [88] combined the derivatization approach of Cao et al. [78] with laser ablation ionization to profile the mono- and diglyceride contents of ovarian cancer cells. The laser also ensured fast derivatization. In this way, they managed to distinguish between cancerous and *para*-cancerous tissues, which might give insight into cancer metabolism.

2.3 | Sample enrichment

After metabolite extraction, sample enrichment can be used to enhance MS sensitivity by removing interferents and increasing analyte concentrations. As mentioned in Section 1, DI–MS suffers from a reduced sensitivity due to ionization suppression, as all analytes enter the ionization chamber simultaneously. This can be reduced by, for example applying a CE, GC or LC separation first, but this generally leads to longer analysis times. Therefore, other methods are also used to provide the analytical workflow with additional orthogonality, for example enriching the analytes after ionization using IMS or by detection of only a selection of the ions generated. This will also be discussed in this section.

2.3.1 | Direct-infusion MS

In DI–MS, the sample is injected into the MS instrument without any prior enrichment and with minimum sample preparation, as is for instance done in live MS (Section 1). The work by Fukano et al. [89] compares the use of LC–MS of bulk samples of hepatocytes with DI–MS of single hepatocytes. Single cells are selected using an optical microscope, ionization solvent is added, and the cellular contents are sucked out and subjected to nano-ESI–MS, proving single-cell heterogeneity. A similar approach is taken in the work by Mizuno et al. [90] to distinguish between cells from different cell lines based on the contents of cytoplasm and different organelles. In both cases, cells were sampled from suspended media. In this method, the throughput is limited by the need for manual selection of the cells. Furthermore, it is disadvantageous that this analysis is not done *in vivo*; it should, however, be possible to execute this analysis on live cells.

DI–MS has also been made compatible with derivatization steps to increase metabolome coverage. In the work by Li et al. [91], the cell lipidome is used to identify four

different subtypes of breast cancer and to identify cells that are sensitive and insensitive to gefitinib treatment. Cells are fixated in glutaraldehyde to prevent cell lysis, and the fatty acids are derivatized using the Paternò–Büchi reaction to localize carbon double-bond positions. The cell suspension was subjected to a capillary to which a voltage was applied: Cells were separated using electromigration, and at the tip of the capillary electroporation was applied to release the cell's metabolites, which were directly led to a nano-ESI source [66]. MS/MS spectra were acquired to determine the position of the double bonds. The fixation step used here most likely will affect the metabolism of the cells. Next to that, the derivatization step and the manual selection of the cells complicate high throughput.

Wide coverage of the metabolome facilitates the identification of biomarkers and their interactions. Hiyama et al. [28] studied the metabolome and lipidome of CTCs and white blood cells. Cells were sucked up using a capillary, after which an ionization solvent was added via the rear end of the tip. Sonication was used for homogenization, enabling the simultaneous detection of metabolites and lipids. Subsequently, the contents of the tip were injected into the MS instrument. A comparison of samples that had and that had not been sonicated showed that sonication led to a vast increase in the number of metabolites and lipids identified, as well as in the peak intensities measured. Throughput in this analysis is reduced by the need for sonication.

Instead of sonication, one can use a hybrid ionization source to widen the metabolomic coverage, as is described in the work by Liu et al. [92] In this work, nano-ESI was combined with dielectric barrier discharge ionization (DBDI) [93] for the simultaneous analysis of polar and apolar metabolites, respectively. Cellular contents were extracted using a microcapillary, after which nano-ESI was used for the ionization of more polar analytes. Subsequently, DBDI was used as a post-ionization source for the ionization of apolar and mainly apolar analytes. This method has been applied successfully both to animal and plant cells and led to a strong increase in the number and variation of metabolites identified [94].

In conclusion, DI–MS has been used for a wide variety of applications, an important advantage being that it allows for real-time analysis as exemplified by live MS. The metabolomic coverage can be increased by making use of a hybrid ionization source, so apolar analytes are more readily ionized, or by the derivatization of apolar analytes.

2.3.2 | Gas and liquid chromatography

A downside of DI–MS is that all analytes enter the ionization chamber simultaneously, which can lead to ionization suppression. Furthermore, the risk of injecting matrix constituents is high, which complicates metabolite analysis. Due to their high separation and identification power, LC–MS and comprehensive GC–MS (GC × GC–MS) are suitable techniques for untargeted analyses and analyte enrichment. Another advantage of the inclusion of a chromatographic separation is that ionization suppression can be reduced by sample dilution. Furthermore, it has been argued that the extraction of the sample can also reduce the extent of ionization suppression, as this reduces matrix effects in general and reduces the amount of salt in the sample [57].

Applying an enrichment step before MS analysis preconcentrates the analytes and separates them from interferents, thereby facilitating MS analysis. However, metabolites are generally too polar to be analysed with GC, which poses the need for derivatization. In the work by Koek et al. [95], GC–MS is used for the analysis of microlitre and nanolitre samples of different biological fluids. To this end, analytes were silylated [96] before injection onto the GC column. This method enabled the analysis of the intracellular contents of a single oocyte. However, oocytes have a much larger volume than most other human cells, which limits the applicability for, for example tumour cells.

In the work by Fairweather et al. [97], GC–MS and LC–MS were used for the metabolomic profiling of frog oocytes. In both cases, the sample was extracted first using a mixture of water, methanol and chloroform, after which samples were lysed and centrifuged. Using GC–MS after derivatization and LC–MS, the authors managed to get more insight into the biochemical pathways underlying and transporters involved in amino acid homeostasis and signalling. However, due to the need for a centrifugation step as well as a derivatization step, together with the need for chromatographic separation, high-throughput analysis using this method is complicated. Furthermore, these cells are large compared to human cells.

In most cases, one needs to resort to LC–MS or HILIC–MS-based methods, as these can handle the high metabolite polarity without derivatization reactions [98–100]. These have, however, less resolving power than GC. A large body of LC–MS-based metabolomics studies is available, but these are generally based on bulk samples [14]. On the other hand, in single-cell proteomics, LC–MS(/MS) is used extensively [101, 102]. Using dedicated platforms incorporating ultrasensitive LC–MS, thousands of proteins can be identified in small populations of single cells [103]. Next to that, important advances have been made through implementation of trapped IMS (TIMS) [104] and data-independent acquisition of MS spectra [41, 105]. However, LC–MS is used much less for the analysis of the metabolomes of single mammalian cells. In the work by Vasiljevic et al. [106], LC hyphenated to HR MS

(LC–HRMS) was used for the analysis of whole blood samples and single caviar eggs. In the case of the single caviar eggs, single cells were analysed, but these cells have a diameter in the order of millimetres. Consequently, a tip was used for metabolite extraction that would be too large for the analysis of single human (tumour) cells. HILIC has the advantage that it allows the separation of more polar analytes than LC does, but it generally suffers from broader peaks. Boelaert et al. [107] used HILIC-TOF-MS to identify biomarkers for chronic kidney disease in blood plasma and urine. Also here, serum samples were analysed instead of single cells.

To summarize, LC–MS and (GC×)GC–MS do not seem suitable for single-cell metabolomics due to their comparably low sensitivity: When single cells are analysed, these are generally much larger than human cells so they meet the volume requirements posed by LC and GC. Furthermore, the use of (comprehensive) GC is limited by the need for derivatization to study polar analytes.

2.3.3 | Capillary electrophoresis

In contrast to LC and GC, CE has received much attention due to its high sensitivity and facile operation with polar analytes. The possibilities of CE–MS for the omics fields have recently been reviewed by DeLaney et al. [108] and Ramautar and Zhang [22]. In the work by Onjiko et al. [109] CE–MS was used to differentiate between dorsal and ventral embryonic frog cells based on their metabolome. Capillaries were used for cell content isolation, after which the sample was extracted in a separate vial. CE–MS was applied to the sample extracts to determine the metabolomes of the different samples. This led to the identification of 52 metabolites as potential biomarkers for different cell types. Furthermore, ca. 200 features were identified that may be targeted in future studies in embryonic metabolomics. This shows that CE–MS can be applied *in situ* but the need for an extraction step prevents it from being real time.

Adjustment of the ionization source can enhance CE–MS sensitivity even further. Huang et al. [110] developed a customized CE capillary – called the spray capillary – that can be used to suck out the cell's contents and directly be used for CE–MS analysis, enabling online single-cell CE–MS. For cell picking, a microscope and a micromanipulator were used. A capillary was used to extract cellular contents, after which it was moved to a vial with background electrolyte when CE–MS was done, or to a vial with column liquid in case DI–MS was done. Not surprisingly, more metabolites were identified when CE–MS was used, but some metabolites were only identified using DI–MS, and DI–MS analysis times were considerably shorter. The setup, however, needs to be operated manually, which limits the throughput.

In the work by Kawai et al. [111], a CE–MS platform is described that can detect amino acids and metabolites with a limit of detection (LOD) of 450 fM. To this end, electrokinetic pre-concentrations (stacking) in conjunction with an in-house developed nano-CESI source (which is based on a sheathless porous tip interface) were used, resulting in dramatic improvement in detection limits as compared to the use of hydrodynamic sample injection only [112]. As a consequence of stacking, larger volumes could be injected without compromising separation efficiency. Single HeLa cells were analysed: Lysis was done separately in methanol in an Eppendorf tube, as in-capillary lysis decreased the throughput and liquid mixing is hindered by laminar flow in microfluidics. Nevertheless, the dilution this lysis step leads to is compensated for by stacking. When applying stacking, one needs to consider that there may be analytes present that hinder the stacking process; this may require (more) sample preparation to remove these, and this in turn reduces the throughput.

A complicating factor in single-cell analyses is that cells are generally emptied completely, making repetitive analyses impossible. Lombard-Banek et al. [113] analysed *in vivo* the proteome–metabolomes of tadpole embryos. Two samples, one for each omics analysis, were taken, and due to the small volumes taken (approximately 10 nL), the survival rate was almost 100%. For metabolomic analysis, ultrasensitive CE–HR MS platforms were used [114]. This showed that approximately 0.01% of the cell's volume sufficed for the detection of ca. 150 features, of which 57 could be identified with a high degree of certainty. Increasing the throughput is, however, complicated, as the cell manipulation is done manually. Furthermore, approximately 5 nL of cell volume was used for chemical analysis, which is large in comparison with human cells. However, only a fraction of this in the order of tens of picolitres was used for metabolomics MS analysis.

Micro-sampling has also been applied successfully in combination with CE–MS. The work by Onjiko et al. [115] showcases capillary micro-sampling, microextraction and CE–MS for fast analysis of frog embryos. Cell contents were sucked out, extracted in a separate microvial, vortexed and centrifuged before CE–MS analysis. Interestingly, this method led to less noise as there were fewer matrix interferents present. Furthermore, judging from the ratio between oxidized and non-oxidized compounds, the authors concluded that this micro-scale approach also exposed the cells to less stress. It was shown that microprobe CE–MS has a higher sensitivity and better reproducibility than regular CE–MS, microprobe CE–MS can be used *in situ* and *in vivo*, and as microprobe CE–MS also allows for repeated analysis of the same cell, cells

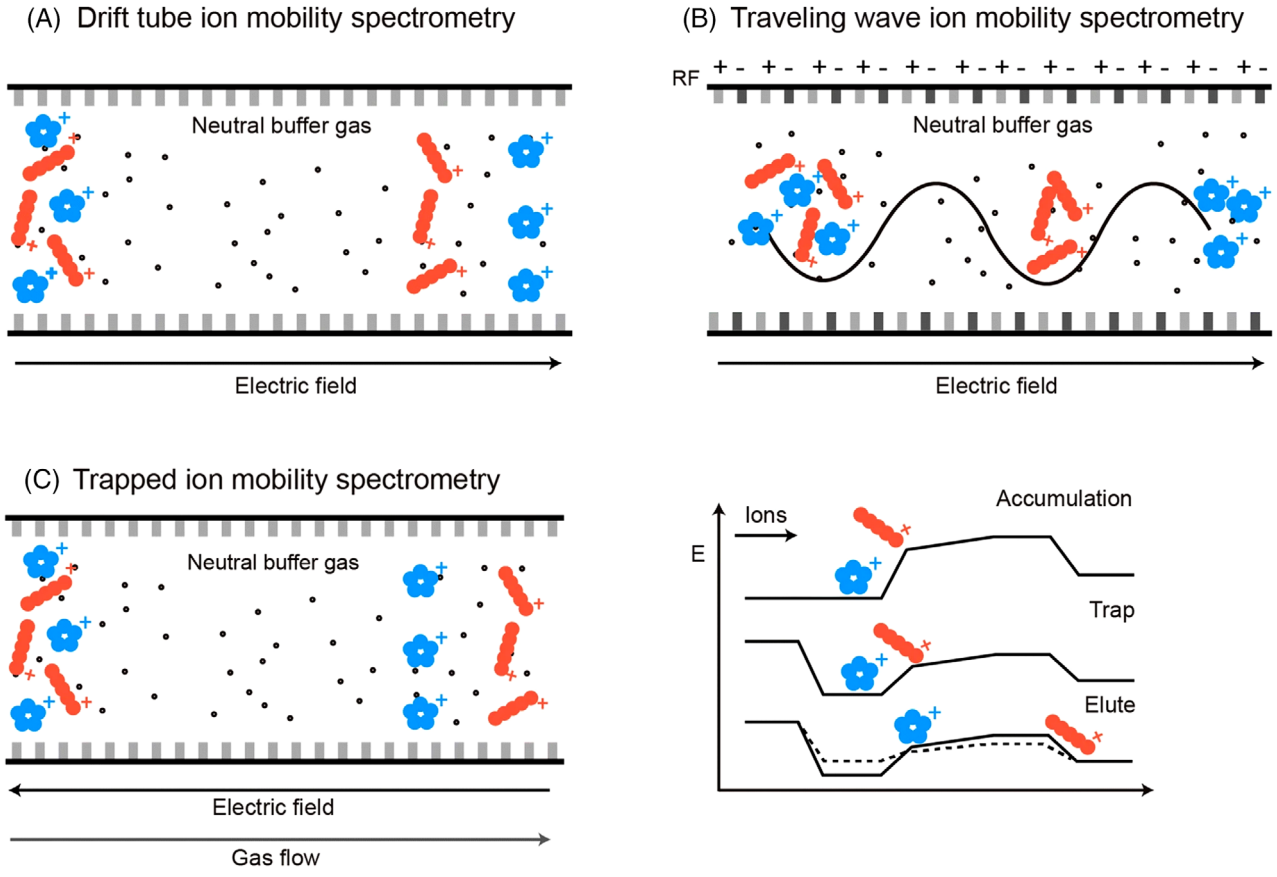

FIGURE 4 Various ion mobility spectrometry (IMS) techniques. (A) Drift time ion mobility spectrometry (DTIMS) makes use of friction between travelling ions (red rods, blue pentagons) and static gas (black dots). (B) Travelling-wave ion mobility spectrometry (TWIMS) makes use of an oscillating electric field, which leads to focusing of ions. (C) Trapped ion mobility spectrometry (TIMS) makes use of an electric field opposing the gas flow, allowing for trapping and ejection of ions. *Source*: Reproduced with permission from Luo et al. [117].

can be analysed multiple times, thereby allowing for the distinction of biological variability. However, the extraction and vortexing steps decrease the throughput of this analysis.

In general, one looks at positively or negatively charged (or chargeable) compounds in CE–MS. In the work by Portero and Nemes [116], a CE–MS workflow is presented that allows for the sequential analysis of anionic and cationic compounds, using the same cell extract. To ensure stability in negative ESI mode, the background electrolyte was exchanged. Analysis in both positive and negative mode led to a strong increase in the number of metabolites identified and wider coverage of the cells' metabolome.

In summary, CE seems a useful addition to the workflow of single-cell MS analysis: It can handle minute volumes and is compatible with ESI–MS as well as other ionization sources. Furthermore, stacking can be employed to increase sensitivity even further. In comparison with DI–MS, however, analyses take longer but more metabolites are identified, so a trade-off is made there. The next step would be to develop a more generic CE–MS-based workflow for single-cell metabolomics and for that an interlaboratory study is needed to assess the reproducibility. With a positive outcome, CE–MS could potentially have a major impact in studies dealing with low amounts of mammalian cells. The developments reported so far are encouraging.

2.3.4 | Ion mobility spectrometry

Although CE separates analytes before they are ionized, IMS separates them after ionization. IMS is a suitable extension of the analytical workflow as it can separate isobaric compounds with different collision cross sections (CCSs) within milliseconds. The CCS of an ion is dependent on its shape, which determines the drag it experiences from the neutral buffer gas it is driven through. Especially in untargeted analysis, IMS can be of use, as it can distinguish between structural isomers. Applications and advances of IMS have recently been reviewed by Luo et al. [117]. An overview of some different IMS techniques is given in Figure 4, a review of these and more techniques can be found in the work by Dodds and Baker [118].

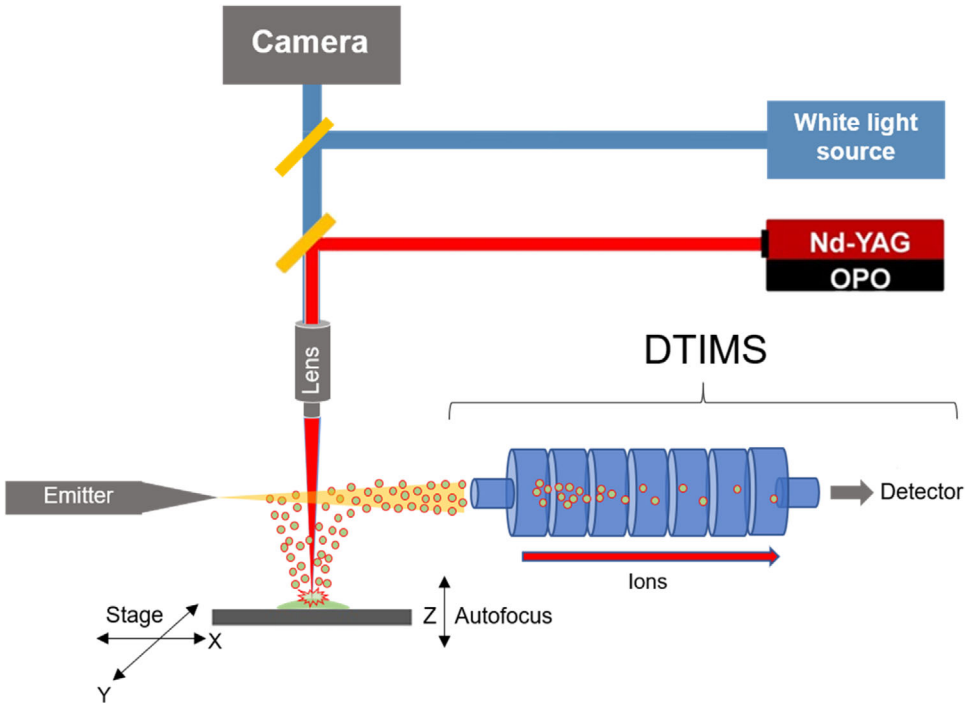

FIGURE 5 Workflow used in the laser ablation electrospray ionization (LAESI) microscope used by Taylor et al. This enables the combination of microscopy with LAESI ionization. Ion mobility spectrometry (IMS) is used to gain additional orthogonality. *Source*: Reproduced with permission from Taylor et al. [120].

Zhang et al. [119] used CE–IMS–MS for the analysis of single plant cells. Single cells of *Arabidopsis thaliana* were selected manually with the help of a micromanipulator, after which the cell's contents were sucked out using a capillary. The capillary was then backfilled with ionization solvent, after which MS and MS/MS analysis were executed. Analyses were executed with and without IMS, which showed that the inclusion of IMS led to a strong increase in the number of metabolites identified and wider coverage of the metabolome. A complicating factor in the analysis of IMS–MS data is that there are – to the best of our knowledge – hardly any tools available that can perform all individual steps of the corresponding data analysis workflow.

In the work by Taylor et al. [120], laser ablation electrospray ionization (LAESI) was used together with drift time IMS (DTIMS) for metabolomic analysis of single onion cells. Image recognition software was used to detect cells, allowing for higher throughput. Furthermore, it allows for *in situ* and spatially resolved analysis, by directing the laser or by using a laser fibre (f-LAESI) [121]. Optical microscopy was combined with LAESI to acquire MS information on manually selected cells, leading to the 'LAESI microscope' depicted in Figure 5. Next to that, a comparison was made between metabolites identified from single cells and metabolites identified from bulk extracts. This showed that single-cell analysis led to the identification of more metabolites, although some compounds were only identified in bulk extracts. In later work of this group [122], the LAESI microscope was used in combination with an Orbitrap mass analyser to obtain higher mass resolutions. Next to that, a smaller LAESI beam width was used, leading to a higher spatial resolution. Furthermore, they showed that single cells can be analysed *in situ* and in high throughput.

The work by Mast et al. [123] shows the combination of CE–MS with TIMS for the targeted separation of stereoisomers of peptides in neural cells. Furthermore, IMS mobility profiles were used to relatively quantify the presence of the different stereoisomers. TIMS works especially well in targeted searches for ions of which the CCS is known, but for untargeted searches, travelling-wave IMS (TWIMS) and DTIMS are more suitable. All in all, IMS seems a very useful addition to the targeted or untargeted MS analysis of single cells, as it provides additional orthogonality that can potentially resolve isomers, and it does so at the millisecond timescale.

2.3.5 | Ion selection techniques in MS

As mentioned earlier, introducing all ions simultaneously into the MS instrument may lead to ionization suppression. This can be reduced using, for example CE or IMS. To reduce the number of ions entering the mass analyser,

ions can also be filtered out based on their mass or fragmentation products. In the work by Nemes et al. [53], SIM was applied in CE–MS. This allows for the sensitivity to be improved without a chemical bias. The authors were able to distinguish among several types of neurons and to quantitatively compare metabolite levels between cells.

Abouleila et al. [124] combined live MS with a microfluidics platform to do untargeted analysis on CTCs obtained from the blood of colorectal and gastric cancer patients. Increasing the dynamic range and reducing the amount of noise SIM were used. This allowed for the identification of more metabolites. A clear distinction was shown among blanks, gastric cancer CTCs and colorectal cancer CTCs, providing a means to find cancer type-specific biomarkers. Besides SIM, one can resort to multiple reaction monitoring (MRM). This has been applied successfully in the field of plant sciences [125], but it is only suitable for targeted analyses, whereas SIM also allows for untargeted analyses. A disadvantage of the methods discussed in this section is that they still may suffer from ionization suppression, due to the large variation in concentrations present.

Sample enrichment has been used successfully to increase the identification rate in single-cell MS. Nonetheless, successes have been booked with DI–MS in the form of live MS, the main advantage being that it provides real-time results. A suitable enrichment technique is CE; this leads to less ionization suppression and the identification of more metabolites than DI–MS. CE capillaries have been used for the extraction of metabolites from single cells. Furthermore, CE–MS has an outstanding sensitivity, which can be improved even further by analyte stacking. Due to the high sensitivity of CE–MS, only a small part of the cell volume is needed, allowing for cells to stay alive so they can be followed over time. IMS can be incorporated to obtain additional orthogonality, especially because it can resolve isomers that differ in shape. Ion selection techniques such as SIM can also be used to reduce ionization suppression and increase MS sensitivity.

2.4 | High-throughput analysis

High-throughput analysis is crucial for large-scale research like population studies, where thousands of samples or more must be analysed. This is particularly challenging in single-cell metabolomics, where there is a high skill threshold and a large amount of hands-on time by the researcher for each sample. The ideal high-throughput single-cell workflow would be comprised automated single-cell sampling, fast sample preparation, prevention of unnecessary sample dilution, efficient metabolome quenching and direct coupling to MS. In the previous section, a diversity of steps in the single-cell analysis workflow have been discussed, some of which can also be applied in high-throughput analysis. Halting the cell metabolism is best done using cold methanol, but the compatibility of this with high-throughput analyses depends on the temperature used. Micro-sampling probes have been used for high-throughput analyses, such as the RS–ESI probe with the Venturi suction for rapid sampling. This method, along with most high-throughput approaches, is not compatible with live single-cell sampling from their native environment. Cell picking is generally performed manually with a microscope and/or using a video camera, which limits the throughput, even though MS spectra can be acquired at high rates. The cell lysis step is more compatible with high throughput: High-throughput cell lysis has been shown using for instance electroporation, which requires only seconds.

The requirement of high-throughput favours the use of DI–MS. GC and LC have generally been applied to bulk samples instead of single cells or cells that are much larger than human cells. On top of that, GC-based analyses generally require the derivatization of the analytes. CE–MS has been used extensively for the analysis of single cells, and extremely low LODs have been achieved. However, an extraction step may be needed, which is difficult to do online. Furthermore, separation in general decreases the throughput. Additional orthogonality and sensitivity can be provided using IMS or ion selection techniques. Application of IMS can provide additional orthogonality at the timescale of milliseconds. This allows for an increase in sensitivity without the need for targeted analyses. In case a targeted analysis is done, one may also resort to MRM or neutral loss scanning. In this section, advances in high-throughput single-cell metabolomics analysis are discussed.

2.4.1 | Automation of single-cell isolation and sample preparation

Single-cell metabolomics critically relies on the analysis of large numbers of cells: In this way, noise and biological variation can be distinguished from one another, and data analysis and biomarker identification become feasible and meaningful. This requires analyses to be done in high throughput. An advantage of using microfluidics for this purpose is that it can be used to simulate the cell's microenvironment *in vitro*, thereby reducing the gap with *in vivo* studies, and that the required equipment can be produced easily [126, 127]. Recently, applications of microfluidics for isolation and sorting of single cells have been reviewed by Luo et al. [128]. Microfluidics has been applied for a variety of steps that constitute single-cell analysis, such as cell sorting, isolation and lysis. Furthermore, Feng et al. [30]

recently reviewed opportunities and challenges in hyphenating microfluidic devices to MS instruments for single-cell metabolomics studies. PDMS, a widely used material for the production of microfluidic chips, is compatible with methanol, even though methanol leads to minor swelling [129]. Microchips are also produced using glass, but this is more labour-intensive and glass can exhibit undesirable surface properties leading to cell adhesion and carryover. Other materials such as cyclic olefin copolymer, PMMA and fluoropolymers have also been employed for biological micro-devices, depending on the application and desired surface chemistry. Due to laminar flow within microchips, proper mixing of quenching solvent, internal standards and the sample is crucial. There are many different approaches to microfluidics mixing, described well in a review by Lee et al. [130].

A conventional way of using microfluidics for the isolation and transport of single cells is by encapsulating them in microdroplets. In the work by Zhang et al. [131], a microfluidics platform is described that allows for the transport of cells, removal of matrix interferents, de-emulsification and coupling with an MS instrument. Cells were extracted online using methanol, after which the extract was subjected to ESI–MS analysis. Differences in lipid profiles were detected between cells that had and that had not been treated with a drug, which also outlined the heterogeneity in the cells' drug responses. This analysis was, however, only applied for the analysis of lipids, not for other metabolites. Furthermore, the use of apolar solvents for encapsulation may interfere with MS analysis.

Various ways have been published to avoid the use of apolar encapsulation solvents. In the work by Huang et al. [132], a microfluidic device with a spiral-shaped channel is used to generate a uniform distribution of the cells along the channel. It makes use of Dean flow, generated due to the spiral-like channel shape: Dean flow leads to differences in flow velocities, with equidistant maxima to which the individual cells are attracted; this is also shown in Figure 6. Subsequently, ESI–MS was used for lipid profiling of single tumour cells; cell lysis was induced by the high ESI voltage. This showed that cells can be classified into subpopulations based on their lipidome. Advantages of this approach are that it allows for easy hyphenation to, for example MS and that it does not require encapsulation. However, due to ionization suppression, the identification of metabolites proved to be difficult. Furthermore, the cells cannot be analysed in their native environment. Xu et al. [133] made use of Dean flow for the high-throughput analysis of cell surface proteins and metabolites. Membrane proteins were tagged to increase specificity and sensitivity, after which the cells were separated and analysed using nano-ESI–MS at a rate of approximately 40 cells per minute. Six surface proteins and approximately 100 metabolites were shown to be important in the differentiation of different breast and ovarian cancer subtypes, as well as in predicting drug resistance.

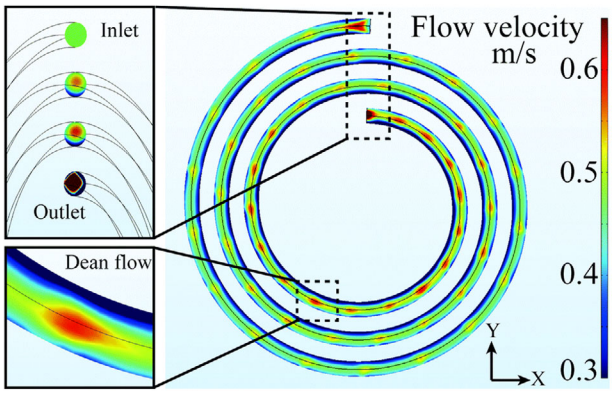

FIGURE 6 Principle of Dean flow: Due to the converging, spiral-like shape of the channel, equidistant maxima in the flow velocity are formed, leading to a uniform distribution of suspended cells. *Source*: Reproduced with permission from Huang et al. [132].

As an alternative to Dean flow, microfluidic chip channels with a serpentine shape can be used for the high-throughput isolation of single cells, as shown by the work by Feng et al. [134]. A high-throughput method is described for ESI-HRMS metabolome analysis of single cells in suspension. A pulsed square wave electric field was used for the online disruption of the cells and ionization of its constituents. This allowed for the analysis of up to 80 cells per minute and the annotation of more than 120 different metabolites in cancer cells, enabling their distinction.

Yao et al. [135] coupled flow cytometry to ESI–MS (CyESI–MS) for highly automated and comprehensive MS analysis of single cells. Cells were isolated, extracted and lysed online, and subsequently, their composition was analysed in real time. In this way, the authors managed to identify a wide range of metabolites, which also resulted in the successful classification of different subtypes of breast cancer. Furthermore, potential biomarkers for the different subtypes could be identified using different data analysis techniques. A downside of this method is that in the cytometer, shear stress is applied to the cells, which may affect the metabolism and hence cloud the results. Furthermore, there is little possibility for the acquisition of MS/MS spectra, which may be disadvantageous in metabolite identification.

An alternative method for the isolation of single cells from a suspension has been proposed by Cahill et al. [136]. In their work, small piezoelectrically ejected droplets and liquid vortex capturing (LVC) are used for high-throughput isolation; this principle is also shown in Figure 7. Video cameras are used to ensure a droplet only contains one cell, and after LVC, the cell is ruptured and extracted using,

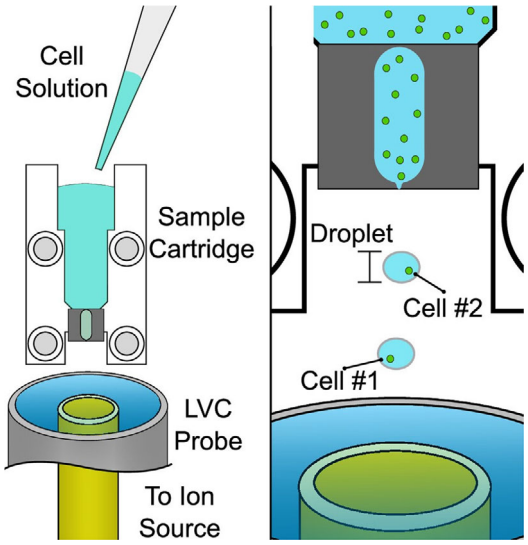

FIGURE 7 Liquid vortex capturing used in the work by Cahill et al. Piezoelectric elements are used to generate droplets containing a single cell, which is confirmed using a video camera. *Source*: Reproduced with permission from Cahill et al. [136].

for example methanol. This allows the cell to stay in its native environment until lysis. Throughputs up to 2.5 cells per minute can be reached, sequential window acquisition of all theoretical-MS spectra can be acquired for the distinction of more complex samples, and chemical analysis is done in real time. A limitation of this method is that it requires the sample to be in suspension and that this suspension must have a viscosity close to that of water.

Besides the separation of cells in suspension, *in situ* image recognition is a viable tool for the identification of single cells in suspension as well as in tissues. This has been described by Brasko et al. [137]. In their work, image analysis algorithms, ML and microscopy are combined into a platform that automatically isolates and extracts single cells from a suspension or a tissue, without disturbing the cell's environment. However, the extract is not used for MS analysis but for RNA sequencing. Furthermore, slices of tissue instead of real-time *in situ* analyses were executed, which might have affected the cell's metabolome.

Lamanna et al. [138] have published a combination of microfluidics, image recognition and MS for proteome analysis of single cells. Cells are stored with 100–300 cells in micro-wells after which a convolutional neural network (CNN) is used for the identification of single cells. Cell identification can be adjusted to the user's needs and wishes or be fully automated using artificial intelligence. Cells are lysed using laser irradiation, leading to the formation of a plasma bubble, that is taken up by a collection droplet. In this work, only genome and transcriptome analysis of the cells were done online, whereas proteome analysis was done offline due to the need for digestion and derivatization steps. In a later work [139], this lysis technique was employed in an all-in-one proteomics workflow comprising cell isolation and counting and sample pre-processing on a microfluidic chip, after which data-independent acquisition MS (DIA-MS) was performed. However, this analysis was done at-line, and only the proteome was analysed. Nonetheless, there are no limiting factors mentioned prohibiting this technique from being used in single-cell metabolomics.

In the work by Zhao et al. [140], a miniaturized platform is described that allows for well-controlled and automated extraction of single cells at a rate of one cell per 3 min. Using this platform, pico-ESI–MS analyses can be executed in an automated fashion. For the detection of individual cells, image recognition is used. Furthermore, this platform can be coupled to different MS instruments and has a higher sensitivity for single-cell metabolite and phospholipid analysis. The authors argue that this approach can also be used for the high-throughput analysis of single cells in metabolomics and lipidomics studies. A downside of this method is that cells are ruptured offline in a vacuum oven, but it allows for continuous sample preparation of thousands of cells.

As mentioned earlier, the analysis of less polar compounds may require derivatization to increase ionization efficiency. Peng et al. [141] showed how double-helix micro-channels on microfluidic chips can be used for online multi-step derivatization of aldehydes. In their work, the online derivatization of aldehydes and their electrophoretic separation were of interest. The helix-shaped channels ensured efficient mixing of the reactants, and reaction and separation were completed within minutes. However, the emphasis was placed on aldehydes, and a different detector than MS was used. Therefore, the applicability of this technique for metabolomics is hard to estimate.

All in all, plenty of ways are available to isolate single cells, be it from suspensions or tissues. Especially microfluidics seems useful, as it allows for efficient isolations but also the derivatization of analytes when needed [142–144]. Furthermore, it has been coupled with CE–MS and pico-ESI–MS. For the isolation of single cells from tissues, image recognition is often used, after which the cell contents can be extracted using a capillary or a micro-sampling probe.

2.4.2 | Reducing the analysis time

For the reduction of analysis time, especially IMS seems suitable: It provides the analytical workflow with additional orthogonality, and it separates ions before they enter the mass analyser, thereby reducing the risk of ionization suppression. This is shown by the work by Taylor et al.

[120]. Furthermore, CE–MS and IMS can be combined, as shown by the work by Mast et al. [123] discussed in section 2.2.5.

CE–MS has been used for fast separations of samples. The work by Li et al. [145] shows the combination of CE–MS and microfluidics, allowing for highly sensitive single-cell analysis at a rate of approximately one cell per minute. Cells were lysed using a double-electrode cell lysis technique monitored by a video camera: Single cells were lysed using two electrodes over which a voltage was applied. Even though only dopamine and glutamine were studied in this work, the authors argue that this approach is feasible for other organic metabolites as well. In this work, only targeted analysis of dopamine and glutamine was done using SRM MS for quantitation. Cells needed to be suspended manually in several steps, decreasing the throughput and prohibiting *in vivo* analysis.

Over the past years, considerable progress was made in the automation of the different steps (including recognition, isolation, lysis, sampling, derivatization and separation) of single-cell MS. MS analyses have also been executed in high throughput, and the distinction of different cell types using the obtained metabolic profiles has been done successfully. Not all of these novel methods, however, (i) have been applied in combination and (ii) have been applied for metabolomics studies. A wide variety of technologies have been published that are suitable for high-throughput analysis of single cells, and some have already been used for that purpose as well. Together, these cover a large part of the steps that are part of single-cell isolation, metabolite extraction and sample enrichment. Separations before MS analysis offer higher identification rates at the cost of longer analysis time, and additional orthogonality can also be provided by IMS or MS/MS analysis. The use of micro-sampling probes seems to minimize the effect of ionization suppression as fewer matrix interferents are present, and they may be combined with automated micromanipulation platforms so they can be used for high-throughput analysis.

3 | DATA ANALYSIS AND SHARING

The increasing sensitivity of instruments, the number of separation dimensions and the number of data sets available together pose a need for data analytical tools that can handle these vast amounts of data. This brings us to the realm of statistics, ML and DL. Statistics and ML differ in that statistics is more aimed at fitting probability models to data, whereas ML is more aimed at finding patterns in the investigated data [146]. ML is suitable for data sets containing more variables than observations, which is not the case for statistical tools such as regression (Section 3.1). An extensive review of the use of statistics in metabolomics can be found in the works by Broadhurst and Kell [147] and Antonelli et al. [148]. ML approaches for metabolomics data have recently been reviewed by Liebal et al. [149]. Next to statistics and ML, DL may also be useful, as it can outperform ML when extremely large data sets need to be processed, but this requires large amounts of data (Section 3.2). To increase the throughput, there may be an interest in more stand-alone methods for data analysis; these are discussed in Section 3.3. Lastly, as the vast amounts of data needed for proper statistics can also be obtained by sharing data between research groups, Section 3.4 is devoted to FAIR data guidelines and steps that need to be taken in this regard.

Besides the algorithm used for sample classification and feature identification, one must also pay attention to the careful pre-processing of the data. This includes for instance missing value imputation, normalization, transformation and so forth, and these can be used in different combinations. To optimize this, an R package called NOREVA has been developed by Fu et al. [150]. Data pre-processing will not be considered further here.

3.1 | Statistical tools and machine learning

In contrast to LC–MS data analysis in metabolomics, there are currently no golden standards for data analysis in single-cell MS studies [14]. Furthermore, the high-dimensionality of the data complicates data analysis. In terms of techniques to handle this high-dimensionality, one can distinguish between univariate techniques (only one variable is considered per test) and multivariate techniques (multiple variables are considered). Univariate techniques include, amongst others, *t*-tests and Kruskal–Wallis tests. Examples of multivariate techniques are dimension-reduction techniques such as PLS-DA and PCA-DA, and classification techniques such as RF classification and logistic regression. When applying statistical tests, one must ensure that the assumptions underlying the test in question are met: Shapiro–Wilk tests can be used to assess the normality of the data, and Levene's test (in case of normally distributed data) or Bartlett's test (in case data is not normally distributed) can be used to check whether the variances of the two groups of samples compared are equal.

When executing large numbers of univariate tests, one must do multiple-test correction: Comparison of many metabolites may lead to chance correlations [151]. Straightforward ways of multiple-test correction are the Bonferroni correction [152] and the Benjamini–Hochberg correction [153], but multiple alternatives exist [147]. Another way

to do multiple-test correction is the use of a so-called *q*-value, a user-chosen percentage that describes the share of false positives considered acceptable. Multiple-test correction may lead to differences being considered insignificant, even though they were significant before correction (for an example, see the work by Clinton et al. [154]).

The work by Gilard et al. [155] shows how metabolites can be identified by looking at the system as a whole, rather than individual metabolites. Gaussian graphical models (GGMs) and RF classification were used to find the most informative variables in predicting the diagnosis and optimal treatment of glioblastoma. GGMs were used for preliminary identification of disease-specific metabolites and lipids: Patient-specific metabolic networks were identified, after which the metabolites that occurred in at least 50% of the patient-specific networks were used for further analysis. Using these variables, RF models were built for all different combinations, so the predictivity of these variables could be determined. In this way, they showed how a systems-based approach can be used to identify biomarkers for characterizing glioblastoma. This way of using RF models, however, necessitates a strict selection of metabolites of interest, as the number of models to be evaluated grows almost exponentially with the number of metabolites considered.

Data analysis in omics studies is generally complicated by high-dimensionality. Yao et al. [135] made use of a variety of dimension-reduction techniques to distinguish among different subtypes of breast cancer. One of these techniques is *t*-distributed stochastic neighbour embedding (*t*-SNE). *t*-SNE reduces a high-dimensional data set to a two-dimensional data set. It makes use of a user-defined 'perplexity', which can be seen as the estimated number of neighbours a datapoint has. A limitation of using *t*-SNE is that one needs to optimize the perplexity value which can influence the clustering and that it does not provide insight into which variables are most predictive in classification. Therefore, Yao et al. also used linear discriminant analysis for classification. In this way, they were able to classify single tumour cells according to their subtype, as well as identify biomarkers useful in this classification.

In later work of this group [156], a data processing workflow is developed to discriminate between different subtypes of leukemic cells. Supervised classification was done using a naïve Bayes classifier (NBC) with kernel density estimation (KDE). An NBC calculates the probability an observation belongs to a particular class: Probability coefficients are calculated independently for all variables, KDE is used to smoothen the probability distributions, and a threshold function is used for the assignment of samples to classes. *t*-SNE and *k*-nearest neighbour classifications are used to visualize the heterogeneity. Based on the results, multiple biomarker candidates for leukaemia subtypes were proposed.

Xie et al. [35] developed a more local RF-based classification algorithm for the classification of different types of brain cells. An RF model was fitted to the MS data to obtain information on the relative importance of the different predictors; instead of the – conventionally used – Gini importance, shapley additive explanation (SHAP) values [157] were used to identify the most important predictors. This method considers local and global patterns in the assignment of importance scores, whereas the Gini importance is a global metric. As the matrix with SHAP values has the same shape as the original matrix, PCA is applied to show that different cells can be grouped according to their subtypes. The most predictive variables can be identified using the mean absolute SHAP values.

A well-known technique for classification based on high-dimensional data is PLS-DA. Liu et al. [12] made use of orthogonal PLS-DA (OPLS-DA) for the classification of HeLa cells that had been exposed to different anticancer treatments. The difference between OPLS-DA and PLS-DA is that the latent variables (LVs) in PLS-DA all explain variation in the dependent and independent data, whereas in OPLS-DA, only the first LV is a predictor of class, and the subsequent LVs explain variation with respect to that first LV [158]. Hence, the predictive information is concentrated in the first LV, which makes the interpretation of the loadings more straightforward. Furthermore, permutation tests were used in their work to assess the model quality: Using permutation tests probability distributions can be determined for performance metrics (e.g. prediction accuracy), so their significance can be evaluated [159].

Lombard-Banek et al. [113] made use of hierarchical clustering analysis (HCA) to study how divisions in tadpole embryos are controlled. First, *t*-tests were applied to find metabolite intensity differences, after which HCA was applied to the 40 metabolites that were most differentially present. Subsequently, they determined which metabolites were more abundant in one of two classes using an intensity heat map. With this approach, they were able to identify metabolites that play a key role in the control of embryonic cleavage, thereby also getting more understanding of the biochemical reactions controlling this process.

In conclusion, a wide variety of data-analytical techniques have been combined to get more insight into whether there are differences between subtypes of cells, which metabolites can be used for classification and if these metabolites correlate. When applying statistical tests, one needs to make sure that the assumptions the test in question relies on are met, or one should use a test that does not make these assumptions in the first place. Furthermore, when doing classification, one needs to make

sure that one is not overfitting the data, for instance using permutation tests. Lastly, after establishing the statistical significance of a particular biomarker, its biological relevance needs to be established as well. For this purpose, a wide variety of pathway analysis tools exists [160].

3.2 | Deep learning

Besides statistics and ML, DL is getting more attention as well in data analysis but also other parts of the workflow. The terms ML and DL are often used interchangeably (and not entirely incorrectly so) but there is an important difference: ML requires the user to define features in the data, whereas DL selects these by itself, therefore making the results more objective and allowing for higher throughput [161]. DL has already found its way to the field of proteomics [40, 162–164], is suggested to be a valuable tool in the integration of omics data [38, 165] and has shown its merit in MS imaging [166]. Furthermore, the increase in ML algorithm performance eventually levels off for increasing data set size, whereas DL algorithm performance keeps improving, thereby surpassing ML algorithm performance. DL has been applied for a variety of steps in the single-cell analysis workflow, some examples of which are discussed in this section.

Lamanna et al. [138] used a CNN for cell recognition; the user can define criteria to determine which cells can be selected, or the process can be done in a fully automated fashion, allowing for a throughput of approximately five cells per minute. CNNs repetitively filter the picture and eventually convert it to an array of numbers, to which a DL classification network is applied. Even though a high accuracy was reached using this CNN, a downside is that they require the user to define the shape of the network, which requires some expertise at the user's end. Furthermore, this optimization is time-consuming.

Liu et al. [167] developed a software framework called 'Trace' that makes use of ML and DL for feature identification and optimization of extraction of trace-level signals from CE-HRMS data. Pattern recognition is done using CNNs, as these do not require pre-defined features. They show that peak picking can be automated, reducing the data analysis time from weeks to less than an hour. Evaluation of the trained model showed that it is robust to changes in migration times and biological replicates, and signal intensities were determined with high reproducibility. Comparison with ML models built for the same data set showed that CNNs for instance outperformed RF models. Limitations are that even though the DL algorithm can learn to recognize patterns by itself, the user still needs to roughly define various parameters describing the network such as the threshold function. Furthermore, a large amount of data is required to obtain reliable models.

In the work by Niu et al. [168] a DL algorithm is used to remove the batch effects from MALDI-MS and CyTOF data and subsequently classify samples based on disease diagnosis. Inspection of the results shows that before batch effect removal (calibration), *t*-SNE clustering takes mainly place based on batch number, whereas after calibration, a clear distinction between samples corresponding to different diagnoses is visible. Furthermore, the classification accuracy of the proposed DL algorithm is compared with that of conventional classification algorithms: This showed that using their newly developed algorithm, higher prediction accuracies for disease diagnosis were obtained. However, in their work, there is no mention of how one can extract information on which variables are most important in prediction; this is relevant as these might provide insights for biomarker identification or drug development.

The relevance of DL has also been investigated for the identification of biomarkers. In the work by Papagiannopoulou et al. [37], both ML and DL models are used to predict the identity of pathogenic micro-organisms based on MALDI-TOF-MS spectra. Optimization of the ML models (logistic regression, RF and *k*-nearest neighbours) and DL models (fully convolutional networks, 2-layer FCN and 3-layer FCN, followed by a concatenation step) was done using a grid search in Python and using the PyTorch library in Python, respectively. To train the DL models properly, the data is augmented to have a large enough training data set. Similar prediction accuracies were obtained for ML and DL models. However, the expectation is raised that, when more observations are available, DL models will outperform ML models. Furthermore, data set augmentation may be feasible for the improvement of image analysis but is complicated for chemical data sets with samples and their compositions.

In conclusion, the advantages of DL in single-cell metabolomics mainly lie in the automation of the analytical workflow, be it by the identification of single cells or data processing. When it comes to feature identification in high-dimensional data sets, currently there may not be enough data available to make DL outperform ML. Furthermore, resolving the most important variables in prediction from a neural network is not straightforward, complicating biomarker identification.

3.3 | Automation of data analysis for higher throughput

Increasing the throughput of single-cell analyses also places a larger burden on data analysis, as this must be able to handle and make sense of larger amounts of data with

less user intervention. Optimization of data pre-processing can be done using the NOREVA R package developed by Fu et al. [150]. As the code is openly available in this and many other cases, it can easily be incorporated into automated data analysis workflows. However, this requires the user to have sufficient knowledge of programming in, for example R or Python. Another advantage of DL with respect to ML is that DL models require less user intervention: Only the network needs to be designed, but the algorithm itself is capable of extracting features to learn the data. For DL models to outperform ML models, however, large amounts of data are required, and these are not always available in the novel field of single-cell metabolomics. Furthermore, the construction of the network can take some time, and the network design and parameters need to be optimized as well.

3.4 | Data handling and FAIR data

Besides increasing the throughput of single-cell analyses, sharing data between research groups can vastly increase the amount of data available so one can draw more substantiated conclusions. FAIR data is especially important in multi-omics studies, so one can compare data from different sample sizes and different measurement techniques [169]. Conversely, Veyel et al. [170] studied protein–metabolite interactions making use of co-fractionation, taking advantage of the size difference between proteins and metabolites; applications of co-fractionation MS to study protein–metabolite interactions have been reviewed recently by Schlossarek et al. [171]. Herein, the focus lies on the combination of multiple (single-cell) metabolomics studies.

Striving for FAIR data puts constrictions on the data format used for publishing – that is raw data should be openly accessible, irrespective of the vendor of the analytical instrument – but also on the metadata, which should include the details needed for reproducibility and intelligible sample annotation. Efforts in this direction have been made over the last two decades in the field of metabolomics [172, 173]; this includes for instance the establishment of shared repositories and concomitant reporting standards [174–176], data formats [177, 178] and data analysis streamlining [176]. However, Spicer et al. [179] investigated the compliance of metabolomics data sets with the guidelines of the public repository they were in and found that the extent of compliance varies enormously but is generally insufficient. Next to FAIR data storage, (partial) method standardization is required to prevent the scattering of data matrices available. Interlaboratory metabolomics studies [180, 181] have shown the importance of using standardized methods and references, but we were not able to find interlaboratory studies in single-cell metabolomics. There are still steps to be taken to make single-cell metabolomics data FAIR, both in terms of storing (meta)data as well as in terms of analytical workflows.

Another complicating factor in making use of FAIR data is that the data processing pipeline is occasionally reported inadequately [182], which hampers the comparison of results from different studies. A wide variety of R packages is openly available for data processing and analysis, providing, in combination with experimental data, a platform for reproducible research [183]. Furthermore, Dekermanjian et al. [184] published an ML-generated catalogue of software tools available for the analysis of metabolomics data. A long list of open-source software tools is available for i.a. data formatting, metabolite annotation and metadata handling in single-cell metabolomics [185]. To benefit to the fullest extent of FAIR data use, authors need to guide their readers through their approach to data analysis, as there is a multitude of data analysis methods available.

As the field of single-cell metabolomics is a relatively young one, to the best of our knowledge, there are no dedicated data set repositories as there are for proteomics and metabolomics. In 2020, Rozenblatt-Rosen et al. [186] set out to create the Human Tumor Atlas Network, to provide a framework for the publication of single-cancer cell data sets, corresponding to different organs, cell types, conditions and states. A wide variety of data analysis platforms for the analysis of single-cell genomics data is available, but the transfer of these platforms to the field of single-cell metabolomics is complicated by the fact that genomics relies on sequencing, whereas single-cell metabolomics typically relies on MS data.

4 | CONCLUDING REMARKS

Cellular heterogeneity in metabolism, drug resistance and understanding disease mechanisms pose the need for analytical tools that can analyse single cells. As the metabolome provides the closest view of the status quo in a cell, this is the most suitable study subject. Due to its high sensitivity and resolution, MS is most suitable for single-cell metabolomic analysis. In this work, a variety of approaches for the different steps in single-cell MS analysis have been discussed, ranging from single-cell isolation to metabolite extraction to sample enrichment. A considerate choice of set-up can also reduce the risk of ionization suppression. Large numbers of cells need to be analysed to do proper statistics. For the high-throughput isolation of single cells in suspension, microfluidic chips employing Dean flow or with serpentine shapes are useful, whereas DL image recognition can be used for the resolution of single cells in tissues. Suspended cells can also be

isolated using LVC: This allows for high throughput, and the cell stays in its natural environment until it is lysed. Lysis of single cells in suspension can be done in high throughput using electroporation or laser irradiation, both of which have already been incorporated into microfluidic chips. CyESI–MS enables the online isolation, extraction and lysis of single cells in suspension, allowing for real-time MS analysis. CE–MS analyses have also been executed in high throughput using microfluidic chips, together with IMS.

A wide variety of strategies have been proposed for the processing, treatment and analysis of single-cell metabolomics data, some of which also allow for more automation of data analysis. However, one needs to carefully check the assumptions that underlie the tests used. Furthermore, one must do multiple-test correction when multiple univariate tests are used and apply permutation testing in the case of classification analysis. DL seems interesting for data processing, image recognition and data analysis, as it handles vast amounts of data more efficiently than ML algorithms and requires less user intervention. As large numbers of cells are required to draw statistically significant conclusions, the field of single-cell metabolomics can strongly benefit from the use of FAIR data and adherence to the concomitant guidelines.

All in all, a lot of progress is being made in the analysis of single cells, reaching higher sensitivities, higher spatial and mass resolutions, higher identification rates, wider coverages and higher throughputs. Individual steps of the workflow have virtually all been automated, and tools for the automation of data processing, treatment and analysis have also become available. However, these methods have generally not been integrated. Furthermore, an automation of biomarker identification will also pose a challenge for molecular pathway analysis: The number of biomarkers to be given a place in metabolic pathways will quickly increase, so there will also be an increasing demand for the automation of biological pathway analysis. To make use of FAIR data to the fullest extent, more standardization of data processing and experimental methods is needed. On a more global scale, databases of single-cell metabolomes are required. Single-cell MS analysis in metabolomics has already led to the (tentative) identification of a variety of biomarkers for a multitude of diseases, emphasizing its value for diagnosis, drug discovery and treatment.

ACKNOWLEDGEMENTS

Open Access funding provided by VSNU.

CONFLICT OF INTEREST STATEMENT

The authors have declared no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data was created or analysed in this study.

ORCID

Rawi Ramautar https://orcid.org/0000-0002-1673-4848

REFERENCES

1. Altschuler SJ, Wu LF. Cellular heterogeneity: do differences make a difference? Cell. 2010;141:559–63.
2. Heppner GH. Tumor heterogeneity. Cancer Res. 1984;44:2259–65.
3. Tellez-Gabriel M, Ory B, Lamoureux F, Heymann M-F, Heymann D. Tumour heterogeneity: the key advantages of single-cell analysis. Int J Mol Sci. 2016;17:2142.
4. Krug K, Jaehnig EJ, Satpathy S, Blumenberg L, Karpova A, Anurag M, et al. Proteogenomic landscape of breast cancer tumorigenesis and targeted therapy. Cell. 2020;183:1436–56.e31.
5. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics. 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.
6. Plaks V, Koopman CD, Werb Z. Circulating tumor cells. Science (80–). 2013;341:1186–88.
7. Liu R, Sun M, Zhang G, Lan Y, Yang Z. Towards early monitoring of chemotherapy-induced drug resistance based on single cell metabolomics: combining single-probe mass spectrometry with machine learning. Anal Chim Acta. 2019;1092:42–8.
8. Blow N. Biochemistry's new look. Nature. 2008;455:697–8.
9. Emara S, Amer S, Ali A, Abouleila Y, Oga A, Masujima T. Single-cell metabolomics. In: Metabolomics: *from fundamentals to clinical applic*ations. Cham: Springer; 2017. p. 323–43.
10. Thiele C, Wunderling K, Leyendecker P. Multiplexed and single cell tracing of lipid metabolism. Nat Methods. 2019;16:1123–30.
11. Wang R, Zhao H, Zhang X, Zhao X, Song Z, Ouyang J. Metabolic discrimination of breast cancer subtypes at the single-cell level by multiple microextraction coupled with mass spectrometry. Anal Chem. 2019;91:3667–74.
12. Liu R, Pan N, Zhu Y, Yang Z. T-Probe: an integrated microscale device for online *in situ* single cell analysis and metabolic profiling using mass spectrometry. Anal Chem. 2018;90:11078–85.
13. Qi M, Philip MC, Yang N, Sweedler JV. Single cell neurometabolomics. ACS Chem Neurosci. 2018;9:40–50.
14. Liu Q, Ge W, Wang T, Lan J, Martínez-Jarquín S, Wolfrum C, et al. High-throughput single-cell mass spectrometry reveals abnormal lipid metabolism in pancreatic ductal adenocarcinoma. Angew Chem. 2021;133:24739–47.
15. Breadmore MC, Grochocki W, Kalsoom U, Alves MN, Phung SC, Rokh MT, et al. Recent advances in enhancing the sensitivity of electrophoresis and electrochromatography in capillaries and microchips (2016–2018). Electrophoresis. 2019;40:17–39.
16. Šlampová A, Malá Z, Gebauer P. Recent progress of sample stacking in capillary electrophoresis (2016–2018). Electrophoresis. 2019;40:40–54.

17. Yu RB, Quirino JP. Pseudophase-aided in-line sample concentration for capillary electrophoresis. TrAC Trends Anal Chem. 2023;161:116914.
18. Miggiels P, Wouters B, van Westen GJP, Dubbelman A-C, Hankemeier T. Novel technologies for metabolomics: more for less. TrAC Trends Anal Chem. 2019;120:115323.
19. Emwas A-H, Roy R, McKay RT, Tenori L, Saccenti E, Gowda GAN, et al. NMR spectroscopy for metabolomics research. Metabolites. 2019;9:123.
20. van Mever M, Hankemeier T, Ramautar R. CE–MS for anionic metabolic profiling: an overview of methodological developments. Electrophoresis. 2019;40:2349–59.
21. Stolz A, Jooß K, Höcker O, Römer J, Schlecht J, Neusüß C. Recent advances in capillary electrophoresis-mass spectrometry: instrumentation, methodology and applications. Electrophoresis. 2019;40:79–112.
22. Zhang W, Ramautar R. CE-MS for metabolomics: developments and applications in the period. 2018–2020. Electrophoresis. 2021;42:381–401.
23. Soga T. Advances in capillary electrophoresis mass spectrometry for metabolomics. TrAC Trends Anal Chem. 2023;158:116883.
24. Zhang L, Vertes A. Single-cell mass spectrometry approaches to explore cellular heterogeneity. Angew Chem Int Ed. 2018;57:4466–77.
25. Draper J, Lloyd AJ, Goodacre R, Beckmann M. Flow infusion electrospray ionisation mass spectrometry for high throughput, non-targeted metabolite fingerprinting: a review. Metabolomics. 2013;9:4–29.
26. Tsuyama N, Mizuno H, Tokunaga E, Masujima T. Live single-cell molecular analysis by video-mass spectrometry. Anal Sci. 2008;24:559–61.
27. Masujima T. Live single-cell mass spectrometry. Anal Sci. 2009;25:953–60.
28. Hiyama E, Ali A, Amer S, Harada T, Shimamoto K, Furushima R, et al. Direct lipido-metabolomics of single floating cells for analysis of circulating tumor cells by live single-cell mass spectrometry. Anal Sci. 2015;31:1215–7.
29. Gross A, Schoendube J, Zimmermann S, Steeb M, Zengerle R, Koltay P. Technologies for single-cell isolation. Int J Mol Sci. 2015;16:16897–919.
30. Feng D, Xu T, Li H, Shi X, Xu G. Single-cell metabolomics analysis by microfluidics and mass spectrometry: recent new advances. J Anal Test. 2020;4:198–209.
31. Naghdi E, Moran GE, Reinau ME, De Malsche W, Neusüß C. Concepts and recent advances in microchip electrophoresis coupled to mass spectrometry: technologies and applications. Electrophoresis. 2023;44:246–67.
32. Xu S, Yang C, Yan X, Liu H. Towards high throughput and high information coverage: advanced single-cell mass spectrometric techniques. Anal Bioanal Chem. 2022;414:219–33.
33. Liu R, Zhang G, Sun M, Pan X, Yang Z. Integrating a generalized data analysis workflow with the single-probe mass spectrometry experiment for single cell metabolomics. Anal Chim Acta. 2019;1064:71–9.
34. Chen X, Sun M, Yang Z. Single cell mass spectrometry analysis of drug-resistant cancer cells: metabolomics studies of synergetic effect of combinational treatment. Anal Chim Acta. 2022;1201:339621.
35. Xie YR, Castro DC, Bell SE, Rubakhin SS, Sweedler JV. Single-cell classification using mass spectrometry through interpretable machine learning. Anal Chem. 2020;92:9338–47.
36. Liu R, Zhang G, Yang Z. Towards rapid prediction of drug-resistant cancer cell phenotypes: single cell mass spectrometry combined with machine learning. Chem Commun. 2019;55:616–9.
37. Papagiannopoulou C, Parchen R, Rubbens P, Waegeman W. Fast pathogen identification using single-cell matrix-assisted laser desorption/ionization-aerosol time-of-flight mass spectrometry data and deep learning methods. Anal Chem. 2020;92:7523–31.
38. Chaudhary K, Poirion OB, Lu L, Garmire LX. Deep learning–based multi-omics integration robustly predicts survival in liver cancer. Clin Cancer Res. 2018;24:1248–59.
39. Asakura T, Date Y, Kikuchi J. Application of ensemble deep neural network to metabolomics studies. Anal Chim Acta. 2018;1037:230–6.
40. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nat Methods. 2020;17:41–4.
41. Messner CB, Demichev V, Bloomfield N, Yu JSL, White M, Kreidl M, et al. Ultra-fast proteomics with scanning SWATH. Nat Biotechnol. 2021;39:846–54.
42. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018.
43. Amantonico A, Urban PL, Zenobi R. Analytical techniques for single-cell metabolomics: state of the art and trends. Anal Bioanal Chem. 2010;398:2493–504.
44. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev. 2007;26:51–78.
45. Annesley TM. Ion suppression in mass spectrometry. Clin Chem. 2003;49:1041–4.
46. Lu W, Su X, Klein MS, Lewis IA, Fiehn O, Rabinowitz JD. Metabolite measurement: pitfalls to avoid and practices to follow. Annu Rev Biochem. 2017;86:277–304.
47. Dietmair S, Timmins NE, Gray PP, Nielsen LK, Krömer JO. Towards quantitative metabolomics of mammalian cells: development of a metabolite extraction protocol. Anal Biochem. 2010;404:155–64.
48. Amantonico A, Urban PL, Fagerer SR, Balabin RM, Zenobi R. Single-cell MALDI-MS as an analytical tool for studying intrapopulation metabolic heterogeneity of unicellular organisms. Anal Chem. 2010;82:7394–400.
49. Guo S, Zhang C, Le A. The limitless applications of single-cell metabolomics. Curr Opin Biotechnol. 2021;71:115–22.
50. Lee AY, Troisi J, Symes SJK. Experimental design in metabolomics. In: Troisi J, editor. Metabolomics perspectives. Amsterdam: Elsevier; 2022. p. 27–61.
51. Wang B, Young JD. 13 C-isotope-assisted assessment of metabolic quenching during sample collection from suspension cell cultures. Anal Chem. 2022;94:7787–94.
52. Onjiko RM, Moody SA, Nemes P. Single-cell mass spectrometry reveals small molecules that affect cell fates in the 16-cell embryo. Proc Natl Acad Sci USA. 2015;112:6545–50.

53. Nemes P, Knolhoff AM, Rubakhin SS, Sweedler JV. Metabolic differentiation of neuronal phenotypes by single-cell capillary electrophoresis–electrospray ionization-mass spectrometry. Anal Chem. 2011;83:6810–17.
54. Duncan KD, Fyrestam J, Lanekoff I. Advances in mass spectrometry based single-cell metabolomics. Analyst. 2019;144:782–93.
55. Palladino GW, Wood JJ, Proctor HJ. Modified freeze clamp technique for tissue assay. J Surg Res. 1980;28:188–90.
56. Canelas AB, ten Pierick A, Ras C, Seifar RM, van Dam JC, van Gulik WM, et al. Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics. Anal Chem. 2009;81:7379–89.
57. Zhang XC, Wei ZW, Gong XY, Si XY, Zhao YY, Yang CD, et al. Integrated droplet-based microextraction with ESI-MS for removal of matrix interference in single-cell analysis. Sci Rep. 2016;6:1–9.
58. Zhang XC, Zang Q, Zhao H, Ma X, Pan X, Feng J, et al. Combination of droplet extraction and pico-ESI-MS allows the identification of metabolites from single cancer cells. Anal Chem. 2018;90:9897–903.
59. Wei Z, Xiong X, Guo C, Si X, Zhao Y, He M, et al. Pulsed direct current electrospray: enabling systematic analysis of small volume sample by boosting sample economy. Anal Chem. 2015;87:11242–8.
60. Onjiko RM, Morris SE, Moody SA, Nemes P. Single-cell mass spectrometry with multi-solvent extraction identifies metabolic differences between left and right blastomeres in the 8-cell frog (*Xenopus*) embryo. Analyst. 2016;141:3648–56.
61. Whitesides GM. The origins and the future of microfluidics. Nature. 2006;442:368–73.
62. Kim D, Campos AR, Datt A, Gao Z, Rycenga M, Burrows ND, et al. Microfluidic-SERS devices for one shot limit-of-detection. Analyst. 2014;139:3227–34.
63. Reece A, Xia B, Jiang Z, Noren B, McBride R, Oakey J. Microfluidic techniques for high throughput single cell analysis. Curr Opin Biotechnol. 2016;40:90–6.
64. Ramautar R, Somsen GW, de Jong GJ. CE-MS for metabolomics: developments and applications in the period 2016–2018. Electrophoresis. 2019;40:165–79.
65. Gong X, Zhao Y, Cai S, Fu S, Yang C, Zhang S, et al. Single cell analysis with probe ESI-mass spectrometry: detection of metabolites at cellular and subcellular levels. Anal Chem. 2014;86:3809–16.
66. Li Z, Wang Z, Pan J, Ma X, Zhang W, Ouyang Z. Single-cell mass spectrometry analysis of metabolites facilitated by cell electro-migration and electroporation. Anal Chem. 2020;92:10138–44.
67. Fang Z, Wang R, Zhao H, Yao H, Ouyang J, Zhang X. Mannose promotes metabolic discrimination of osteosarcoma cells at single-cell level by mass spectrometry. Anal Chem. 2020;92:2690–6.
68. Wu J, Zhang W, Ouyang Z. On-demand mass spectrometry analysis by miniature mass spectrometer. Anal Chem. 2021;93:6003–7.
69. Pan N, Rao W, Kothapalli NR, Liu R, Burgett AWG, Yang Z. The single-probe: a miniaturized multifunctional device for single cell mass spectrometry analysis. Anal Chem. 2014;86:9376–80.
70. Nguyen TD, Lan Y, Kane SS, Haffner JJ, Liu R, McCall LI, et al. Single-cell mass spectrometry enables insight into heterogeneity in infectious disease. Anal Chem. 2022;94:10567–72.
71. Pan N, Standke SJ, Kothapalli NR, Sun M, Bensen RC, Burgett AWG, et al. Quantification of drug molecules in live single cells using the single-probe mass spectrometry technique. Anal Chem. 2019;91:9018–24.
72. Standke SJ, Colby DH, Bensen RC, Burgett AWG, Yang Z. Mass spectrometry measurement of single suspended cells using a combined cell manipulation system and a single-probe device. Anal Chem. 2019;91:1738–42.
73. Yoshimura K, Yamada Y, Ninomiya S, Chung WY, Chang Y-T, Dennison AR, et al. Real-time analysis of living animals and rapid screening of human fluid samples using remote sampling electrospray ionization mass spectrometry. J Pharm Biomed Anal. 2019;172:372–8.
74. Yamada Y, Ninomiya S, Hiraoka K, Chen LC. Development of remote sampling ESI mass spectrometry for the rapid and automatic analysis of multiple samples. Mass Spectrom. 2017;5:S0068.
75. Liu N, Liu Y, Yang Y, He L, Ouyang J. A simpler sampling interface of venturi easy ambient sonic-spray ionization mass spectrometry for high-throughput screening enzyme inhibitors. Anal Chim Acta. 2016;913:86–93.
76. Krylov SN, Dovichi NJ. Single-cell analysis using capillary electrophoresis: influence of surface support properties on cell injection into the capillary. Electrophoresis. 2000;21:767–73.
77. Zhu Y, Liu R, Yang Z. Redesigning the T-probe for mass spectrometry analysis of online lysis of non-adherent single cells. Anal Chim Acta. 2019;1084:53–9.
78. Cao Y-Q, Zhang L, Zhang J, Guo Y-L. Single-cell on-probe derivatization–noncontact nanocarbon fiber ionization: unraveling cellular heterogeneity of fatty alcohol and sterol metabolites. Anal Chem. 2020;92:8378–85.
79. Wu M-X, Wang H-Y, Zhang J-T, Guo Y-L. Multifunctional carbon fiber ionization mass spectrometry. Anal Chem. 2016;88:9547–53.
80. Ali A, Davidson S, Fraenkel E, Gilmore I, Hankemeier T, Kirwan JA, et al. Single cell metabolism: current and future trends. Metabolomics. 2022;18:77.
81. Taylor MJ, Lukowski JK, Anderton CR. Spatially resolved mass spectrometry at the single cell: recent innovations in proteomics and metabolomics. J Am Soc Mass Spectrom. 2021;32:872–94.
82. Hansen RL, Lee YJ. High-spatial resolution mass spectrometry imaging: toward single cell metabolomics in plant tissues. Chem Rec. 2018;18:65–77.
83. Passarelli MK, Pirkl A, Moellers R, Grinfeld D, Kollmer F, Havelund R, et al. The 3D OrbiSIMS—label-free metabolic imaging with subcellular lateral resolution and high mass-resolving power. Nat Methods. 2017;14:1175–83.
84. Krismer J, Sobek J, Steinhoff RF, Brönnimann R, Pabst M, Zenobi R. A MALDI-MS methodology for studying metabolic heterogeneity of single cells in a population. Methods Mol Biol. 2020;2064:113–24.
85. Berthelot H, Duhamel S, L'Helguen S, Maguer J-F, Wang S, Cetinić I, et al. NanoSIMS single cell analyses reveal the contrasting nitrogen sources for small phytoplankton. ISME J. 2019;13:651–62.

86. Bruinen AL, Fisher GL, Balez R, van der Sar AM, Ooi L, Heeren RMA. Identification and high-resolution imaging of α-tocopherol from human cells to whole animals by TOF-SIMS tandem mass spectrometry. J Am Soc Mass Spectrom. 2018;29:1571–81.
87. Ruiz-Rodado V, Lita A, Larion M. Advances in measuring cancer cell metabolism with subcellular resolution. Nat Methods. 2022;19:1048–63.
88. Lu Y, Cao Y, Zhang L, Lv Y, Zhang Y, Su Y, et al. Online quaternized derivatization mapping and glycerides profiling of cancer tissues by laser ablation carbon fiber ionization mass spectrometry. Anal Chem. 2022;94:3756–61.
89. Fukano Y, Tsuyama N, Mizuno H, Date S, Takano M, Masujima T. Drug metabolite heterogeneity in cultured single cells profiled by pico-trapping direct mass spectrometry. Nanomedicine. 2012;7:1365–74.
90. Mizuno H, Tsuyama N, Harada T, Masujima T. Live single-cell video-mass spectrometry for cellular and subcellular molecular detection and cell classification. J Mass Spectrom. 2008;43:1692–700.
91. Li Z, Cheng S, Lin Q, Cao W, Yang J, Zhang M, et al. Single-cell lipidomics with high structural specificity by mass spectrometry. Nat Commun. 2021;12:2869.
92. Liu Q, Lan J, Wu R, Begley A, Ge W, Zenobi R. Hybrid ionization source combining nanoelectrospray and dielectric barrier discharge ionization for the simultaneous detection of polar and nonpolar compounds in single cells. Anal Chem. 2022;94:2873–81.
93. Ahmed E, Xiao D, Dumlao MC, Steel CC, Schmidtke LM, Fletcher J, et al. Nanosecond pulsed dielectric barrier discharge ionization mass spectrometry. Anal Chem. 2020;92:4468–74.
94. Liu R, Yang Z. Single cell metabolomics using mass spectrometry: techniques and data analysis. Anal Chim Acta. 2021;1143:124–34.
95. Koek MM, Bakels F, Engel W, van den Maagdenberg A, Ferrari MD, Coulier L, et al. Metabolic profiling of ultrasmall sample volumes with GC/MS: from microliter to nanoliter samples. Anal Chem. 2010;82:156–62.
96. Koek MM, Muilwijk B, van der Werf MJ, Hankemeier T. Microbial metabolomics with gas chromatography/mass spectrometry. Anal Chem. 2006;78:1272–81.
97. Fairweather SJ, Okada S, Gauthier-Coles G, Javed K, Bröer A, et al. A GC-MS/single-cell method to evaluate membrane transporter substrate specificity and signaling. Front Mol Biosci. 2021;8:646574. https://doi.org/10.3389/fmolb.2021.646574
98. Tang D-Q, Zou L, Yin X-X, Ong CN. HILIC-MS for metabolomics: an attractive and complementary approach to RPLC-MS. Mass Spectrom Rev. 2016;35:574–600.
99. Ramautar R. Advanced mass spectrometry-based analytical separation techniques for probing the polar metabolome. Cambridge: Royal Society of Chemistry; 2021.
100. Cui L, Lu H, Lee YH. Challenges and emergent solutions for LC-MS/MS based untargeted metabolomics in diseases. Mass Spectrom Rev. 2018;37:772–92.
101. Mansuri MS, Williams K, Nairn AC. Uncovering biology by single-cell proteomics. Commun Biol. 2023;6:381.
102. Budnik B, Levy E, Harmange G, Slavov N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. Genome Biol. 2018;19:161.
103. Zhu Y, Piehowski PD, Zhao R, Chen J, Shen Y, Moore RJ, et al. Nanodroplet processing platform for deep and quantitative proteome profiling of 10–100 mammalian cells. Nat Commun. 2018;9:882.
104. Brunner A, Thielert M, Vasilopoulou C, Ammar C, Coscia F, Mund A, et al. Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. Mol Syst Biol. 2022;18:e10798.
105. Ludwig C, Gillet L, Rosenberger G, Amon S, Collins BC, Aebersold R. Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. Mol Syst Biol. 2018;14:e8126.
106. Vasiljevic T, Singh V, Pawliszyn J. Miniaturized SPME tips directly coupled to mass spectrometry for targeted determination and untargeted profiling of small samples. Talanta. 2019;199:689–97.
107. Boelaert J, Lynen F, Glorieux G, Schepers E, Neirynck N, Vanholder R. Metabolic profiling of human plasma and urine in chronic kidney disease by hydrophilic interaction liquid chromatography coupled with time-of-flight mass spectrometry: a pilot study. Anal Bioanal Chem. 2017;409:2201–11.
108. DeLaney K, Sauer C, Vu N, Li L. Recent advances and new perspectives in capillary electrophoresis-mass spectrometry for single cell "omics". Molecules. 2018;24:42.
109. Onjiko RM, Plotnick DO, Moody SA, Nemes P. Metabolic comparison of dorsal versus ventral cells directly in the live 8-cell frog embryo by microprobe single-cell CE-ESI-MS. Anal Methods. 2017;9:4964–70.
110. Huang L, Fang M, Cupp-Sutton KA, Wang Z, Smith K, Wu S. Spray-capillary-based capillary electrophoresis mass spectrometry for metabolite analysis in single cells. Anal Chem. 2021;93:4479–87.
111. Kawai T, Ota N, Okada K, Imasato A, Owa Y, Morita M, et al. Ultrasensitive single cell metabolomics by capillary electrophoresis-mass spectrometry with a thin-walled tapered emitter and large-volume dual sample preconcentration. Anal Chem. 2019;91:10564–72.
112. Liao H-W, Rubakhin SS, Philip MC, Sweedler JV. Enhanced single-cell metabolomics by capillary electrophoresis electrospray ionization-mass spectrometry with field amplified sample injection. Anal Chim Acta. 2020;1118:36–43.
113. Lombard-Banek C, Li J, Portero EP, Onjiko RM, Singer CD, Plotnick DO, et al. *In vivo* subcellular mass spectrometry enables proteo-metabolomic single-cell systems biology in a chordate embryo developing to a normally behaving tadpole (*X. laevis*). Angew Chem Int Ed. 2021;60:12852–58.
114. Lombard-Banek C, Moody SA, Manzini MC, Nemes P. Microsampling capillary electrophoresis mass spectrometry enables single-cell proteomics in complex tissues: developing cell clones in live *Xenopus laevis* and zebrafish embryos. Anal Chem. 2019;91:4797–805.
115. Onjiko RM, Portero EP, Moody SA, Nemes P. *In situ* microprobe single-cell capillary electrophoresis mass spectrometry: metabolic reorganization in single differentiating cells in the live vertebrate (*Xenopus laevis*) embryo. Anal Chem. 2017;89:7069–76.

116. Portero EP, Nemes P. Dual cationic–anionic profiling of metabolites in a single identified cell in a live *Xenopus laevis* embryo by microprobe CE-ESI-MS. Analyst. 2019;144:892–900.
117. Luo M-D, Zhou Z-W, Zhu Z-J. The application of ion mobility-mass spectrometry in untargeted metabolomics: from separation to identification. J Anal Test. 2020;4:163–74.
118. Dodds JN, Baker ES. Ion mobility spectrometry: fundamental concepts, instrumentation, applications, and the road ahead. J Am Soc Mass Spectrom. 2019;30:2185–95.
119. Zhang L, Foreman DP, Grant PA, Shrestha B, Moody SA, Villiers F, et al. *In situ* metabolic analysis of single plant cells by capillary microsampling and electrospray ionization mass spectrometry with ion mobility separation. Analyst. 2014;139:5079–85.
120. Taylor MJ, Mattson S, Liyu A, Stopka SA, Ibrahim YM, Vertes A, et al. Optical microscopy-guided laser ablation electrospray ionization ion mobility mass spectrometry: ambient single cell metabolomics with increased confidence in molecular identification. Metabolites. 2021;11:200.
121. Samarah LZ, Khattar R, Tran TH, Stopka SA, Brantner CA, Parlanti P, et al. Single-cell metabolic profiling: metabolite formulas from isotopic fine structures in heterogeneous plant cell populations. Anal Chem. 2020;92:7289–98.
122. Taylor MJ, Liyu A, Vertes A, Anderton CR. Ambient single-cell analysis and native tissue imaging using laser-ablation electrospray ionization mass spectrometry with increased spatial resolution. J Am Soc Mass Spectrom. 2021;32:2490–4.
123. Mast DH, Liao H-W, Romanova EV, Sweedler JV. Analysis of peptide stereochemistry in single cells by capillary electrophoresis–trapped ion mobility spectrometry mass spectrometry. Anal Chem. 2021;93:6205–13.
124. Abouleila Y, Onidani K, Ali A, Shoji H, Kawai T, Lim CT, et al. Live single cell mass spectrometry reveals cancer-specific metabolic profiles of circulating tumor cells. Cancer Sci. 2019;110:697–706.
125. Jin X, Wang R-S, Zhu M, Jeon BW, Albert R, Chen S, et al. Abscisic acid–responsive guard cell metabolomes of *Arabidopsis* wild-type and gpa1 G-protein mutants. Plant Cell. 2014;25:4789–811.
126. Lecault V, VanInsberghe M, Sekulovic S, Knapp DJHF, Wohrer S, Bowden W, et al. High-throughput analysis of single hematopoietic stem cell proliferation in microfluidic cell culture arrays. Nat Methods. 2011;8:581–6.
127. Huang Q, Mao S, Khan M, Li W, Zhang Q, Lin J-M. Single-cell identification by microfluidic-based *in situ* extracting and online mass spectrometric analysis of phospholipids expression. Chem Sci. 2020;11:253–6.
128. Luo T, Fan L, Zhu R, Sun D. Microfluidic single-cell manipulation and analysis: methods and applications. Micromachines. 2019;10:104.
129. Lee JN, Park C, Whitesides GM. Solvent compatibility of poly(dimethylsiloxane)-based microfluidic devices. Anal Chem. 2003;75:6544–4.
130. Lee C-Y, Chang C-L, Wang Y-N, Fu L-M. Microfluidic mixing: a review. Int J Mol Sci. 2011;12:3263–87.
131. Zhang W, Li N, Lin L, Huang Q, Uchiyama K, Lin J. Concentrating single cells in picoliter droplets for phospholipid profiling on a microfluidic system. Small. 2020;16:1903402.
132. Huang Q, Mao S, Khan M, Zhou L, Lin J-M. Dean flow assisted cell ordering system for lipid profiling in single-cells using mass spectrometry. Chem Commun. 2018;54:2595–8.
133. Xu S, Liu M, Bai Y, Liu H. Multi-dimensional organic mass cytometry: simultaneous analysis of proteins and metabolites on single cells. Angew Chem Int Ed. 2021;60:1806–12.
134. Feng D, Li H, Xu T, Zheng F, Hu C, Shi X, et al. High-throughput single cell metabolomics and cellular heterogeneity exploration by inertial microfluidics coupled with pulsed electric field-induced electrospray ionization-high resolution mass spectrometry. Anal Chim Acta. 2022;1221:340116.
135. Yao H, Zhao H, Zhao X, Pan X, Feng J, Xu F, et al. Label-free mass cytometry for unveiling cellular metabolic heterogeneity. Anal Chem. 2019;91:9777–83.
136. Cahill JF, Riba J, Kertesz V. Rapid, untargeted chemical profiling of single cells in their native environment. Anal Chem. 2019;91:6118–26.
137. Brasko C, Smith K, Molnar C, Farago N, Hegedus L, Balind A, et al. Intelligent image-based *in situ* single-cell isolation. Nat Commun. 2018;9:226.
138. Lamanna J, Scott EY, Edwards HS, Chamberlain MD, Dryden MDM, Peng J, et al. Digital microfluidic isolation of single cells for -Omics. Nat Commun. 2020;11:5632.
139. Gebreyesus ST, Siyal AA, Kitata RB, Chen ES-W, Enkhbayar B, Angata T, et al. Streamlined single-cell proteomics by an integrated microfluidic chip and data-independent acquisition mass spectrometry. Nat Commun. 2022;13:37.
140. Zhao P, Cheng S, Feng Y, Liang F, Zhang X, Ma X, et al. Automated and miniaturized pico-liter metabolite extraction system for single-cell mass spectrometry. IEEE Trans Biomed Eng. 2023;70(2):470–8.
141. Peng X, Zhao L, Guo J, Yang S, Ding H, Wang X, et al. Double-helix micro-channels on microfluidic chips for enhanced continuous on-chip derivatization followed by electrophoretic separation. Biosens Bioelectron. 2015;72:376–82.
142. Wu H, Zhu J, Huang Y, Wu D, Sun J. Microfluidic-based single-cell study: current status and future perspective. Molecules. 2018;23:2347.
143. Vitorino R, Guedes S, da Costa JP, Kašička V. Microfluidics for peptidomics, proteomics, and cell analysis. Nanomaterials. 2021;11:1118.
144. Václavek T, Foret F. Microfluidic device integrating single-cell extraction and electrical lysis for mass spectrometry detection of intracellular compounds. Electrophoresis. 2023;44:313–22.
145. Li X, Zhao S, Hu H, Liu Y-M. A microchip electrophoresis-mass spectrometric platform with double cell lysis nano-electrodes for automated single cell analysis. J Chromatogr A. 2016;1451:156–63.
146. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods. 2018;15:233–4.
147. Broadhurst DI, Kell DB. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics. 2007;2:171–96.
148. Antonelli J, Claggett BL, Henglin M, Kim A, Ovsak G, Kim N, et al. Statistical workflow for feature selection in human metabolomics data. Metabolites. 2019;9:143.
149. Liebal UW, Phan ANT, Sudhakar M, Raman K, Blank LM. Machine learning applications for mass spectrometry-based metabolomics. Metabolites. 2020;10:243.

150. Fu J, Zhang Y, Wang Y, Zhang H, Liu J, Tang J, et al. Optimization of metabolomic data processing using NOREVA. Nat Protoc. 2022;17:129–51.
151. Saccenti E, Hoefsloot HCJ, Smilde AK, Westerhuis JA, Hendriks MMWB. Reflections on univariate and multivariate analysis of metabolomics data. Metabolomics. 2014;10:361–74.
152. VanderWeele TJ, Mathur MB. Some desirable properties of the Bonferroni correction: is the Bonferroni correction really so bad? Am J Epidemiol. 2019;188:617–8.
153. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B. 1995;57:289–300.
154. Clinton CM, Bain JR, Muehlbauer MJ, Li Y, Li L, O'Neal SK, et al. Non-targeted urinary metabolomics in pregnancy and associations with fetal growth restriction. Sci Rep. 2020;10:5307.
155. Gilard V, Ferey J, Marguet F, Fontanilles M, Ducatez F, Pilon C, et al. Integrative metabolomics reveals deep tissue and systemic metabolic remodeling in glioblastoma. Cancers (Basel). 2021;13:5157.
156. Yao H, Zhao H, Pan X, Zhao X, Feng J, Yang C, et al. Discriminating leukemia cellular heterogeneity and screening metabolite biomarker candidates using label-free mass cytometry. Anal Chem. 2021;93:10282–91.
157. Lundberg SM, Erion GG, Lee S-I. Consistent individualized feature attribution for tree ensembles. arXiv. 2018. https://doi.org/10.48550/arXiv.1802.03888
158. Westerhuis JA, van Velzen EJJ, Hoefsloot HCJ, Smilde AK. Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. Metabolomics. 2010;6:119–28.
159. Westerhuis JA, Hoefsloot HCJ, Smit S, Vis DJ, Smilde AK, van Velzen EJJ, et al. Assessment of PLSDA cross validation. Metabolomics. 2008;4:81–9.
160. Minakshi P, Ghosh M, Kumar R, Patki HS, Saini HM, Ranjan K, et al. Single-cell metabolomics: technology and applications. In: Barh D, Azevedo V, editors. Single-cell omics. Amsterdam: Elsevier; 2019. p. 319–53.
161. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436–44.
162. Yang Y, Lin L, Qiao L. Deep learning approaches for data-independent acquisition proteomics. Expert Rev Proteomics. 2021;18:1031–43.
163. Cadow J, Manica M, Mathis R, Guo T, Aebersold R, Rodríguez Martínez M. On the feasibility of deep learning applications using raw mass spectrometry data. Bioinformatics. 2021;37:245–53.
164. Sun B, Kovatch JR, Badiong A, Merbouh N. Optimization and modeling of quadrupole orbitrap parameters for sensitive analysis toward single-cell proteomics. J Proteome Res. 2017;16:3711–21.
165. Erfanian N, Heydari AA, Iañez P, Derakshani A, Ghasemigol M, Farahpour M, et al. Deep learning applications in single-cell omics data analysis. bioRxiv. 2021. https://doi.org/10.1101/2021.11.26.470166
166. Ma X, Fernández FM. Advances in mass spectrometry imaging for spatial cancer metabolomics. Mass Spectrom Rev. 2022. https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21804
167. Liu Z, Portero EP, Jian Y, Zhao Y, Onjiko RM, Zeng C, et al. Trace, machine learning of signal images for trace-sensitive mass spectrometry: a case study from single-cell metabolomics. Anal Chem. 2019;91:5768–76.
168. Niu J, Yang J, Guo Y, Qian K, Wang Q. Joint deep learning for batch effect removal and classification toward MALDI MS based metabolomics. BMC Bioinformatics. 2022;23:270.
169. Krassowski M, Das V, Sahu SK, Misra BB. State of the field in multi-omics research: from computational needs to data mining and sharing. Front Genet. 2020;11:610798. https://doi.org/10.3389/fgene.2020.610798
170. Veyel D, Kierszniowska S, Kosmacz M, Sokolowska EM, Michaelis A, Luzarowski M, et al. System-wide detection of protein-small molecule complexes suggests extensive metabolite regulation in plants. Sci Rep. 2017;7:42387.
171. Schlossarek D, Zhang Y, Sokolowska EM, Fernie AR, Luzarowski M, Skirycz A. Don't let go: co-fractionation mass spectrometry for untargeted mapping of protein–metabolite interactomes. Plant J. 2023;113:904–14.
172. Mendez KM, Pritchard L, Reinke SN, Broadhurst DI. Toward collaborative open data science in metabolomics using Jupyter Notebooks and cloud computing. Metabolomics. 2019;15:125.
173. Kumar A, Misra BB. Challenges and opportunities in cancer metabolomics. Proteomics. 2019;19:1900042.
174. Haug K, Salek RM, Conesa P, Hastings J, de Matos P, Rijnbeek M, et al. MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. Nucleic Acids Res. 2013;41:D781–6.
175. Sud M, Fahy E, Cotter D, Azam K, Vadivelu I, Burant C, et al. Metabolomics Workbench: an international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. Nucleic Acids Res. 2016;44:D463–70.
176. Domingo-Almenara X, Montenegro-Burke JR, Ivanisevic J, Thomas A, Sidibé J, Teav T, et al. XCMS-MRM and METLIN-MRM: a cloud library and public resource for targeted analysis of small molecules. Nat Methods. 2018;15:681–4.
177. Hoffmann N, Rein J, Sachsenberg T, Hartler J, Haug K, Mayer G, et al. mzTab-M: a data standard for sharing quantitative results in mass spectrometry metabolomics. Anal Chem. 2019;91:3302–10.
178. Larralde M, Lawson TN, Weber RJM, Moreno P, Haug K, Rocca-Serra P, et al. mzML2ISA and nmrML2ISA: generating enriched ISA-Tab metadata files from metabolomics XML data. Bioinformatics. 2017;33:2598–600.
179. Spicer RA, Salek R, Steinbeck C. Compliance with minimum information guidelines in public metabolomics repositories. Sci Data. 2017;4:170137.
180. Bowden JA, Heckert A, Ulmer CZ, Jones CM, Koelmel JP, Abdullah L, et al. Harmonizing lipidomics: NIST Interlaboratory comparison exercise for lipidomics using SRM 1950–Metabolites in Frozen Human Plasma. J Lipid Res. 2017;58:2275–88.
181. Siskos AP, Jain P, Römisch-Margl W, Bennett M, Achaintre D, Asad Y, et al. Interlaboratory reproducibility of a targeted metabolomics platform for analysis of human serum and plasma. Anal Chem. 2017;89:656–65.
182. Considine EC, Thomas G, Boulesteix AL, Khashan AS, Kenny LC. Critical review of reporting of the data analysis step in metabolomics. Metabolomics. 2018;14:7.

183. Stanstrup J, Broeckling C, Helmus R, Hoffmann N, Mathé E, Naake T, et al. The metaRbolomics toolbox in bioconductor and beyond. Metabolites. 2019;9:200.
184. Dekermanjian J, Labeikovsky W, Ghosh D, Kechris K. MSCAT: a machine learning assisted catalog of metabolomics software tools. Metabolites. 2021;11:678.
185. Misra BB. Open-source software tools, databases, and resources for single-cell and single-cell-type metabolomics. Methods Mol Biol. 2020;2064:191–217.
186. Rozenblatt-Rosen O, Regev A, Oberdoerffer P, Nawy T, Hupalowska A, Rood JE, et al. The human tumor atlas network: charting tumor transitions across space and time at single-cell resolution. Cell. 2020;181:236–49.

How to cite this article: Wevers D, Ramautar R, Clark C, Hankemeier T, Ali A. Opportunities and challenges for sample preparation and enrichment in mass spectrometry for single-cell metabolomics. Electrophoresis. 2023;44:2000–2024. https://doi.org/10.1002/elps.202300105