

# Single-Cell Metabolic Profiling: Metabolite Formulas from Isotopic Fine Structures in Heterogeneous Plant Cell Populations

Laith Z. Samarah,<sup>†</sup> Rikkita Khattar,<sup>†</sup> Tina H. Tran,<sup>†</sup> Sylwia A. Stopka,<sup>†</sup> [Christine A. Brantner,<sup>‡</sup>](#) [Paola Parlanti,<sup>‡</sup>](#) Dušan Veličković,<sup>§</sup> Jared B. Shaw,<sup>§</sup> Beverly J. Agtuca,<sup>⊥</sup> Gary Stacey,<sup>⊥</sup> Ljiljana Paša-Tolić,<sup>§</sup> Nikola Tolić,<sup>§</sup> Christopher R. Anderton,<sup>§</sup> and Akos Vertes<sup>\*,†</sup>

<sup>†</sup>Department of Chemistry, George Washington University, Washington, DC 20052

<sup>‡</sup>[Nanofabrication and Imaging Center, George Washington University, Washington, DC 20052](#)

<sup>§</sup>Environmental Molecular Sciences Laboratory and Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99354

<sup>⊥</sup>Divisions of Plant Sciences and Biochemistry, C. S. Bond Life Sciences Center, University of Missouri, Columbia, MO 65211

<sup>\*</sup>[Email: vertes@gwu.edu](mailto:vertes@gwu.edu)

---

**ABSTRACT:** Characterization of the metabolic heterogeneity in cell populations requires the analysis of single cells. Most current methods in single-cell analysis rely on cell manipulation, potentially altering the abundance of metabolites in individual cells. Small sample volume and the chemical diversity of metabolites are additional challenges in single-cell metabolomics. Here, we describe the combination of fiber-based laser ablation electrospray ionization (f-LAESI) with 21 tesla Fourier transform ion cyclotron resonance mass spectrometry (21TFTICR-MS) for in situ single-cell metabolic profiling in plant tissue. Single plant cells infected by bacteria were selected and sampled directly from the tissue without cell manipulation through mid-infrared ablation with a fine optical fiber tip for ionization by f-LAESI. Ultrahigh performance 21 T-FTICR-MS enabled the simultaneous capture of isotopic fine structures (IFSSs) for 47 known and 11 unknown compounds, thus elucidating their elemental compositions from single cells, and providing information on metabolic heterogeneity in the cell population.

---

In multicellular systems, different cell types can be closely interspersed, obscuring differences in their metabolite composition and active biochemical pathways. Common tissue-scale or cell-population-level analysis yields averaged abundance data that masks critical information about the metabolic state of individual cells, an important aspect of cellular heterogeneity.<sup>1-</sup>

<sup>2</sup> Mass spectrometry (MS) has become an essential technique in single cell analysis of metabolites due to its high sensitivity, broad molecular coverage, wide dynamic range, and its ability to provide structural information.<sup>3-4</sup> Assigning molecular formulas for metabolites, solely based on the accurate masses of the produced ions is often ambiguous, as many compounds with different elemental compositions have close to identical accurate masses. Ultrahigh performance Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) can reveal the elemental composition of compounds by exquisite mass accuracy and mass resolution, and by capturing isotopic fine structures (IFSSs), for increased confidence in their identification.<sup>5</sup> State-of-the-art 21 tesla (21 T) FTICR-MS offers ultrahigh mass resolution and exceptional mass accuracy. Owing to the complexity of their chemical compositions, characterization of natural organic matter, e.g., petroleum crude oil and soil, has been an important application for such high performance instruments.<sup>6-9</sup> The field of top-down proteomics has also profited

from ultrahigh resolution MS to tackle several challenges including identifying highly similar human proteoforms.<sup>10</sup> Recently, the application of ultrahigh resolution MS in metabolomics has been reported.<sup>11</sup>

Current methods that enable high-throughput single-cell analysis target the transcriptome, and rely on PCR amplification.<sup>12</sup> In addition, many high-throughput single-cell studies involve some degree of cell sorting, labeling and manipulation using techniques such as fluorescence- and magnetic-activated cell sorting. MS-based proteomic studies on single cells are emerging both for animal and human cell types.<sup>13-14</sup> Single-cell metabolomics by MS is constrained by the low sample volume (e.g., ~1 pL for a mammalian cell),<sup>15</sup> high turnover rates, and the structural complexity of certain metabolites.<sup>2</sup> The first two constraints limit the number of observed spectral features that can be obtained by MS from single cells, whereas structural complexity limits unambiguous interpretation of fragmentation in tandem MS. These challenges are more pronounced for the analysis of rare cells, e.g., circulating tumor cells, where obtaining maximum information from a limited sample volume is especially critical.<sup>16-17</sup>

To address the obstacles associated with single cell metabolomics, we combined fiber-based laser ablation electrospray ionization (f-LAESI) with a 21T-FTICR mass spectrometer for

*in situ* analysis of single plant cells. In our model system, the soybean (*Glycine max*) root nodule, a specialized organ associated with biological nitrogen fixation, plant cells infected by soil bacteria (*Bradyrhizobium japonicum*) and uninfected cells are intimately mixed.<sup>18</sup> Thus, this tissue presents an appropriate test case for *in situ* metabolic analysis of single cells in highly heterogeneous systems. Mid-infrared (IR) laser ablation by a sharp optical fiber tip enabled direct ambient analysis of cells without the need for sample processing, thus minimizing external perturbation to the cells prior to analysis.<sup>19-21</sup> The ultrahigh mass resolution, mass accuracy, and high dynamic range of the 21T-FTICR mass spectrometer were exploited to capture IFS for numerous metabolites simultaneously from single cells, increasing the confidence in their identification. The IFS were in turn computationally processed to rapidly determine the corresponding elemental compositions. In contrast, only one compound at a time could be fragmented for identification by tandem MS from single cells. This feature confers special benefits for the identification of unknown compounds in single and rare cells when tandem MS provides limited information and facilitates the characterization of cellular heterogeneity.

## EXPERIMENTAL SECTION

**Rhizobia cultures.** Cultures of wild-type *B. japonicum* USDA110 strain were incubated for 2 days at 30 °C in HM medium.<sup>22</sup> To prevent infection by other microorganisms, 25 mg/L of tetracycline and 100 mg/L of spectinomycin were added to the medium for the wild-type rhizobia. Optical density at a wavelength of 600 nm measured by a spectrophotometer (DeNovix DS-C, DeNovix Inc., Wilmington, DE, USA) was used to estimate the bacterial cell counts in the cultures. When the optical density reached 0.8 (or 10<sup>8</sup> cells/mL), the culture tubes were removed from the incubator, centrifuged at 3,000 rpm for 10 min and washed three times with sterile deionized water. Finally, sterile deionized water was added to the rhizobia pellet and vortexed to form a homogenous solution that was used later for inoculating soybean seeds.

**Soybean growth, inoculation with *B. japonicum* and root nodule harvesting.** Soybean (*Glycine max*) seeds of 'Williams 82' were surface-sterilized with 20% (v/v) bleach for 10 min and rinsed five times with sterile water. The sterile seeds were planted into pots containing a mixture of sterilized 3:1 vermiculite:perlite. For inoculation, 500 µL of the solution containing wild-type *B. japonicum* were added to each sterilized soybean seed. After covering the seeds with the potting material, the pots were placed in the Bond Life Sciences Center's greenhouse at the University of Missouri at 30 °C with a 16-h light/8-h dark cycle. These nodules were used in the IFS analysis, and the comparative studies for the three sample types. After 21 days of growth, the nodules attached to the primary root were harvested, immediately frozen in liquid nitrogen, and stored at -80 °C. For analysis of soybean root nodule cells by LAESI-MS, harvested nodules were embedded in 2.5% carboxymethylcellulose (CMC) in a mounting tray and placed in a cryostat microtome (Cryostar NX70, ThermoFisher Scientific, Hillsboro, OR, USA) at -10 °C until the embedding medium was frozen. For single-cell analysis, 30 µm sections were prepared and thaw-mounted on microscope glass slides.

**Optical fiber preparation for single cell analysis by f-LAESI-MS.** The preparation of GeO<sub>2</sub>-based glass optical fiber

for f-LAESI was described in a previous publication.<sup>19</sup> Briefly, the outer thermoplastic and polyamide coatings at the ends of a GeO<sub>2</sub>-based glass optical fiber (250 µm core diameter, HP Fiber, Infrared Fiber Systems, Inc., Silver Spring, MD, USA) of ~1 m in length were stripped from 2 cm sections by placing the ends of the fiber in a glass vial containing 10 mL of 1-methyl-2-pyrrolidinone and heated to 200 °C. When the outer coating was dissolved and the glass cladding was exposed, the optical fiber ends were rinsed with isopropanol. The two ends were then cleaved with a sharp, sterile blade to form straight, uniform surfaces. One of the fiber ends was etched to a sharp tip of ~10 µm radius of curvature by vertically immersing it into a 4% (v/v) nitric acid solution to a depth of ~500 µm. After the tip spontaneously detached from the acid solution surface, it was rinsed with deionized water to remove any acid residue.

**f-LAESI-MS.** Single-cell analysis based on f-LAESI-MS had been described in an earlier publication.<sup>19</sup> Briefly, the non-etched end of the optical fiber was held by a bare fiber chuck (BFC300, Siskiyou Corporation, Grants Pass, OR, USA) and secured in a single-mode fiber coupler (F-915, Newport Corporation, Irvine, CA, USA). Mid-IR laser radiation at 2940 nm was produced by a diode-pumped Nd:YAG laser-driven optical parametric oscillator (OPO) (Opolette 100, Oportek, Carlsbad, CA, USA), and the laser beam was coupled to the optical fiber by a plano-convex calcium fluoride lens (50-mm focal length, LA5763, Thorlabs Inc., Newton, NJ, USA). The energy of each 7-ns long laser pulse after attenuation and before coupling it to the optical fiber was 1.3±0.2 mJ. The etched end of the fiber was secured on a ceramic probe holder (MXP-150, Siskiyou Corporation, Grants Pass, OR) that was mounted on a five-axis micromanipulator (MX110, Siskiyou Corporation, Grants Pass, OR, USA).

To generate an electrospray, the Velos Pro syringe pump was used to supply the electrospray solution (50% methanol in water with 0.1% (v/v) acetic acid for analysis in positive ion mode, and 2:1 methanol:chloroform in negative ion mode analysis) at 500 nL/min to a pulled fused silica emitter (inner diameter 50 µm) that was mounted on a manual translation stage. High voltage (+3.3 kV for positive ion mode and -2.5 kV for negative ion mode) generated by a power supply inside the Velos Pro mass spectrometer, was applied to the emitter. The electrospray emitter and the mass spectrometer orifice (temperature 325 °C) were aligned on the same axis, whereas the etched optical fiber tip was in the same plane just above the sample surface, sitting at ~8 mm below the spray axis. The electrospray emitter was positioned 10 mm from the mass spectrometer orifice, whereas the fiber tip was ~2 mm ahead of the emitter tip position.

Thaw-mounted soybean nodule sections on microscope glass slides were placed on a Peltier cooling stage with the temperature set to 2.5 °C to preserve sample moisture and reduce degradation. This assembly was mounted on a motorized high-resolution XYZ translation stage (Zaber Technologies, Vancouver, BC, Canada) for positioning. The sample was located under the fiber tip which was tilted at a 45° zenith angle. In order to visualize the individual cells and monitor their ablation, a long distance video microscope (InFocus Model KC, Infinity, Boulder CO, USA) with a 5× infinity corrected objective lens (M Plan Apo 5×, Mitutoyo Co., Kanagawa, Japan) and a CCD camera were mounted at a right angle to the sample surface using a dynamically damped post (DP14A, Thorlabs Inc., Newton, NJ, USA) to reduce transmission of mechanical vibrations from the

environment. The focal plane of the microscope was fixed to keep the optical fiber tip in focus at all times. Using custom LabVIEW software, the sample was raised in the Z-direction towards the fiber tip until distinct cells became visible, but before the fiber tip came in contact with the sample. The stage was then moved in the XY plane to position the targeted cell under the optical fiber tip, and then raised further until the fiber tip made light contact with the targeted cell.

After the targeted cell was ablated, the stage was lowered until the fiber tip detached from the sample, which was then moved in the XY plane for targeting the next cell.

**Conventional LAESI.** For the analysis of large cell populations by conventional LAESI, mid-IR laser beam was steered using gold-protected mirrors, and was focused on the surface of ~ 100  $\mu\text{m}$ -thick cryosections using a plano-convex calcium fluoride lens (50 mm focal length, LA5763, Thorlabs Inc., Newton, NJ).

**f-LAESI-21T FTICR-MS.** The 21T FTICR mass spectrometer, and the technical parameters for the combination of a conventional LAESI source to the mass spectrometer, including external mass calibration and signal processing, were described in previous publications.<sup>5, 11</sup> Briefly, to synchronize the ablation with the collection of ions by the mass spectrometer, a TTL pulse was sent from the Velos Pro to trigger the laser flash lamp of the mid-IR laser. After 675 ns delay, another pulse was sent by the delay generator that opened the Q-switch, thus delivering a laser shot to the cell through the optical fiber tip. Ion accumulation time in the Velos Pro was set to 50 ms in order to fill the trap with ions resulting from a single laser shot.

**Direct metabolite assignments by ultrahigh mass accuracy and IFS.** Separate mass spectra for consecutive intervals of 50  $m/z$  units were collected between  $m/z$  180 and 1500. The mass resolution of the instrument was gradually increased as higher  $m/z$  values were recorded. The resulting mass spectra from all recorded  $m/z$  bands ( $180 \leq m/z \leq 1500$ ) were 'stitched' together to produce a comprehensive mass spectrum. Using the software

Formularity (<https://omics.pnl.gov/software/formularity>), internal calibration of the mass spectrum was performed based on previously assigned molecular formulas from soybean root nodules.<sup>23</sup> A mean post-calibration mass measurement accuracy of  $145 \pm 60$  ppb was achieved (see Table S1). To find potential matches for the detected peaks in the mass spectrum, an in-house database was built by combining all molecular formulas for metabolites from the PlantCyc (<https://www.plantcyc.org/data/search>) and Human Metabome Database (<http://www.hmdb.ca/>) using the software Eci-pex (<https://cran.r-project.org/web/packages/ecipex/index.html>).<sup>24</sup> Molecular formulas for common electrospray contaminants were added to the database. Theoretical IFSs were generated using the software Mercury for assigning the measured IFSs for metabolites.<sup>25-26</sup>

The experimental peak list was then compared with the database to search for  $m/z$  peaks that matched with endogenous compounds in *G. max* and *B. japonicum* (mass tolerance  $\leq 1$  ppm), and for IFS matching. The following Formularity search parameters were adopted: elemental composition ( $N \leq 10$ ,  $S \leq 5$  and  $P \leq 5$ ), for major peaks  $S/N \geq 3$ , for minor peaks minimum  $S/N \geq 1.5$ , minimum peak presence/absence score = 1, and minimum peak probability to score was set to 0.001. Formula score was

based on relative abundance matching between theoretical and experimental spectra, and the lowest mass error. A manual inspection of all the mass matches was performed for verification and electrospray background peaks were eliminated. Theoretical IFSs in all figures were generated using Thermo Xcalibur 2.2 (ThermoFisher Scientific, San Jose, CA, USA).

For unknown metabolites that returned no hits in the database search (see Table S3), the monoisotopic masses were used to generate possible molecular formulas using ChemCalc ([http://www.chemcalc.org/mf\\_finder](http://www.chemcalc.org/mf_finder)) with  $< 1$  ppm mass accuracy and the following possible elemental compositions:  $C \leq 100$ ,  $H \leq 100$ ,  $N \leq 10$ ,  $O \leq 20$ ,  $S \leq 5$ ,  $P \leq 5$ ,  $Cl \leq 3$ . Simulated IFS for the generated chemical formulas were produced using Thermo Xcalibur 2.2 (ThermoFisher Scientific, San Jose, CA, USA), and the matching between the recorded and simulated IFS, which took into account the mass accuracy and the relative isotopic abundance, was performed manually.

**Scanning electron microscopy (SEM) of root nodules.** Harvested nodules (21 days old) were put in a fixative solution comprising of 2.5% glutaraldehyde, 0.01% triton X-100 in 0.1 M sodium cacodylate buffer for 2 days. The nodules were rinsed 3 times with 0.1 M sodium cacodylate buffer solution for 15 minutes, then kept in 1% osmium tetroxide solution overnight. After rinsing with water, the nodules were kept for 1 hour in 10% ethanol containing 3% uranyl acetate. The nodules were then rinsed with water and dehydrated by a series of 15-min ethanol rinses as follows: 15%, 30%, 50%, 70%, 80%, and 90% ethanol, and finally 3 $\times$ 15-min rinses with 100% ethanol. After critical-point drying (Tousimis Autosamdri 931.GL, Rockville, MD, USA), the samples were mounted on SEM stubs using adhesive carbon tabs. The outer surface of the nodules was removed by applying adhesive tape onto the sample surface and pulling, thus exposing cells within the infection zone, and peeling off their cell walls and membranes. The surface of the samples was then sputter-coated using a high-resolution sputter coater (Cressington 208 HR, Cressington Scientific Instruments Limited, Watford, UK) with a 7 nm gold layer by placing the samples at a 0° angle relative to the sputtering cathode then with another 7 nm of gold at ~45°. The samples were imaged using a FEI Teneo scanning electron microscope (FEI Teneo LV SEM, ThermoFisher Scientific, Hillsboro, OR, USA).

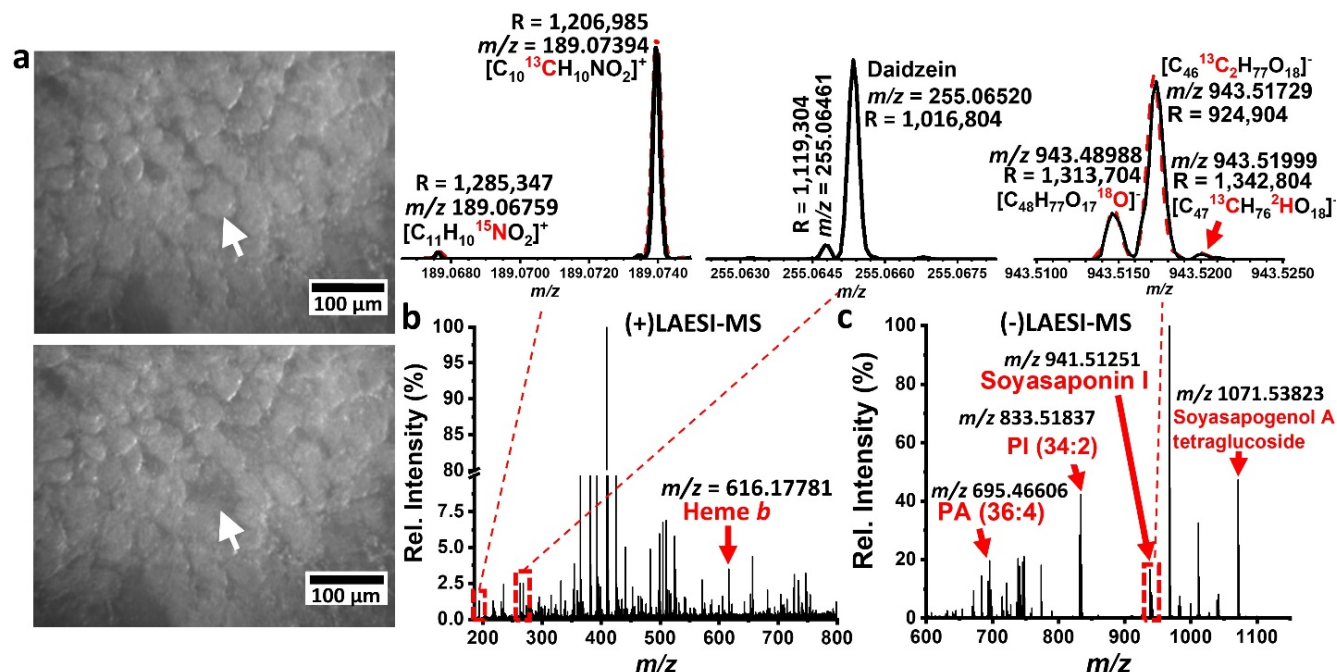
**Metabolite set enrichment analysis.** In order to determine metabolites that were up- and downregulated in the infected cells compared to the uninfected cell clusters, peak lists and intensities for the two groups were uploaded to MetaboAnalyst 4.0 (<https://www.metaboanalyst.ca/>) and multivariate statistical analysis was performed. Metabolites with fold changes  $FC \geq 2$  and  $FC \leq 0.5$ , and  $p$ -values (determined by  $t$ -test) of  $p \leq 0.05$ , were selected, and metabolite set enrichment analysis (MSEA) was performed. The regulated metabolite names were searched against an in-house library comprising of 138 metabolic pathways specific to *G. max* and *B. japonicum* that were compiled from KEGG (<https://www.genome.jp/kegg/kegg2.html>) and PlantCyc (<https://www.plantcyc.org/>). The  $p$ -values reported in the MSEA were determined by phenotype-based permutation tests adopted from gene set enrichment analysis.<sup>27</sup> The cutoff for  $p$ -values in MSEA was  $\leq 0.05$ .

## RESULTS AND DISCUSSION

In order to achieve the spatial resolution required for targeting individual infected cells, a mid-IR laser beam ( $\lambda = 2940$  nm) was coupled to an optical fiber that was etched to a tip with  $\sim 10$   $\mu\text{m}$  radius of curvature (see Figure S1 and the Experimental section in the Supporting Information (SI)). Due to the strong absorption of energy by the water content of the sample at this wavelength, a single laser shot (with a pulse energy of  $1.3 \pm 0.2$  mJ) discharged into the cell through the optical fiber tip resulted in ablation creating a plume of cellular material. The ablation plume, consisting of mostly neutrals, was ionized by electrospray and introduced into the 21T-FTICR-MS for mass analysis (see Figure S1 in the SI). An advantage of the f-LAESI-21T-FTICR-MS approach is the ability to analyze individual cells in their native tissue environment. Perturbation of the molecular composition of a cell can be significant during

tissue dissociation often used for the analysis of non-circulating cells. Compared to proteins and nucleic acids, this can be especially problematic for metabolites with short half-lives (high turnover rates). For cells analyzed by the f-LAESI system this problem is mitigated as they are still alive right before the analysis.

Optimal signal was recorded at an ion accumulation time of 50 ms, and the resulting mass spectrum contained  $\sim 180$  spectral features for  $180 \leq m/z \leq 800$  (Figure 1). Many peaks that corresponded to close-to-isobaric compounds, e.g., daidzein at  $m/z$  255.06520 and an unidentified species at  $m/z$  255.06461 separated by only 0.00059 Da, i.e. by little over the mass of an electron, were resolved (see Figure 1).



**Figure 1.** Analysis of a single soybean root nodule cell infected with *B. japonicum* by f-LAESI with 21T-FTICR-MS. (a) Ablation of an infected cell in a 30  $\mu\text{m}$ -thick root nodule section. Arrows show the targeted infected cell in the top image and the resulting ablation spot in the bottom image. Panels (b) and (c) show resulting mass spectra in positive and negative ion modes, respectively. Mass spectrum in (b) was acquired at a transient length of 3.072 s, whereas in (c) it was acquired at 6.144 s. The top left inset shows the captured IFS (black) for N-acetylglutamic acid overlaid on top of its simulated mass spectrum (red dashed line). The middle inset shows the resolving of two peaks just 0.59 mDa apart that correspond to two different metabolites. The top right inset shows the IFS of dehydrosoyasaponin I (black) captured at higher transient length (6.144 s). The peaks drawn with a red dashed line correspond to the simulated mass spectrum of the same ion.

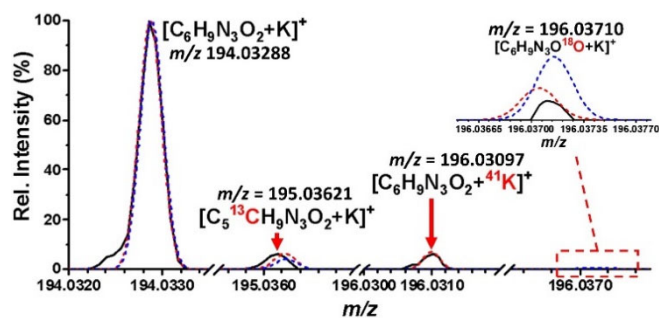
To capture IFSs for metabolites and lipids from single cells, the transient length of the ICR cell was set to 3.072 s (see Figure 1b) or 6.144 s (see Figure 1c), depending on the recorded  $m/z$  range. This resulted in mass resolutions that ranged between  $R = 500,000$  and  $2,000,000$ . The IFSs aided in identifying close-to-isobaric compounds with different elemental compositions. For example, the quasi-molecular ion,  $[\text{M}+\text{H}-\text{H}_2\text{O}]^+$ , of phosphocreatine, and the potassium adduct of histidine have exact masses that are only 0.11 mDa apart. Comparing the calculated IFS patterns for the two ions and their  $\text{M}+1$  peaks with the experimental data indicated a better match for histidine at  $m/z$

194.03288 and its  $^{41}\text{K}$  and  $^{18}\text{O}$  isotopologues than the dehydrated phosphocreatine (see Figure 2).

A critical factor in capturing IFS is the dynamic range of the FTICR mass analyzer. As high ion numbers in an ICR cell can lead to ion cloud coalescence, the high intensity end of the dynamic range can be truncated.<sup>28</sup> Thus, for the determination of elemental formulas based on high quality IFS, we recorded separate mass spectra from individual cells at 26 bands spanning  $m/z$  50 units in the  $m/z$  180 to 1,500 range. The major advantage of recording in  $m/z$  50 bands compared to a wide band is illustrated in Figure S2. The narrow band recording captured the IFS for the  $\text{M}+2$  ion of a glucoside in agreement with the calculated



profile, a clear difference from the wide band recording. For peaks that fell within  $600 \leq m/z \leq 1,500$ , the transient length was increased to 6.144 s. Initially, the molecular formulas of more than 120 metabolites and lipids were tentatively assigned based on accurate mass (see Table S1). For 47 of these compounds, the elemental compositions were determined based on their single-cell IFS using the software Formularity,<sup>29</sup> followed by manual verification of the mass differences (see Table S2). All of these compounds were also found in database (DB) searches. However, for another 11 metabolites, the elemental compositions were determined based on ultrahigh mass accuracy and IFS but no matches were found by DB searches (see Table S3). This demonstrates the utility of IFS both for the identification of known molecular species and the discovery of unknown compounds in single cells.



**Figure 2.** Peak assignment for [histidine+K]<sup>+</sup> by IFS in an expanded mass spectrum from an infected cell. Measured peaks are in black, whereas dashed peaks in red and blue correspond to simulated spectra for the [M+H-H<sub>2</sub>O]<sup>+</sup> ion of phosphocreatine and the potassium adduct of histidine, respectively. The measured M+1 peak, which corresponded to the <sup>13</sup>C isotopologue, matched closely with both simulated peaks. The measured peak at  $m/z$  196.03097 matched with the <sup>41</sup>K isotopologue of the potassium adduct of histidine. The [M+H-H<sub>2</sub>O]<sup>+</sup> ion of phosphocreatine does not have an IFS feature at this  $m/z$  value. The inset shows the closer match between the relative abundances of the measured peak and the simulated peak of the <sup>18</sup>O isotopologue of histidine.

Whereas IFS for small ions could be captured by high-resolution Orbitrap mass spectrometers, IFSs for larger biomolecules, such as lipids, oligosaccharides and peptides, require higher mass resolution that currently cannot be achieved with Orbitraps. The 21T-FTICR-MS is capable of reaching mass resolution levels above 1,000,000 that can resolve IFSs for many of these biomolecules.<sup>5</sup> The IFSs were captured from single cells for 26 metabolite ions, comprising of lipids, oligosaccharides, and soyasaponins, for which the  $m/z$  values ranged from 600 to 1200 (see Table S2). Overall tentative peak assignments were established for 46% of the spectral features based on ultrahigh mass accuracy alone. However, many compounds with different elemental compositions can have close to identical accurate masses. This issue can be mitigated by acquiring IFSs, which were present for an additional 33% of the total number of peaks assigned from single cells.

The dependence on ultrahigh mass resolution, mass accuracy and high dynamic range for fast elemental composition assignment from IFS, although a steppingstone in single-cell metabolomics, falls short of providing structural information that can help discern between constitutional isomers. The ability to elu-

cidate such information is critical when structural isomers perform different biological functions. Some of the techniques that are capable of providing structural information for metabolites, and that are coupled to MS include ion mobility separation (IMS) and gas-phase ion IR spectroscopy. IMS-MS enables distinguishing gas-phase isobaric and isomeric ions based on their mass, charge, size, and shape on a timescale of milliseconds.<sup>30-33</sup> Collision cross-section (CCS) values can be measured by IMS-MS and used to enhance confidence in metabolite assignment.<sup>15, 34-35</sup> Tandem MS or sequential MS (MS<sup>n</sup>), sometimes combined with IMS, as orthogonal methods can provide additional structural information through induced fragmentation of selected ions.<sup>33</sup> Gas phase ion IR spectroscopy coupled with MS exploits resonant photon energy absorption by certain functional groups in the ion, leading to photodissociation and the generation of multiple fragments that can be used to decipher the molecular structure.<sup>36-38</sup>

Comparing single cell data with bulk analysis revealed that in some cases the former resulted in enhanced metabolite identification through isotopologues, as some of the corresponding peaks were not recorded by the latter method. For example, a peak at  $m/z$  of 399.14451, identified as the monoisotopic peak of S-adenosylmethionine (SAM), was detected from the analysis of cell clusters by conventional LAESI from a 100  $\mu\text{m}$  section and the same peak was also recorded from a single cell by f-LAESI (see Figure S3). In plant metabolite databases, this particular mass is consistent with compounds of different elemental compositions. However, the M+2 peak for SAM was only revealed by single-cell analysis. The mass accuracy of the 21T-FTICR-MS enabled the identification of the M+2 peak as the <sup>34</sup>S bearing isotopologue of SAM, thus aiding in assigning the elemental composition of this compound (see Figure S3).

In some cases, the detection of a chemical species exclusively by single-cell analysis can be attributed to the effect of dilution during bulk analysis by other cells that do not contain that compound. Similar observations were made in single-neuron analysis, where two neuropeptides were detected from specific neurons that were absent in the analysis of extracts from the whole central nervous system.<sup>39</sup> Similarly, heterogeneity among infected cells, asymmetry in the spatial distribution of a number of metabolites within a root nodule,<sup>40</sup> and the close interspersed of infected and uninfected cells can collectively mask information about the biochemical profile of individual cells during large cell-population analysis.

To characterize cellular heterogeneity, initially we used descriptive statistical metrics without any assumption about the abundance distribution of the metabolites. Measurements of the metabolic noise, defined as  $\eta^2 = \sigma^2/\mu^2$ , where  $\mu$  is the mean abundance and  $\sigma$  is its standard deviation for a cell population, can provide information on the variance of metabolite levels from cell to cell. The total measured noise,  $\eta_m^2 = \eta_t^2 + \eta_b^2$ , is a combination of technical,  $\eta_t^2$ , and biological (metabolic) noise,  $\eta_b^2$ . In order to determine the technical noise, 23 replicates of 1.0  $\mu\text{L}$  aliquots from 500  $\mu\text{M}$  glucarate standard solution were ablated with the optical fiber tip under the same experimental conditions that were used for single-cell analysis. Intensities for glucarate ions at  $m/z$  209.03039, and ions at  $m/z$  191.01986 resulting from in-source fragmentation, were summed and normalized, and used to evaluate the technical noise. The normalized intensities from 23 replicates exhibited a normal distribution, with a mean normalized intensity,  $\mu_i$ , and a standard deviation,  $\sigma_i$ , of 13.36 and 2.43, respectively. This corresponded to a technical noise of  $\eta_t^2 = \sigma_i^2/\mu_i^2 = 0.033$  (see Figure S4a in the SI).

Signal variability in the single cell measurements,  $\eta_m^2$ , for  $n = 124$  infected cells was determined and compared to signal fluctuations caused by the technique,  $\eta_t^2$ . For example, the measured mean normalized intensity for glucarate was  $\mu_m = 2.41$ , with a standard deviation of  $\sigma_m = 2.77$ , which yielded a measured noise of  $\eta_m^2 = 1.31$  (see Figure S4b in the SI). Thus, the biological noise (reflecting phenotypic differences),  $\eta_b^2 = \eta_m^2 - \eta_t^2 = 1.28$ , for glucarate is close to the measured noise for this metabolite,  $\eta_m^2 = 1.31$ , with a small correction due to the technical noise,  $\eta_t^2 = 0.033$ . [To minimize cross-contamination and altering the cellular composition by ablating a neighboring cell, we avoided sampling adjacent cells. Cells sampled for analysis were 7-10 cells apart.](#)

Biological noise levels for 47 metabolites are listed in Table S4. Several primary metabolites, such as saccharides and  $\text{NAD}^+$ , and some secondary metabolites endogenous to plants, such as dihydroxydimethoxyisoflavone glucoside and furanocoumarin, exhibited relatively low metabolic noise ( $\eta_b^2 \leq 0.50$ ). Conversely, lipids, including phosphatidic acids (PAs), phosphatidylglycerols (PGs) and phosphatidylethanolamines (PEs), showed greater noise levels ( $\eta_b^2 \geq 0.50$ ).

Biological noise originates from a variety of sources and it is manifested at the transcript, protein, and metabolite levels. The potential sources of cell-to-cell abundance fluctuations depend on the copy numbers of the particular biomolecules. For example, abundance fluctuations for low copy number proteins are sensitive to intrinsic factors, e.g., the production and breakdown of mRNA. In contrast, the levels of species with high copy numbers, e.g., some proteins and most metabolites, vary due to extrinsic factors, such as variations in the local environment, fluctuations in ribosome and enzyme concentrations,<sup>41</sup> and can be affected by the tightness of the regulatory network. Measuring metabolic noise provides insight into the potential sources behind cell-to-cell variations in metabolite levels.

For any molecule, it is expected that fluctuations in the copy number are lower when the molecule is present in large numbers and become more significant as the count decreases. For example, single-cell global profiling of proteins in *Escherichia coli*, for which the protein copy numbers range between 1 and  $3 \times 10^5$  molecules/cell,<sup>4</sup> showed that the protein expression noise,  $\eta_p^2$ , depends on the mean copy number,  $\mu_p$ , and that the relative contributions of the two components that make up the biological noise, i.e., intrinsic and extrinsic noise, are also dependent on the mean.<sup>42</sup> At low protein copy numbers ( $\mu_p < 10$  molecules/cell), the intrinsic component dominates protein noise, and it decays as  $1/\mu_p$ . As  $\mu_p$  increases, the intrinsic component becomes less dominant, and it decreases until the mean reaches a certain value ( $\mu_p > 10$  molecules/cell), beyond which the extrinsic noise takes over with a value independent of  $\mu_p$ .<sup>42</sup>

For metabolites, however, the noise levels,  $\eta_m^2$ , were independent of the mean metabolite ion intensities,  $\mu_m$ , that represent metabolite abundances (see Figure S5 in the SI). This result can be attributed to the higher metabolite copy numbers, for example in *E. coli*,  $10^2 \leq n \leq 10^8$  molecules/cell, compared to transcripts at  $1 \leq n \leq 10^2$  molecules/cell, and proteins at  $1 \leq n \leq 3 \times 10^5$  molecules/cell.<sup>4</sup> Thus, metabolic noise levels are more likely to be influenced by factors other than metabolite copy numbers.

For metabolic noise measurements in soybean root nodules, the low metabolic noise for primary metabolites in infected cells indicates low signal variability from cell to cell, which can be attributed to the tighter regulation of primary metabolite levels. This control may be essential for cellular growth and

maintenance supported by many primary metabolites. On the other hand, higher metabolic noise is observed for lipids associated with membrane synthesis in bacteria and plants, such as PAs, PGs, and PEs. Lipids of these classes play an important role in root nodule development, as they are required for membrane biogenesis in the growing plant cells, bacteroids, infection threads, and symbiosomes.<sup>43</sup> As the degree of infection, i.e., the percentage of plant cell volume occupied by bacteroids, can vary between cells, lipid assembly for membrane biosynthesis varies significantly throughout the nodule. Accordingly, we suggest that this contributes to the presence of higher metabolic noise.

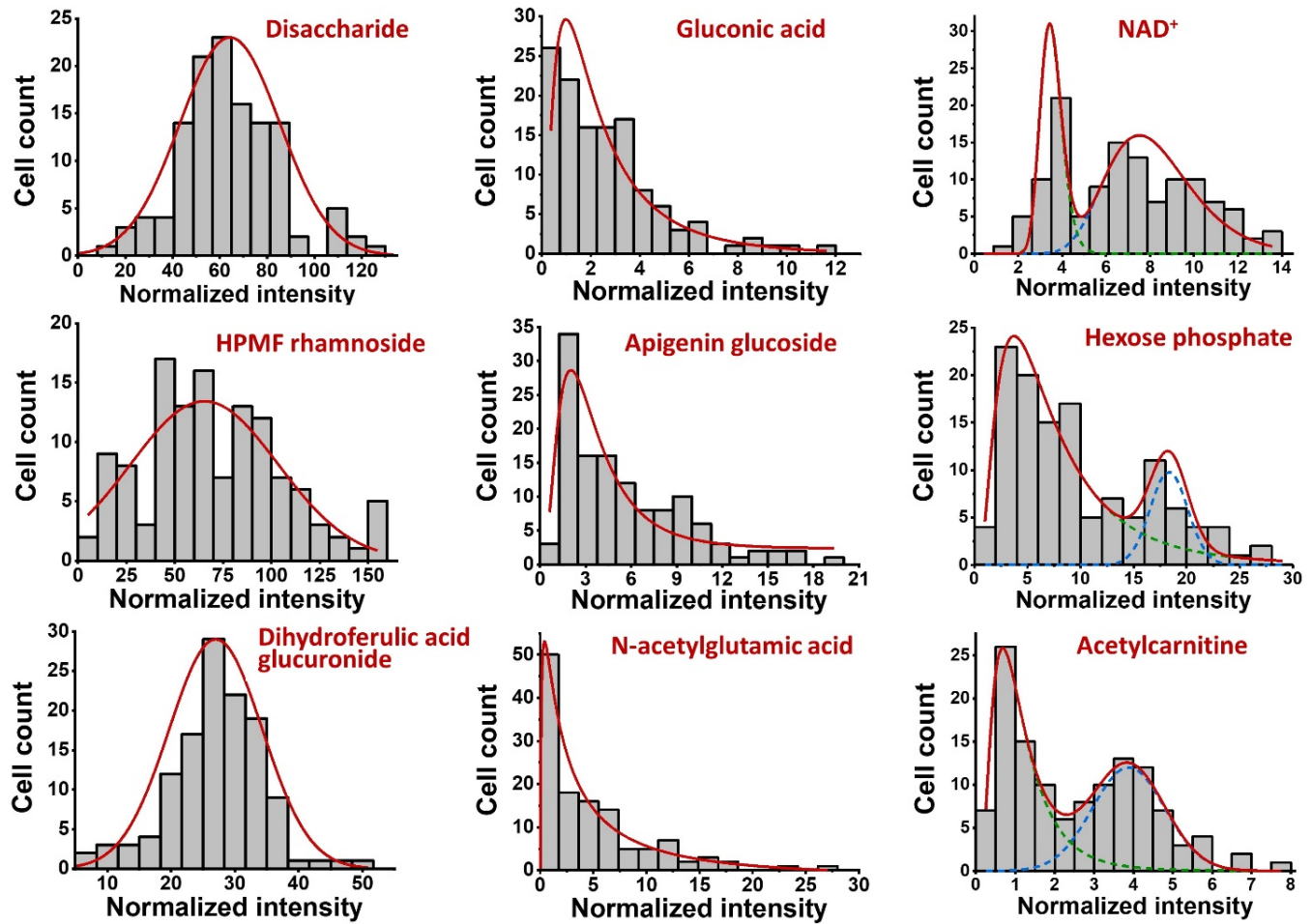
[To explore the variation in the degree of infection between cells, we used SEM as an orthogonal technique to obtain information on the population of bacteroids in infected plant cells. The SEM images revealed that, for many of these cells, the population density of the bacteroids was high, where the nitrogen-fixing microorganisms appeared to occupy close to the entire volume of the plant cell \(see Figure S6\). For other infected cells, however, the population density of the bacteroids appeared significantly lower \(Figure S6\). Previous light-microscope images of infected root nodule cells in another legume species, \*Medicago truncatula\*, showed similar observations.<sup>44</sup> These two types of infection stages, proliferating and volume limited, can be linked to different metabolic states of the bacteroids. These states can be associated with different activity levels of certain pathways and different abundances of some metabolites manifesting in their bimodal distributions.](#)

Additional insight can be gleaned by determining the distribution of metabolite abundances over the studied cell population. Within a randomly selected population of 124 infected cells, lognormal, normal and bimodal distributions of certain metabolite abundances were observed. For example, N-acetylglutamic acid and gluconic acid exhibited lognormal distributions, whereas hydroxy-pentamethoxyflavanone rhamnoside and disaccharide distributed normally (see Figure 3).

Abundances for other metabolites, such as acetylcarnitine, hexose phosphate and  $\text{NAD}^+$ , displayed bimodal distributions (see Figure 3) that were deconvoluted into two components. For  $\text{NAD}^+$  abundance distribution, each of the individual subpopulations exhibited a lognormal distribution. Of the two subpopulations that composed hexose phosphate abundance distributions, the larger subpopulation characterized by a larger area under the component green dashed curve in Figure 3 exhibited a lognormal distribution, whereas the smaller subpopulation (marked by the blue dashed curve) exhibited a normal distribution. The bimodal distribution for acetylcarnitine abundances was adequately represented by a linear combination of either one lognormal and one normal distribution, or two lognormal distributions. There are several approaches to establish the existence of two distinct components in a bimodal distribution. For example, for two underlying normal distributions, Ashman's  $D > 2$  criterion can be applied.<sup>45</sup> We also applied this criterion for the combination of a lognormal and a normal distribution to establish the presence of distinct subpopulations for certain metabolites.

Chemical analysis of single cells can be the basis for *post hoc* detection of additional subpopulations within a prespecified classification. For example, the prespecified groups of uninfected and infected cells can be further differentiated *post hoc* based on the presence of bimodal abundance distributions for some chemical species in the infected group.<sup>46</sup> For the distribution of  $\text{NAD}^+$  levels in infected cells, the mean values for normalized intensities for the two hypothetical subpopulations,  $\mu_1$

and  $\mu_2$ , were 3.53 and 8.25, respectively, and the standard deviations,  $\sigma_1$  and  $\sigma_2$ , were 0.49 and 2.09, respectively. For these



**Figure 3.** Distributions of normalized ion intensities for nine metabolites in a population of 124 infected cells. The first, second, and third columns show ion abundances with normal, lognormal, and bimodal distributions, respectively. For the latter, green and blue dashed curves represent the deconvoluted distributions for each subpopulation. The following mean values and standard deviations were determined for the deconvoluted distributions in the third column: for  $\text{NAD}^+$   $\mu_1 = 3.53$ ,  $\mu_2 = 8.25$ ,  $\sigma_1 = 0.49$ , and  $\sigma_2 = 2.09$ , for hexose phosphate  $\mu_1 = 8.18$ ,  $\mu_2 = 18.33$ ,  $\sigma_1 = 6.75$  and  $\sigma_2 = 1.76$ , and for acetylcarnitine  $\mu_1 = 1.21$ ,  $\mu_2 = 3.88$ ,  $\sigma_1 = 0.83$ , and  $\sigma_2 = 0.96$ . Abbreviation: HPMF, hydroxy-pentamethoxyflavanone.

deconvoluted distributions,  $D = 3.11$ , which satisfied the  $D > 2$  condition, and supported the *post hoc* grouping of the infected cells into two subpopulations with relatively low and high abundances of  $\text{NAD}^+$ . Similarly,  $D = 2.97$  and  $D = 2.06$  was found for the bimodal distributions of acetylcarnitine and hexose phosphate abundances, respectively, which indicated the potential of two distinct subpopulations based on these metabolites.

Bimodal metabolite abundance distributions can identify two subsets of cells with different mean abundance values caused by intrinsic or extrinsic factors. For example, a subpopulation of cells that arises from intrinsic changes in gene expression can be identified by distinct levels for certain metabolites, whereas cells that are located in a different microenvironment from the rest of the cells in the tissue can exhibit separate levels for certain metabolites due to their exposure to different extrinsic stimuli. Throughout our analysis, the spatial location for each cell was recorded in order to determine any potential correlation between metabolite abundances and cell locations in a root nodule. Such a correlation was not observed.

A potential explanation behind the bimodal distribution observed for  $\text{NAD}^+$  levels is provided by the phenotypic divergence between two subpopulations of infected cells. At 21 days post-inoculation, the root nodule is considered mature. However, [as our SEM images have revealed](#), the individual infected plant cells can exist at different infection stages.<sup>47</sup> These can be defined by the plant cellular volume, bound by the cell wall, fully or partially occupied by the bacteroids. At the point when bacteroids fully colonize a cell, they can still be metabolically active, though unable to grow and divide due to spatial restrictions, resulting in a metabolic state similar to that of quiescent cells. The subpopulation of cells that are not yet fully colonized by bacteroids could still host more of them, thus they are able to proliferate inside of these cells. In some bacteria (and in cancer), quiescent and proliferating cells express high and low  $\text{NAD}^+/\text{NADH}$  ratios, respectively.<sup>48-49</sup> The bimodal distribution of  $\text{NAD}^+$  abundances the infected cells of root nodules indicated the detection of two subpopulations with high and low  $\text{NAD}^+$  levels. These may correspond to two subpopulations of plant cells, one that hosts proliferating bacteroids, and another



that contains non-dividing, yet metabolically active quiescent bacteroids.

There is an interesting correlation between bimodality and the primary source of some metabolites that can be the plant, the bacteroids, or both. Disaccharides (e.g., sucrose) are primarily synthesized by the plant and transported to the nodule to support the bacteroids. Indeed, disaccharide levels in the root are orders of magnitude higher than in the nodule, which itself exceeds the levels in free living rhizobia by additional orders of magnitude.<sup>23</sup> The distribution of disaccharide levels, established by the plant, follows a normal distribution. In contrast, hexose phosphate levels in the nodule are orders of magnitude higher than in the root. This means that the quiescent and proliferating bacteroids have a significant contribution to the levels of this metabolite. Correspondingly, hexose phosphate levels, similar to NAD<sup>+</sup>, exhibit a bimodal distribution. When the production of a metabolite by the plant and the bacteroids is comparable, no prediction can be made about the nature of the distribution.

To explore potential correlations between regulated metabolites in the infected cells and certain biochemical pathways, we analyzed small clusters of uninfected cells (~5 cells/cluster) and compared their metabolic compositions to the infected ones. Uninfected cells interspersed the infected cells and were identified based on their significantly smaller sizes. Regulated metabolites in the infected cells, with  $FC \geq 2$  and  $FC \leq 0.5$ , and  $p$ -values of  $p \leq 0.05$ , were selected, and MSEA was performed to determine enriched metabolic pathways based on their fold enrichment and  $p$ -values.

Nine and one metabolic pathways were enriched and depleted, respectively, in the infected cells compared to the uninfected group (see Figure S7). Of the enriched pathways, zeatin biosynthesis had the highest statistical significance based on the associated  $p$ -value ( $p = 0.001$ ), and ~14-fold enrichment in the infected cells compared to the uninfected ones. On the other hand, a ~24-fold depletion of starch and sucrose metabolism ( $p = 0.003$ ) was observed in the infected cells compared to the uninfected group (see Fig. S7).

MSEA indicated that the upregulation of PAs, PGs, and PEs in the infected cells lead to ~6-fold enrichment in glycerophospholipid metabolism compared to the uninfected cells (Fig. S7). As glycerophospholipids are integral components of membranes in rhizobia and plant cells, the higher abundances of PAs, PGs, and PEs in infected cells may be related to the need for membrane synthesis in infection, e.g., for proliferating bacteroids. Additionally, enrichment of amino sugar and nucleotide sugar metabolism, arginine and proline metabolism, zeatin biosynthesis, and homogluthathione biosynthesis in the infected cells (Fig S7) indicated upregulation of metabolites associated with biological nitrogen fixation.<sup>50</sup>

## CONCLUSION

The combination of f-LAESI and 21T-FTICR-MS enabled *in situ* single-cell analysis with direct molecular formula assignments for numerous known and yet unidentified metabolites simultaneously from single cells by ultrahigh mass accuracy and IFS. This demonstrates the utility of this approach for untargeted MS-based metabolic analysis of cells. Discerning elemental formulas for unknown metabolites from single cells is especially relevant for molecules only produced by rare cells, where bulk analysis is not feasible.

Single-cell measurements enabled *post hoc* identification of latent subpopulations based on the bimodal abundance distributions for some chemical species, suggesting the presence of cell subpopulations in proliferating and quiescent phases. This observation, that is only accessible through analysis at the cellular level, describes an unexpected stratification within a complex multicellular symbiotic organ. In addition to revealing cellular heterogeneity among the infected cells, metabolic noise measurements provided possible correlations between the degree of infection, the tightness of regulation for selected metabolites, and their metabolic noise levels.

Although our work demonstrates the improvement in the quality of the data generated from single-cell metabolomics and the throughput of data interpretation through exploitation of superior instrumental performance, it should be noted that the throughput of data generation in single-cell metabolomics is still low compared to single-cell transcriptomics (~100,000 cells/study) and proteomics (~1,000 cells/study). A major factor that limits the current throughput in single-cell metabolomics that potentially can be resolved by f-LAESI-MS is the low efficiency of cell targeting. The duration of this process can be shortened by incorporating image processing for the determination of cell centroids and implementing a motorized translation stage for automated ablation of single cells with the optical fiber tip. Enabling high throughput single-cell metabolomics can aid the development of multi-omics for individual cells and help reveal the link between genotype and phenotype for rare cells.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

f-LAESI-21T FTICR-MS design, bandwidth effect on IFS, comparison of conventional and f-LAESI in peak assignment, technical and measured glucarate noise, metabolic noise levels vs intensities, SEM, MSEA, peak assignment based on ultrahigh mass accuracy, elemental composition based on IFS, IFS for unknown compounds, biological noise levels

Additional information as noted in the text (PDF)

## AUTHOR INFORMATION

### Corresponding Author

Akos Vertes – Department of Chemistry, George Washington University, Washington, DC 20052; ORCID 0000-0001-5186-5352; E-mail: [vertes@gwu.edu](mailto:vertes@gwu.edu). Phone: +1 (202) 994-2717. Fax: +1 (202) 994-5873

### Author Contributions

A.V., C.R.A., and G.S. conceived the research. L.Z.S., R.K., T.T., and D.V. performed the experiments. [P.P.](#), and [C.A.B.](#) guided the collection of SEM images. L.Z.S., S.A.S., R.K., and A.V. participated in the design and development of the f-LAESI ion source. J.B.S. assisted with retrofitting the f-LAESI source to the 21T-FTICR-MS and provided maintenance for the MS. N.T. contributed to the programming for the Formularity software, and compiled the databases of accurate masses and predicted IFSs. L.Z.S., T.T., and B.J.A. grew the soybean plants and provided nodules for the analysis. L.Z.S. performed the data analysis and wrote the manuscript with input from C.R.A., L.P.-T., G.S., and A.V.

### Notes

The authors declare that they have no conflict of interest.



## ACKNOWLEDGMENT

The design parameters and component list for the Peltier stage were kindly provided by Jarod A. Fincher of GWU. We would like to thank Rosalie Chu from PNNL for her help with the periodic mass calibration of the mass spectrometer. We also want to thank Sarai Williams of PNNL for her help in preparing electrospray and cleaning solutions. **Funding:** This material was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (OBER) under award number DE-SC0013978. Additional support was provided by University of Missouri's Gus T. Ridgel Fellowship and George Washington Carver Fellowship (B.J.A.), as well as grant no. IoS-1734145 from the National Science Foundation Plant Genome Program (to G. S.). Much of this research was performed at the Environmental Molecular Sciences Laboratory, a DOE Office of Science User Facility sponsored by the OBER and located at PNNL. PNNL is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830.

## REFERENCES

- (1) Altschuler, S. J.; Wu, L. F., Cellular Heterogeneity: Do Differences Make a Difference?, *Cell*, **2010**, *141*, 559-563.
- (2) Zenobi, R., Single-Cell Metabolomics: Analytical and Biological Perspectives, *Science*, **2013**, *342*, 1201-1211.
- (3) Svatos, A., Single-cell metabolomics comes of age: new developments in mass spectrometry profiling and imaging, *Anal. Chem.*, **2011**, *83*, 5037-5044.
- (4) Zhang, L. W.; Vertes, A., Single-Cell Mass Spectrometry Approaches to Explore Cellular Heterogeneity, *Angew. Chem. Int. Ed.*, **2018**, *57*, 4466-4477.
- (5) Shaw, J. B.; Lin, T. Y.; Leach, F. E.; Tolmachev, A. V.; Tolic, N.; Robinson, E. W.; Koppenaal, D. W.; Pasa-Tolic, L., 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer Greatly Expands Mass Spectrometry Toolbox, *J. Am. Soc. Mass Spectrom.*, **2016**, *27*, 1929-1936.
- (6) Marshall, A. G.; Rodgers, R. P., Petroleomics: Chemistry of the underworld, *Proc. Natl. Acad. Sci. U. S. A.*, **2008**, *105*, 18090-18095.
- (7) Tfaily, M. M.; Chu, R. K.; Tolic, N.; Roscioli, K. M.; Anderton, C. R.; Pasa-Tolic, L.; Robinson, E. W.; Hess, N. J., Advanced Solvent Based Methods for Molecular Characterization of Soil Organic Matter by High-Resolution Mass Spectrometry, *Anal. Chem.*, **2015**, *87*, 5206-5215.
- (8) Choi, J. H.; Kim, Y. G.; Lee, Y. K.; Pack, S. P.; Jung, J. Y.; Jang, K. S., Chemical characterization of dissolved organic matter in moist acidic tussock tundra soil using ultra-high resolution 15T FT-ICR mass spectrometry, *Biotechnology and Bioprocess Engineering*, **2017**, *22*, 637-646.
- (9) Krajewski, L. C.; Rodgers, R. P.; Marshall, A. G., 126 264 Assigned Chemical Formulas from an Atmospheric Pressure Photoionization 9.4 T Fourier Transform Positive Ion Cyclotron Resonance Mass Spectrum, *Anal. Chem.*, **2017**, *89*, 11318-11324.
- (10) Anderson, L. C.; DeHart, C. J.; Kaiser, N. K.; Fellers, R. T.; Smith, D. F.; Greer, J. B.; LeDuc, R. D.; Blakney, G. T.; Thomas, P. M.; Kelleher, N. L.; Hendrickson, C. L., Identification and Characterization of Human Proteoforms by Top-Down LC-21 Tesla FT-ICR Mass Spectrometry, *J. Proteome Res.*, **2017**, *16*, 1087-1096.
- (11) Stopka, S. A.; Samarah, L. Z.; Shaw, J.; Liyu, A.; Velickovic, D.; Agtuca, B.; Kukolj, C.; Koppenaal, D.; Stacey, G.; Pasa-Tolic, L.; Anderton, C.; Vertes, A., Ambient Metabolic Profiling and Imaging of Biological Samples with Ultrahigh Molecular Resolution Using Laser Ablation Electrospray Ionization 21 Tesla FTICR Mass Spectrometry, *Anal. Chem.*, **2019**, *91*, 5028-5035.
- (12) Ziegenhain, C.; Vieth, B.; Parekh, S.; Reinius, B.; Guillaumet-Adkins, A.; Smets, M.; Leonhardt, H.; Heyn, H.; Hellmann, I.; Enard, W., Comparative Analysis of Single-Cell RNA Sequencing Methods, *Mol. Cell*, **2017**, *65*, 631-643.
- (13) Lombard-Banek, C.; Moody, S. A.; Manzin, M. C.; 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*, 4797-4805.
- (14) Zhu, Y.; Clair, G.; Chrisler, W. B.; Shen, Y. F.; Zhao, R.; Shukla, A. K.; Moore, R. J.; Misra, R. S.; Pryhuber, G. S.; Smith, R. D.; Ansong, C.; Kelly, R. T., Proteomic Analysis of Single Mammalian Cells Enabled by Microfluidic Nanodroplet Sample Preparation and Ultrasensitive NanoLC-MS, *Angew. Chem. Int. Ed.*, **2018**, *57*, 12370-12374.
- (15) Zhang, L. W.; Foreman, D. P.; Grant, P. A.; Shrestha, B.; Moody, S. A.; Villiers, F.; Kwake, J. M.; Vertes, A., In Situ metabolic analysis of single plant cells by capillary microsampling and

- electrospray ionization mass spectrometry with ion mobility separation, *Analyst*, **2014**, *139*, 5079-5085.
- (16) Nagrath, S.; Sequist, L. V.; Maheswaran, S.; Bell, D. W.; Irimia, D.; Ulkus, L.; Smith, M. R.; Kwak, E. L.; Digumarthy, S.; Muzikansky, A.; Ryan, P.; Balis, U. J.; Tompkins, R. G.; Haber, D. A.; Toner, M., Isolation of rare circulating tumour cells in cancer patients by microchip technology, *Nature*, **2007**, *450*, 1235-1239.
- (17) Chen, S. K.; El-Heliebi, A.; Tauber, G.; Langsenlehner, T.; Potscher, M.; Kashofer, K.; Czyz, Z. T.; Polzer, B.; Riethdorf, S.; Kuske, A.; Leitinger, G.; Pantel, K.; Kroneis, T.; Sedlmayr, P., Catch and Release: rare cell analysis from a functionalised medical wire, *Sci. Rep.*, **2017**, *7*, 43424.
- (18) Cohn, J.; Day, R. B.; Stacey, G., Legume nodule organogenesis, *Trends Plant Sci.*, **1998**, *3*, 105-110.
- (19) Shrestha, B.; Vertes, A., In Situ Metabolic Profiling of Single Cells by Laser Ablation Electrospray Ionization Mass Spectrometry, *Anal. Chem.*, **2009**, *81*, 8265-8271.
- (20) Shrestha, B.; Patt, J. M.; Vertes, A., In Situ Cell-by-Cell Imaging and Analysis of Small Cell Populations by Mass Spectrometry, *Anal. Chem.*, **2011**, *83*, 2947-2955.
- (21) Li, H.; Smith, B. K.; Shrestha, B.; Mark, L.; Vertes, A., in *Mass Spectrometry Imaging of Small Molecules*, ed. L. He. 2015, vol. 1203, pp 117-127.
- (22) Cole, M. A.; Elkan, G. H., Transmissible Resistance To Penicillin-G, Neomycin, And Chloramphenicol In *Rhizobium-Japonicum*, *Antimicrob. Agents Chemother.*, **1973**, *4*, 248-253.
- (23) Stopka, S. A.; Agtuca, B. J.; Koppenaar, D. W.; Pasa-Tolic, L.; Stacey, G.; Vertes, A.; Anderton, C. R., Laser-ablation electrospray ionization mass spectrometry with ion mobility separation reveals metabolites in the symbiotic interactions of soybean roots and rhizobia, *Plant J.*, **2017**, *91*, 340-354.
- (24) Ipsen, A., Efficient Calculation of Exact Fine Structure Isotope Patterns via the Multidimensional Fourier Transform, *Anal. Chem.*, **2014**, *86*, 5316-5322.
- (25) Rockwood, A. L.; Vanorden, S. L.; Smith, R. D., Rapid calculation of isotope distributions, *Anal. Chem.*, **1995**, *67*, 2699-2704.
- (26) Rockwood, A. L.; VanOrden, S. L.; Smith, R. D., Ultrahigh resolution isotope distribution calculations, *Rapid Commun. Mass Spectrom.*, **1996**, *10*, 54-59.
- (27) Subramanian, A.; Tamayo, P.; Mootha, V. K.; Mukherjee, S.; Ebert, B. L.; Gillette, M. A.; Paulovich, A.; Pomeroy, S. L.; Golub, T. R.; Lander, E. S.; Mesirov, J. P., Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles, *Proc. Natl. Acad. Sci. U. S. A.*, **2005**, *102*, 15545-15550.
- (28) Vladimirov, G.; Hendrickson, C. L.; Blakney, G. T.; Marshall, A. G.; Heeren, R. M. A.; Nikolaev, E. N., Fourier Transform Ion Cyclotron Resonance Mass Resolution and Dynamic Range Limits Calculated by Computer Modeling of Ion Cloud Motion, *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 375-384.
- (29) Tolic, N.; Liu, Y.; Liyu, A.; Shen, Y. F.; Tfaily, M. M.; Kujawinski, E. B.; Longnecker, K.; Kuo, L. J.; Robinson, E. W.; Pasa-Tolic, L.; Hess, N. J., Formularity: Software for Automated Formula Assignment of Natural and Other Organic Matter from Ultrahigh-Resolution Mass Spectra, *Anal. Chem.*, **2017**, *89*, 12659-12665.
- (30) Fenn, L. S.; McLean, J. A., Structural resolution of carbohydrate positional and structural isomers based on gas-phase ion mobility-mass spectrometry, *PCCP*, **2011**, *13*, 2196-2205.
- (31) Hamid, A. M.; Prabhakaran, A.; Garimella, S. V. B.; Ibrahim, Y. M.; Smith, R. D., Characterization of applied fields for ion mobility separations in traveling wave based structures for lossless ion manipulations (SLIM), *Int. J. Mass spectrom.*, **2018**, *430*, 8-13.
- (32) Lanucara, F.; Holman, S. W.; Gray, C. J.; Evers, C. E., The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics, *Nat. Chem.*, **2014**, *6*, 281-294.
- (33) Shrestha, B.; Vertes, A., High-Throughput Cell and Tissue Analysis with Enhanced Molecular Coverage by Laser Ablation Electrospray Ionization Mass Spectrometry Using Ion Mobility Separation, *Anal. Chem.*, **2014**, *86*, 4308-4315.
- (34) Paglia, G.; Angel, P.; Williams, J. P.; Richardson, K.; Olivos, H. J.; Thompson, J. W.; Menikarachchi, L.; Lai, S.; Walsh, C.; Moseley,

- A.; Plumb, R. S.; Grant, D. F.; Palsson, B. O.; Langridge, J.; Geromanos, S.; Astarite, G., Ion Mobility-Derived Collision Cross Section As an Additional Measure for Lipid Fingerprinting and Identification, *Anal. Chem.*, **2015**, *87*, 1137-1144.
- (35) Paglia, G.; Williams, J. P.; Menikarachchi, L.; Thompson, J. W.; Tyldesley-Worster, R.; Halldorsson, S.; Rolfsson, O.; Moseley, A.; Grant, D.; Langridge, J.; Palsson, B. O.; Astarita, G., Ion Mobility Derived Collision Cross Sections to Support Metabolomics Applications, *Anal. Chem.*, **2014**, *86*, 3985-3993.
- (36) Nieckarz, R. J.; Oomens, J.; Berden, G.; Sagulenko, P.; Zenobi, R., Infrared multiple photon dissociation (IRMPD) spectroscopy of oxazine dyes, *PCCP*, **2013**, *15*, 5049-5056.
- (37) Polfer, N. C., Infrared multiple photon dissociation spectroscopy of trapped ions, *Chem. Soc. Rev.*, **2011**, *40*, 2211-2221.
- (38) Polfer, N. C.; Valle, J. J.; Moore, D. T.; Oomens, J.; Eyler, J. R.; Bendiak, B., Differentiation of isomers by wavelength-tunable infrared multiple-photon dissociation-mass spectrometry: Application to glucose-containing disaccharides, *Anal. Chem.*, **2006**, *78*, 670-679.
- (39) Zhang, L. W.; Khattar, N.; Kemenes, I.; Kemenes, G.; Zrinyi, Z.; Pirger, Z.; Vertes, A., Subcellular Peptide Localization in Single Identified Neurons by Capillary Microsampling Mass Spectrometry, *Sci. Rep.*, **2018**, *8*, 10.
- (40) Velickovic, D.; Agtuca, B. J.; Stopka, S. A.; Vertes, A.; Koppenaal, D. W.; Paga-Tolic, L.; Stacey, G.; Anderton, C. R., Observed metabolic asymmetry within soybean root nodules reflects unexpected complexity in rhizobacteria-legume metabolite exchange, *Isme Journal*, **2018**, *12*, 2335-2338.
- (41) Bennett, B. D.; Kimball, E. H.; Gao, M.; Osterhout, R.; Van Dien, S. J.; Rabinowitz, J. D., Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*, *Nat. Chem. Biol.*, **2009**, *5*, 593-599.
- (42) Taniguchi, Y.; Choi, P. J.; Li, G.-W.; Chen, H.; Babu, M.; Hearn, J.; Emili, A.; Xie, S., Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells, *Science*, **2010**, *329*, 533-538.
- (43) de la Pena, T. C.; Fedorova, E.; Pueyo, J. J.; Lucas, M. M., The Symbiosome: Legume and Rhizobia Co-evolution toward a Nitrogen-Fixing Organelle?, *Frontiers in Plant Science*, **2018**, *8*, 2229.
- (44) Wang, D.; Griffiths, J.; Starker, C.; Fedorova, E.; Limpens, E.; Ivanov, S.; Bisseling, T.; Long, S. R., A Nodule-Specific Protein Secretory Pathway Required for Nitrogen-Fixing Symbiosis, *Science*, **2010**, *327*, 1126-1129.
- (45) Ashman, K. M.; Bird, C. M.; Zepf, S. E., Detecting Bimodality in Astronomical Datasets, *Astronomical Journal*, **1994**, *108*, 2348-2361.
- (46) Stopka, S. A.; Khattar, R.; Agtuca, B. J.; Anderton, C. R.; Pasa-Tolic, L.; Stacey, G.; Vertes, A., Metabolic Noise and Distinct Subpopulations Observed by Single Cell LAESI Mass Spectrometry of Plant Cells in situ, *Frontiers in Plant Science*, **2018**, *9*, 1646.
- (47) Libault, M., Transcriptional Reprogramming of Legume Genomes: Perspective and Challenges Associated With Single-Cell and Single Cell-Type Approaches During Nodule Development, *Frontiers in Plant Science*, **2018**, *9*, 1600.
- (48) DeBerardinis, R. J.; Lum, J. J.; Hatzivassiliou, G.; Thompson, C. B., The biology of cancer: Metabolic reprogramming fuels cell growth and proliferation, *Cell Metab.*, **2008**, *7*, 11-20.
- (49) Ruhl, M.; Le Coq, D.; Aymerich, S.; Sauer, U., C-13-flux Analysis Reveals NADPH-balancing Transhydrogenation Cycles in Stationary Phase of Nitrogen-starving *Bacillus subtilis*, *J. Biol. Chem.*, **2012**, *287*, 27959-27970.
- (50) Resendis-Antonio, O.; Reed, J. L.; Encarnacion, S.; Collado-Vides, J.; Palsson, B. O., Metabolic reconstruction and Modeling of nitrogen fixation in *rhizobium etli*, *PLoS Comp. Biol.*, **2007**, *3*, 1887-1895.



# Table of Contents graphic

