

AutonoMS: Automated ion mobility metabolomic fingerprinting

Gabriel K. Reder*, Erik Y. Bjurström, Daniel Brunnsåker, Filip Kronström, Praphapan Lasin, Ievgeniia Tiukova, Otto I. Savolainen, James N. Dodds, Jody C. May, John P. Wikswo, John A. McLean, Ross D. King

AUTHOR ADDRESS: Data Science and AI division, Department of Computer Science and Engineering, Elektrogården 1, Göteborg 412 58, Sweden

ABSTRACT: Automation is dramatically changing the nature of laboratory life science. Robotic lab hardware able to perform manual operations with greater speed, endurance, and reproducibility opens an avenue for faster scientific discovery with less time spent on laborious repetitive tasks. A major bottleneck remains in integrating cutting-edge laboratory equipment into automated workflows, notably specialized analytical equipment which is designed for human usage. Here we present AutonoMS, a platform for automatically running, processing, and analyzing high-throughput mass spectrometry experiments. AutonoMS is currently written around an ion mobility-mass spectrometry (IM-MS) platform and can be adapted to additional analytical instruments and data processing flows. AutonoMS enables automated software agent-controlled end-to-end measurement and analysis runs from experimental specification files that can be produced by human users or upstream software processes. We demonstrate the use and abilities of AutonoMS in a high-throughput flow-injection ion mobility configuration with 5 second sample analysis time, processing robotically-prepared chemical standards and cultured yeast samples in targeted and untargeted metabolomics applications. The platform exhibited consistency, reliability, and ease of use while eliminating the need for human intervention in the process of sample injection, data processing, and analysis. The platform paves the way towards a more fully automated mass spectrometry analysis and ultimately closed-loop laboratory workflows involving automated experimentation and analysis coupled to AI-driven experimentation utilizing cutting-edge analytical instrumentation. AutonoMS documentation is available at <https://autonoms.readthedocs.io>

INTRODUCTION

Compared to traditional benchtop experimentation, modern life science laboratories are high-throughput and data-centric discovery platforms. This transformation is largely supported by two pillars (1) experimental hardware automation and (2) informatic and control software integration. Automated robotic laboratory equipment can increasingly perform labor-intensive physical experimental processes including sample preparation, maintenance, and assay execution.¹⁻³ In addition to increasing the quantity and quality of data produced, the use of automated labware also produces metadata audit trails at every step of the experimental process to increase data reusability.⁴ A shift is underway from low-throughput manual laboratory operation towards high-throughput screens generating large quantities of raw data which can only be understood through informatic analysis. This creates a new relationship between the scientist, the benchtop, and software. Experimental platforms that can be run through software calls without human supervision can generate large amounts of high-quality data at lower cost to the human scientist to greatly improve the rate of discovery, especially in screening applications in fields such as drug development and metabolic engineering in which combing through experimental space is often the rate-limiting factor. Such high-throughput screening platforms almost always rely on an analytical measurement of samples of interest. These instrumental ‘omics’ measurements provide the crucial biochemical readout of the system of interest. Despite the promises of integrated hardware-software automation for life science discovery, there remains a great need for further development of automated analytical platforms at the granular end of the omics scale, namely proteomics

and metabolomics. In these realms, analytical instrumentation often remains manually operated and therefore underutilized. As experimentation becomes increasingly automated and high-throughput, analytical instrumentation must keep pace.

Mass spectrometry (MS) is a valuable and broadly used analytical technique in the life sciences. The ability to sensitively and broadly detect the molecular components of biochemistry has made it an essential technique for biomarker discovery, drug development, bioprocess development, and basic discovery.⁵⁻⁸ A key driver of the technique’s utility has been the continuous development and refinement of MS instrumentation providing increased sensitivity, resolution, reproducibility, and throughput. However, with these benefits comes a high cost. The vast number of acquisition parameters, diverse instrumentation, and varied applications of MS make it a time-intensive technique requiring multiple iteration cycles and substantial hands-on intervention.

Ion mobility-mass spectrometry (IM-MS), an especially promising MS variant, couples a high-throughput separation dimension coupled with MS instrumentation offering analyte separation on the basis of ion structure (ion mobility) in addition to standard mass separation in complex samples.^{9,10} The drift tube implementation of ion mobility MS (DTIMS) involves usage of a linear ion mobility tube to separate ionized molecules before mass-to-charge measurement. Ions exhibit different transit times through the drift tube determined by their size, shape, charge, and instrument acquisition parameters. This measured drift time can then be converted to a collision cross section (CCS) value, which is a function the molecule’s structural

properties.¹¹ One of the primary advantages of IM-MS separations is the resolution of isomers on the basis of structure, which in some cases can serve as a replacement to slower liquid

chromatography separations.^{12,13} However, ion mobility introduces further experimental complexity to already labor-intensive MS workflows.

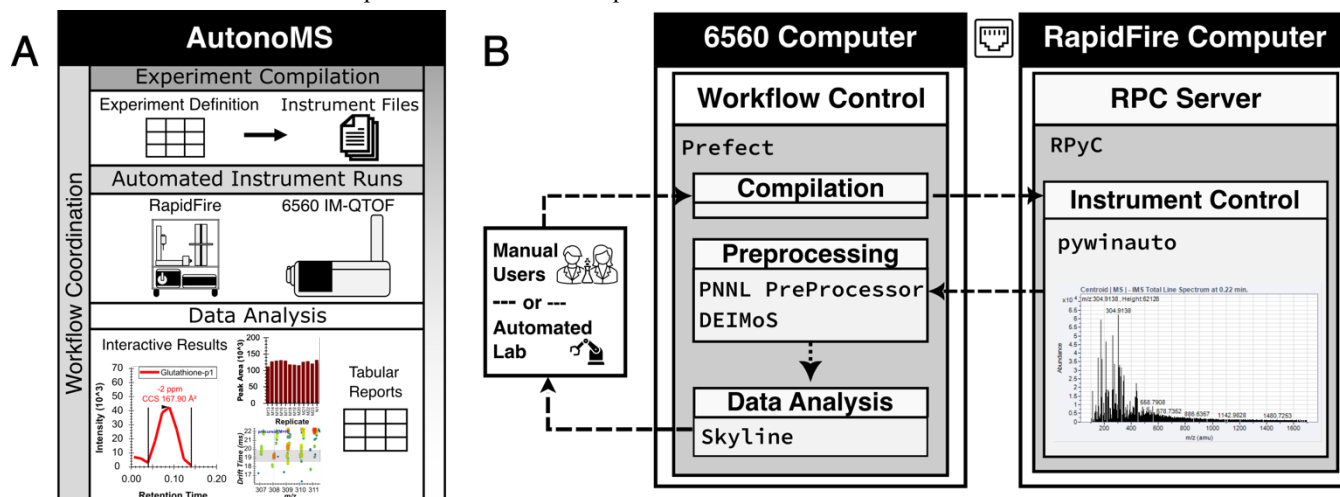


Figure 1. AutonoMS offers walkaway automation of ion mobility mass spectrometry data collection and analysis. (A) AutonoMS integrates software control layers with the Agilent RapidFire - 6560 ion mobility mass spectrometry system to provide fully automated data acquisition, raw data handling, data processing, and metabolomic end-to-end analysis resulting in tabular metabolite reports and interactive Skyline documents. (B) The AutonoMS software stack is hosted on a shared drive between the 6560 and RapidFire control computers. Human laboratory users or an upstream software agent may trigger AutonoMS runs using a tabular experiment definition file. The AutonoMS workflow control is written using Prefect¹⁴ which coordinates the event-triggered actions of modules responsible for instrument file compilation, instrument control (pywinauto¹⁵), post-acquisition raw data handling including ion mobility demultiplexing and CCS calibration (PNNL PreProcessor¹⁶ and DEIMoS¹⁷), and metabolomic data analysis (Skyline^{18,19}). Additional modules may be written and incorporated into the workflow to accommodate different instruments or analysis workflows.

Here we introduce AutonoMS, which offers end-to-end automated runs of MS instrumentation involving sample injection, raw data processing, and metabolomic analysis with little user intervention, and is currently written around the Agilent RapidFire²⁰ and 6560²¹ DTIMS-QTOF system (Figure 1A). AutonoMS coordinates instrument control, resource allocation, and data processing across the RapidFire and 6560 control computers using a collection of open-source software libraries (Figure 1B). Sample runs can be automatically triggered from experiment plan files, so that either a human user or upstream software agent may design and run experiments. This means that the AutonoMS platform can be integrated into a larger automated laboratory setting in which software agents control and coordinate multiple experimental, analytical, and informatic modules. Runs may involve multiple acquisition modes, sequences, and variable run parameters. After sample acquisition, data is automatically prepared, processed, and then analyzed via Skyline,^{18,19} producing both interactive results and tabular metabolite summaries. To demonstrate the use of the AutonoMS platform, we analyzed a set of chemical standards chosen from the yeast metabolic network. These standards, serially diluted by an Agilent Bravo liquid handling robot, exhibited the expected dynamic measurement responses as autonomously collected and detected by AutonoMS (Figure 2). We also processed extracted intracellular yeast samples through the platform, indicating its potential utility in automated untargeted and discovery applications (Figure 3 and Figure 4).

EXPERIMENTAL SECTION

Workflow Control. The AutonoMS control flow automation and workflow logic was implemented using Prefect (version 2.10.12).¹⁴ Prefect is an open-source Python-native workflow

manager for orchestrating complex code workflows from modularly-defined tasks, smoothly turning Python functions into workflow steps. We chose Prefect because of its open-source nature, active developer community (currently over 200 GitHub contributors), and increasing adoption across industrial data science teams. Its native Python implementation obviates the need to learn a separate workflow domain specific language. The workflow begins by compiling an input tabular experiment definition file into RapidFire XML files. AutonoMS then sequentially triggers an IM-MS acquisition run for each sequence in the experiment definition using the instrument control utilities described below. In our configuration, the workflow currently runs only one sequence at a time in order of appearance in the experiment file. However, the workflow parameters may be modified for laboratories with capabilities of running parallel data acquisition on multiple instruments. A sample experiment file is available in the AutonoMS GitHub repository (<https://github.com/gkreder/autonomms>).

Upon completion of data acquisition for all sequences, the workflow then automatically performs post-run data processing. The RapidFire 365 - 6560 creates a single Agilent .D raw data file that must be split into .D files corresponding to injections from the individual microplate wells. Currently for RapidFire instruments configured in BLAZE (direct injection) mode, the RapidFire software can detect injection boundaries but does not correctly split data into the corresponding well files. AutonoMS automatically shifts the file split times according to the detected injection boundaries and assigns them to the correct well plate. A description of configuration of the RapidFire for BLAZE mode is provided below.

Data multiplexing is a powerful ion mobility technique for increasing sensitivity.²² Our configuration utilizes 4-bit IM demultiplexing, but this requires post-acquisition data demultiplexing. AutoMS performs ion mobility data demultiplexing on the individual well files using the PNNL PreProcessor utility.¹⁶ Before conversion to collision cross section (CCS) values, raw measurements in DTIMS files from the Agilent 6560 must have their drift time values calibrated according to standard measurements with known CCS values. AutoMS automatically performs this CCS calibration for each injection within a given sequence using nearest prior injection occurring in the same sequence with sample type “TUNE”. The demultiplexed tune file is converted to mzML²³ format using the msconvert utility.^{24,25} CCS correction coefficients are then calculated from the mzML tune file using the DEIMoS library for ion mobility data processing¹⁷ and user-specified reference CCS values. A CCS calibration XML file is generated and copied to each injection .D file in the given sequence. For the experiments outlined in this work, the Agilent ESI tuning mix was used for CCS calibration. A sample sheet containing the Agilent tune ions and their CCS values is available in the AutoMS GitHub repository. We note that the workflow can be configured to perform this CCS calibration on manually calculated calibration coefficients from separate runs; however, it is recommended to include a tune injection in each sequence for the sake of automation simplicity and data robustness. AutoMS then performs peak detection and quantification on the demultiplexed CCS calibrated injection files using the Skyline method (described below) via the Skyline Command-Line interface.

Instrument Control. The pywinauto library (version 0.6.8)¹⁵ was used to write automation control wrappers around the Agilent MassHunter Workstation Data Acquisition (version 11.0) and RapidFire UI (version 6.1.1.2114) software used to control the 6560 mass spectrometer and RapidFire sampler. Functionalities necessary for automatically running plates according to user-supplied acquisition method and experimental parameters were implemented. These include reading instrument state, loading files and methods, starting and stopping runs, running the calibrant line, checking RapidFire vacuum pump pressure, setting run mode, and data file splitting. Of special note is the standard Agilent hardware configuration of separate control desktops for the 6560 and RapidFire connected via ethernet (Figure 1B). To work in this configuration, the codebase is hosted on the network drive shared between the two computers, and the AutoMS workflow is run from the 6560 control computer. A Remote Procedure Call server using the RPyC library (version 5.3.1)²⁶ is hosted on the RapidFire computer to execute RapidFire control functions from the 6560 computer. The codebase for RapidFire - 6560 instrument control is available in the `agilent_methods` modules of AutoMS.

Data Processing. Peak detection and area quantification were performed using Skyline (version 22.2.0.351).^{18,19} Skyline was chosen because of its open-source nature, ability to handle ion mobility data, and its combination of command-line and GUI functionality. Skyline natively supports RapidFire ion mobility data and has previously been used for acquisition workflows such as those described below.¹² Its command line utilities can be integrated into a fully automated workflow as in AutoMS and results can later be loaded into the GUI to be verified by the end user. Peak detection was run using the TOF mass analyzer settings at a resolving power of 30,000,

an ion mobility resolving power window of 30, and a maximum m/z of 1700. The full set of Skyline processing parameters in Skyline document format is available in the AutoMS repository.

Metabolite CCS Library. Yeast metabolites were taken from the Yeast Metabolome Database (YMDB)^{27,28} and compared against the Cross Collision Section Database (CCSDB) hosted by the Erin Baker Lab and available at <https://brwebportal.cos.ncsu.edu/baker/>. These CCS values were of particular interest since they were also measured on an Agilent 6560 mass spectrometer for comparison purposes. Metabolites appearing in both YMDB and CCSDB were compiled together with their experimentally observed CCS values for their $[M+H]^+$ and $[M-H]^-$ adducts. The compiled list of YMDB CCS metabolites in Skyline transition list format is available in the AutoMS repository.

RapidFire Sample Injection. The RapidFire was operated in BLAZE flow injection mode to achieve rapid sample injection and analysis times.²⁹ This involves configuring the RapidFire valve tubing such that the sample loop feeds directly into to the mass spectrometer outlet rather than through the solid phase extraction (SPE) cartridges. The RapidFire configuration files must also be modified for the valve positions to correctly correspond to sample sipping with this connection configuration. We note that the RapidFire is capable of some automated in-line sample preparation, for example desalting, via its built-in SPE functionality. This functionality was not utilized for our demonstration experiments since our sample preparation was performed prior to injection. RapidFire BLAZE mode configuration instructions are provided in the supplementary information. Operating the RapidFire using the SPE functionality via AutoMS can be done by simply reverting the instrument back to its standard configuration and changing the cartridge and sipper parameters in the input experimental file.

The RapidFire method used involved a sample sipping time of 600ms followed by 4400ms of sample elution into the MS. As seen previously¹², sipping (aspiration) time is reported rather than injection volume as this is the instrument’s controllable parameter given its mechanism of sample aspiration driven by a vacuum pump. The RapidFire’s sample loop holds roughly 30 μ L. This elution time was chosen to ensure baseline peak separation at higher sample concentrations, and we note that this can be reduced for faster cycle times. The mobile phase (pump 1) consisted of 50/50 water/ methanol with 0.1% formic acid at a flow rate of 1.25 mL/min. The full RapidFire parameter set are provided in the supplementary information.

6560 Mass Spec Data Acquisition. Mass spectrometry data was collected in IM-QTOF mode with 4-bit multiplexed introduction of the ion packets into the drift tube. Multiplexing has been shown to improve ion utilization and resolving power in IMS;¹² however, it creates the requirement for additional data post-processing as described in the workflow management section. The Agilent 6560 was operated in the 100-1700 m/z range at a frame rate of 1.1 frames/sec, and a gas temperature of 325°C. A full description of the QTOF and IM acquisition parameters are available in the supplementary information.

Standards Preparation. Dry chemical stocks of glutathione, L-histidine, L-valine, and L-arginine were mixed at room temperature with a stock solution of 50/50 water/methanol to a concentration of 10 mg/mL. These stocks were dispensed into

standard 384 well microplates using a Thermo Fisher Combi Multidrop reagent dispenser and 10-fold serial dilutions were robotically performed using an Agilent Bravo liquid handling robot controlled with the VWorks Automation Control software (version 8.0.0.335, Agilent Technologies).

Yeast Culturing. *S. cerevisiae* wildtype strain BY4741 (Accession number: Y00000) from the EUROSCARF deletant library was revived from -80 °C glycerol stocks by overnight cultivation in YPD media (10 g/L yeast extract, 20 g/L peptone from meat, 20 g/L dextrose) at 30 °C, 220 rpm. The strain was then streaked out on a YPD agar plate and incubated at 30 °C for 3 days. A YPD preculture was inoculated using multiple colonies from the agar plate and then incubated at 30 °C, 220 rpm for 14 hours. The cells from the YPD preculture were washed twice (centrifugation at 5,000 xg, 5 minutes) with YNB media (6.7 g/L YNB without amino acids and with ammonium sulfate, 1x Amino acid mix, 20 g/L dextrose). Cells were resuspended in 1 mL YNB media and used as inoculum for the main culture with an initial OD600 of 0.05. The main cultivations were performed in 4x250 mL wide-necked baffled shake flasks sealed with cotton stoppers, each with a working volume of 40 mL YNB media. The shake flasks were incubated at 30 °C, 220 rpm. Cultivations were stopped after 24 hours post-inoculation and the flasks were pooled. Final OD600 was measured to be 3.56 and was used to adjust the 2-propanol alcohol volume in the ensuing extraction method.

Yeast Quenching and Extraction. Sample preparation and intracellular metabolite extraction followed a previously established protocol.^{30,31} Samples were quickly transferred to 15 mL centrifuge tubes (5mL per tube) containing absolute methanol (99% purity) prechilled to -80 °C. The ratio between sample and methanol was kept at 1:1 v/v. Tubes were kept in dry ice during the process and were transferred to a centrifuge and spun for 5 min at 3000g and -9 °C. The supernatant was then discarded, and the pellet was transferred and stored at -80 °C. Samples were lyophilized (-40 °C, 0.1 mbar) overnight and kept at -80 °C pending extraction. Metabolite extraction was performed by adding 75% 2-propanol (preheated to boiling temperatures, with a ratio of 1 mL 2-propanol per 1 mg of sample) to the lyophilized yeast biomass in 15mL centrifuge tubes. Sample weight was estimated from optical density (0.34mgDCW/mL per 1 OD600). Samples were placed on a heating block for 1 minute at 100 °C, then shaken and vortexed for 2 minutes, followed by an additional 3 minutes on the heat block. Samples were then cooled for 15 min at 4 °C before centrifuging for 20 min at 3200 g and 4 °C. The supernatant was filtered through a 0.45 µm nylon-filter and transferred to 50 mL centrifuge tubes and stored at -20 °C until analysis. Samples were transferred to a standard 384 well microplate for AutonoMS injection using a Thermo Fisher Combi Multidrop reagent dispenser.

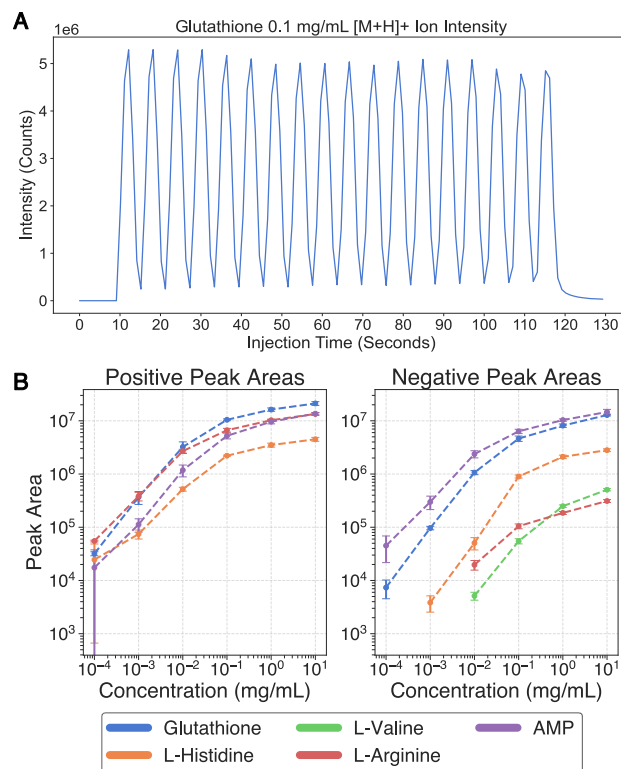


Figure 2. Automated targeted data analysis of standards with AutonoMS. (A) Detected glutathione [M+H]⁺ ion intensity from an automated AutonoMS analysis of glutathione in 50/50 methanol/water at 0.1 mg/mL injected from separate wells in a robotically dispensed 384 well microplate. (B) Detected peak areas from AutonoMS analysis of robotically-prepared triplicate serial dilutions of 5 chemical standards in positive and negative ionization modes robotically dispensed into a 384 well microplate.

RESULTS AND DISCUSSION

Walkaway Automation. Over the course of our experimentation, we found that AutonoMS enabled reliable automated runs and analysis of data from simple experiment definition files without any need for human intervention (Figure 1). The AutonoMS platform integrated smoothly into our automated laboratory workflows and provides an event-triggered software-compatible interface to a larger automated environment. Resources can be scaled automatically through the control workflow in which specific tasks are granted user-defined resource usage and concurrency rights. We found this was crucial for automating this combination of data acquisition, data preprocessing, and data analysis tasks with varying degrees of interdependencies and resource demands. Through the control workflow, the informatic steps of the platform can be run on remote or distributed resources to cut down on computation time. Instrument data acquisition for a given experiment definition runs in its entirety before initiating the informatic portions of the pipeline. As such, multiple sequences can be run from the same microwell plate while minimizing sample evaporation time. The targeted data and untargeted assays described below involved laboratory robotics and multiple data sequences and acquisition modes per experiment. We found that aside from hardware maintenance and physical sample transfer, usage of the AutonoMS platform eliminated the need for human intervention or even presence during the process of metabolomic data acquisition and analysis.

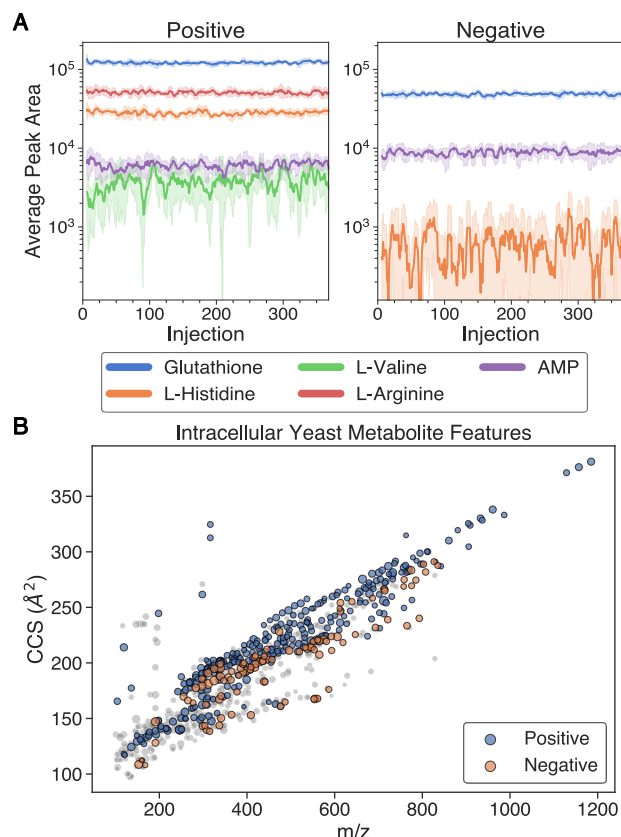


Figure 3. Automated analysis of extracted intracellular yeast samples with AutonoMS. (A) Detected peak areas in extracted yeast samples across plate injections (368 over 38 minutes per mode) of the 5 ions used in the chemical standards analysis. Ions correspond to the $[\text{M}+\text{H}]^+$ and $[\text{M}-\text{H}]^-$ adducts in positive and negative modes, respectively. Peak areas shown as the 6-injection moving average (solid lines) together with 6-injection standard deviation (shaded areas). (B) Untargeted metabolite features found across all extracted yeast samples across positive (blue) and negative (orange) ionization modes. Displayed features were present in at least 2/3 of samples in a given mode and had Agilent quality scores greater than 70. 812 total features were found of which 404 involved multiple ions in various ionization states. Single ionization state ($z = 1$) ion features are shown in grey, marker size is scaled according to $\log_{10}(\text{abundance})$.

Targeted Data Analysis of Standards. We tested the use of AutonoMS in a targeted data analysis metabolomics application by automated sampling and analysis of a chemical standard, glutathione, at a set concentration of 10^{-1} mg/mL in a stock solution of 50/50 methanol/water (Figure 2A). AutonoMS, running the direct injection ion mobility method described in the experimental section, enabled the automated analysis of the known standard with an injection time of five seconds per sample with consistent performance. The platform autonomously injected, processed, and detected the $[\text{M}+\text{H}]^+$ adduct in a positive mode run without human intervention. We also tested the AutonoMS platform on robotically-prepared triplicate 10-fold serial dilutions of five standards: glutathione, L-histidine, L-valine, L-arginine, and adenosine monophosphate, from a starting concentration of 10 mg/mL (Figure 2B). Peaks areas were filtered to include only detections at levels 5x higher than those in 50/50 water/methanol blanks. The platform similarly autonomously collected data and reproducibly detected the standards with a linear

peak area response range in the 10^{-4} - 10^{-1} mg/mL range. In both cases, no human intervention was required other than cleaning of the instruments and transfer of the prepared sample microplates to the RapidFire. These panels lead us to conclude that (1) the AutonoMS platform automates existing targeted data analytical workflows and (2) using the automated workflow produces robust and consistent results useful to downstream human users or software processes.

Metabolomic Fingerprinting of Yeast. To investigate the utility of the AutonoMS platform in systems biology discovery applications, we cultured and prepared yeast samples to study their intercellular metabolomic content via untargeted metabolomic fingerprinting (Figure 3 and Figure 4). Extracted intracellular yeast samples were pooled and dispensed into a standard 384 well microplate after which AutonoMS was used to automatically run whole-plate sequences of injections in both positive and negative ionization modes. 368 consecutive injections were run in each mode (leaving the first plate column reserved for tune ions) for a total injection time of 38 minutes per mode. Ion behavior in this complex sample matrix largely agreed with exhibited behavior in the targeted data analysis standards test (Figure 3A). Ions with higher intensities had relatively consistent peak areas across the plate injections, with peak area consistency decaying dramatically around the 10^4 mark due to poor counting statistics.

The pre-processed data produced by AutonoMS was also run through untargeted feature finding with Mass Profiler (version 10.0.195, Agilent Technologies), yielding 812 metabolite features across positive and negative mode with (1) a Q-Score greater than 70 and (2) occurrence in at least 2/3 of a given ionization mode's injections (Figure 3B). Of these features, 404 were higher-quality features involving multiple ions in various ionization states. This Mass Profiler analysis was performed manually but could be automated and integrated into AutonoMS using DEIMoS¹⁷ if it suits the end-user's needs. AutonoMS automatically ran this yeast injection data through the Skyline pipeline using the YMDB CCS library (described in the experimental section). The reports generated from the AutonoMS runs facilitated compilation of a panel of 35 detectable YMDB intracellular metabolites with publicly available CCS values filtered according to the earlier tests (displaying an average peak area greater than 10^4 across injections) using the ion mobility direct injection method described in the experimental section. These metabolites are shown in

Figure 4 grouped by their ClassyFire/ChemOnt³² superclass labels together with their mean TIC-normalized peak areas across injections. We note that further automated data collection and analysis modules can be incorporated via AutonoMS to improve the application-specific detection performance of such a workflow. Usage of the AutonoMS platform enabled hands-off automated metabolomic profiling in the context of existing background knowledge of yeast metabolism. In this case, usage of the platform dramatically lowered the time and effort required to characterize yeast samples and produce interpretable results in the context of this background knowledge. This opens the door towards incorporation of a cutting-edge analytical platform into more fully autonomous discovery applications in which AI software agents control experimentation, interpret results, and run further rounds of experimentation as has been proposed and demonstrated previously.³³⁻³⁵ Metabolomics-based biological discovery, especially in yeast, is a rich field. Previous work has demonstrated

the utility of applying rigorous mass spectrometry methods towards measurements of yeast metabolites.^{36,37} Integrating these with automated culturing, sampling, additional data modalities, and modeling techniques has yielded powerful approaches.^{38,39} Existing approaches have demonstrated the potential of in-line SPE-IM-MS exometabolome analysis in live cultures.⁴⁰ AutonoMS should facilitate such multifaceted approaches, allowing for granular control logic and flexibility to new techniques. We also note that in addition to downstream discovery, the automated characterization of well-behaved metabolites for a given instrument acquisition method immediately opens the door to closed-loop automated MS acquisition methods development.

CONCLUSIONS

Analytical instrumentation must keep pace with the increasing ability of life science laboratories to quickly produce large quantities of experimental samples. Here we introduced AutonoMS, a platform combining cutting-edge ion mobility-mass spectrometry instrumentation with software layers capable of end-to-end instrument control, data processing, and metabolomic analysis. From our findings, we conclude that the resulting configuration can be immediately utilized in labs already using the same instrumentation to dramatically improve workload and improve results, especially as related to human burden. Even when using human-compiled experiment definition files, the streamlined AutonoMS workflow produced actionable results in about 1/4th of the time compared to the same tasks using the already-optimized conventional Agilent RapidFire - 6560 vendor workflow (25 minutes compared to an hour for the targeted data analysis standards workflow) with much less manual intervention. We also note the extensibility of this platform, since the software workflow coordinates the actions of modular instrument control and informatic components, each of which can be replaced with a new module fitting a given laboratory's configuration. For example, AutonoMS could be used for untargeted metabolomic feature discovery or proteomic profiling by swapping out the current Skyline metabolomics module. The current RapidFire - 6560 instrumentation setup is already capable of a large

range of analysis including proteomics. A new instrument could be controlled by AutonoMS by calling its instrument control wrappers in the workflow. Multiple instruments can be run in parallel simply by modifying the workflow control configuration parameters. This process can be made easier through instrument vendor cooperation, particularly with regard to whether high-level users will have access to an application programming interface (API) to allow direct command-line control of all instrument functions, rather than requiring all commands to be passed manually to the instrument through a graphical user interface (GUI) by means of interactive mouse-clicks and pull-down menus on a screen. As described in the experimental section, instrument control for the RapidFire - 6560 required GUI automation. General laboratory automation outside of mass spectrometry remains a challenge. For example, our yeast quenching and extraction protocol was performed manually save for the final step of robotically dispensing samples into the injection plate using the Multidrop dispenser given logistical and resource constraints. With the proper automated liquid handling equipment and laboratory space, this protocol can be fully automated and the capabilities of AutonoMS would allow for completely hands-off analysis of samples from culture to output results. Looking further ahead, we believe AutonoMS opens the door to closed-loop automation utilizing cutting-edge analytical instrumentation in which automated laboratories produce samples, transfer them to the instrument, trigger AutonoMS runs, then use the processed results to perform the next round of experimentation. For example, a setup combining automated software-controlled chemostats, liquid handling robotics, centrifuges, heat plates, shakers, freezers, and the RapidFire - 6560 could be used to culture yeast samples in various conditions, prepare them, then metabolically analyze them all via software calls with AutonoMS handling the IM-MS portion of the workflow. An AI system sitting on top of this configuration could take produced results and compare with existing background knowledge to find discrepancies between observed and expected output. Such systems could utilize database systems under development⁴¹ designed for AI software agent unified access to experimental data, protocol metadata, results, and background knowledge.

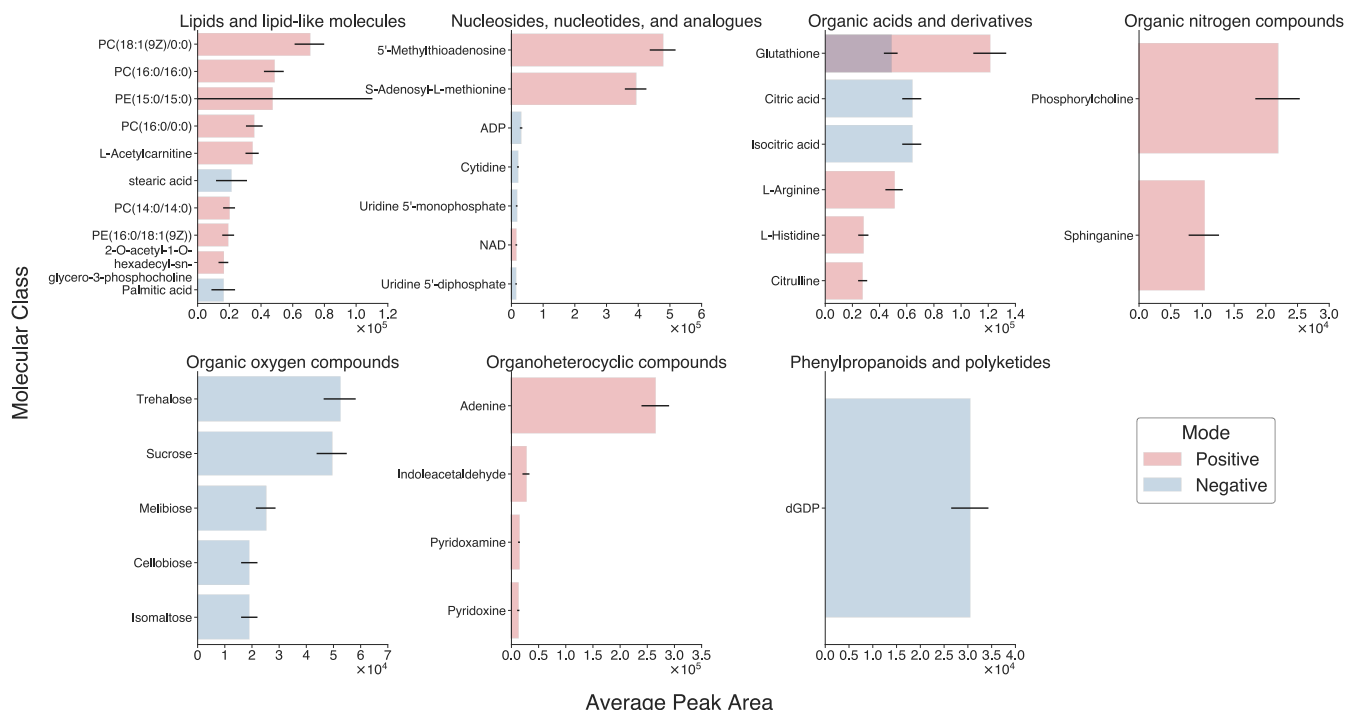


Figure 4. Use of AutoMS for automated data collection and integration with background knowledge. Panel of 35 metabolites from the Yeast Metabolome Database (YMDB^{27,28}) with publicly available collision cross section (CCS) values from the Baker Lab Cross Collision Section Database (CCSDB). Chosen metabolites were detected by the acquisition method described in the experimental section and exhibited mean peak areas across extracted yeast injections greater than 10⁴. Metabolites are grouped by their ClassyFire/ChemOnt³² superclass labels and are plotted against their mean peak areas across extracted yeast injections per ionization mode. Error bars display the intensity standard deviations.

ASSOCIATED CONTENT

Supporting Information

The file “supplementary-information” contains 6560 MS acquisition parameters, RapidFire method parameters, BLAZE mode configuration instructions, and MassProfiler feature finding parameters. (PDF)

The file “autonoms-demo” is a video demonstration of AutoMS usage demonstrating a complete run through the RapidFire – 6560 automated workflow on two short sequences, one in positive ionization mode and the other in negative mode. (MP4)

AutoMS source code, example experimental, configuration, and data processing files are available via GitHub at <https://github.com/gkreder/autonoms>

AutoMS experimental run data, manuscript visualization notebooks, and source code are available via Zenodo at <https://doi.org/10.5281/zenodo.8318982>

AUTHOR INFORMATION

Corresponding Author

*Gabriel K. Reder – Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg, Sweden 412 96; Department of Applied Physics, SciLifeLab, KTH Royal Institute of Technology, Solna, Sweden 171 21; ORCID: 0000-0001-8918-0789; Email: rederg@kth.se

Authors

Erik Y. Bjurström – Department of Life Sciences, Chalmers University of Technology, Gothenburg, Sweden 412 96; ORCID: 0000-0002-8863-1207

Daniel Brunnsåker – Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg, Sweden 412 96; ORCID: 0000-0002-5167-0536

Filip Kronström – Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg, Sweden 412 96; ORCID: 0000-0002-3011-5541

Praphapan Lasin – Department of Life Sciences, Chalmers University of Technology, Gothenburg, Sweden 412 96

Ievgeniia Tiukova – Department of Life Sciences, Chalmers University of Technology, Gothenburg, Sweden 412 96

Otto I. Savolainen – Department of Life Sciences, Chalmers University of Technology, Gothenburg, Sweden 412 96; Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland 70211

James N. Dodds – Chemistry Department, The University of North Carolina at Chapel Hill, Chapel Hill, NC USA 27599; ORCID: 0000-0002-9702-2294

Jody C. May – Department of Chemistry, Vanderbilt University, Nashville, TN, USA 37235; Center for Innovative Technology, Vanderbilt University, Nashville, TN, USA 37235 ORCID: 0000-0003-4871-5024

John P. Wikswo – Vanderbilt Institute for Integrative Biosystems Research and Education, Vanderbilt University, Nashville, TN, USA 37235; Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, USA 37235; Department of Physics and Astronomy, Vanderbilt University, Nashville, TN, USA 37235; Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA 37240; ORCID: 0000-0003-2790-1530

John A. McLean – Department of Chemistry, Vanderbilt University, Nashville, TN, USA 37235; Center for Innovative Technology, Vanderbilt University, Nashville, TN, USA 37235; Vanderbilt Institute for Integrative Biosystems Research and Education, Vanderbilt University, Nashville, TN, USA 37235; ORCID: 0000-0001-8918-6419

Ross D. King – Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg, 412 96; Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, UK CB3 0AS; The Alan Turing Institute, London, UK NW1 2DB; ORCID: 0000-0001-7208-4387

Author Contributions

Conceptualization: G.K.R., J.C.M., J.A.M., J.P.W., R.D.K.; Methodology: G.K.R., J.C.M., J.N.D.; Software: G.K.R., F.K. Investigation: G.K.R., E.Y.B., D.B., P.L.; Resources: R.D.K., J.A.M., J.P.W.; Writing: G.K.R.; Supervision: O.S., J.P.W., J.A.M., I.T., R.D.K. Funding acquisition: J.P.W., J.A.M., R.D.K.

ACKNOWLEDGMENT

The authors gratefully acknowledge the members of the Ross King Group at Chalmers University for their thoughtful insights and discussions. We would also like to thank the members of John Wikswo's group, John McLean's group, and the Center for Innovative Technology (CIT) at Vanderbilt University. This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. This work was also supported in part by the National Institutes of Health grant R01 GM141277, National Institute of Environmental Health Sciences grant P42 ES027704, and National Science Foundation grant 2117782.

REFERENCES

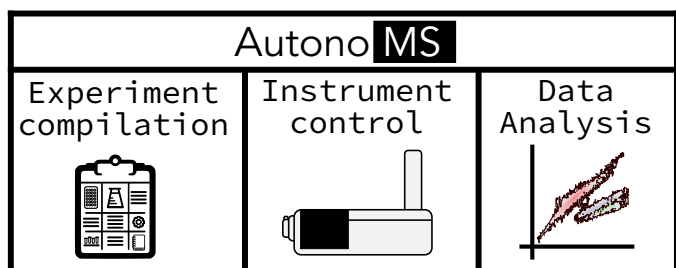
- Holland, I.; Davies, J. A. Automation in the Life Science Research Laboratory. *Front. Bioeng. Biotechnol.* **2020**, *8*. <https://doi.org/10.3389/fbioe.2020.571777>.
- Bai, J.; Cao, L.; Mosbach, S.; Akroyd, J.; Lapkin, A. A.; Kraft, M. From Platform to Knowledge Graph: Evolution of Laboratory Automation. *JACS Au* **2022**, *2* (2), 292–309. <https://doi.org/10.1021/jacsau.1c00438>.
- Chory, E. J.; Gretton, D. W.; DeBenedictis, E. A.; Esvelt, K. M. Enabling High-Throughput Biology with Flexible Open-Source Automation. *Mol. Syst. Biol.* **2021**, *17* (3), e9942. <https://doi.org/10.15252/msb.20209942>.
- Check Hayden, E. The Automated Lab. *Nature* **2014**, *516* (7529), 131–132. <https://doi.org/10.1038/516131a>.
- Crutchfield, C. A.; Thomas, S. N.; Sokoll, L. J.; Chan, D. W. Advances in Mass Spectrometry-Based Clinical Biomarker Discovery. *Clin. Proteomics* **2016**, *13* (1), 1. <https://doi.org/10.1186/s12014-015-9102-9>.
- Carneiro, G.; Radcenco, A. L.; Evaristo, J.; Monnerat, G. Novel Strategies for Clinical Investigation and Biomarker Discovery: A Guide to Applied Metabolomics. *Horm. Mol. Biol. Clin. Invest.* **2019**, *38* (3). <https://doi.org/10.1515/hmbci-2018-0045>.
- Ren, J.-L.; Zhang, A.-H.; Kong, L.; Wang, X.-J. Advances in Mass Spectrometry-Based Metabolomics for Investigation of Metabolites. *RSC Adv.* **2018**, *8* (40), 22335–22350. <https://doi.org/10.1039/C8RA01574K>.
- Zhang, X.; Li, Q.; Xu, Z.; Dou, J. Mass Spectrometry-Based Metabolomics in Health and Medical Science: A Systematic Review. *RSC Adv.* **2020**, *10* (6), 3092–3104. <https://doi.org/10.1039/C9RA08985C>.
- Paglia, G.; Smith, A. J.; Astarita, G. Ion Mobility Mass Spectrometry in the Omics Era: Challenges and Opportunities for Metabolomics and Lipidomics. *Mass Spectrom. Rev.* **2022**, *41* (5), 722–765. <https://doi.org/10.1002/mas.21686>.
- 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* (4), 281–294. <https://doi.org/10.1038/nchem.1889>.
- Stow, S. M.; Causon, T. J.; Zheng, X.; Kurulugama, R. T.; Mairinger, T.; May, J. C.; Rennie, E. E.; Baker, E. S.; Smith, R. D.; McLean, J. A.; Hann, S.; Fjeldsted, J. C. An Interlaboratory Evaluation of Drift Tube Ion Mobility–Mass Spectrometry Collision Cross Section Measurements. *Anal. Chem.* **2017**, *89* (17), 9048–9055. <https://doi.org/10.1021/acs.analchem.7b01729>.
- Dodds, J. N.; Baker, E. S. Improving the Speed and Selectivity of Newborn Screening Using Ion Mobility Spectrometry–Mass Spectrometry. *Anal. Chem.* **2021**, *93* (51), 17094–17102. <https://doi.org/10.1021/acs.analchem.1c04267>.
- Zhang, X.; Romm, M.; Zheng, X.; Zink, E. M.; Kim, Y.-M.; Burnum-Johnson, K. E.; Orton, D. J.; Apffel, A.; Ibrahim, Y. M.; Monroe, M. E.; Moore, R. J.; Smith, J. N.; Ma, J.; Renslow, R. S.; Thomas, D. G.; Blackwell, A. E.; Swinford, G.; Sausen, J.; Kurulugama, R. T.; Eno, N.; Darland, E.; Stafford, G.; Fjeldsted, J.; Metz, T. O.; Teeguarden, J. G.; Smith, R. D.; Baker, E. S. SPE-IMS-MS: An Automated Platform for Sub-Sixty Second Surveillance of Endogenous Metabolites and Xenobiotics in Biofluids. *Clin. Mass Spectrom.* **2016**, *2*, 1–10. <https://doi.org/10.1016/j.clinms.2016.11.002>.
- Prefect, 2023. <https://github.com/PrefectHQ/prefect> (accessed 2023-06-05).
- Pywinauto, 2023. <https://github.com/pywinauto/pywinauto> (accessed 2023-06-05).
- Bilbao, A.; Gibbons, B. C.; Stow, S. M.; Kyle, J. E.; Bloodsworth, K. J.; Payne, S. H.; Smith, R. D.; Ibrahim, Y. M.; Baker, E. S.; Fjeldsted, J. C. A Preprocessing Tool for Enhanced Ion Mobility–Mass Spectrometry-Based Omics Workflows. *J. Proteome Res.* **2022**, *21* (3), 798–807. <https://doi.org/10.1021/acs.jproteome.1c00425>.
- Colby, S. M.; Chang, C. H.; Bade, J. L.; Nunez, J. R.; Blumer, M. R.; Orton, D. J.; Bloodsworth, K. J.; Nakayasu, E. S.; Smith, R. D.; Ibrahim, Y. M.; Renslow, R. S.; Metz, T. O. DEIMoS: An Open-Source Tool for Processing High-Dimensional Mass Spectrometry Data. *Anal. Chem.* **2022**, *94* (16), 6130–6138. <https://doi.org/10.1021/acs.analchem.1c05017>.
- MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments.

- Bioinformatics* **2010**, *26* (7), 966–968. <https://doi.org/10.1093/bioinformatics/btq054>.
- (19) Adams, K. J.; Pratt, B.; Bose, N.; Dubois, L. G.; John-Williams, L. St.; Perrott, K. M.; Ky, K.; Kapahi, P.; Sharma, V.; MacCoss, M. J.; Moseley, M. A.; Colton, C. A.; MacLean, B. X.; Schilling, B.; Thompson, J. W. Skyline for Small Molecules: A Unifying Software Package for Quantitative Metabolomics. *J. Proteome Res.* **2020**, *19* (4), 1447–1458. <https://doi.org/10.1021/acs.jproteome.9b00640>.
 - (20) Veach, B. T.; Mudalige, T. K.; Rye, P. RapidFire Mass Spectrometry with Enhanced Throughput as an Alternative to Liquid–Liquid Salt Assisted Extraction and LC/MS Analysis for Sulfonamides in Honey. *Anal. Chem.* **2017**, *89* (6), 3256–3260. <https://doi.org/10.1021/acs.analchem.6b04889>.
 - (21) May, J. C.; Goodwin, C. R.; Lareau, N. M.; Leaptrot, K. L.; Morris, C. B.; Kurulugama, R. T.; Mordehai, A.; Klein, C.; Barry, W.; Darland, E.; Overney, G.; Imatani, K.; Stafford, G. C.; Fjeldsted, J. C.; McLean, J. A. Conformational Ordering of Biomolecules in the Gas Phase: Nitrogen Collision Cross Sections Measured on a Prototype High Resolution Drift Tube Ion Mobility-Mass Spectrometer. *Anal. Chem.* **2014**, *86* (4), 2107–2116. <https://doi.org/10.1021/ac4038448>.
 - (22) May, J. C.; Knochenmuss, R.; Fjeldsted, J. C.; McLean, J. A. Resolution of Isomeric Mixtures in Ion Mobility Using a Combined Demultiplexing and Peak Deconvolution Technique. *Anal. Chem.* **2020**, *92* (14), 9482–9492. <https://doi.org/10.1021/acs.analchem.9b05718>.
 - (23) Martens, L.; Chambers, M.; Sturm, M.; Kessner, D.; Levander, F.; Shofstahl, J.; Tang, W. H.; Römpf, A.; Neumann, S.; Pizarro, A. D.; Montecchi-Palazzi, L.; Tasman, N.; Coleman, M.; Reisinger, F.; Souda, P.; Hermjakob, H.; Binz, P.-A.; Deutsch, E. W. mzML—a Community Standard for Mass Spectrometry Data*. *Mol. Cell. Proteomics* **2011**, *10* (1), R110.000133. <https://doi.org/10.1074/mcp.R110.000133>.
 - (24) Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open Source Software for Rapid Proteomics Tools Development. *Bioinformatics* **2008**, *24* (21), 2534–2536. <https://doi.org/10.1093/bioinformatics/btn323>.
 - (25) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruder- man, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egerton, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat. Biotechnol.* **2012**, *30* (10), 918–920. <https://doi.org/10.1038/nbt.2377>.
 - (26) Tomerfiliba-Org/Rpyc, 2023. <https://github.com/tomerfiliba-org/rpyc> (accessed 2023-06-05).
 - (27) Jewison, T.; Knox, C.; Neveu, V.; Djoumbou, Y.; Guo, A. C.; Lee, J.; Liu, P.; Mandal, R.; Krishnamurthy, R.; Sinelnikov, I.; Wilson, M.; Wishart, D. S. YMDB: The Yeast Metabolome Database. *Nucleic Acids Res.* **2012**, *40* (D1), D815–D820. <https://doi.org/10.1093/nar/gkr916>.
 - (28) Ramirez-Gaona, M.; Marcu, A.; Pon, A.; Guo, A. C.; Sajed, T.; Wishart, N. A.; Karu, N.; Djoumbou Feunang, Y.; Arndt, D.; Wishart, D. S. YMDB 2.0: A Significantly Expanded Version of the Yeast Metabolome Database. *Nucleic Acids Res.* **2017**, *45* (D1), D440–D445. <https://doi.org/10.1093/nar/gkw1058>.
 - (29) Bretschneider, T.; Ozbal, C.; Holstein, M.; Winter, M.; Buettner, F. H.; Thamm, S.; Bischoff, D.; Luippold, A. H. RapidFire BLAZE-Mode Is Boosting ESI-MS Toward High-Throughput-Screening. *SLAS Technol.* **2019**, *24* (4), 386–393. <https://doi.org/10.1177/2472630318822449>.
 - (30) Brunnsäker, D.; Reder, G. K.; Soni, N. K.; Savolainen, O. I.; Gower, A. H.; Tiukova, I. A.; King, R. D. High-Throughput Metabolomics for the Design and Validation of a Diauxic Shift Model. *Npj Syst. Biol. Appl.* **2023**, *9* (1), 1–9. <https://doi.org/10.1038/s41540-023-00274-9>.
 - (31) Campbell, K.; Westholm, J.; Kasvandik, S.; Di Bartolomeo, F.; Mormino, M.; Nielsen, J. Building Blocks Are Synthesized on Demand during the Yeast Cell Cycle. *Proc. Natl. Acad. Sci.* **2020**, *117* (14), 7575–7583. <https://doi.org/10.1073/pnas.1919535117>.
 - (32) Djoumbou Feunang, Y.; Eisner, R.; Knox, C.; Chepelev, L.; Hastings, J.; Owen, G.; Fahy, E.; Steinbeck, C.; Subramanian, S.; Bolton, E.; Greiner, R.; Wishart, D. S. ClassyFire: Automated Chemical Classification with a Comprehensive, Computable Taxonomy. *J. Cheminformatics* **2016**, *8* (1). <https://doi.org/10.1186/s13321-016-0174-y>.
 - (33) King, R. D.; Whelan, K. E.; Jones, F. M.; Reiser, P. G. K.; Bryant, C. H.; Muggleton, S. H.; Kell, D. B.; Oliver, S. G. Functional Genomic Hypothesis Generation and Experimentation by a Robot Scientist. *Nature* **2004**, *427* (6971), 247–252. <https://doi.org/10.1038/nature02236>.
 - (34) Williams, K.; Bilsland, E.; Sparkes, A.; Aubrey, W.; Young, M.; Soldatova, L. N.; De Grave, K.; Ramon, J.; de Clare, M.; Sirawaraporn, W.; Oliver, S. G.; King, R. D. Cheaper Faster Drug Development Validated by the Repositioning of Drugs against Neglected Tropical Diseases. *J. R. Soc. Interface* **2015**, *12* (104), 20141289. <https://doi.org/10.1098/rsif.2014.1289>.
 - (35) Coutant, A.; Roper, K.; Trejo-Banos, D.; Bouthinon, D.; Carpenter, M.; Grzebyta, J.; Santini, G.; Soldano, H.; Elati, M.; Ramon, J.; Rouveiro, C.; Soldatova, L. N.; King, R. D. Closed-Loop Cycles of Experiment Design, Execution, and Learning Accelerate Systems Biology Model Development in Yeast. *Proc. Natl. Acad. Sci.* **2019**, *116* (36), 18142–18147. <https://doi.org/10.1073/pnas.1900548116>.
 - (36) Ewald, J. C.; Matt, T.; Zamboni, N. The Integrated Response of Primary Metabolites to Gene Deletions and the Environment. *Mol. Biosyst.* **2013**, *9* (3), 440. <https://doi.org/10.1039/c2mb25423a>.
 - (37) Crutchfield, C. A.; Lu, W.; Melamud, E.; Rabinowitz, J. D. Chapter 16 - Mass Spectrometry-Based Metabolomics of Yeast. In *Methods in Enzymology*; Guide to Yeast Genetics: Functional Genomics, Proteomics, and Other Systems Analysis; Academic Press, 2010; Vol. 470, pp 393–426. [https://doi.org/10.1016/S0076-6879\(10\)70016-1](https://doi.org/10.1016/S0076-6879(10)70016-1).
 - (38) Link, H.; Fuhrer, T.; Gerosa, L.; Zamboni, N.; Sauer, U. Real-Time Metabolome Profiling of the Metabolic Switch between Starvation and Growth. *Nat. Methods* **2015**, *12* (11), 1091–1097. <https://doi.org/10.1038/nmeth.3584>.
 - (39) Hackett, S. R.; Zanutelli, V. R. T.; Xu, W.; Goya, J.; Park, J. O.; Perlman, D. H.; Gibney, P. A.; Botstein, D.; Storey, J. D.; Rabinowitz, J. D. Systems-Level Analysis of Mechanisms Regulating Yeast Metabolic Flux. *Science* **2016**, *354* (6311), aaf2786. <https://doi.org/10.1126/science.aaf2786>.
 - (40) Marasco, C. C.; Enders, J. R.; Seale, K. T.; McLean, J. A.; Wikswo, J. P. Real-Time Cellular Exometabolome Analysis with a Microfluidic-Mass Spectrometry Platform. *PLOS ONE* **2015**, *10* (2), e0117685. <https://doi.org/10.1371/journal.pone.0117685>.
 - (41) Reder, G. K.; Gower, A. H.; Kronström, F.; Halle, R.; Mahamuni, V.; Patel, A.; Hayatnagarkar, H.; Soldatova, L. N.; King, R. D. Genesis-DB: A Database for Autonomous Laboratory Systems. *Bioinforma. Adv.* **2023**, *3* (1), vbad102. <https://doi.org/10.1093/bioadv/vbad102>.

For Table of Contents Use Only

AutonoMS: Automated ion mobility metabolomic fingerprinting

Gabriel K. Reder*, Erik Y. Bjurström, Daniel Brunnsåker, Filip Kronström, Praphapan Lasin, Ievgeniia Tiukova, Otto I. Savolainen, James N. Dodds, Jody C. May, John P. Wikswo, John A. McLean, Ross D. King



Synopsis: The authors introduce a software suite, AutonoMS, for automating mass spectrometry runs including sample run compilation, instrument run control, and data analysis. Current functionality is written around the Agilent RapidFire – 6560 ion mobility platform. Modules for these functionalities are integrated using workflow control software in a manner allowing users to add modules for their own instruments and data workflows.
