

Metabolomics across scales: from single cells to population studies

<https://doi.org/10.1038/s41586-026-10277-1>

Theodore Alexandrov^{1,2,3}✉ & Nicola Zamboni^{4,5}✉

Received: 3 March 2025

Accepted: 12 February 2026



Check for updates

Metabolomics has matured into a powerful approach for probing metabolism, offering readouts that closely reflect cellular and organismal function in health and disease. Here we highlight two rapidly advancing frontiers: single-cell metabolomics and population-scale metabolomics. Single-cell metabolomics resolves the metabolic states of individual cells, uncovering cell-to-cell heterogeneity and spatial organization within tissues. Population-scale profiling profiles metabolites across large cohorts, enabling the discovery of markers of disease, environmental exposures and genetic variation. Although these approaches operate at different scales, they face shared challenges—including metabolite identification, quantification and multimodal data integration—and offer common advantages, such as the ability to capture non-genetic influences on phenotype and to scale to high throughput. We propose that continued advances in scalability will bring these domains together, enabling the construction of comprehensive metabolic atlases that chart cellular and interindividual variation and provide training data for foundation models of metabolism. By integrating cellular and population-level insights, single-cell and population-scale metabolomics promise to advance our understanding of metabolism across biology, medicine and pharmacology.

Over the past decade, metabolomics has had a pivotal role in reshaping our understanding of metabolism as a central regulator in biology, medicine and pharmacology¹. Once associated solely with biosynthesis and energy production, metabolism is now recognized for its crucial role in cellular regulation, influencing nearly every aspect of cellular and organismal phenomena^{2,3}. Beyond directly probing metabolism, metabolomics complements other omics disciplines by offering integrated, phenotypic insights into cellular states, characterizing metabolic fluxes^{4,5} and detecting protein–metabolite interactions. Complementing its roles in biology and medicine, metabolomics is uniquely positioned to elucidate the molecular effects of the exposome⁶—including dietary constituents, microbiome products and pollutants—and provides functional, mechanistically grounded readouts in pharmacology for identifying disease biomarkers and drug targets, uncovering a drug's mode of action and assessing toxicity^{7–11}. The introduction of high-resolution mass spectrometry (MS), coupled with metabolomics' distinct advantages in throughput and scalability, compared with sequencing-based omics and proteomics, has expanded its practice across the life sciences and biomedicine.

These technological advances have fostered the emergence of specialized subfields that operate on a wide range of scales. Single-cell metabolomics examines the metabolic profiles of individual cells in heterogeneous populations, revealing metabolic states and spatial organization at cellular resolution^{12–14}. Population-scale metabolomics characterizes the metabolic profiles of individuals from large human

cohorts, identifying metabolic markers of disease, environmental exposures and genetic variation^{15,16}. Despite their distinct applications, both approaches share common technological foundations and a unified goal: to map the complexity and heterogeneity of metabolism.

In this perspective, we explore the technological capabilities, promises and challenges of single-cell and population-scale metabolomics. We examine their synergistic opportunities for investigating metabolism and addressing shared challenges in metabolite identification, data integration and quantification. We predict that the scalability of both approaches will drive their convergence, starting from the creation of comprehensive metabolic atlases integrating data from single cells to large cohorts, and ultimately transforming our understanding of metabolism across biological scales.

Single-cell metabolomics

Single-cell metabolomics (SCM) focuses on detecting metabolic profiles of individual cells within tissues or from heterogeneous cell populations^{12,14} (Fig. 1). It follows in the footsteps of other single-cell omics, particularly single-cell transcriptomics or single-cell RNA sequencing, (scRNA-seq), which revolutionized our understanding of biological systems. Single-cell transcriptomics paved the way for other single-cell omics by revealing the profound heterogeneity of cells and demonstrating the vast opportunities such technologies can unlock in biology, medicine and pharmacology.

¹Department of Pharmacology, University of California, San Diego, La Jolla, CA, USA. ²Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA. ³DeepCyte, San Diego, CA, USA. ⁴Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland. ⁵PHRT Swiss Multi-Omics Center, Zurich, Switzerland. ✉e-mail: talexandrov@health.ucsd.edu; zamboni@imsb.biol.ethz.ch

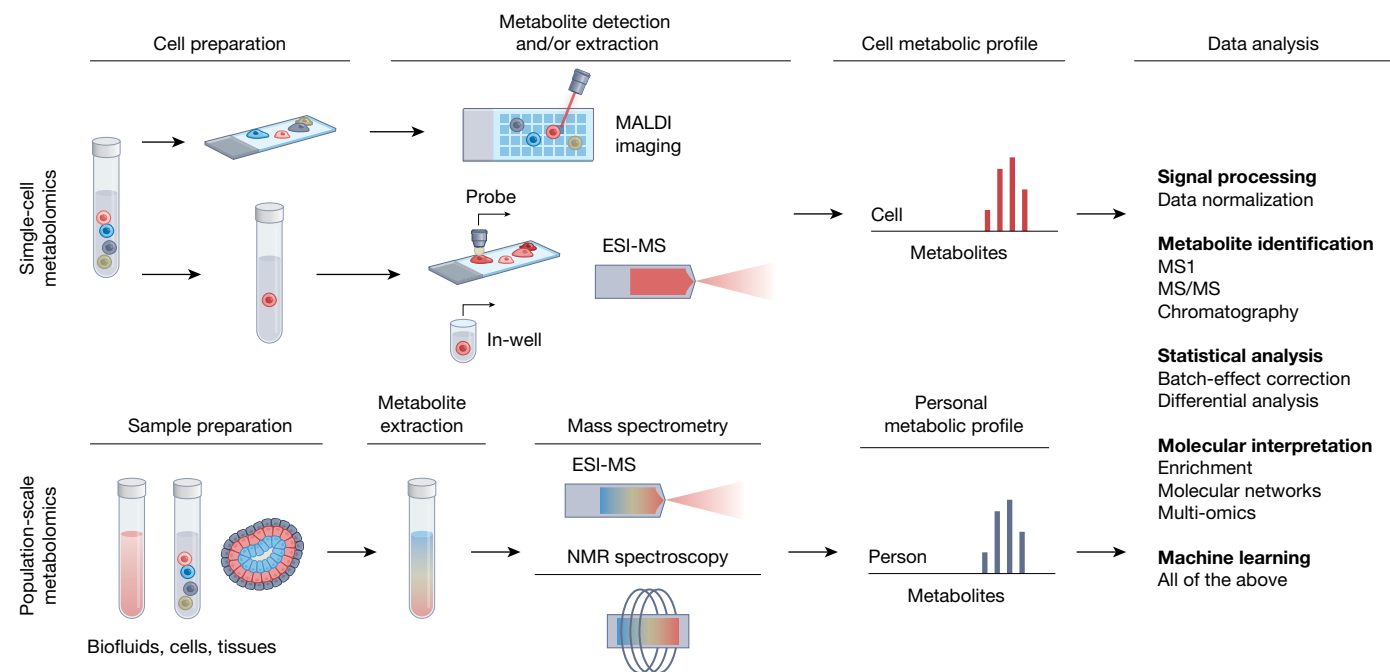


Fig. 1 | Single-cell metabolomics and population-scale metabolomics.

Single-cell metabolomics (SCM) can be performed following various approaches, either by interrogating cells on a glass slide or by extracting metabolites from

individual cells. Population-scale metabolomics (PSM) can be performed using MS (focus of this perspective) or NMR spectroscopy. The data analysis for both SCM and MS-based PSM generally follows similar steps.

Technology view

After early demonstrations in the late 2000s, SCM expanded rapidly, owing to advancements in the sensitivity and robustness of electrospray ionization (ESI)-based metabolomics and improvements in the spatial resolution of matrix-assisted laser desorption/ionization (MALDI) imaging MS, with recent studies achieving a sub-1- μm pixel size¹⁷. Today, various SCM approaches exist, which rely mainly on these two ionization technologies, both of which were recognized with the 2002 Nobel Prize in Chemistry^{13,14}.

SCM has been successfully applied to individual cells from cell lines as well as to primary cells, including neurons¹⁸ or stem cells¹⁹. Although it has proved useful in classifying cells by type, its most compelling potential lies in identifying, characterizing and spatially mapping metabolic states—elucidating how these states are established in homeostasis, influenced by cellular heterogeneity, linked to specific functions and dysregulated in disease^{19–23}. Just as scRNA-seq enabled the discovery and refinement of cellular states and multicellular transcriptional programs that were previously obscured by bulk transcriptomics, SCM holds the promise of uncovering metabolic states that have remained elusive because they are represented among small subpopulations of cells or blurred by the averaging effect of bulk analyses. Beyond merely identifying metabolic states, SCM can advance our understanding of metabolism by interpreting metabolic states in the context of vast knowledge about cell types, subtypes and transcriptional programs through single-cell multi-omics co-detecting metabolites with protein^{24,25} and gene markers^{26,27}.

At present, SCM is performed mainly on isolated or cultured cells¹³. Using SCM on a tissue section represents a major frontier in the field, because tissues exhibit high cellular heterogeneity and complex spatial organization, including metabolite gradients²⁸. It is, however, challenging and out of reach even for most specialists, because many current SCM methods lack the spatial resolution and/or sensitivity required to reliably detect metabolites from individual cells within tissue sections. However, promising studies in recent years^{17,24,25} indicate that such capabilities will soon be achieved more broadly, supported by rapid advancements in spatial sampling precision, desorption techniques and MS sensitivity.

Promise

SCM holds promise in addressing various open questions in biology. Beyond identifying metabolic states, SCM offers a perfect tool to map and reveal cell–cell metabolic heterogeneity, a phenomenon of increasing interest in cancer yet considerably unexplored in other fields of biology^{29–31}. The ability to profile metabolism and simultaneously phenotype individual cells, as already demonstrated²², would enable novel discoveries about how metabolism controls and regulates cell fate and function^{32,33}.

A key strength of SCM is its ability to integrate with other single-cell technologies and omics. Early demonstrations show its feasibility for the co-detection of gene markers^{26,27}, using ESI-MS in single cells and MALDI at the near-single-cell resolution in tissues^{34,35}, as well as for the co-detection of metabolites and proteins in tissues^{24,25}. In parallel, other imaging MS techniques are beginning to approach single-cell resolution, thus further expanding the SCM toolbox³⁶. Single-cell multi-omics integrating metabolomics holds tremendous potential for mapping cellular programs, from DNA to phenotype, and for revealing mechanistic links between cellular programs that encompass different molecular entities. In particular, it could be used to identify interactions between metabolites and proteins³ or RNA³⁷, with the potential to reveal the roles of metabolite signalling in cell differentiation during embryonic development; interpret the effects of nutrient gradients and oncometabolites on cellular heterogeneity and drug resistance in tumours; and identify how metabolic pathways are regulated during CAR-T-cell differentiation and their effects on therapeutic efficacy^{38,39}. These advances can be facilitated by mapping SCM data onto single-cell transcriptomic atlases, thereby opening new avenues for interpreting SCM data in the context of large, diverse and biologically and disease-relevant annotated cell atlases.

By enabling the construction of pseudotime trajectories for heterogeneous cells, SCM can provide insights into the dynamic nature of metabolic processes. Pseudotime trajectory inference in single-cell transcriptomics has been shown to be a useful framework for elucidating dynamical cellular processes from a single time point or multiple snapshots—although it needs to be used with caution, and requires

Q10

Q11

Q14

validation⁴⁰. Complementing data-driven pseudotime modelling from time-resolved snapshots, isotope-tracing and labelling approaches were proposed for single-cell metabolomics and lipidomics^{41–44}. Such approaches could enable the discovery of metabolic factors that are affected by and control cell differentiation, early development, the cell cycle and immune-cell activation and polarization.

Performing SCM in tissues and overlaying it with other spatial omics would open a new chapter in understanding the cell specificity of metabolism and how it is compartmentalized and spatially regulated across cell types in anatomically and pathologically defined regions. Such technology would enable metabolic cell–cell interactions within tissues⁴⁵—such as the division of metabolic roles between cells in a spatial niche⁴⁶ or an organism⁴⁷—to be profiled in a direct way, rather than using indirect methods. Complementing discoveries of extracellular ATP hydrolysis and creatine phosphorylation in the tumour interstitium^{48,49}, and building on previous examples of interrogating extracellular vesicles⁵⁰ and ovarian follicular fluid using imaging MS⁵¹, SCM could provide insights into extracellular metabolism.

In pharmacology, SCM offers a unique advantage over scRNA-seq⁵², in that it detects both endogenous and exogenous metabolites in the same cells. Finding endogenous metabolites that are correlated with a drug or drug metabolites can identify dose-dependent effects with a higher sensitivity than would be possible with bulk assays. Furthermore, by providing mechanistically grounded readouts that closely reflect cell function, SCM can reveal the functional effects of treatments, identify the mode of action of drugs and reveal cell-type-specific toxicity.

In medicine, SCM can aid in the analysis of rare or valuable cells, such as CAR T cells or circulating tumour cells, for biomarker discovery, therapy prediction, response monitoring and patient stratification in clinical trials. Moreover, SCM could enhance disease understanding and patient phenotyping, particularly when performed in tissues and integrated with microscopy-based pathology or emerging spatial omics atlases of clinical samples, such as the Kidney Precision Medicine Project (KPMP)⁵³.

Challenges

Yet, as an emerging technology, SCM faces considerable challenges. As for bulk measurements, changes from the physiological state owing to unwanted enzymatic and non-enzymatic reactions, as well as metabolite delocalization, can affect the cell metabolome and bias the analysis, in particular for metabolites with high turnover. These factors can and should be mitigated by using proper controls, validating results with orthogonal technologies and designing protocols to minimize exposure to stress factors before the cells are fixed or desiccated. This challenge requires particular consideration when SCM is performed on cells that have been sampled or prepared externally; for example, on primary peripheral blood mononuclear cells (PBMCs) sampled and isolated clinically. In such a case, protocols for blood sampling, cell isolation, cell storage and revival need to be developed or reviewed by metabolism specialists. One of the primary technical challenges in SCM is metabolite sensitivity and reproducibility, although recent advances have enabled the detection of approximately 100 metabolites per cell, with high reproducibility and a coefficient of variation below 20%, as required for the validation of a bioanalytical method by the US Food and Drug Administration (FDA)⁵⁴. Achieving broad metabolite coverage remains difficult, even in bulk metabolomics. Detecting small-molecule metabolites, particularly with MALDI imaging, remains challenging as compared with lipids, which are more readily detected, but recent studies have demonstrated feasibility^{23,54}.

The evaluation of SCM methods is at present complicated by the lack of reference values for validating metabolite intensities in single cells. Using internal standards, as done in bulk metabolomics, is challenging, because achieving known quantities of a metabolite in individual cells is difficult. Dual-detection molecules, detectable by both MS and fluorescence, represent a powerful single-cell approach⁵⁵, but this technique needs to be further expanded to endogenous metabolites.

Batch-to-batch variability remains a concern; however, methods that were developed for scRNA-seq data integration⁵⁶ can be adapted to SCM data, leading to comparability between datasets within and across experiments. Distinguishing biological heterogeneity from technical noise is another challenge, hindering the development of methods to minimize unwanted variability. On the other hand, modulating metabolic processes through pharmacological inhibitors or CRISPR-based perturbations, and detecting expected changes, such as the accumulation of upstream metabolites⁵⁴, could be a powerful approach to validate SCM.

The scalability of SCM in terms of the number of samples and cells remains a bottleneck, because this requires streamlined methods for sample preparation, instruments designed and optimized for SCM, commercial sample-to-data solutions and democratization of the software for data analysis. The lack of dedicated data repositories for SCM limits data sharing; established metabolomics and single-cell transcriptomics repositories are not tailored for SCM datasets. Open questions in data processing—such as optimal normalization, scaling and filtering strategies—need to be addressed. Furthermore, the absence of SCM-specific data standards and metadata guidelines impedes data interoperability and reproducibility.

Population-scale metabolomics

Population-scale metabolomics (PSM) aims to characterize the metabolic profiles of large, diverse human cohorts, thereby providing insights into the complex relationships between metabolism, environment, lifestyle, diseases and other factors across populations. In contrast to SCM, which focuses on primary metabolism to capture the metabolic state of individual cells, PSM seeks to profile a broader chemical space of compounds that might affect health, including dietary products, microbiome products, drugs and pollutants. PSM has emerged from the convergence of large-scale epidemiology, conventionally limited to a small number of biomarkers, and high-throughput metabolomics, which enables the profiling of hundreds to thousands of compounds in parallel. This synergy has been driven by the growing recognition that metabolic signatures can be accurate biomarkers for disease risk, progression and therapeutic outcomes, as shown by several studies encompassing up to 700,000 individuals^{57–60}.

One fundamental question in PSM is how individual metabolites or metabolic patterns can serve as early biomarkers for disease risk, including the prediction of cardiovascular diseases, diabetes or neurodegenerative disorders. Another key inquiry centres on how environmental factors, such as diet, lifestyle and exposure to pollution, influence metabolic profiles and contribute to disease susceptibility across populations. PSM also examines the genetic basis of metabolic variability, investigating how genetic variants affect metabolic pathways and predispose individuals to particular health conditions^{61,62}. In all of these queries, the enhanced statistical power of PSM studies—stemming from large cohorts and reproducibly measurable features—allows the detection of subtle patterns and correlations that would otherwise be lost in noise. These associations are essential for hypothesis generation, but mechanistic understanding requires complementary approaches, such as genetic integration, flux analysis and experimental validation.

Technology view

PSM can analyse a wide variety of biological samples (Fig. 1). These include blood (plasma, serum and whole blood), urine, saliva, tissues and faecal samples. Blood remains the most commonly used sample type, owing to its ease of collection and the rich metabolic information it provides, followed by urine. Blood is normally collected venously in the clinical laboratory, and this approach supports reproducibility through rapid processing, standardized protocols and use of internal standards. Future studies could benefit from microsampling and

Q12

Q13

Q15

Perspective

other more user-friendly approaches, including self-collection⁶³. The highest-impact studies used samples from biobanks such as the UK Biobank, the Tohoku Medical Megabank Project, the Framingham Heart Study or the THL (Finnish Institute for Health and Welfare) Biobank.

To profile metabolites, PSM relies on either MS or nuclear magnetic resonance (NMR) spectroscopy. NMR offers excellent quantitation and high reproducibility with minimal sample preparation, making it well-suited to multi-site and large cohorts. MS has a lower inter-batch reproducibility but excels in sensitivity and coverage. The sensitivity is crucial because signalling metabolites might be present at low concentrations. The coverage is important for detecting thousands of metabolites, lipids, natural products or xenobiotics per sample. Compared with NMR, the deeper coverage and higher sensitivity of MS increases the chances of identifying causal links over merely correlated biomarkers.

For PSM, MS is coupled with a fast, generic separation technique, typically liquid chromatography (LC)⁶⁴. Increasing LC–MS throughput comes at the cost of incomplete separation of isomers. This is acceptable because it doesn't affect the unique capability of MS to provide a holistic profile that integrates non-heritable factors, such as the environment, lifestyle, diet and microbiome.

The initial analysis of PSM data often begins with metabolite profiling alone, but integration with genomic data has enabled major advances, including metabolome-wide association studies (MWASs) that identify genetic variants associated with metabolic traits.

Promise

PSM is applied both to large cohorts of healthy individuals and to patient cohorts. In the former case, it helps establish the baseline metabolic levels and diversity, gene–metabolite interactions and the effects of lifestyle and environmental exposures. By identifying metabolic markers of disease risk before clinical symptoms appear, PSM can enhance early diagnosis and prevention strategies. MWASs reveal metabolic traits linked to genetic variants, refining our understanding of disease mechanisms⁶⁵. Most importantly, PSM captures metabolic responses to diet, physical activity and environmental factors, offering a potentially more functional perspective on health. When combined with genome-wide association studies (GWASs) on the same samples, PSM helps bridge genetic predisposition and biochemical function. When applied to analyses of patient cohorts, PSM can support therapy selection by identifying metabolic biomarkers that predict individual responses to specific drugs and by training predictive models that map metabolic profiles to treatment outcomes.

Aside from its medical applications, PSM serves as a robust resource for fundamental biological discoveries. Large-scale metabolic profiling facilitates the examination of human metabolism in its natural variation, revealing how genetic and environmental factors shape metabolic networks. By integrating PSM with GWASs, researchers can map the genetic architecture of metabolism, identifying key enzymes and transporters that regulate metabolite levels. This will help uncover previously unrecognized metabolic pathways and biochemical interactions. PSM can also highlight metabolic adaptations to diet, geography and lifestyle. Furthermore, large-scale metabolomics can illuminate metabolic individuality, revealing population-level patterns in nutrient processing, microbiome-derived metabolites and endogenous signalling molecules. Notably, PSM provides a unique perspective on the exposome—comprising the entirety of environmental exposures such as diet, pollutants and microbiome-derived compounds—and will help researchers to decipher how these factors influence metabolism and long-term health.

PSM offers substantial opportunities for drug discovery and development, by linking metabolic profiles to genetic, environmental and disease-related factors. Integrating MWASs with GWASs can reveal metabolic pathways and enzymes linked to disease risk, thus highlighting potential drug targets⁶¹. By analysing pre-disease conditions, PSM could identify pathways that can be therapeutically adjusted before

irreversible harm occurs. In clinical research, metabolic profiling can predict subgroups of patients with specific treatment responses^{7,66,67} with the potential to improve patient stratification in clinical trials. In addition, monitoring metabolomic shifts over time in future clinical trials and large therapeutic groups will allow PSM to identify early biomarkers of drug effectiveness or toxicity, facilitating improved dose optimization and personalized treatment plans⁶⁸.

Challenges

PSM faces several key challenges. The major hurdle is to ensure high throughput. At present, analysing tens of thousands of samples is beyond the reach of regular laboratories and facilities, which normally operate with 10–100 samples. This requires special infrastructure for storing, barcoding and processing such samples, often with the help of robotics or automation systems, as well as a sufficient number and availability of instruments (MS or NMR) that are able to handle such numbers of samples. Managing such numbers of samples poses large logistical challenges, yet sufficient continuous demand is needed to support investments for setting up such special high-throughput facilities.

The most crucial factor for success in PSM is reproducibility. This is already a key factor in low-throughput experiments, in which various experimental approaches (for example, using internal standards, technical replicates, pooled quality control samples and normalization) are established to ensure and monitor reproducibility by identifying and eliminating outlier samples and quantifying the variability of the metabolite levels. For PSM, these requirements are more stringent because of the need to compare data collected over a long time period and, possibly, by several laboratories. Regular life science kits, reagents and instrumentation are not designed to support stability over tens of thousands of samples. Hence, large studies must be split into smaller batches. The long duration of analyses introduces the possibility of sample deterioration over time, requiring additional validation of protocols. Heavy internal standards and quality controls must be introduced at all steps to monitor and possibly correct batch effects, signal drifts and other deterioration phenomena by machine learning^{69–71}. This remains a major challenge for MS-based metabolomics because of the limited coverage of standards compared with the sheer number of features that can be profiled. Comprehensive systems suitability tests are being developed to prevent variability in multi-batch comparisons^{70,72}.

The complexity of data processing and management increases with scale. A common challenge in MS-based PSM is detecting low-abundance or rare features, such as drugs that occur in only a single sample, and distinguishing them from noise. This requires properly parameterizing data processing to balance sensitivity and reproducibility^{73,74}. The vast amount of metabolic information requires robust systems for data storage, sharing and retrieval, alongside strong data governance protocols to ensure consistency and accuracy. Finally, automated data analysis is essential for extracting meaningful insights from large, complex datasets.

Bridging the scales

Although they differ in scale, sample types and current focus, SCM and PSM have many similarities (Table 1), and offer synergistic opportunities for investigating metabolism (Fig. 2).

SCM and PSM are uniquely positioned to provide crucial insights into how metabolic states are regulated at the individual cellular level and, at the same time, orchestrated within tissues and across the entire organism. SCM could help researchers to outline possible metabolic states for particular cell types studied *in vitro*²⁰, and to map metabolic states of cells within the spatial architecture of tissues from animal models, tissue resections or tissue biopsies, as well as providing a window into the metabolic states of organisms by interrogating circulating cells from peripheral blood. On the other end of the spectrum,

Q16

Table 1 | Similarities and differences between single-cell metabolomics and population-scale metabolomics

Similarities in	Single-cell metabolomics and population-scale metabolomics	
Goals	Map and explain metabolism, discover metabolic mechanisms, identify metabolic factors of disease	
Complementing other omics	Enable phenotyping and systems-level interpretation, provide insights into non-genetic factors	
Underlying technology	Mass spectrometry; in particular, high-resolution mass spectrometry for untargeted profiling	
Technological benefits	Simple sample preparation, scalability, rapid analysis and high throughput compared with other omics	
Common challenges in quantitation	Sample complexity, ionization variability, matrix effects	
Common challenges in data analysis	Large scale of the data, metabolite identification, batch effects, normalization	
Potential for molecular discovery	Novel metabolites, metabolic markers, metabolic 'dark matter'	
Contribution to AI/ML efforts	Generating large-scale datasets for foundation and predictive models of metabolism and virtual cells	
Differences in	Single-cell metabolomics	Population-scale metabolomics
Scale	Single cells from large cell populations	Patients and healthy individuals from large cohorts
Sample types	Mainly cells; expansion to tissues is expected	Mainly biofluids
Focus of current developments	Sensitivity, integration with other omics	Reproducibility, clinical interpretation
Focus of current applications	Biology and pharmacology; expansion to clinical is expected	Clinical
Biological focus	Metabolic heterogeneity and cell states	Exposome, associations with genetic factors and disease risk
Metabolite coverage	100–200 metabolites	100–1,000 metabolites
Spatial resolution	Approaching single-cell resolution in tissues	N/A
Technological maturity	Workflows are still emerging and are not automated	Well-developed in specialist laboratories and providers
Data interpretation	Benefiting from single-cell transcriptomics frameworks, will further benefit from integrating cell typing using other omics	Often used independently or together with genomics
Current throughput limits	10+ samples (regular methods) to 100+ samples (high-throughput methods)	10,000+ samples

PSM could provide information about possible metabolic states of an organism, particularly by studying blood-circulating metabolites⁷⁵, metabolic diversity of individuals and changes relating to exercise⁷⁶, diet⁷⁷ and disease⁷⁸. Metabolomics studies in the past few years have shown the complementary value of interrogating both circulating cells and plasma^{79,80}.

The concept of metabolic age—reflecting an organism's biological age^{60,81,82} through its metabolic profile—has gained increasing attention. Metabolomics, particularly through PSM, offers a powerful approach

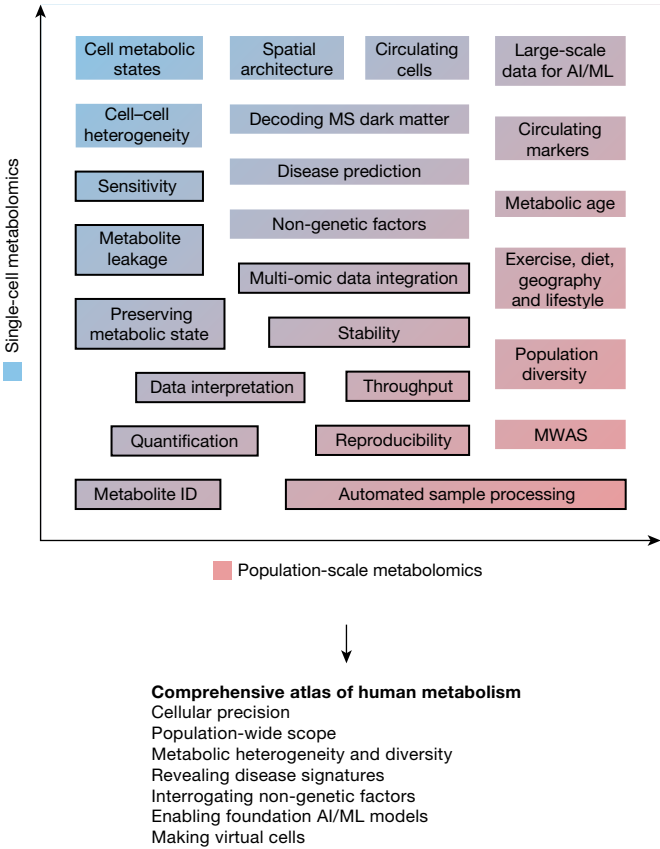


Fig. 2 | Bridging metabolomics across scales: from single cells to population studies. Single-cell and population-scale metabolomics provide unique capabilities for addressing key questions in biology, medicine and pharmacology (boxes without borders). Overcoming current challenges (boxes with black borders) will further advance these approaches, enabling the development of a comprehensive atlas of human metabolism that links cellular processes to organismal health and disease progression. AI/ML, artificial intelligence and machine learning.

to interrogate metabolic changes associated with biological ageing and to establish a 'metabolic clock' that predicts an organism's biological age⁸³. Beyond ageing, SCM and PSM can also reveal dynamic metabolic shifts during other key processes such as cell differentiation, the cell cycle and disease progression. For example, metabolic reprogramming during differentiation can be tracked at the single-cell level, providing insights into how metabolic states evolve in different cell types. Similarly, the metabolic alterations that accompany disease progression, such as cancer or metabolic disorders, can be captured through PSM to elucidate how metabolic profiles shift in response to pathophysiological changes. Integrating SCM and PSM data thus allows us to investigate the complex relationships between metabolism, ageing, differentiation and disease, offering a comprehensive view of metabolic dynamics across cellular and organismal scales.

Metabolic heterogeneity exists on several levels, including cell–cell variability, spatial heterogeneity within tissues and diversity between individuals³⁰. Despite the complexity of metabolism, there are underlying principles that govern distinct metabolic organizations, such as specific metabolic pathways, states or programs that are conserved across systems. SCM and PSM provide unique opportunities to study this heterogeneity, enabling the exploration of how metabolic diversity is structured and maintained within these invariant frameworks. These technologies allow us to uncover the mechanisms behind metabolic heterogeneity, its role in maintaining homeostasis and how it contributes to disease progression and therapeutic resistance. By analysing cell–cell heterogeneity and metabolic diversity across human

populations, SCM and PSM are poised to offer crucial insights into how metabolic variability influences health outcomes.

Shared challenges

SCM and PSM share several common challenges that can be addressed synergistically. Metabolite identification in untargeted analyses remains a major challenge for both SCM and MS-based PSM, with distinct but related obstacles. In SCM, the limited cellular material often precludes tandem MS (MS/MS)-based structural analysis. In PSM, although sample quantity is not a constraint, the need for rapid data acquisition limits the number of MS/MS scans per sample. As a result, both fields can strongly benefit from ultra-fast, ultra-sensitive MS platforms and advanced computational approaches that are capable of handling high noise levels and inferring metabolite identities with minimal or no MS/MS data. In such scenarios, it is becoming increasingly common to use identification methods that retain some ambiguity to balance sensitivity and throughput. For example, an ion with m/z 259.0224 in negative-ion mode can be identified with high confidence as a $[M-H]^-$ ion of a hexose phosphate, with possible isomers reported, yet without resolving whether it corresponds to glucose-6-phosphate, glucose-1-phosphate or fructose-6-phosphate. This breaks with the purist view that only molecules with fully elucidated structures should be considered in the analysis, but allows rapid and automated annotation. The ambiguity in metabolite assignments can and should be explicitly reported and can be accounted for in downstream analyses⁸⁴. Furthermore, we emphasize that fair reporting of ambiguous features with partial structural or class elucidation promotes the interoperability of metabolomics data acquired with different technologies or instruments. Going forward, similarly to proteomics⁸⁵, both SCM and PSM can benefit from representing data and results in a knowledge graph that contains information about studies, samples, experiments and detected metabolites, and linking them to existing ontologies, databases, other molecular entities and literature. This can facilitate the interpretation of identified metabolites in light of biological, molecular and clinical information.

Both SCM and PSM, when performed in an untargeted manner, generate large datasets that encompass a wide range of metabolites, including some that are not catalogued in existing databases. This creates opportunities for the discovery of novel metabolites, even in well-studied systems, as exemplified by the discovery of itaconate in human cells⁸⁶. PSM, with its additional complexity arising from microbial metabolites, diet and the exposome, holds even greater potential for uncovering previously unknown metabolic products⁸⁷. These discoveries could help address the challenge in metabolomics of ‘dark matter’—the vast proportion of detected features lacking structural annotations. Recently, a debate arose around the claim that MS metabolomics dark matter can, by and large, be accounted for by unwanted MS artefacts, such as unwanted in-source fragmentation (ISF). This led to a response from a large collective of authors, who emphasized the limited role of ISF and the potential for discovery hidden in the ‘dark metabolome’⁸⁸. The advances in machine learning and computational metabolite annotation are expected to considerably enhance the identification of metabolites, including compounds that have not been observed before, or that are not available as pure standards^{89,90}. This massively increases the chemical space that can be tested⁹¹ and has pronounced effects on the confidence of annotation. Moving forward, it will become essential to report not only which structures are compatible with the measured data, but also which isomers can be excluded.

Metabolite quantification is essential for cross-study comparability and for using metabolomics modelling approaches, but it remains a challenge in MS-based metabolomics. Both SCM and PSM face similar obstacles, including variations in ionization efficiency, matrix effects and limited dynamic range. In PSM, quantification can be achieved by

spiking samples with stable-isotope-labelled metabolites at known concentrations. Machine-learning methods can reduce artefacts and make data more comparable across a large number of samples⁷¹. In SCM, tailored strategies are required depending on the ionization method: for ESI-based SCM, different approaches are required depending on whether online or offline lysis is used⁹², whereas for MALDI-based SCM, quantitation approaches that are used for spatial metabolomics can be adopted. Machine learning can potentially enhance stable-isotope-labelled-based quantitation by extrapolating results from a few labelled metabolites towards a larger set of metabolites. Advancements in quantification strategies in either SCM or PSM could benefit the other field, highlighting the potential for cross-disciplinary innovation.

Biological, cellular and functional interpretation of metabolomics data remains challenging owing to our limited knowledge of metabolite functions. By using flux analysis technologies^{93–95}, considerable progress has been achieved in understanding metabolic pathways⁹⁶ and charting the space of possible metabolic reactions, combining them into metabolic networks⁹⁷ and simulating fluxes. However, the regulatory and signalling roles of metabolites in how they affect gene expression³⁷, enzyme activity⁹⁸ and other cellular programs remain only fractionally mapped³. MS-based approaches have emerged for systematic mapping of protein–metabolite interactions^{99–101} that, combined with knowledge about metabolic networks, principles of flux–balance analysis and kinetics, could transform our understanding of metabolism. Such protein–metabolite interactions and other metabolite functions need to not only be mapped but also systematically stored in public databases, which are currently lacking despite recent efforts¹⁰². At present, even for well-characterized metabolites, functional descriptions are presented mainly in scientific publications, rather than in a standardized format. This gap necessitates the development of experimental approaches, computational tools, standardized resources and knowledge graphs¹⁰³.

Future outlook

A primary goal for the coming years is to chart human metabolism both at cellular resolution and at the population scale. Reflecting on the impact of single-cell transcriptomic atlases^{104,105}, we suggest that a cross-scale metabolic atlas should be developed as an integrative resource that combines data from both SCM and PSM (Fig. 2). Such an atlas would link cellular metabolic states to systemic variation across individuals, providing a reference for comparing cell types, tissues and individuals.

Building this atlas will require systematic data collection: large-scale PSM to capture metabolic signatures of disease risk, ageing and lifestyle factors; and SCM to resolve how these signatures arise in specific cell types, using paired PSM–SCM samples such as blood and PBMCs, and, when available, tissues obtained from donation or surgery. Realizing this goal will depend on advances in ultra-sensitive, high-throughput MS and on computational methods that can integrate heterogeneous data. To maximize its utility, the atlas should provide queryable and interoperable information, enabling integration across studies, technologies and clinical contexts. Ultimately, its value will lie both in capturing metabolic diversity across cells and individuals and in revealing invariant features of metabolism—such as conserved states, recurrent programs across tissues and robust signatures of disease and therapy response—that provide generalizable biological knowledge.

The large-scale and physiologically diverse data in such an atlas would also provide a unique substrate for developing foundation models—general-purpose machine-learning models trained on massive datasets that capture reusable representations for analysis and prediction^{106,107}. Trained directly on raw MS data¹⁰⁸, these models could advance technical tasks such as data processing and metabolite identification. Trained on metabolic profiles enriched with cellular or clinical context, they could enable a deeper understanding of metabolism, discovery of novel

Q17

metabolic states and characterization of cell-type-specific reprogramming in disease.

Building on this, the emerging concept of virtual cells—artificial intelligence (AI)-driven representational and predictive models of cellular function¹⁰⁹—naturally extends from the atlas and foundation models. By incorporating the fine-grained resolution of SCM and the systemic breadth of PSM within the atlas, virtual cells could integrate both genetic and non-genetic influences on metabolism. The scalability and affordability of metabolomics, compared with other omics, might be a decisive advantage, enabling the generation of large, diverse training datasets including large-scale perturbations for causal inference and validation.

The singular advances and convergence of SCM and PSM will enable integrative resources and AI models that connect cellular biology with human health. A cross-scale atlas of metabolism will enable generalizable insights into metabolism, support predictive AI and help realize the vision of virtual cells, revealing the principles that govern life across scales.

1. Johnson, C. H., Ivanisevic, J. & Siuzdak, G. Metabolomics: beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* **17**, 451–459 (2016).
2. Wellen, K. E. & Thompson, C. B. A two-way street: reciprocal regulation of metabolism and signalling. *Nat. Rev. Mol. Cell Biol.* **13**, 270–276 (2012).
3. Baker, S. A. & Rutter, J. Metabolites as signalling molecules. *Nat. Rev. Mol. Cell Biol.* **24**, 355–374 (2023).
4. Zamboni, N., Saghatelian, A. & Patti, G. J. Defining the metabolome: size, flux, and regulation. *Mol. Cell* **58**, 699–706 (2015).
5. Jang, C., Chen, L. & Rabinowitz, J. D. Metabolomics and isotope tracing. *Cell* **173**, 822–837 (2018).
6. Blozies, C. S. & Fiehn, O. Using untargeted metabolomics for detecting exposome compounds. *Curr. Opin. Toxicol.* **8**, 87–92 (2018).
7. Wishart, D. S. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discov.* **15**, 473–484 (2016).
8. Saigusa, D., Matsukawa, N., Hishinuma, E. & Koshiba, S. Identification of biomarkers to diagnose diseases and find adverse drug reactions by metabolomics. *Drug Metab. Pharmacokinet.* **37**, 100373 (2021).
9. Pang, H. & Hu, Z. Metabolomics in drug research and development: the recent advances in technologies and applications. *Acta Pharm. Sin.* **B 13**, 3238–3251 (2023).
10. Kirwan, J. A. Translating metabolomics into clinical practice. *Nat. Rev. Bioeng.* **1**, 228–229 (2023).
11. Schuhknecht, L. et al. A human metabolic map of pharmacological perturbations reveals drug modes of action. *Nat. Biotechnol.* **43**, 1996–2008 (2025).
- A large-scale study focusing on how metabolomics can reveal the metabolic mode of action of drugs.**
12. Ali, A. et al. Single-cell metabolomics by mass spectrometry: advances, challenges, and future applications. *Trends Anal. Chem.* **120**, 115436 (2019).
13. Saunders, K. D. G., Lewis, H.-M., Beste, D. J. V., Cexus, O. & Bailey, M. J. Spatial single cell metabolomics: current challenges and future developments. *Curr. Opin. Chem. Biol.* **75**, 102327 (2023).
14. Petrova, B. & Guler, A. T. Recent developments in single-cell metabolomics by mass Spectrometry—A perspective. *J. Proteome Res.* **24**, 1493–1518 (2025).
15. Hajjar, G. et al. Scaling-up metabolomics: current state and perspectives. *Trends Anal. Chem.* **167**, 117225 (2023).
16. Plekhova, V., De Windt, K., De Spiegeleer, M., De Graeve, M. & Vanhaecke, L. Recent advances in high-throughput biofluid metabolotyping by direct infusion and ambient ionization mass spectrometry. *Trends Anal. Chem.* **168**, 117287 (2023).
17. Young, R. S. E. et al. Subcellular mass spectrometry imaging of lipids and nucleotides using transmission geometry ambient laser desorption and plasma ionisation. *Nat. Commun.* **16**, 9130 (2025).
18. Castro, D. C., Chan-Andersen, P., Romanova, E. V. & Sweedler, J. V. Probe-based mass spectrometry approaches for single-cell and single-organelle measurements. *Mass Spectrom. Rev.* **43**, 888–912 (2024).
19. Cao, J. et al. Deciphering the metabolic heterogeneity of hematopoietic stem cells with single-cell resolution. *Cell Metab.* **36**, 209–221 (2024).
20. Rappez, L. et al. SpaceM reveals metabolic states of single cells. *Nat. Methods* **18**, 799–805 (2021).
21. Capolupo, L. et al. Sphingolipids control dermal fibroblast heterogeneity. *Science* **376**, eabh1623 (2022).
22. Wang, Z. et al. Integrative single-cell metabolomics and phenotypic profiling reveals metabolic heterogeneity of cellular oxidation and senescence. *Nat. Commun.* **16**, 2740 (2025).
23. Cairns, J. L. et al. Mass-guided single-cell MALDI imaging of low-mass metabolites reveals cellular activation markers. *Adv. Sci.* **12**, e2410506 (2025).
24. Hu, T. et al. Single-cell spatial metabolomics with cell-type specific protein profiling for tissue systems biology. *Nat. Commun.* **14**, 8260 (2023).
- An integrative approach highlighting the presence of unique metabolic cell states within cell types.**
25. Nunes, J. B. et al. Integration of mass cytometry and mass spectrometry imaging for spatially resolved single-cell metabolic profiling. *Nat. Methods* **21**, 1796–1800 (2024).
- A study demonstrating single-cell metabolic heterogeneity in tissue for cancer cells and immune cells.**

26. Mao, X. et al. Single-cell simultaneous metabolome and transcriptome profiling revealing metabolite-gene correlation network. *Adv. Sci.* **12**, e2411276 (2025).
27. Kang, M. et al. Single-cell metabolome and RNA-seq multiplexing on single plant cells. *Proc. Natl Acad. Sci. USA* **122**, e2512828122 (2025).
28. Samarah, L. Z. et al. Spatial metabolic gradients in the liver and small intestine. *Nature* **648**, 182–190 (2025).
29. Christoff, H. et al. Metabolic heterogeneity in humans. *Cell* **187**, 3821–3823 (2024).
30. Kim, J. & DeBerardinis, R. J. Mechanisms and implications of metabolic heterogeneity in cancer. *Cell Metab.* **30**, 434–446 (2019).
31. Demicco, M., Liu, X.-Z., Leithner, K. & Fendt, S.-M. Metabolic heterogeneity in cancer. *Nat. Metab.* **6**, 18–38 (2024).
32. Ghosh-Choudhary, S., Liu, J. & Finkel, T. Metabolic regulation of cell fate and function. *Trends Cell Biol.* **30**, 201–212 (2020).
33. Stegen, S. & Carmeliet, G. Metabolic regulation of skeletal cell fate and function. *Nat. Rev. Endocrinol.* **20**, 399–413 (2024).
34. Zhang, H. et al. Mass spectrometry imaging for spatially resolved multi-omics molecular mapping. *Npj Imaging* **2**, 20 (2024).
35. Vicari, M. et al. Spatial multimodal analysis of transcriptomes and metabolomes in tissues. *Nat. Biotechnol.* **42**, 1046–1050 (2024).
- First paper demonstrating the feasibility of detecting both transcriptomes and metabolomes from the same tissue section.**
36. Colwell, N., Chen, D. & Yang, Z. Achieving single-cell resolution via desorption electrospray ionization mass spectrometry imaging (DESI-MSI). Preprint at *ChemRxiv* <https://doi.org/10.26434/chemrxiv-2024-826pr> (2024).
37. Kavita, K. & Breaker, R. R. Discovering riboswitches: the past and the future. *Trends Biochem. Sci.* **48**, 119–141 (2023).
38. Trefny, M. P., Kroemer, G., Zitvogel, L. & Kobold, S. Metabolites as agents and targets for cancer immunotherapy. *Nat. Rev. Drug Discov.* **24**, 764–784 (2025).
39. Kawalekar, O. U. et al. Distinct signaling of coreceptors regulates specific metabolism pathways and impacts memory development in CAR T cells. *Immunity* **44**, 380–390 (2016).
40. Hutton, A. & Meyer, J. G. Trajectory inference for single cell omics. Preprint at <https://doi.org/10.48550/arXiv.2502.09354> (2025).
41. Jost, P. J., Weindl, D., Wunderling, K., Thiele, C. & Hasenauer, J. Pseudo-time trajectory of single-cell lipidomics: suggestion for experimental setup and computational analysis. Preprint at *bioRxiv* <https://doi.org/10.1101/2025.04.11.648323> (2025).
42. Zhang, Y. et al. Dynamic single-cell metabolomics reveals cell-cell interaction between tumor cells and macrophages. *Nat. Commun.* **16**, 4582 (2025).
43. Buglakova, E. et al. Spatial single-cell isotope tracing reveals heterogeneity of de novo fatty acid synthesis in cancer. *Nat. Metab.* **6**, 1695–1711 (2024).
44. Wang, G. et al. Analyzing cell-type-specific dynamics of metabolism in kidney repair. *Nat. Metab.* **4**, 1109–1118 (2022).
45. Lysiotis, C. A. & Kimmelman, A. C. Metabolic interactions in the tumor microenvironment. *Trends Cell Biol.* **27**, 863–875 (2017).
46. Rodriguez-Colman, M. J. et al. Interplay between metabolic identities in the intestinal crypt supports stem cell function. *Nature* **543**, 424–427 (2017).
47. Hui, S. et al. Glucose feeds the TCA cycle via circulating lactate. *Nature* **551**, 115–118 (2017).
48. Di Virgilio, F. & Adinolfi, E. Extracellular purines, purinergic receptors and tumor growth. *Oncogene* **36**, 293–303 (2017).
49. Loo, J. M. et al. Extracellular metabolic energetics can promote cancer progression. *Cell* **160**, 393–406 (2015).
50. Bamford, S. E. et al. High resolution imaging and analysis of extracellular vesicles using mass spectral imaging and machine learning. *J. Extracell. Biol.* **2**, e110 (2023).
51. Maugrion, E. et al. Extracellular vesicles contribute to the difference in lipid composition between ovarian follicles of different size revealed by mass spectrometry imaging. *Metabolites* **13**, 1001 (2023).
52. Van de Sande, B. et al. Applications of single-cell RNA sequencing in drug discovery and development. *Nat. Rev. Drug Discov.* **22**, 496–520 (2023).
53. Hansen, J. et al. A reference tissue atlas for the human kidney. *Sci. Adv.* **8**, eabn4965 (2022).
54. Delafiori, J. et al. HT SpaceM: a high-throughput and reproducible method for small-molecule single-cell metabolomics. *Cell* **188**, 6028–6043 (2025).
- Paper demonstrating the detection of small-molecule metabolites from single cells at high throughput.**
55. Molenaar, M. R. et al. Increasing quantitation in spatial single-cell metabolomics by using fluorescence as ground truth. *Front. Mol. Biosci.* **9**, 1021889 (2022).
56. Luecken, M. D. et al. Benchmarking atlas-level data integration in single-cell genomics. *Nat. Methods* **19**, 41–50 (2021).
57. Yu, B. et al. The Consortium of Metabolomics Studies (COMETS): metabolomics in 47 prospective cohort studies. *Am. J. Epidemiol.* **188**, 991–1012 (2019).
58. Begzati, A. et al. Plasma lipid metabolites, clinical glycemic predictors, and incident type 2 diabetes. *Diabetes Care* **48**, 473–480 (2025).
59. Nightingale Health Biobank Collaborative Group. Metabolomic and genomic prediction of common diseases in 700,217 participants in three national biobanks. *Nat. Commun.* **15**, 10092 (2024).
- The largest-scale metabolomics study so far, which, moreover, shows that metabolomics is more predictive than a polygenic score in predicting disease outcomes, even for non-metabolic diseases.**
60. Zhang, S. et al. A metabolomic profile of biological aging in 250,341 individuals from the UK Biobank. *Nat. Commun.* **15**, 8081 (2024).
61. Wang, N. et al. Genetic architecture and analysis practices of circulating metabolites in the NHLBI Trans-Omics for Precision Medicine Program. *Am. J. Hum. Genet.* **122**, 2720–2738 (2025).
62. Chu, X. et al. Integration of metabolomics, genomics, and immune phenotypes reveals the causal roles of metabolites in disease. *Genome Biol.* **22**, 198 (2021).
63. Bossi, E. et al. Revolutionizing blood collection: Innovations, applications, and the potential of microsampling technologies for monitoring metabolites and lipids. *Metabolites* **14**, 46 (2024).

64. Correia, M. S. P., Othman, A. & Zamboni, N. Fast, general-purpose metabolome analysis by mixed-mode liquid chromatography-mass spectrometry. *Analyst* **150**, 4955–4961 (2025).
65. Suhre, K. et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* **477**, 54–60 (2011).
Pioneering paper demonstrating the power of combining genetics with metabolomics to reveal genetic determinants of metabolic variation in human populations.
66. Antcliffe, D. B. et al. Metabolic septic shock sub-phenotypes, stability over time and association with clinical outcome. *Intensive Care Med.* **51**, 529–541 (2025).
67. Buerge, T. et al. Metabolomic profiles predict individual multidisease outcomes. *Nat. Med.* **28**, 2309–2320 (2022).
The largest study at that time showing that metabolomics can predict multi-disease outcomes.
68. Kohler, I., Hankemeier, T., van der Graaf, P. H., Knibbe, C. A. J. & van Hasselt, J. G. C. Integrating clinical metabolomics-based biomarker discovery and clinical pharmacology to enable precision medicine. *Eur. J. Pharm. Sci.* **109**, S15–S21 (2017).
69. Dunn, W. B. et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **6**, 1060–1083 (2011).
70. Broadhurst, D. et al. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* **14**, 72 (2018).
71. Dmitrenko, A., Reid, M. & Zamboni, N. Regularized adversarial learning for normalization of multi-batch untargeted metabolomics data. *Bioinformatics* **39**, btad096 (2023).
72. Dmitrenko, A., Reid, M. & Zamboni, N. A system suitability testing platform for untargeted, high-resolution mass spectrometry. *Front. Mol. Biosci.* **9**, 1026184 (2022).
73. Delabriere, A., Warmer, P., Brennstetter, V. & Zamboni, N. SLAW: a scalable and self-optimizing processing workflow for untargeted LC–MS. *Anal. Chem.* **93**, 15024–15032 (2021).
74. Li, S., Siddiqua, A., Thapa, M., Chi, Y. & Zheng, S. Trackable and scalable LC–MS metabolomics data processing using asari. *Nat. Commun.* **14**, 4113 (2023).
75. Wang, W. et al. Cancer metabolites: promising biomarkers for cancer liquid biopsy. *Biomark. Res.* **11**, 66 (2023).
76. Murphy, R. M., Watt, M. J. & Febbraio, M. A. Metabolic communication during exercise. *Nat. Metab.* **2**, 805–816 (2020).
77. Sieber, M. H. & Spradling, A. C. The role of metabolic states in development and disease. *Curr. Opin. Genet. Dev.* **45**, 58–68 (2017).
78. Chi, H. Immunometabolism at the intersection of metabolic signaling, cell fate, and systems immunology. *Cell. Mol. Immunol.* **19**, 299–302 (2022).
79. Park, S. J., Park, M. J., Park, S., Lee, E.-S. & Lee, D. Y. Integrative metabolomics of plasma and PBMCs identifies distinctive metabolic signatures in Behçet's disease. *Arthritis Res. Ther.* **25**, 5 (2023).
80. Sen, P. et al. Metabolic alterations in immune cells associate with progression to type 1 diabetes. *Diabetologia* **63**, 1017–1031 (2020).
81. Huang, H. et al. Decoding aging clocks: new insights from metabolomics. *Cell Metab.* **37**, 34–58 (2025).
82. Kim, H.-H. & Dixit, V. D. Metabolic regulation of immunological aging. *Nat. Aging* **5**, 1425–1440 (2025).
83. Sebastiani, P. et al. Metabolite signatures of chronological age, aging, survival, and longevity. *Cell Rep.* **43**, 114913 (2024).
84. Wadie, B., Molenaar, M. R., Vieira, L. M. & Alexandrov, T. Enrichment analysis for spatial and single-cell metabolomics accounting for molecular ambiguity. *Bioinform. Adv.* **5**, vbaf100 (2025).
85. Santos, A. et al. A knowledge graph to interpret clinical proteomics data. *Nat. Biotechnol.* **40**, 692–702 (2022).
86. Cordes, T., Michelucci, A. & Hiller, K. Itaconic acid: the surprising role of an industrial compound as a mammalian antimicrobial metabolite. *Annu. Rev. Nutr.* **35**, 451–473 (2015).
87. Dorrestein, P. C., Mazmanian, S. K. & Knight, R. Finding the missing links among metabolites, microbes, and the host. *Immunity* **40**, 824–832 (2014).
88. El Abiead, Y. et al. Discovery of metabolites prevails amid in-source fragmentation. *Nat. Metab.* **7**, 435–437 (2025).
89. Heinonen, M., Shen, H., Zamboni, N. & Rousu, J. Metabolite identification and molecular fingerprint prediction through machine learning. *Bioinformatics* **28**, 2333–2341 (2012).
90. Wadie, B. et al. METASPACE-ML: context-specific metabolite annotation for imaging mass spectrometry using machine learning. *Nat. Commun.* **15**, 9110 (2024).
91. Peets, P. et al. Chemical characteristics vectors map the chemical space of natural biomes from untargeted mass spectrometry data. *J. Cheminform.* **17**, 82 (2025).
92. Heininen, J. et al. Targeted and untargeted Amine metabolite quantitation in single cells with isobaric multiplexing. *Chemistry* **30**, e202403278 (2024).
93. Orth, J. D., Thiele, I. & Palsson, B. Ø What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010).
94. Bartman, C. R., TeSlaa, T. & Rabinowitz, J. D. Quantitative flux analysis in mammals. *Nat. Metab.* **3**, 896–908 (2021).
95. Bartman, C. R., Faubert, B., Rabinowitz, J. D. & DeBerardinis, R. J. Metabolic pathway analysis using stable isotopes in patients with cancer. *Nat. Rev. Cancer* **23**, 863–878 (2023).
96. Lee, K. S., Su, X. & Huan, T. Metabolites are not genes—avoiding the misuse of pathway analysis in metabolomics. *Nat. Metab.* **7**, 858–861 (2025).
97. Badur, M. G. & Metallo, C. M. Reverse engineering the cancer metabolic network using flux analysis to understand drivers of human disease. *Metab. Eng.* **45**, 95–108 (2018).
98. Wegner, A., Meiser, J., Weindl, D. & Hiller, K. How metabolites modulate metabolic flux. *Curr. Opin. Biotechnol.* **34**, 16–22 (2015).
99. Gruber, C. H., Noor, E., Buffing, M. F. & Sauer, U. Systematic identification of allosteric effectors in *Escherichia coli* metabolism. *Proc. Natl. Acad. Sci. USA* **122**, e2423767122 (2025).
100. Hicks, K. G. et al. Protein–metabolite interactomics of carbohydrate metabolism reveal regulation of lactate dehydrogenase. *Science* **379**, 996–1003 (2023).
Large-scale demonstration of the extent of previously unknown interactions between enzymes and metabolites, even across distant pathways.
101. Piazza, I. et al. A map of protein–metabolite interactions reveals principles of chemical communication. *Cell* **172**, 358–372 (2018).
102. Farr, E. et al. MetalinksDB: a flexible and contextualizable resource of metabolite–protein interactions. *Brief. Bioinform.* **25**, bbae347 (2024).
103. Amara, A. et al. Networks and graphs discovery in metabolomics data analysis and interpretation. *Front. Mol. Biosci.* **9**, 841373 (2022).
104. Rood, J. E., Maartens, A., Hupalowska, A., Teichmann, S. A. & Regev, A. Impact of the Human Cell Atlas on medicine. *Nat. Med.* **28**, 2486–2496 (2022).
105. Quake, S. R. A decade of molecular cell atlases. *Trends Genet.* **38**, 805–810 (2022).
106. Guo, F. et al. Foundation models in bioinformatics. *Natl. Sci. Rev.* **12**, nwa028 (2025).
107. Cui, H. et al. Towards multimodal foundation models in molecular cell biology. *Nature* **640**, 623–633 (2025).
108. Bushuiev, R. et al. Self-supervised learning of molecular representations from millions of tandem mass spectra using DreaMS. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-025-02663-3> (2025).
109. Bunne, C. et al. How to build the virtual cell with artificial intelligence: priorities and opportunities. *Cell* **187**, 7045–7063 (2024).

Q26

Acknowledgements We thank Joshua Rabinowitz and Gary Siuzdak for their constructive and detailed feedback, which helped us to improve the manuscript.

Q18

Author contributions T.A. and N.Z. wrote the manuscript.

Competing interests T.A. is a co-founder, shareholder and board member of DeepCyte and Metacloud and is an author on a patent on single-cell metabolomics. N.Z. is a co-founder of General Metabolics.

Additional information

Correspondence and requests for materials should be addressed to Theodore Alexandrov or Nicola Zamboni.

Q25

Peer review information Nature thanks Gary Siuzdak and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2026